

A Macroeconomic Approach to Optimal Unemployment Insurance: Theory

Camille Landais, Pascal Michaillat, Emmanuel Saez *

August 11, 2016

Abstract

This paper develops a theory of optimal unemployment insurance (UI) in matching models. The optimal UI replacement rate is the conventional Baily-Chetty replacement rate, which solves the tradeoff between insurance and job-search incentives, plus a correction term, which is positive when an increase in UI pushes the labor market tightness toward its efficient level. In matching models most wage mechanisms do not ensure efficiency so tightness is generally inefficient. The effect of UI on tightness depends on the model: increasing UI may raise tightness by alleviating the rat race for jobs or lower tightness by increasing wages through bargaining. (JEL: E24, E32, H21)

*Camille Landais: Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, UK (email: c.landais@lse.ac.uk). Pascal Michaillat: Department of Economics, Brown University, Box B, Providence, RI 02906, USA (email: pascal_michaillat@brown.edu). Emmanuel Saez: Department of Economics, University of California–Berkeley, 530 Evans Hall #3880, Berkeley, CA 94720, USA (email: saez@econ.berkeley.edu). We thank Daron Acemoglu, George Akerlof, Varanya Chaubey, Raj Chetty, Peter Diamond, Yuriy Gorodnichenko, Philipp Kircher, Etienne Lehmann, Robert Shimer, Johannes Spinnewijn, and numerous seminar and conference participants for very helpful discussions and comments. This work was supported by the Center for Equitable Growth at the University of California–Berkeley, the British Academy, the Economic and Social Research Council [grant number ES/K008641/1], the Institute for New Economic Thinking, and the Sandler Foundation.

Unemployment insurance (UI) is a key component of social insurance in modern welfare states. The microeconomic theory of optimal UI, developed by Baily (1978) and Chetty (2006), is well understood. It is an insurance-incentive tradeoff in the presence of moral hazard. UI helps workers smooth consumption when they are unemployed, but it also increases unemployment by discouraging job search. The Baily-Chetty formula resolves this tradeoff.

But the microeconomic theory only provides a partial description of the effects of UI on unemployment because, while it accounts for workers' labor-supply behavior, it ignores firms' labor-demand behavior. For instance, UI may exert upward pressure on wages by raising the outside option of unemployed workers, thereby discouraging job creation by firms. In that case, UI increases unemployment more than in the microeconomic theory. Alternatively, the labor market may operate like a rat race: firms offer a fixed number of jobs and jobseekers queue for these jobs. A jobseeker who reduces her search effort is less likely to find a job, but by moving down the queue, she improves the job prospects of the jobseekers who move ahead of her. In that case, although UI discourages job search, UI increases unemployment less than in the microeconomic theory. In these two examples, the microeconomic theory misses important labor-demand channels through which UI affects unemployment and hence social welfare.

In this paper we develop a macroeconomic theory of optimal UI that extends the microeconomic theory by accounting for firms' labor-demand behavior. To that end we embed the Baily-Chetty model of UI into a matching model with generic production function and wage mechanism. The matching model is well suited for our purpose because it includes both workers searching for jobs and firms creating jobs.

The labor market tightness—the ratio of firms' aggregate number of vacancies to workers' aggregate job-search effort—is central to our theory. Tightness is determined in equilibrium to equalize labor demand and labor supply. Tightness is important to workers because it influences their probability of finding a job: a higher tightness implies a higher job-finding rate per unit of effort. Tightness is important to firms because it determines their recruiting costs: a higher tightness imposes that a larger share of the workforce is allocated to recruiting instead of producing.

The microeconomic theory of optimal UI considers the effect of UI on labor supply taking tightness as given; in that, it is only partial equilibrium. In contrast, our macroeconomic theory is general equilibrium because it considers the effects of UI on labor supply, labor demand, and

tightness. As UI generally affects tightness and tightness generally affects social welfare, the results of our macroeconomic theory generally differ from those of the microeconomic theory.

There is a situation in which our optimal UI replacement rate is given by the Baily-Chetty formula: when the level of tightness is efficient. In that case, tightness maximizes social welfare for a given UI. Therefore, when a change in UI leads to a variation in tightness, this variation has no first-order effect on welfare. At the same time, in the Baily-Chetty model, tightness is constant so UI has no effect on welfare through tightness. It is because UI has no effect on welfare through tightness both in the Baily-Chetty model and when tightness is efficient in our model that optimal UI follows the same principles in the two cases.

However, the level of tightness is usually not efficient in matching models, since most wage mechanisms do not ensure efficiency. When tightness is inefficiently low, increasing tightness raises social welfare; in that case, unemployment is inefficiently high and too few jobseekers find a job. When tightness is inefficiently high, reducing tightness raises social welfare; in that case, unemployment is inefficiently low and firms devote too many workers to recruiting.

And when the level of tightness is not efficient, the replacement rate given by the Baily-Chetty formula is no longer optimal. This important result has a simple intuition. Imagine for example that tightness is inefficiently low—that is, increasing tightness raises welfare. If increasing UI raises tightness, UI is more desirable than the insurance-incentive tradeoff suggests, and the optimal replacement rate is higher than the Baily-Chetty replacement rate. Conversely, if increasing UI lowers tightness, UI is less desirable than the insurance-incentive tradeoff suggests, and the optimal replacement rate is lower than the Baily-Chetty replacement rate.

Formally, we show that the optimal UI replacement rate equals the Baily-Chetty replacement rate plus a correction term which equals the effect of UI on tightness times the effect of tightness on social welfare. The correction term is positive if an increase in UI brings tightness closer to its efficient level, and negative otherwise. Hence, the optimal replacement rate is above the Baily-Chetty replacement rate when an increase in UI brings tightness closer to its efficient level, and below the Baily-Chetty replacement rate otherwise.

In matching models, increasing UI may lower or raise tightness. It may lower tightness through a *job-creation channel*: when UI increases, the outside option of unemployed workers rises, wages rise through bargaining, firms create fewer jobs, and tightness falls. Alternatively, it may raise

tightness through a *rat-race channel*. Suppose to simplify that the number of jobs is fixed. In equilibrium, job-search effort times the job-finding rate per unit of effort is equal to the number of jobs and thus fixed. By discouraging job search, an increase in UI raises the job-finding rate and therefore tightness. When the number of jobs available is somewhat limited instead of completely fixed, the same logic applies and an increase in UI raises tightness. The overall effect of UI on tightness depends on which channel dominates. In the standard model of Pissarides (2000), only the job-creation channel operates and increasing UI lowers tightness. But in the job-rationing model of Michailat (2012), only the rat-race channel operates and increasing UI raises tightness.

To facilitate empirical applications of the theory, we express our optimal UI formula with estimable statistics, as in Chetty (2006). We obtain two empirical criteria: one to evaluate whether tightness is inefficiently high or low, and another one to evaluate whether an increase in UI raises or lowers tightness. The first criterion is that tightness is inefficiently low when the value of having a job relative to being unemployed is high compared to the share of the workforce devoted to recruiting. The second criterion is that an increase in UI raises tightness when the microelasticity of unemployment with respect to UI is larger than the macroelasticity of unemployment with respect to UI. The microelasticity measures the increase in unemployment caused by an increase in UI when the reduction in job-search effort is accounted for but tightness is kept constant. The macroelasticity measures the increase in unemployment caused by an increase in UI when both the reduction in job-search effort and the equilibrium response of tightness are accounted for. The second criterion is simple to understand. Imagine for instance that increasing UI raises tightness. After an increase in UI, the job-finding rate increases, which partially offsets the increase in unemployment caused by the reduction in job-search effort; therefore, the macroelasticity is smaller than the microelasticity.

The paper is organized as follows. Section I develops a generic matching model for the analysis of UI. Section II expresses social welfare as a function of UI and tightness, and computes the derivatives of the social welfare function with respect to the two variables. These derivatives are the building blocks of the optimal UI formula derived in Section III. Section IV studies the effect of UI on tightness in three specific matching models illustrating the range of possibilities. Section V shows that our optimal UI formula continues to hold when workers use home production to partially insure themselves against unemployment and when workers suffer a nonpecuniary cost

from unemployment. Section VI concludes by discussing empirical applications of the theory.

I. A Generic Matching Model

This section develops a generic matching model for the analysis of UI. This model embeds the model of UI by Baily (1978) and Chetty (2006) into a matching model that uses the formalism from Michaillat and Saez (2015a). The model is generic in that it accommodates a broad range of labor demands, arising from diverse production functions and wage mechanisms. For simplicity we consider a static model; Landais, Michaillat and Saez (2015) present a dynamic extension of the model more adapted to quantitative analysis. Table 1 summarizes the notation.

A. The Labor Market

There is a measure 1 of identical workers and a measure 1 of identical firms. Initially, all workers are unemployed and search for a job, and all firms post vacancies to recruit workers. The matching function m determines the number of worker-firm matches formed: $l = m(e, v)$, where l is the number of workers who find a job, e is the aggregate job-search effort, and v is the aggregate number of vacancies. The function m has constant returns to scale, is differentiable and increasing in both arguments, and satisfies $m(e, v) \leq 1$.

The labor market tightness θ is defined by the ratio of aggregate vacancies to aggregate search effort: $\theta = v/e$. Since the matching function has constant returns to scale, the labor market tightness determines the probabilities to find a job and fill a vacancy. A jobseeker finds a job at a rate $f(\theta) = m(e, v)/e = m(1, \theta)$ per unit of search effort; hence, a jobseeker searching with effort e finds a job with probability $e \cdot f(\theta)$. A vacancy is filled with probability $q(\theta) = m(e, v)/v = m(1/\theta, 1) = f(\theta)/\theta$. Since the matching function is increasing in its two arguments, $f(\theta)$ is increasing in θ and $q(\theta)$ is decreasing in θ . Accordingly, when the labor market is tighter, workers are more likely to find a job but vacancies are less likely to be filled. We denote by $1 - \eta$ and $-\eta$ the elasticities of $f(\theta)$ and $q(\theta)$ with respect to θ .

Table 1: Notation

	Description	Definition
<i>Panel A. The labor market</i>		
e	Job-search effort	
v	Number of vacancies	
$m(e, v)$	Matching function	
θ	Labor market tightness	$\theta = v/e$
$f(\theta)$	Job-finding rate per unit of effort	$f(\theta) = m(1, \theta)$
$q(\theta)$	Vacancy-filling probability	$q(\theta) = f(\theta)/\theta$
$1 - \eta$	Elasticity of $f(\theta)$ wrt θ	$1 - \eta = \theta \cdot f'(\theta)/f(\theta)$
ρ	Recruiting cost	
$\tau(\theta)$	Recruiter-producer ratio	$\tau(\theta) = \rho/(q(\theta) - \rho)$
<i>Panel B. Firms</i>		
l	Number of employees	
n	Number of producers	$n = l/(1 + \tau(\theta))$
w	Real wage	
$y(n)$	Production function	
$l^d(\theta, w)$	Labor demand	Equation (1)
<i>Panel C. The unemployment insurance program</i>		
c^e	Consumption of employed workers	
c^u	Consumption of unemployed workers	
Δc	Consumption gain from work	$\Delta c = c^e - c^u$
R	Unemployment insurance replacement rate	$R = 1 - \Delta c/w$
ΔU	Utility gain from work	$\Delta U = U(c^e) - U(c^u)$
$SW(\theta, \Delta U)$	Social welfare function	Equation (6)
ε^m	Microelasticity of unemployment wrt ΔU	Equation (8)
ε^f	Discouraged-worker elasticity	Equation (9)
ϕ	Harmonic mean of $U'(c^e)$ and $U'(c^u)$	Equation (12)
ε^M	Macroelasticity of unemployment wrt ΔU	Equation (20)
ε_R^m	Microelasticity of unemployment duration wrt R	Equation (25)
<i>Panel D. Workers</i>		
$U(c)$	Utility from consumption	
$\psi(e)$	Disutility from job-search effort	
$e^s(f, \Delta U)$	Effort supply	Equation (3)
$l^s(\theta, \Delta U)$	Labor supply	Equation (4)
h	Home production	
$\lambda(h)$	Disutility from home production	
$h^s(c^u)$	Home-production supply	Equation (32)
c^h	Consumption of unemployed workers with home production	$c^h = c^u + h^s(c^u)$
ΔU^h	Utility gain from work with home production	$\Delta U^h = U(c^e) - U(c^h) + \lambda(h^s(c^u))$
z	Nonpecuniary cost of unemployment	
<i>Panel E. Equilibrium</i>		
$w(\theta, \Delta U)$	Wage mechanism	
$\theta(\Delta U)$	Equilibrium tightness	Equation (5)

B. Firms

The representative firm employs l workers paid a real wage w to produce a consumption good. The firm has two types of employees: n are producing output while $l - n$ are recruiting workers by posting vacancies. The firm's production function is $y(n)$. The function y is differentiable, increasing, and concave.

Posting a vacancy requires $\rho \in (0, 1)$ recruiters. Since recruiting l employees requires $l/q(\theta)$ vacancies, the numbers of producers and recruiters in a firm with l employees are $n = l \cdot (1 - \rho/q(\theta))$ and $l - n = l \cdot \rho/q(\theta)$. Accordingly, the firm's recruiter-producer ratio is $\tau(\theta) = (l - n)/n = \rho/(q(\theta) - \rho)$. Furthermore, the numbers of employees and producers are related by $l = (1 + \tau(\theta)) \cdot n$. Since $q(\theta) > \rho$ and $q(\theta)$ is decreasing in θ , $\tau(\theta)$ is positive and increasing in θ .¹ The recruiter-producer ratio is increasing in tightness because when tightness is higher, the probability to fill a vacancy is lower, so firms must post more vacancies and thus employ more recruiters to recruit a given number of workers. The elasticity of $\tau(\theta)$ with respect to θ is $\eta \cdot (1 + \tau(\theta))$.

The firm sells its output on a perfectly competitive market. Given θ and w , the firm chooses l to maximize profits $y(l/(1 + \tau(\theta))) - w \cdot l$. The labor demand $l^d(\theta, w)$ gives the optimal number of employees. It is implicitly defined by the first-order condition of the maximization:

$$(1) \quad y' \left(\frac{l^d(\theta, w)}{1 + \tau(\theta)} \right) = (1 + \tau(\theta)) \cdot w.$$

This equation says that a profit-maximizing firm hires producers to the point where their marginal product, $y'(n)$, equals their marginal cost, which is the wage, w , plus the recruiting cost, $\tau(\theta) \cdot w$. Since $y'(n)$ is decreasing in n and $\tau(\theta)$ is increasing in θ , $l^d(\theta, w)$ is decreasing in w and, for common production functions, in θ .² Intuitively, firms hire fewer workers when the wage is higher and when tightness is higher because in these conditions the marginal cost of hiring is higher.

¹The condition $q(\theta) > \rho$ is necessary to have a positive number of producers. It limits tightness to a range $[0, \theta^m)$, where $\theta^m = q^{-1}(\rho)$.

²The function $l^d(\theta, w)$ is decreasing in θ when the elasticity of $y'(n)$ with respect to n is in $(-1, 0)$. This condition is satisfied with the standard specification $y(n) = n^\alpha$, $\alpha \in (0, 1)$.

C. *The Unemployment Insurance Program*

The government's UI program provides employed workers with consumption c^e and unemployed workers with consumption $c^u < c^e$. UI benefits and taxes are not contingent on search effort because effort is not observable. The generosity of UI is measured by the replacement rate

$$R \equiv 1 - \frac{\Delta c}{w},$$

where $\Delta c \equiv c^e - c^u$ is the consumption gain from work. When a jobseeker finds work, she keeps a fraction $\Delta c/w = 1 - R$ of the wage and gives up a fraction R , corresponding to the UI benefits that are lost.³ The government must satisfy the budget constraint

$$(2) \quad y(n) = (1 - l) \cdot c^u + l \cdot c^e.$$

If firms' profits are equally distributed, the UI program can be implemented with a benefit b funded by an income tax t such that $c^u = \text{profits} + b$ and $c^e = \text{profits} + w - t$. If profits are unequally distributed, they can be taxed fully and rebated lump sum to implement the UI program.

D. *Workers*

Initially, workers are unemployed and search for a job with effort e . The disutility from job search is $\psi(e)$. The function ψ is differentiable, increasing, and convex. The probability of becoming employed is $e \cdot f(\theta)$ and the probability of remaining unemployed $1 - e \cdot f(\theta)$.

Workers cannot insure themselves against unemployment in any way, so they consume c^e if employed and c^u if unemployed. The utility from consumption is $U(c)$. The function U is differentiable, increasing, and concave. We denote by $\Delta U \equiv U(c^e) - U(c^u)$ the utility gain from work. The utility gain from work is higher when UI is less generous.

³Consider a UI program providing a benefit b funded by a tax t so that $\Delta c = w - t - b$. Our replacement rate is $R = (t + b)/w$. The conventional replacement rate is b/w : it ignores the tax t and is not the same as R . But in practice, the conventional replacement rate is approximately equal to R because t is small relative to b . (Indeed, $t = b \cdot (1 - l)/l$ and in practice unemployment $1 - l$ is small relative to employment l).

Given θ , c^e , and c^u , the representative worker chooses e to maximize its expected utility

$$e \cdot f(\theta) \cdot U(c^e) + (1 - e \cdot f(\theta)) \cdot U(c^u) - \psi(e).$$

The effort supply $e^s(f(\theta), \Delta U)$ gives the optimal job-search effort. It is implicitly defined by the first-order condition of the maximization:

$$(3) \quad \psi'(e^s(f(\theta), \Delta U)) = f(\theta) \cdot \Delta U.$$

This equation says that a utility-maximizing jobseeker searches to the point where the marginal cost of search, $\psi'(e)$, equals the marginal benefit of search, which is the rate at which search leads to a job, $f(\theta)$, times the utility gain from work, ΔU . Since $\psi'(e)$ is increasing in e , $e^s(f(\theta), \Delta U)$ is increasing in $f(\theta)$ and ΔU . Intuitively, jobseekers search more when the job-finding rate is higher and when UI is less generous because in these conditions the marginal benefit of search is higher.

The labor supply $l^s(\theta, \Delta U)$ gives the number of workers who find a job when they search optimally. It is defined by

$$(4) \quad l^s(\theta, \Delta U) = e^s(f(\theta), \Delta U) \cdot f(\theta).$$

Since $e^s(f(\theta), \Delta U)$ is increasing in $f(\theta)$ and ΔU , and $f(\theta)$ is increasing in θ , $l^s(\theta, \Delta U)$ is increasing in θ and ΔU . Intuitively, more workers find a job when tightness is higher because a higher tightness encourages job search and increases the job-finding rate per unit of effort; more workers find a job when UI is less generous because a lower UI encourages job search.

E. The Wage Mechanism

As in any matching model, we need to specify a wage mechanism. Common wage mechanisms include Nash bargaining or a fixed wage. Here we assume that the wage mechanism is a generic function of tightness and the utility gain from work:

$$w = w(\theta, \Delta U).$$

With this mechanism, wages depend on labor market conditions and the generosity of UI. Additionally, we show in the next section that the pair $(\theta, \Delta U)$ determines all the other variables in a feasible allocation. Hence, the wage mechanism could be any function of any variable: it is the most generic wage mechanism possible.

II. Equilibrium and Social Welfare

In this section we describe the equilibrium of the model and express social welfare in an equilibrium as a function of the generosity of UI and the labor market tightness. We compute the derivatives of the social welfare function with respect to UI and tightness. These derivatives are key building blocks of the optimal UI formula derived in Section III. To facilitate empirical applications of the theory, we follow Chetty (2006) and express the derivatives with estimable statistics.

A. Feasible Allocation and Equilibrium

Before defining an equilibrium, we introduce a concept of feasible allocation that helps separating issues of insurance and efficiency when we analyze social welfare.

A feasible allocation parameterized by a labor market tightness θ and a utility gain from work ΔU is a collection of five variables $\{e, l, n, c^e, c^u\}$ that satisfy five constraints: (i) the job-search effort maximizes workers' utility: $e = e^s(f(\theta), \Delta U)$; (ii) the matching function determines employment: $l = l^s(\theta, \Delta U)$; (iii) the recruiting cost imposes a wedge between the numbers of producers and employees: $n = l/(1 + \tau(\theta))$; (iv) the consumption levels c^e and c^u satisfy the resource constraint, given by (2); and (v) the definition of ΔU is satisfied: $U(c^e) - U(c^u) = \Delta U$.

Typically a feasible allocation is a collection of quantities satisfying a resource constraint. Here we have additional constraints because of moral hazard (constraint (i)) and because of the matching structure on the labor market (constraints (ii) and (iii)). A feasible allocation does not contain prices, but it is convenient to define a notional wage and a notional replacement rate: $w \equiv y'(n)/(1 + \tau(\theta))$ and $R \equiv 1 - (c^e - c^u)/[y'(n)/(1 + \tau(\theta))]$. Of course, in an equilibrium, notional wage and replacement rate equal actual wage and replacement rate.

An equilibrium parameterized by a utility gain from work ΔU is a collection of variables $\{e, l, n, c^e, c^u, w, \theta\}$ such that (i) $\{e, l, n, c^e, c^u\}$ is the feasible allocation parameterized by θ and

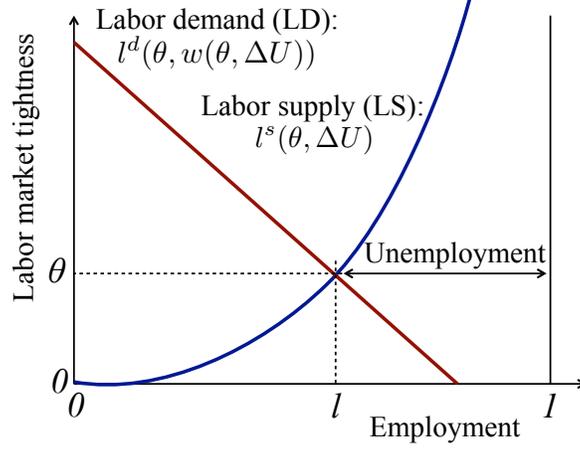


Figure 1: An Equilibrium

ΔU ; (ii) the wage is given by the wage mechanism: $w = w(\theta, \Delta U)$; and (iii) the labor market tightness equalizes labor supply and labor demand:

$$(5) \quad l^s(\theta, \Delta U) = l^d(\theta, w(\theta, \Delta U)).$$

This equation defines θ as an implicit function of ΔU , denoted $\theta(\Delta U)$. This function describes the equilibrium level of tightness for a given ΔU . All the variables in an equilibrium can be expressed as a function of ΔU and $\theta(\Delta U)$.

Figure 1 displays an equilibrium in a (l, θ) plane. The intersection of the labor supply and demand curves gives labor market tightness, employment, and unemployment. The labor supply curve shifts inward when UI increases. The labor demand curve responds to UI if wages do: when UI increases, the labor demand curve shifts inward if the wage mechanism is increasing in UI and it shifts outward if the wage mechanism is decreasing in UI.

B. The Social Welfare Function

DEFINITION 1. The social welfare function SW is defined by

$$(6) \quad SW(\theta, \Delta U) = e^s(f(\theta), \Delta U) \cdot f(\theta) \cdot \Delta U + U(c^u(\theta, \Delta U)) - \psi(e^s(f(\theta), \Delta U)),$$

where $c^u(\theta, \Delta U)$ is implicitly defined by

$$(7) \quad y \left(\frac{l^s(\theta, \Delta U)}{1 + \tau(\theta)} \right) = l^s(\theta, \Delta U) \cdot U^{-1}(U(c^u(\theta, \Delta U)) + \Delta U) + (1 - l^s(\theta, \Delta U)) \cdot c^u(\theta, \Delta U).$$

The function SW gives the social welfare in a feasible allocation parameterized by θ and ΔU . The consumption level $c^u(\theta, \Delta U)$ in (6) ensures that the government's budget constraint is satisfied. The term $U^{-1}(U(c^u(\theta, \Delta U)) + \Delta U)$ in (7) gives the consumption of employed workers when unemployed workers consume $c^u(\theta, \Delta U)$ and the utility gain from work is ΔU . The function SW plays a central role in the analysis because it allows us to compute the social welfare in an equilibrium parameterized by ΔU : in that equilibrium, the social welfare is $SW(\theta(\Delta U), \Delta U)$, where $\theta(\Delta U)$ is the equilibrium level of tightness.

To facilitate the analysis of the social welfare function, we define two elasticities that measure the response of job-search effort to UI and labor market conditions.

DEFINITION 2. *The microelasticity of unemployment with respect to UI is*

$$(8) \quad \varepsilon^m \equiv - \frac{\Delta U}{1-l} \cdot \frac{\partial(1-l^s)}{\partial \Delta U} \Big|_{\theta} = \frac{\Delta U}{1-l} \cdot \frac{\partial l^s}{\partial \Delta U} \Big|_{\theta}.$$

The microelasticity measures the percentage increase in unemployment when the utility gain from work decreases by 1 percent, the change in jobseekers' search effort is accounted for, but the equilibrium adjustment of tightness is ignored. Jobseekers reduce their search effort when UI increases so $\varepsilon^m > 0$. Because it keeps tightness constant, the microelasticity measures a partial-equilibrium response of unemployment to UI. In Figure 2, an increase in UI reduces search effort, which shifts the labor supply curve in; the microelasticity measures this shift. The ideal way to estimate the microelasticity is to offer higher or longer UI benefits to a randomly selected, small subset of jobseekers within a labor market and compare unemployment durations between treated and nontreated jobseekers.

DEFINITION 3. *The discouraged-worker elasticity is*

$$(9) \quad \varepsilon^f = \frac{f(\theta)}{e} \cdot \frac{\partial e^s}{\partial f(\theta)} \Big|_{\Delta U}.$$

The discouraged-worker elasticity measures the percentage decrease in job-search effort when

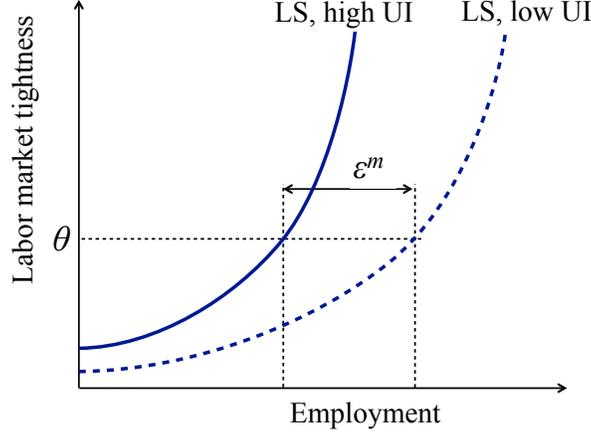


Figure 2: The Microelasticity of Unemployment With Respect to UI (ϵ^m)

the job-finding rate per unit of effort decreases by 1 percent, keeping UI constant. Jobseekers search less when the job-finding rate decreases so $\epsilon^f > 0$. The discouraged-worker elasticity can be estimated by comparing the search effort of jobseekers facing different labor market conditions but receiving similar UI. Search effort can be measured directly using time-use surveys or indirectly using the number of job-application methods reported in household surveys.

Equipped with the elasticities ϵ^m and ϵ^f , we differentiate the social welfare function:

LEMMA 1. *The social welfare function admits the following partial derivatives:*

$$(10) \quad \left. \frac{\partial SW}{\partial \theta} \right|_{\Delta U} = \frac{l}{\theta} \cdot (1 - \eta) \cdot \phi \cdot w \cdot \left[\frac{\Delta U}{\phi \cdot w} + R \cdot (1 + \epsilon^f) - \frac{\eta}{1 - \eta} \cdot \tau(\theta) \right]$$

$$(11) \quad \left. \frac{\partial SW}{\partial \Delta U} \right|_{\theta} = (1 - l) \cdot \frac{\phi \cdot w}{\Delta U} \cdot \epsilon^m \cdot \left[R - \frac{l}{\epsilon^m} \cdot \frac{\Delta U}{w} \cdot \left(\frac{1}{U'(c^e)} - \frac{1}{U'(c^u)} \right) \right],$$

where ϕ is the harmonic mean of workers' marginal consumption utilities:

$$(12) \quad \frac{1}{\phi} = \frac{l}{U'(c^e)} + \frac{1 - l}{U'(c^u)}.$$

Proof. We first compute (10). Since workers choose job-search effort to maximize expected utility, the envelope theorem says that changes in the effort $e^s(f(\theta), \Delta U)$ resulting from changes in θ have no impact on social welfare. The effect of θ on welfare therefore is

$$(13) \quad \frac{\partial SW}{\partial \theta} = \frac{l}{\theta} \cdot (1 - \eta) \cdot \Delta U + U'(c^u) \cdot \frac{\partial c^u}{\partial \theta}.$$

The first right-hand-side term is obtained by noting that $1 - \eta = \theta \cdot f'(\theta)/f(\theta)$ so $e \cdot f'(\theta) = (l/\theta) \cdot (1 - \eta)$. This term is the welfare gain created by the employment gain following an increase in tightness. Higher employment is beneficial to welfare because it allows more workers to enjoy the high level of consumption c^e instead of the low level c^u . As we apply the envelope theorem, this term only accounts for the change in employment resulting from a change in job-finding rate, not for that resulting from a change in job-search effort. The second right-hand-side term is the welfare change caused by the adjustment of unemployed workers' consumption following a change in tightness. This consumption adjustment is required to satisfy the government's budget constraint.

Next, we compute the consumption change $\partial c^u/\partial \theta$. To do that, we implicitly differentiate $c^u(\theta, \Delta U)$ with respect to θ in equation (7). We need a few preliminary result. First, the definition of the notional wage imposes $y'(n)/(1 + \tau(\theta)) = w$. Second, since $l^s(\theta, \Delta U) = e^s(f(\theta), \Delta U) \cdot f(\theta)$, the elasticity of $e^s(f(\theta), \Delta U)$ with respect to $f(\theta)$ is ε^f , and the elasticity of $f(\theta)$ with respect to θ is $1 - \eta$, the elasticity of $l^s(\theta, \Delta U)$ with respect to θ is

$$(14) \quad \frac{\theta}{l} \cdot \frac{\partial l^s}{\partial \theta} \Big|_{\Delta U} = (1 - \eta) \cdot (1 + \varepsilon^f).$$

Third, the definition of ΔU implies that $U^{-1}(U(c^u(\theta, \Delta U)) + \Delta U) - c^u = \Delta c$. Fourth, the elasticity of $1 + \tau(\theta)$ with respect to θ is $\eta \cdot \tau(\theta)$ so the derivative of $1/(1 + \tau(\theta))$ with respect to θ is $-\eta \cdot \tau(\theta)/[\theta \cdot (1 + \tau(\theta))]$. Fifth, the derivative of $c^e(c^u, \Delta U) \equiv U^{-1}(U(c^u) + \Delta U)$ with respect to c^u is $\partial c^e/\partial c^u = U'(c^u)/U'(c^e)$. Hence, the implicit differentiation of (7) with respect to θ gives

$$(15) \quad \frac{l}{\theta} \cdot (1 - \eta) \cdot (1 + \varepsilon^f) \cdot (w - \Delta c) - \frac{l}{\theta} \cdot \eta \cdot \tau(\theta) \cdot w = \left(\frac{l}{U'(c^e)} + \frac{1 - l}{U'(c^u)} \right) \cdot U'(c^u) \cdot \frac{\partial c^u}{\partial \theta}.$$

The first left-hand-side term is the budgetary gain coming from the new jobs created by an increase in tightness. Each new job increases government revenue by $w - \Delta c$. The new jobs result from a higher job-finding rate and a higher job-search effort; the term $(1 + \varepsilon^f)$ captures the combination of these two forces. The second left-hand-side term is the loss of resources due to a higher tightness. Indeed, a higher tightness forces firms to allocate more workers to recruiting and fewer to producing. The entire left-hand side is the change in the resources available to fund the UI program after a change in tightness. This change dictates the consumption change $\partial c^u/\partial \theta$.

Finally, we obtain (10) by inserting into (13) the value of $\partial c^u / \partial \theta$ obtained in (15) and introducing the variable ϕ defined by (12).

We follow similar steps to compute (11). The effect of ΔU on welfare is

$$(16) \quad \frac{\partial SW}{\partial \Delta U} = l + U'(c^u) \cdot \frac{\partial c^u}{\partial \Delta U}.$$

The first term on the right-hand side is the welfare gain enjoyed by employed workers after a reduction in UI contributions. The second term is the welfare change caused by the adjustment of unemployed workers' consumption following a reduction in UI contributions. This consumption adjustment is required to satisfy the government's budget constraint.

Next, we compute the consumption change $\partial c^u / \partial \Delta U$. To do that, we implicitly differentiate $c^u(\theta, \Delta U)$ with respect to ΔU in equation (7). We need two preliminary results in addition to those above. First, the definition of the microelasticity implies that $\partial l^s / \partial \Delta U = [(1-l)/\Delta U] \cdot \varepsilon^m$. Second, the derivative of $c^e(c^u, \Delta U) = U^{-1}(U(c^u) + \Delta U)$ with respect to ΔU is $\partial c^e / \partial \Delta U = 1/U'(c^e)$. Hence, the implicit differentiation of (7) with respect to ΔU gives

$$(17) \quad \frac{1-l}{\Delta U} \cdot \varepsilon^m \cdot (w - \Delta c) - \frac{l}{U'(c^e)} = \left(\frac{l}{U'(c^e)} + \frac{1-l}{U'(c^u)} \right) \cdot U'(c^u) \cdot \frac{\partial c^u}{\partial \Delta U}.$$

The first term on the left-hand side is the budgetary gain coming from the new jobs created by reducing the generosity of UI. The new jobs result from a higher job-search effort. This budgetary gain is a behavioral effect. The second term is the budgetary loss following the reduction of the UI contributions paid by employed workers. This budgetary loss is a mechanical effect. These budgetary changes determine the consumption change $\partial c^u / \partial \Delta U$.

Inserting into (16) the value of $\partial c^u / \partial \Delta U$ obtained in (17) and introducing the variable ϕ defined by (12), we obtain

$$(18) \quad \frac{\partial SW}{\partial \Delta U} = (1-l) \cdot \phi \cdot \left[\frac{w}{\Delta U} \cdot \varepsilon^m \cdot R + \frac{l}{1-l} \cdot \left(\frac{1}{\phi} - \frac{1}{U'(c^e)} \right) \right].$$

Equation (12) implies that $1/\phi - 1/U'(c^e) = -(1-l) \cdot (1/U'(c^e) - 1/U'(c^u))$. Combining this expression with (18) yields (11). □

C. A Criterion for Efficiency

It is well understood that the equilibrium is generally inefficient in matching models. The reason is that most wage mechanisms cannot ensure efficiency. The implication is that a small tightness change triggered by a small wage change generally has a first-order effect on social welfare. When the wage is inefficiently high, tightness is inefficiently low and a small increase in tightness enhances welfare. On the other hand, when the wage is inefficiently low, tightness is inefficiently high and a small increase in tightness reduces welfare. Last, when the wage and tightness are efficient, a small increase in tightness has no first-order effect on welfare.

The following proposition provides a criterion to determine whether tightness is efficient:

DEFINITION 4. *The efficiency term is*

$$(19) \quad \frac{\Delta U}{\phi \cdot w} + R \cdot (1 + \varepsilon^f) - \frac{\eta}{1 - \eta} \cdot \tau(\theta).$$

PROPOSITION 1. *Consider a feasible allocation parameterized by a utility gain from work ΔU and a tightness θ . A marginal increase in θ raises social welfare when the efficiency term is positive; it has no first-order effect on social welfare when the efficiency term is zero; and it lowers social welfare when the efficiency term is negative.*

Proof. The result directly follows from (10) because social welfare in a feasible allocation parameterized by ΔU and θ is given by $SW(\theta, \Delta U)$. □

The proposition shows that an increase in tightness has a positive effect on welfare when the value of having a job relative to being unemployed (ΔU) is high compared to the share of the workforce devoted to recruiting ($\tau(\theta)$). Conversely, an increase in tightness has a negative effect on welfare when the value of having a job relative to being unemployed is low compared to the share of the workforce devoted to recruiting. Intuitively, the efficient level of unemployment is positive because some unemployment allows firms to devote fewer workers to recruiting and more to production, thus increasing output. While some unemployment is desirable, too much unemployment is costly because it makes too many workers unproductive, thus reducing output, and because unemployed workers are worse off than employed workers. Hence, the efficient level of

unemployment and tightness balances the amount of labor devoted to recruiting with the cost of being unemployed.

Proposition 1 offers an empirical criterion for the efficiency of tightness: tightness is inefficiently low if the efficiency term is positive, efficient if the efficiency term is zero, and inefficiently high if the efficiency term is negative. Besides, (5) shows that by choosing the wage, one can manipulate the labor demand and select the equilibrium level of tightness: the higher the wage, the lower the labor demand, and the lower the tightness. Therefore, the proposition also offers a criterion for the efficiency of the wage: the wage is inefficiently high if the efficiency term is positive, efficient if the efficiency term is zero, and inefficiently low if the efficiency term is negative. In that, the proposition is closely related to the famous Hosios (1990) condition, which says that the wage is efficient when workers' bargaining power is equal to η . There is one major difference, however. The Hosios condition only applies to the standard matching model with risk-neutral workers, no unemployment insurance, Nash bargaining over wages, and a linear production function. Our criterion is much more general: it applies to matching models with risk-averse workers, unemployment insurance, any wage mechanism, and any production function.

III. The Optimal Unemployment Insurance Formula

The government chooses UI to maximize social welfare subject to the equilibrium relationship between tightness and UI. Formally, the government chooses ΔU to maximize $SW(\theta(\Delta U), \Delta U)$, where $SW(\theta, \Delta U)$ is defined by (6) and $\theta(\Delta U)$ is implicitly defined by (5). This section derives a formula characterizing the optimal replacement rate of the UI program. The formula is expressed with estimable statistics.

To obtain the formula, we need an elasticity measuring the general-equilibrium response of unemployment to UI:

DEFINITION 5. *The macroelasticity of unemployment with respect to UI is*

$$(20) \quad \varepsilon^M \equiv -\frac{\Delta U}{1-l} \cdot \frac{d(1-l)}{d\Delta U} = \frac{\Delta U}{1-l} \cdot \frac{dl}{d\Delta U}.$$

The macroelasticity measures the percentage increase in unemployment when the utility gain

from work decreases by 1 percent and both the change in jobseekers' search effort and the equilibrium adjustment of tightness are accounted for. In particular, the macroelasticity takes into account the response of wages to UI; in fact, the response of wages conditions the equilibrium adjustment of tightness. Because it accounts for the response of tightness to UI, the macroelasticity measures the general-equilibrium response of unemployment to UI.

When wages are weakly increasing in UI, $\epsilon^M > 0$ —that is, raising UI increases unemployment. This assumption is natural as higher UI improves workers' outside option. (The assumption is satisfied in the three specific models considered in Section IV.) The macroelasticity is positive under this assumption because raising UI depresses both the labor supply (by reducing search effort) and the labor demand (by increasing wages). Figure 3 illustrates the result: in both panels, employment is lower when UI is higher so the macroelasticity is positive.

Estimating the macroelasticity is inherently more difficult than estimating the microelasticity because the estimation requires exogenous variations in UI benefits across comparable labor markets, not exogenous variations in UI benefits across comparable individuals within a labor market. The ideal way to estimate the macroelasticity is to offer higher benefits to all individuals in a randomly selected subset of labor markets and compare unemployment rates between treated and nontreated labor markets.

The macroelasticity matters because the wedge between the macroelasticity and the microelasticity measures the effect of UI on tightness:

DEFINITION 6. *The elasticity wedge is $1 - \epsilon^M / \epsilon^m$. The elasticity wedge is positive if $\epsilon^M < \epsilon^m$, zero if $\epsilon^M = \epsilon^m$, and negative if $\epsilon^M > \epsilon^m$.*

PROPOSITION 2. *The elasticity wedge measures the equilibrium response of tightness to UI:*

$$(21) \quad \frac{\Delta U}{\theta} \cdot \frac{d\theta}{d\Delta U} = -\frac{1-l}{l} \cdot \frac{1}{1-\eta} \cdot \frac{\epsilon^m}{1+\epsilon^f} \cdot \left(1 - \frac{\epsilon^M}{\epsilon^m}\right).$$

The elasticity wedge is positive when tightness increases with the generosity of UI, negative when tightness decreases with the generosity of UI, and zero when tightness does not respond to UI.

Proof. Since $l = l^s(\theta, \Delta U)$, we have

$$\epsilon^M = \frac{\Delta U}{1-l} \cdot \frac{dl}{d\Delta U} = \frac{\Delta U}{1-l} \cdot \frac{\partial l^s}{\partial \Delta U} + \left[\frac{\theta}{1-l} \cdot \frac{\partial l^s}{\partial \theta} \right] \cdot \left[\frac{\Delta U}{\theta} \cdot \frac{d\theta}{d\Delta U} \right].$$

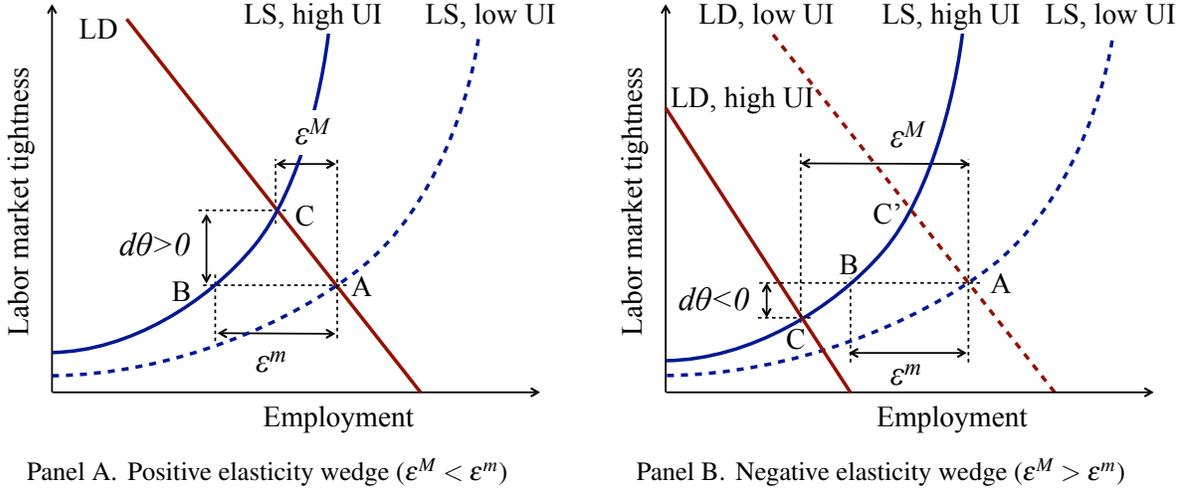


Figure 3: The Sign of the Elasticity Wedge $1 - \epsilon^M / \epsilon^m$ Gives the Effect of UI on Tightness

Notes: In Panel A, wages and thus the labor demand curve do not respond to UI. In Panel B, wages are increasing in UI so the labor demand curve shifts inward when UI increases.

Using (8) and (14), we obtain

$$(22) \quad \epsilon^M = \epsilon^m + \left[\frac{l}{1-l} \cdot (1-\eta) \cdot (1+\epsilon^f) \right] \cdot \left[\frac{\Delta U}{\theta} \cdot \frac{d\theta}{d\Delta U} \right].$$

Dividing this equation by ϵ^m and rearranging the terms yields (21). □

The proposition shows that a wedge appears between macroelasticity and microelasticity when UI affects tightness, and that this wedge has the same sign as the effect of UI on tightness. Figure 3 illustrates this result. The horizontal distance A–B measures the microelasticity and the horizontal distance A–C measures the macroelasticity. In Panel A, the labor demand curve is downward sloping, and it does not shift with a change in UI. After an increase in UI, the labor supply shifts inward (A–B) and tightness increases along the new labor supply curve (B–C). Hence, tightness rises after the increase in UI, and the macroelasticity is smaller than the microelasticity. In Panel B, the labor demand also shifts inward after an increase in UI. Tightness falls along the new labor supply curve after the shift in labor demand (C'–C). Overall, tightness may rise or fall depending on the amplitude of the labor-demand shift. In Panel B, tightness falls after the increase in UI, and the macroelasticity is larger than the microelasticity. In Section IV we will consider several specific models to explore further the channels through which UI affects tightness.

Having described the effect of UI on tightness, we can derive the optimal UI formula:

PROPOSITION 3. *The optimal UI replacement rate satisfies*

$$(23) \quad R = \frac{l}{\varepsilon^m} \frac{\Delta U}{w} \left[\frac{1}{U'(c^e)} - \frac{1}{U'(c^u)} \right] + \left[1 - \frac{\varepsilon^M}{\varepsilon^m} \right] \frac{1}{1 + \varepsilon^f} \left[\frac{\Delta U}{w\phi} + (1 + \varepsilon^f) R - \frac{\eta}{1 - \eta} \tau(\theta) \right].$$

The first term in the right-hand side is the Baily-Chetty replacement rate and the second term is the correction term.

Proof. The first-order condition of the government's problem is $0 = dSW/d\Delta U$. The total derivative of social welfare with respect to ΔU satisfies $dSW/d\Delta U = \partial SW/\partial \Delta U + (\partial SW/\partial \theta) \cdot (d\theta/d\Delta U)$. The equation $0 = \partial SW/\partial \Delta U$ is the Baily-Chetty formula. The term $\partial SW/\partial \theta$ is proportional to the efficiency term (Proposition 1). The term $d\theta/d\Delta U$ is proportional to the elasticity wedge (Proposition 2). Hence, the optimal UI formula is the Baily-Chetty formula plus a correction term proportional to the efficiency term times the elasticity wedge.

More precisely, we compute (23) by plugging into the first-order condition $0 = \partial SW/\partial \Delta U + (\partial SW/\partial \theta) \cdot (d\theta/d\Delta U)$ the expressions for $\partial SW/\partial \Delta U$, $\partial SW/\partial \theta$, and $d\theta/d\Delta U$ given by equations (11), (10), and (21). Then we divide the resulting equation by $(1 - l) \cdot \phi \cdot w \cdot \varepsilon^m / \Delta U$. \square

COROLLARY 1. *If labor market tightness is efficient, the optimal UI replacement rate satisfies the Baily-Chetty formula:*

$$(24) \quad R = \frac{l}{\varepsilon^m} \cdot \frac{\Delta U}{w} \cdot \left[\frac{1}{U'(c^e)} - \frac{1}{U'(c^u)} \right].$$

Proof. The result directly follows from Propositions 1 and 3. \square

Formula (23) shows that the optimal UI replacement rate is the Baily-Chetty replacement rate plus a correction term. The Baily-Chetty replacement rate solves the tradeoff between the need for insurance, measured by $1/U'(c^e) - 1/U'(c^u)$, and the need for incentives to search, measured by ε^m , as in the work of Baily (1978) and Chetty (2006). The correction term is the product of the effect of UI on tightness, measured by the elasticity wedge, and the effect of tightness on welfare, measured by the efficiency term. Hence, the correction term is positive when an increase in UI pushes tightness toward its efficient level.

Our formula is similar to the optimal government purchases formula derived by Michailat and Saez (2015b) in another matching model: both formulas correct a standard formula from

public economics with a term that is positive when the policy brings tightness closer to its efficient level. Our formula is also similar to numerous optimal tax formulas obtained in the presence of externalities: all these formulas have an additive structure with a standard term plus a correction term. The similarity arises because the response of tightness to UI is akin to a pecuniary externality; indeed, tightness acts as a price influencing welfare when the equilibrium is inefficient.

As in many optimal tax formulas, the right-hand side of formula (23) is endogenous to UI. Even though the formula characterizes optimal UI only implicitly, it is useful because it transparently shows the economic forces at play and because it gives the conditions under which the optimal UI replacement rate is above or below the Baily-Chetty replacement rate.

There are two situations in which the correction term is zero and the optimal UI replacement rate is given by the Baily-Chetty formula. The first situation is when UI has no effect on tightness such that the elasticity wedge is zero. In that case, our model and the Baily-Chetty model are isomorphic so they have the same optimal UI. The second situation is when tightness is efficient such that the efficiency term is zero. This is the situation described by Corollary 1. In that case, the marginal effect of UI on tightness has no first-order effect on welfare. Hence, in our model as in the Baily-Chetty model, UI has no effect on welfare through tightness, and optimal UI is governed by the same principles.

It may not be apparent that formula (24) is equivalent to the traditional Baily-Chetty formula, but the equivalence becomes clear once we introduce the microelasticity of unemployment duration with respect to the replacement rate. This elasticity is denoted ε_R^m and defined by

$$(25) \quad \varepsilon_R^m = - \left. \frac{\partial \ln(e^s \cdot f(\theta))}{\partial \ln(R)} \right|_{\theta, c^e}.$$

The elasticity ε_R^m can be estimated by measuring the change in average unemployment duration ($1/(e^s \cdot f(\theta))$) generated by changing unemployment benefits (c^u) but keeping tightness (θ) and the consumption of employed workers (c^e) constant. In the empirical literature the microelasticity ε_R^m is more frequently estimated than the microelasticity ε^m found in our formulas; but ε^m is more convenient for theoretical work.

The elasticities ε_R^m to ε^m are closely related. Consider a change dR keeping c^e and θ constant. Since $1 - l^s = s/(s + e^s \cdot f(\theta))$, we have $d \ln(1 - l^s) = -l \cdot d \ln(e^s \cdot f(\theta))$. As $\Delta c = (1 - R) \cdot w$,

we have $c^u = c^e - (1 - R) \cdot w$ and the change dR implies a change $dc^u = w \cdot dR$, which implies a change $d\Delta U = -U'(c^u) \cdot dc^u = -U'(c^u) \cdot w \cdot dR$. Since $\varepsilon^m = -\Delta U \cdot d\ln(1 - l^s)/d\Delta U$ and $\varepsilon_R^m = R \cdot d\ln(e^s \cdot f(\theta))/dR$, we obtain $\varepsilon^m = \varepsilon_R^m \cdot l \cdot \Delta U / (U'(c^u) \cdot w \cdot R)$. Using this relationship, we can rewrite (24) as the traditional Baily-Chetty formula:

$$(26) \quad \varepsilon_R^m = \left[\frac{U'(c^u)}{U'(c^e)} - 1 \right].$$

A weakness of the traditional expression of the Baily-Chetty formula is that the elasticity ε_R^m cannot be stable with R . The elasticity has to be zero when $R = 0$, so it has to increase with R for R small enough. In contrast, the elasticity ε^m found in our expression of the Baily-Chetty formula can potentially be stable with R . For example, if the disutility from job-search effort is a power function $\psi(e) = \delta \cdot e^{1+1/\kappa}$ with parameters δ and κ , then we see from equations (3) and (4) that ε^m is constant and equal to κ .

When the correction term is nonzero, the optimal UI replacement rate departs from the Baily-Chetty replacement rate. For instance, when increasing UI pushes tightness toward its efficient level, the correction term is positive and the optimal replacement rate is above the Baily-Chetty replacement rate. This happens either if tightness is inefficiently low and increasing UI raises tightness or if tightness is inefficiently high and increasing UI lowers tightness. In terms of estimable statistics, this happens if the efficiency term and elasticity wedge are both positive or both negative. Table 2 summarizes all the possibilities.

One result requires additional explanation. Consider the case in which an increase in UI raises tightness. We find that the optimal replacement rate is above the Baily-Chetty replacement rate when tightness is inefficiently low. Besides, we know that unemployment is inefficiently high when tightness is inefficiently low, and that an increase in UI raises unemployment (under the natural assumption that wages are weakly increasing in UI). Hence, our finding implies that raising unemployment through UI is desirable when unemployment is inefficiently high. How can this be the case? Imagine that UI is at the Baily-Chetty replacement rate and that tightness is inefficiently low. A small increase in UI raises insurance, which is good for welfare; it reduces job-search effort, which is bad for welfare; and it raises tightness, which is good for welfare. At the Baily-Chetty replacement rate, the cost of lower effort offsets the benefit of higher insurance; the only remaining

Table 2: Optimal UI Replacement Rate Compared to Baily-Chetty Replacement Rate

	Elasticity wedge < 0	Elasticity wedge $= 0$	Elasticity wedge > 0
Efficiency term > 0	lower	same	higher
Efficiency term $= 0$	same	same	same
Efficiency term < 0	higher	same	lower

Notes: The UI replacement rate is $R = 1 - (c^e - c^u)/w$. The Baily-Chetty replacement rate is given by (24). The efficiency term is $\Delta U / (\phi \cdot w) + (1 + \varepsilon^f) \cdot R - [\eta / (1 - \eta)] \cdot \tau(\theta)$. The elasticity wedge is $1 - \varepsilon^M / \varepsilon^m$. Compared to the Baily-Chetty replacement rate, the optimal replacement rate is higher if the correction term in formula (23) is positive, equal if the correction term is zero, and lower if the correction term is negative.

effect on welfare is the positive effect from higher tightness. Overall, as the effect of the UI reform on welfare is positive, it is optimal to increase UI above the Baily-Chetty replacement rate. At the same time, the increase in UI raises unemployment because the increase in unemployment due to lower job-search effort dominates the decrease in unemployment due to higher tightness (see Figure 3, Panel A). But because the rise in unemployment due to lower job-search effort is already internalized in the Baily-Chetty formula, it is the decrease in unemployment due to higher tightness that determines deviations from the Baily-Chetty formula. Since unemployment is inefficiently high, the decrease in unemployment pushes the optimal replacement rate above the Baily-Chetty replacement rate.

An implication of formula (23) is that even if UI is provided by private insurers, the public provision of UI is justified. Indeed, small private insurers do not internalize the effect of UI on tightness and offer insurance at the Baily-Chetty replacement rate. It is therefore optimal for the government to correct privately provided UI by a quantity equal to the correction term, which may be positive or negative.

Formula (23) reveals some interesting special cases. A first case is when UI has no adverse effect on unemployment ($\varepsilon^M = 0$). Maybe surprisingly, complete insurance is undesirable in that case. In fact, if tightness is efficient, UI is given by the Baily-Chetty formula and the magnitude of ε^M is irrelevant. It is true that UI redistributes consumption from employed to unemployed workers without destroying jobs, but because $\varepsilon^M = 0$ while $\varepsilon^m > 0$, an increase in UI raises tightness and forces firms to devote more workers to recruiting, thus reducing output available to consumption. Hence, the optimal replacement rate is below 1. To see this formally, set $\varepsilon^M = 0$ and $R = 1$ (which implies $\Delta U = 0$ and $c^e = c^u$) in formula (23). The resulting equation is

$1 = 1 - \eta \cdot \tau(\theta) / [(1 + \varepsilon^f) \cdot (1 - \eta)]$, which never holds. Hence, $R = 1$ is never optimal. Moreover, the right-hand side is always smaller than the left-hand side, so reducing R at $R = 1$ has a positive impact on welfare. Hence, the optimal R is strictly below 1.

A second case is when workers are risk-neutral ($U(c) = c$). Although there is no need for insurance, some UI may be desirable in that case. It is true that the Baily-Chetty replacement rate is 0; nevertheless, formula (23) shows that when the correction term is positive, the optimal replacement rate is positive. Furthermore, since $U(c) = c$, the efficiency term simplifies to $1 + R \cdot e^f + \eta \cdot \tau(\theta) / (1 - \eta)$, so it is easy to determine whether tightness is inefficiently high or low: tightness is inefficiently low when $\tau(\theta) < (1 - \eta) \cdot (1 + R \cdot e^f) / \eta$.

The third case is when UI has no adverse effect on job search ($\varepsilon^m = 0$). Although there is no need to provide job-search incentives, complete insurance may be undesirable in that case. It is true that the Baily-Chetty replacement rate is 1; nevertheless, formula (23) shows that when the correction term is negative, the optimal replacement rate is below 1. Since UI has no effect on job search, UI has an effect on tightness only if it affects wages; accordingly, the correction term is nonzero only if UI affects wages. Imagine for instance that an increase in UI raises wages. In that case, $\varepsilon^M > \varepsilon^m = 0$ so an increase in UI lowers tightness. This implies that the optimal replacement rate is below 1 when tightness is inefficiently low.

Last, formula (23) can be empirically implemented because it is expressed with observable variables (R , l , w) and estimable statistics. The statistics from the Baily-Chetty formula (ε^m , ΔU , $U'(c)$, ϕ) have been estimated in many studies (see Krueger and Meyer (2002) and Chetty and Finkelstein (2013) for surveys). The statistics from the matching framework (η , $\tau(\theta)$, ε^f) have also been estimated in the literature: many studies measure η (see Petrongolo and Pissarides (2001) for a survey), and a few papers estimate ε^f and $\tau(\theta)$ (for instance, Shimer (2004), DeLoach and Kurt (2013), Mukoyama, Patterson and Sahin (2014), and Gomme and Lkhagvasuren (2015) measure ε^f , while Barron, Berger and Black (1997), Silva and Toledo (2009), and Villena Roldan (2010) measure $\tau(\theta)$ or related quantities). Finally, the elasticity wedge ($1 - \varepsilon^M / \varepsilon^m$) is new to our formula, but a growing number of papers estimate it (for example, Lalive, Landais and Zweimüller 2015; Marinescu 2016; Johnston and Mas 2016). Exploiting all this empirical evidence, Landais, Michailat and Saez (2015) implement the formula for the United States.

Table 3: Three Specific Matching Models

	Model		
	Standard	Fixed-wage	Job-rationing
Production function	linear	linear	concave
Wage mechanism	bargaining	fixed	fixed
Elasticity wedge $1 - \varepsilon^M / \varepsilon^m$	negative	zero	positive
Reference	Pissarides (2000)	Hall (2005)	Michaillat (2012)

IV. The Elasticity Wedge in Three Specific Matching Models

Section III showed that the optimal generosity of UI depends on the effect of UI on tightness. It also showed that this effect is measured by the wedge between macroelasticity and microelasticity of unemployment with respect to UI. But Section III remained vague on the channels through which UI affects tightness and creates an elasticity wedge. To illustrate possible channels, this section considers three matching models with different wage mechanisms and production functions (see Table 3). The models illustrate a job-creation channel that creates a negative elasticity wedge and a rat-race channel that creates a positive elasticity wedge.

A. *The Standard Model of Pissarides (2000)*

The production function is linear: $y(n) = n$. When a worker and a firm are matched, they bargain over the wage. The worker's bargaining power is $\beta \in (0, 1)$.⁴ The surplus from the match is shared, with the worker keeping a fraction β of the surplus.⁵

We begin by determining the bargained wage. The worker's surplus from the match is ΔU . Once a worker is recruited, she produces 1 unit of good and receives a real wage w ; hence the firm's surplus from the match is $1 - w$. The total surplus from the match is $1 - w + \Delta U$. Worker and firm split the total surplus so $\Delta U = \beta \cdot (1 - w + \Delta U)$ and $1 - w = (1 - \beta) \cdot (1 - w + \Delta U)$. Thus, the wage satisfies

$$w = 1 - \frac{1 - \beta}{\beta} \cdot \Delta U.$$

⁴To obtain a positive wage, we impose $\beta / (1 - \beta) > \Delta U$.

⁵In a seminal paper, Diamond (1982) also assumes a surplus-sharing solution to the bargaining problem. If workers and firms are risk neutral, the surplus-sharing solution coincides with the generalized Nash solution. Under risk aversion, the two solutions generally differ. We use the surplus-sharing solution for simplicity.

Increasing UI lowers ΔU and thus raises wages. Intuitively, an increase in UI raises jobseekers' outside option, which enables them to bargain higher wages.

We combine the wage equation with (1) and $y'(n) = 1$ to obtain the labor demand:

$$(27) \quad \frac{\tau(\theta)}{1 + \tau(\theta)} = \frac{1 - \beta}{\beta} \cdot \Delta U.$$

The labor demand is perfectly elastic with respect to tightness. In the (l, θ) plane of Figure 4, Panel A, the labor demand curve is horizontal. Since $\tau(\theta)$ increases with θ and ΔU decreases with UI, the labor demand curve shifts downward when UI increases. Intuitively, the labor demand is depressed when UI is higher because an increase in UI pushes wages up through bargaining, which makes it less profitable to hire workers.

Having obtained the labor demand, we describe the elasticity wedge:

PROPOSITION 4. *In the standard model the elasticity wedge is negative:*

$$1 - \frac{\varepsilon^M}{\varepsilon^m} = -\frac{l}{1-l} \cdot \frac{1-\eta}{\eta} \cdot \frac{1+\varepsilon^f}{\varepsilon^m} < 0.$$

Proof. We differentiate (27) with respect to ΔU . Since the elasticities of $\tau(\theta)$ and $1 + \tau(\theta)$ with respect to θ are $\eta \cdot (1 + \tau(\theta))$ and $\eta \cdot \tau(\theta)$, we obtain $(\Delta U / \theta) \cdot (d\theta / d\Delta U) = 1/\eta$. Then, we use (21) to obtain the expression for the elasticity wedge. \square

Panel A of Figure 4 illustrates the proposition. After an increase in UI, jobseekers search less, shifting the labor supply curve inward by a distance A–B. In addition, wages rise through bargaining, shifting the labor demand curve downward and further increasing unemployment by the horizontal distance B–C. The total increase in unemployment is given by the horizontal distance A–C. Since A–C is larger than A–B, the macroelasticity is larger than the microelasticity and the elasticity wedge is negative.

The increase A–B in unemployment occurs through the moral-hazard channel: an increase in UI reduces job-search effort—a source of moral hazard because it is unobservable—and thus raises unemployment. The increase B–C in unemployment occurs through the job-creation channel, which operates as follows. An increase in UI raises wages and thus reduces the profitability of creating jobs. This reduction leads to a decline in tightness and a further increase in unemployment.

Proposition 4 implies that an increase in UI reduces tightness. Thus, in the standard model, the optimal replacement rate is below the Baily-Chetty rate when tightness is inefficiently low and above it when tightness is inefficiently high.

In the standard model, tightness is inefficiently low when workers' bargaining power, β , is inefficiently high. Namely, tightness is inefficiently low when the efficiency term is positive, or $\tau(\theta) < (1 - \eta) \cdot [\Delta U / (\phi \cdot w) + R \cdot (1 + \varepsilon^f)] / \eta$. Using equation (27), we can transform this inequality into a condition on β . We find that tightness is inefficiently low if and only if

$$(28) \quad \beta > \frac{\eta + (1 - \eta) \cdot \left[\frac{\Delta U}{\phi \cdot w} + R \cdot (1 + \varepsilon^f) \right]}{\eta + (1 - \eta) \cdot \left[\frac{\Delta U}{\phi \cdot w} + R \cdot (1 + \varepsilon^f) \right] \cdot \left(1 + \frac{1}{\Delta U} \right)}.$$

This condition generalizes the Hosios condition to situations with risk aversion and unemployment insurance.⁶ If workers are risk neutral ($U(c) = c$ so $\Delta U = \Delta c$ and $\phi = 1$) and there is no unemployment insurance ($R = 0$ and $\Delta c = w$), the bargained wage simply is $w = \beta$ and inequality (28) reduces to $\beta > \eta$, which is the traditional Hosios condition.

B. The Fixed-Wage Model of Hall (2005)

The production function is linear: $y(n) = n$. The wage is fixed: $w = \omega$, where $\omega \in (0, 1)$ is a parameter. Unlike the bargained wage from the standard model, the fixed wage does not respond to UI. We combine the wage equation with (1) and $y'(n) = 1$ to obtain the labor demand:

$$(29) \quad 1 = \omega \cdot (1 + \tau(\theta)).$$

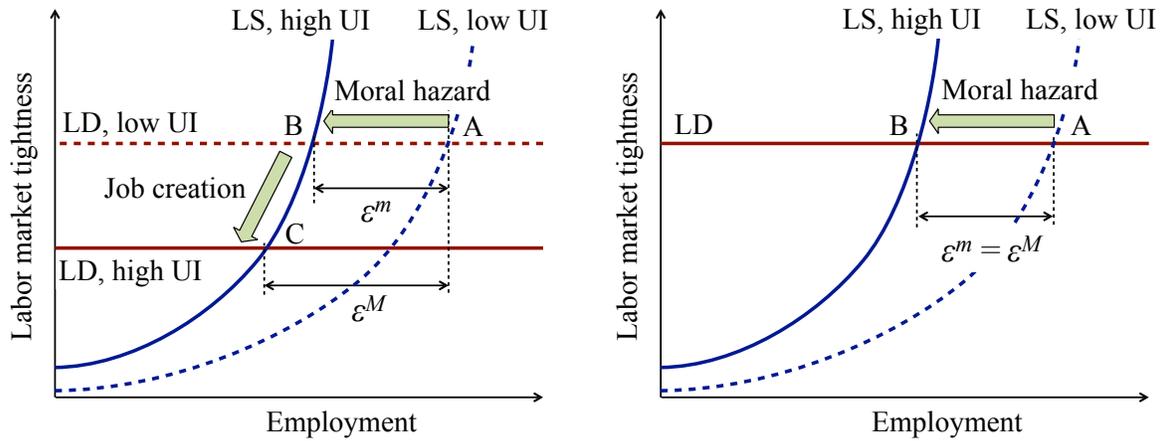
The labor demand is perfectly elastic with respect to tightness. In the (l, θ) plane of Figure 4, Panel B, the labor demand curve is horizontal. The labor demand curve is unaffected by UI because wages do not respond to UI.

Having obtained the labor demand, we describe the elasticity wedge:

PROPOSITION 5. *In the fixed-wage model the elasticity wedge is 0.*

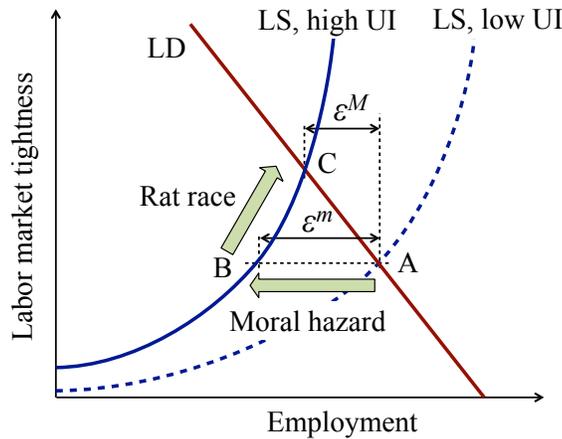
Proof. In equilibrium, θ is determined by (29). Equation (29) is independent of ΔU so $d\theta / \Delta U = 0$.

⁶Since the right-hand side of (28) depends on β and ΔU , the constraint on β imposed by (28) is only implicit.



Panel A. Negative wedge in the standard model

Panel B. Zero wedge in the fixed-wage model



Panel C. Positive wedge in the job-rationing model

Figure 4: The Elasticity Wedge $1 - \epsilon^M/\epsilon^m$ in Three Specific Matching Models

Using (21), we conclude that $\epsilon^M = \epsilon^m$. □

Panel B of Figure 4 illustrates the result. In the fixed-wage model, UI affects employment only through the moral-hazard channel; the wage does not respond to UI so the job-creation channel is eliminated. Hence, the macroelasticity equals the microelasticity and the elasticity wedge is zero.

Since macroelasticity and microelasticity are equal, UI has no effect on tightness and optimal UI is always given by the Baily-Chetty formula, even if tightness is inefficient. Basically, the fixed-wage model is isomorphic to the Baily-Chetty model as tightness is constant in both models.

C. The Job-Rationing Model of Michailat (2012)

The production function is concave: $y(n) = n^\alpha$, where $\alpha \in (0, 1)$ parameterizes decreasing marginal returns to labor. The wage is fixed: $w = \omega$, where $\omega \in (0, 1)$ is a parameter. Unlike the bargained wage from the standard model, the fixed wage does not respond to UI.

We combine the wage equation with (1) to obtain the labor demand:

$$(30) \quad l^d(\theta, \omega) = \left(\frac{\omega}{\alpha}\right)^{-\frac{1}{1-\alpha}} \cdot (1 + \tau(\theta))^{-\frac{\alpha}{1-\alpha}}.$$

Since $\tau(\theta)$ is increasing in θ , $l^d(\theta, \omega)$ is decreasing in θ . Intuitively, when tightness is higher, hiring producers requires more recruiters and is less profitable, so firms employ fewer workers. Moreover, the labor demand does not respond to UI because the wage does not. In the (l, θ) plane of Figure 4, Panel C, the labor demand curve is downward sloping and unaffected by UI.

As the labor demand is downward sloping, there is job rationing when the wage is high enough. Indeed, we have $l^d(\theta = 0, \omega) < 1$ when $\omega > \alpha$. In Figure 4, Panel C, the labor demand curve crosses the x-axis at $l < 1$ when $\omega > \alpha$. This means that jobs are rationed: firms would not hire all the jobseekers even if the jobseekers searched infinitely hard and tightness were zero.

Having characterized the labor demand, we describe the elasticity wedge:

PROPOSITION 6. *In the job-rationing model the elasticity wedge is positive:*

$$(31) \quad 1 - \frac{\varepsilon^M}{\varepsilon^m} = \frac{1}{1 + \frac{\eta}{1-\eta} \cdot \frac{\alpha}{1-\alpha} \cdot \frac{1}{1+\varepsilon^f} \cdot \tau(\theta)} > 0.$$

Proof. The elasticity of $1 + \tau(\theta)$ with respect to θ is $\eta \cdot \tau(\theta)$. Hence, (30) implies that the elasticity of $l^d(\theta, \omega)$ with respect to θ is $-\eta \cdot \tau(\theta) \cdot \alpha / (1 - \alpha)$. Since ε^M is $l / (1 - l)$ times the elasticity of l with respect to ΔU and $l = l^d(\theta, \omega)$ in equilibrium, we infer that

$$\varepsilon^M = -\frac{l}{1-l} \cdot \eta \cdot \frac{\alpha}{1-\alpha} \cdot \tau(\theta) \cdot \frac{\Delta U}{\theta} \cdot \frac{d\theta}{d\Delta U}.$$

We plug into the equation the expression of $(\Delta U / \theta) \cdot (d\theta / d\Delta U)$ given by (21) and obtain

$$\varepsilon^M = \frac{\eta}{1-\eta} \cdot \frac{\alpha}{1-\alpha} \cdot \frac{1}{1+\varepsilon^f} \cdot \tau(\theta) \cdot (\varepsilon^m - \varepsilon^M).$$

Dividing this equation by ε^m and rearranging the terms yields (31). □

Panel C of Figure 4 illustrates the proposition. After an increase in UI, jobseekers search less, shifting the labor supply curve inward by a distance A–B. Since the labor demand curve is downward sloping and does not respond to UI, the initial increase in unemployment is attenuated by a horizontal distance B–C. The total increase in unemployment is given by the horizontal distance A–C. Since A–C is smaller than A–B, the macroelasticity is smaller than the microelasticity and the elasticity wedge is positive. However, since A–C is positive, the macroelasticity remains positive.

The increase A–B in unemployment occurs through the moral-hazard channel. The reduction B–C in unemployment occurs through the rat-race channel. This channel is absent from the standard and fixed-wage models. It operates as follows. The number of jobs available is somewhat limited because of decreasing marginal returns to labor in production. When a jobseeker searches less, she reduces her probability of finding a job but mechanically increases others' probability of finding one of the few available jobs. Thus, by discouraging job search, UI alleviates the rat-race for jobs and increases the job-finding rate per unit of effort.⁷ This increase in job-finding rate leads to a reduction in unemployment that partially offsets the increase in unemployment occurring through the moral-hazard channel.

Proposition 6 implies that an increase in UI raises tightness. Thus, in the job-rationing model, the optimal replacement rate is above the Baily-Chetty rate when tightness is inefficiently low and below it when tightness is inefficiently high. In the job-rationing model, tightness is inefficiently low when the level of the fixed wage, ω , is inefficiently high. Following the same steps as in the standard model, we could derive a condition on ω for tightness to be inefficiently low.

Here the rat-race channel appears because of decreasing marginal returns to labor in production, but alternative assumptions could give rise to it. In fact, the channel operates as soon as the labor demand is downward sloping in a (l, θ) plane such that the number of jobs is limited for a given tightness. For instance, an aggregate demand on the product market as in the model by Michailat

⁷To see this formally, consider an increase in UI and assume that the job-finding rate per unit of effort, $f(\theta)$, remains constant. This assumption implies that tightness, θ , and therefore the marginal recruiting cost, $\tau(\theta)$, remain constant. As the wage, w , is constant, the marginal cost of labor, $w \cdot (1 + \tau(\theta))$, also remain constant. Simultaneously, because jobseekers search less, firms recruit fewer workers and the marginal product of labor is higher because of decreasing marginal returns to labor. Hence, firms face the same marginal cost but a higher marginal product of labor. This is suboptimal: firms could increase profits by posting more vacancies and recruiting more workers. Consequently, θ and $f(\theta)$ are higher after the increase in UI.

and Saez (2015a, Section 3) would give rise to the rat-race channel even under constant marginal returns to labor in production.

V. Robustness of the Formula

This section shows that the optimal UI formula derived in Section III is robust. The formula continues to hold when workers partially insure themselves against unemployment through home production and when workers suffer a nonpecuniary cost from unemployment.

A. *The Formula With Partial Self-Insurance Against Unemployment*

Section III assumes that workers cannot insure themselves against unemployment. In reality workers are able to partially insure themselves against unemployment using saving, spousal income, and home production (Gruber 1997; Aguiar and Hurst 2005). Here we assume that workers partially insure themselves against unemployment with home production. Home production is a convenient representation of all the means of self-insurance available to workers.

The UI program provides employed workers with consumption c^e and unemployed workers with consumption c^u . In addition to consuming c^u , unemployed workers consume an amount h produced at home at a utility cost $\lambda(h)$. The function λ is differentiable, increasing, and convex, with $\lambda(0) = 0$.

Unemployed workers choose h to maximize their utility $U(c^u + h) - \lambda(h)$. The home-production supply $h^s(c^u)$ gives the optimal level of home production. It is implicitly defined by the first-order condition of the maximization:

$$(32) \quad \lambda'(h^s(c^u)) = U'(c^u + h^s(c^u)).$$

This equation says that a utility-maximizing unemployed worker produces to the point where the marginal cost of home production, $\lambda'(h)$, equals the marginal benefit of home production, $U'(c^u + h)$. Since $\lambda'(h)$ is increasing in h and $U'(c)$ is decreasing in c , $h^s(c^u)$ is decreasing in c^u . Intuitively, unemployed workers produce less at home when UI benefits are more generous because higher benefits reduce the marginal value of home production. When home production is optimal, the

consumption of unemployed workers is $c^h = c^u + h^s(c^u)$ and the utility gain from work is $\Delta U^h = U(c^e) - U(c^h) + \lambda(h^s(c^u))$.

With home production, formula (23) carries over once the utility gain from work is adjusted from ΔU to ΔU^h and the marginal utility of unemployed workers is adjusted from $U'(c^u)$ to $U'(c^h)$. Indeed, the optimal UI replacement rate satisfies

$$R = \frac{l}{\varepsilon^m} \frac{\Delta U^h}{w} \left[\frac{1}{U'(c^e)} - \frac{1}{U'(c^h)} \right] + \left[1 - \frac{\varepsilon^M}{\varepsilon^m} \right] \frac{1}{1 + \varepsilon^f} \left[\frac{\Delta U^h}{w\phi} + (1 + \varepsilon^f) R - \frac{\eta}{1 - \eta} \tau(\theta) \right].$$

In this formula, ϕ is redefined using c^h instead of c^u and the elasticities ε^m , ε^M , and ε^f are redefined using ΔU^h instead of ΔU . The formula is derived like formula (23).⁸ In a similar way, Chetty (2006) generalizes the work of Baily (1978) to account for partial self-insurance.

Self-insurance affects both the Baily-Chetty replacement rate and the efficiency term in the formula. In the Baily-Chetty replacement rate, the value of insurance is reduced to $1/U'(c^e) - 1/U'(c^h) < 1/U'(c^e) - 1/U'(c^u)$; hence, the Baily-Chetty replacement rate is lower. In the efficiency term, the utility gain from work is reduced to $\Delta U^h = \min_h \{U(c^e) - U(c^u + h) + \lambda(h)\} < U(c^e) - U(c^u) = \Delta U$; hence, the efficiency term is lower. Accordingly, the correction term is lower if the elasticity wedge is positive but higher if the elasticity wedge is negative. Overall, with partial self-insurance, the optimal UI replacement rate is unambiguously lower if the elasticity wedge is positive, but it may be higher or lower if the elasticity wedge is negative.

⁸Repeating the derivation of (23) is simple except for two steps. First, the social welfare function admits a slightly different expression:

$$SW(\theta, \Delta U^h) = e^s(f(\theta), \Delta U^h) \cdot f(\theta) \cdot \Delta U^h + U(c^u(\theta, \Delta U^h) + h^s(c^u(\theta, \Delta U^h))) - \lambda(h^s(c^u(\theta, \Delta U^h))) - \psi(e^s(f(\theta), \Delta U^h)).$$

But, because unemployed workers choose home production to maximize their utility, the envelope theorem says that changes in home production $h^s(c^u(\theta, \Delta U^h))$ resulting from changes in θ and ΔU^h have no impact on social welfare. Therefore, (13) and (16) remain valid once ΔU and $U'(c^u)$ are replaced by ΔU^h and $U'(c^h)$. Second, since $\Delta U^h = U(c^e) - U(c^u + h^s(c^u)) + \lambda(h^s(c^u))$, the consumption of employed workers is given by

$$c^e(\theta, \Delta U^h) = U^{-1} \left(U(c^u(\theta, \Delta U^h) + h^s(c^u(\theta, \Delta U^h))) - \lambda(h^s(c^u(\theta, \Delta U^h))) + \Delta U^h \right).$$

But, because unemployed workers choose home production to maximize $U(c^u + h) - \lambda(h)$, changes in $h^s(c^u(\theta, \Delta U^h))$ resulting from changes in θ and ΔU^h have no impact on $c^e(\theta, \Delta U^h)$. Therefore, (15) and (17) remain valid once ΔU and $U'(c^u)$ are replaced by ΔU^h and $U'(c^h)$.

B. The Formula With a Nonpecuniary Cost of Unemployment

Section III assumes that the well-beings of unemployed and employed workers differ only because unemployed workers consume less. But unemployment has detrimental effects on mental and physical health that go beyond what lower consumption would induce.⁹ Some of the early studies on unemployment and health suffered from two issues. First, they were not able to separate between causality (unemployment causes low health) and selection (people who have low health become unemployed). But recent studies, such as Burgard, Brand and House (2007) and Sullivan and von Wachter (2009), are able to identify the causal effect from unemployment to low health. Second, they were not able to control for the loss of income associated with unemployment and thus separate between the pecuniary and nonpecuniary costs of unemployment. But recent studies, such as Winkelmann and Winkelmann (1998), Di Tella, MacCulloch and Oswald (2003), Blanchflower and Oswald (2004), and Helliwell and Huang (2014), find that unemployed workers report much lower well-being than employed workers even after controlling for household income and other personal characteristics. This lower well-being seems to stem from higher anxiety, lower self-esteem, and lower life satisfaction (Darity and Goldsmith 1996; Theodossiou 1998; Krueger and Mueller 2011).

Here we assume that unemployed workers have utility $U(c^u) - z$, where the parameter z captures the nonpecuniary cost of unemployment. The utility gain from work therefore is $\Delta U^z = U(c^e) - U(c^u) + z$. Given the amount of evidence that unemployment entails significant nonpecuniary costs, it is likely that $z > 0$. In theory, however, it is possible that $z < 0$. In that case, workers enjoy nonpecuniary benefits from unemployment—for instance, additional time for leisure.

When unemployed workers suffer a nonpecuniary cost z , formula (23) carries over once the utility gain from work is adjusted from ΔU to $\Delta U^z = \Delta U + z$. Indeed, the optimal UI replacement rate satisfies

$$R = \frac{l}{\varepsilon^m} \frac{\Delta U^z}{w} \left[\frac{1}{U'(c^e)} - \frac{1}{U'(c^u)} \right] + \left[1 - \frac{\varepsilon^M}{\varepsilon^m} \right] \frac{1}{1 + \varepsilon^f} \left[\frac{\Delta U^z}{w\phi} + \left(1 + \varepsilon^f \right) R - \frac{\eta}{1 - \eta} \tau(\theta) \right].$$

⁹The detrimental effects of unemployment on mental and physical health are documented by a large literature. See Dooley, Fielding and Levi (1996), Hawton and Platt (2000), Frey and Stutzer (2002), and Winkelmann (2014) for surveys and Murphy and Athanasou (1999) and McKee-Ryan et al. (2005) for meta-analyses.

In this formula, the elasticities ε^m , ε^M , and ε^f are redefined using ΔU^z instead of ΔU . The formula is obtained like formula (23).

The nonpecuniary cost of unemployment affects the efficiency term but not the Baily-Chetty replacement rate in the formula. (In the Baily-Chetty replacement rate, the ΔU^z in the numerator cancels out with the ΔU^z in the numerator of ε^m .) Hence, as already noted by Chetty (2006), the Baily-Chetty replacement rate is independent of the level of well-being of unemployment workers. Furthermore, since $\Delta U^z = \Delta U + z$, the efficiency term is higher when $z > 0$. Accordingly, when $z > 0$, the correction term is higher if the elasticity wedge is positive but lower if the elasticity wedge is negative. Overall, when $z > 0$, the optimal UI replacement rate is higher if the elasticity wedge is positive but lower if the elasticity wedge is negative.

VI. Conclusion

This paper proposes a theory of optimal UI in matching models. The optimal UI replacement rate is the sum of the conventional Baily-Chetty replacement rate, which solves the tradeoff between insurance and job-search incentives, and a correction term, which is positive when an increase in UI pushes labor market tightness toward its efficient level. Hence, the optimal replacement rate is higher than the Baily-Chetty replacement rate if tightness is inefficiently low and an increase in UI raises tightness, or if tightness is inefficiently high and an increase in UI lowers tightness.

In some countries, including the United States, the generosity of UI depends on the unemployment rate. Our formula provides guidance to link the generosity of UI to labor market conditions. It indicates that if UI has an influence on tightness and if tightness is inefficiently high or low, UI should be adjusted to bring tightness closer to its efficient level.

We propose empirical criteria to determine whether an increase in UI raises or lowers tightness and whether tightness is inefficiently high or low. Landais, Michailat and Saez (2015) apply these criteria to the United States. They find that an increase in UI raises tightness and that tightness is inefficiently low in slumps and inefficiently high in booms. Our theory combined with their empirical findings suggests that in the United States, the optimal UI replacement rate is above the Baily-Chetty replacement rate in slumps but below it in booms.

References

- Aguiar, Mark, and Erik Hurst.** 2005. "Consumption versus Expenditure." *Journal of Political Economy*, 113(5): 919–948.
- Baily, Martin N.** 1978. "Some Aspects of Optimal Unemployment Insurance." *Journal of Public Economics*, 10(3): 379–402.
- Barron, John M., Mark C. Berger, and Dan A. Black.** 1997. "Employer Search, Training, and Vacancy Duration." *Economic Inquiry*, 35(1): 167–92.
- Blanchflower, David G., and Andrew J. Oswald.** 2004. "Well-Being Over Time in Britain and the USA." *Journal of Public Economics*, 88(7–8): 1359–1386.
- Burgard, Sarah A., Jennie E. Brand, and James S. House.** 2007. "Toward a Better Estimation of the Effect of Job Loss on Health." *Journal of Health and Social Behavior*, 48(4): 369–384.
- Chetty, Raj.** 2006. "A General Formula for the Optimal Level of Social Insurance." *Journal of Public Economics*, 90(10–11): 1879–1901.
- Chetty, Raj, and Amy Finkelstein.** 2013. "Social Insurance: Connecting Theory to Data." In *Handbook of Public Economics*. Vol. 5, , ed. Alan J. Auerbach, Raj Chetty, Martin Feldstein and Emmanuel Saez, 111–193. Amsterdam:Elsevier.
- Darity, William, and Arthur H. Goldsmith.** 1996. "Social Psychology, Unemployment and Macroeconomics." *Journal of Economic Perspectives*, 10(1): 121–140.
- DeLoach, Stephen, and Mark Kurt.** 2013. "Discouraging Workers: Estimating the Impacts of Macroeconomic Shocks on the Search Intensity of the Unemployed." *Journal of Labor Research*, 34(4): 433–454.
- Diamond, Peter A.** 1982. "Wage Determination and Efficiency in Search Equilibrium." *Review of Economic Studies*, 49(2): 217–227.
- Di Tella, Rafael, Robert J. MacCulloch, and Andrew J. Oswald.** 2003. "The Macroeconomics of Happiness." *Review of Economics and Statistics*, 85(4): 809–827.
- Dooley, David, Jonathan Fielding, and Lennart Levi.** 1996. "Health and Unemployment." *Annual Review of Public Health*, 17: 449–465.
- Frey, Bruno S., and Alois Stutzer.** 2002. "What Can Economists Learn From Happiness Research?" *Journal of Economic Literature*, 40(2): 402–435.
- Gomme, Paul, and Damba Lkhagvasuren.** 2015. "Worker Search Effort as an Amplification Mechanism." *Journal of Monetary Economics*, 75: 106–122.
- Gruber, Jonathan.** 1997. "The Consumption Smoothing Benefits of Unemployment Insurance." *American Economic Review*, 87(1): 192–205.
- Hall, Robert E.** 2005. "Employment Fluctuations with Equilibrium Wage Stickiness." *American Economic Review*, 95(1): 50–65.
- Hawton, Keith, and Stephen Platt.** 2000. "Suicidal Behaviour and the Labour Market." In *The International Handbook of Suicide and Attempted Suicide*. , ed. Keith Hawton and Kees van Heeringen, 309–384. Chichester, UK:John Wiley & Sons.
- Helliwell, John F., and Haifang Huang.** 2014. "New Measures of the Costs of Unemployment: Evidence from the Subjective Well-Being of 3.3 Million Americans." *Economic Inquiry*,

52(4): 1485–1502.

- Hosios, Arthur J.** 1990. “On the Efficiency of Matching and Related Models of Search and Unemployment.” *Review of Economic Studies*, 57(2): 279–298.
- Johnston, Andrew, and Alexandre Mas.** 2016. “Potential Unemployment Insurance Duration and Labor Supply: The Individual and Market-Level Response to a Benefit Cut.” National Bureau of Economic Research (NBER) Working Paper 22411.
- Krueger, Alan B., and Andreas Mueller.** 2011. “Job Search, Emotional Well-Being, and Job Finding in a Period of Mass Unemployment: Evidence from High-Frequency Longitudinal Data.” *Brookings Papers on Economic Activity*, 42(1): 1–57.
- Krueger, Alan B., and Bruce D. Meyer.** 2002. “Labor Supply Effects of Social Insurance.” In *Handbook of Public Economics*. Vol. 4, , ed. Alan J. Auerbach and Martin Feldstein, 2327–2392. Amsterdam:Elsevier.
- Lalive, Rafael, Camille Landais, and Josef Zweimüller.** 2015. “Market Externalities of Large Unemployment Insurance Extension Programs.” *American Economic Review*, 105(12): 3564–3596.
- Landais, Camille, Pascal Michailat, and Emmanuel Saez.** 2015. “A Macroeconomic Approach to Optimal Unemployment Insurance: Empirical Applications.”
- Marinescu, Ioana.** 2016. “The General Equilibrium Impacts of Unemployment Insurance: Evidence from a Large Online Job Board.” <http://www.marinescu.eu/Marinescu.UI.pdf>.
- McKee-Ryan, Frances, Zhaoli Song, Connie R. Wanberg, and Angelo J. Kinicki.** 2005. “Psychological and Physical Well-Being During Unemployment: A Meta-Analytic Study.” *Journal of Applied Psychology*, 90(1): 53–76.
- Michailat, Pascal.** 2012. “Do Matching Frictions Explain Unemployment? Not in Bad Times.” *American Economic Review*, 102(4): 1721–1750.
- Michailat, Pascal, and Emmanuel Saez.** 2015a. “Aggregate Demand, Idle Time, and Unemployment.” *Quarterly Journal of Economics*, 130(2): 507–569.
- Michailat, Pascal, and Emmanuel Saez.** 2015b. “The Optimal Use of Government Purchases for Stabilization.” National Bureau of Economic Research (NBER) Working Paper 21322.
- Mukoyama, Toshihiko, Christina Patterson, and Aysegül Sahin.** 2014. “Job Search Behavior over the Business Cycle.” Federal Reserve Bank of New York Staff Report 689.
- Murphy, Gregory C., and James A. Athanasou.** 1999. “The Effect of Unemployment on Mental Health.” *Journal of Occupational and Organizational Psychology*, 72(1): 83–99.
- Petrongolo, Barbara, and Christopher A. Pissarides.** 2001. “Looking into the Black Box: A Survey of the Matching Function.” *Journal of Economic Literature*, 39(2): 390–431.
- Pissarides, Christopher A.** 2000. *Equilibrium Unemployment Theory*. . 2nd ed., Cambridge, MA:MIT Press.
- Shimer, Robert.** 2004. “Search Intensity.” <http://sites.google.com/site/robertshimer/intensity.pdf>.
- Silva, José I., and Manuel Toledo.** 2009. “Labor Turnover Costs and the Behavior of Vacancies and Unemployment.” *Macroeconomic Dynamics*, 13(1): 76–96.
- Sullivan, Daniel G., and Till M. von Wachter.** 2009. “Job Displacement and Mortality: An Analysis using Administrative Data.” *Quarterly Journal of Economics*, 124(3): 1265–1306.

- Theodossiou, I.** 1998. “The Effects of Low-Pay and Unemployment on Psychological Well-Being: A Logistic Regression Approach.” *Journal of Health Economics*, 17(1): 85–104.
- Villena Roldan, Benjamin.** 2010. “Aggregate Implications of Employer Search and Recruiting Selection.” Center for Applied Economics (CEA), University of Chile Working Paper 271.
- Winkelmann, Liliana, and Rainer Winkelmann.** 1998. “Why are the Unemployed So Unhappy? Evidence from Panel Data.” *Economica*, 65(257): 1–15.
- Winkelmann, Rainer.** 2014. “Unemployment and Happiness.” *IZA World of Labor*, 94: 1–10.