Optimal Unemployment Insurance Reconsidered: The Role of Fiscal Externalities

Nicholas Lawson*
Princeton University
nlawson@princeton.edu
JOB MARKET PAPER
September 17, 2012

Abstract

The literature on optimal unemployment insurance abstracts from any role of government in the economy other than UI. This abstraction hides the importance of fiscal externalities: external effects of individual actions such as job search intensity which operate through the tax system. I use a standard dynamic job search model and a simple two-period model of unemployment to answer the question: how much does abstracting from government activities other than UI matter for optimal UI calculations? My numerical results indicate that it can matter a lot, and in the baseline case in which UI benefits have no effect on subsequent wages, the optimal replacement rate drops from around 0.4 to zero; a large effect of benefits on wages, however, could significantly increase the optimal replacement rate. Theoretical analysis indicates that allowing for a personal income tax increases the estimated optimal benefit level if and only if higher benefits increase total earnings.

*I am very grateful to David Lee and Richard Rogerson for many helpful comments and suggestions, as well as to the participants of the Industrial Relations Sections Graduate Lunch and Public Finance Working Group seminars at Princeton University. Any errors or omissions are the responsibility of the author.
1 Introduction

For several decades, economists have been studying the problem of optimal unemployment insurance, and the resulting literature has at its centre the following question: how generous should UI benefits be? Two basic approaches have been taken in this literature: the “sufficient statistics” approach and the macro job search approach. The former involves deriving broadly applicable formulas for the welfare consequences of UI which depend only on a few reduced-form empirical values, while the macro approach calibrates or estimates the parameters of a dynamic job search model, allowing the model to be simulated and solved numerically for the optimal benefit level. Numerical results in the literature have often tended to indicate that benefits should be more generous, but this is far from a universal conclusion.

Despite the variety of approaches and conclusions, the literature has at least one important characteristic in common: all optimal UI papers abstract from any role of government in the economy other than the UI system. In reality, however, UI accounts for a very small portion of total government spending and revenue in all developed countries; for example, in the United States, spending on UI represented less than 1% of total government spending as recently as 2008.

This abstraction is a problem because when we ignore other roles of government, we hide the importance of fiscal externalities; this refers to externalities which operate through the tax system, when individual actions affect that individual’s taxable income and thus the amount of taxes they pay to the government. The tax rate or amount of public goods provided must then change to maintain government budget balance, thereby imposing costs or benefits on all workers which are not internalized by the individual in question. This becomes relevant when government policy, such as UI, aggravates or mitigates such an externality, generating an indirect revenue effect of the policy that must be taken into account. For example, if higher UI benefits reduce search intensity and cause a worker to spend longer unemployed, that worker pays less in taxes, requiring the government to raise taxes on all workers. This phenomenon is already at work in the standard analysis,\textsuperscript{1} it is greatly mag-

\textsuperscript{1}Even with a UI tax that is lump-sum conditional on employment, if UI causes individuals to stay unemployed longer, the reduction in time spent employed and paying the tax represents a negative fiscal
nified when we take the full amount of government spending and taxation into account. Alternatively, if more generous UI raises reservation wages and thus allows workers to find jobs with better wages, the resulting increase in tax revenue benefits other workers, and this effect is also much larger when we capture the full size of government.

Therefore, the question which I wish to answer in this paper is: how much does abstracting from government activities other than UI matter for optimal UI calculations? In particular, what effect does this omission have on the optimal benefit level? The answer is that it can matter quite a lot: if UI benefits have no effect on subsequent wages, as is often assumed, the optimal replacement rate drops from around 0.4 to zero, and I estimate welfare gains from moving from what was believed to be the optimum with no non-UI spending to the new estimated optimum that are around 7% of current government spending on UI. However, a positive elasticity of post-unemployment wages with respect to benefits can reverse this finding, leading to an increase in the optimal benefit level, and perhaps a replacement rate near one. This wage elasticity is therefore an important quantity, as the direction of the fiscal externality is highly sensitive to it, and the sparseness and inconsistency of the literature estimating it prevents us from drawing definite conclusions. Analytically, I show in a simple model that allowing for an income tax that pays for a large amount of other government spending increases the estimated optimal benefit level if and only if higher benefits increase total earnings.

My analysis uses both of the two main approaches in the optimal UI literature. I begin with a standard structural dynamic job search model from Lentz (2009), which I calibrate to match a number of real-world moments, first using an estimate of the real value of government spending and then ignoring all government activities other than UI; I then solve numerically for the optimal replacement rate in each case. I then switch to a simpler model, specifically Baily (1978), which is both the seminal paper in the optimal UI literature and the original paper in the sufficient statistics approach. The simplicity of this model allows for a more transparent demonstration of the central results, permitting us to get inside the “black box” of the dynamic job search model. I solve the model for the derivative of social welfare with respect to the UI benefit level, and provide a series of analytical results about the effect of externality that decreases the appeal of insurance.
fiscal externalities on this derivative and the resulting estimate of the optimal benefit level. I subsequently use the standard method of statistical extrapolations to numerically estimate the optimal benefit level, providing results which are generally quite similar to those obtained from the structural approach. The use of two very different models from opposite technical extremes of the literature is intended to show the generality of my results.

The rest of the paper is organized as follows. Section 2 provides further detail on the optimal UI literature and the concept of fiscal externalities. Section 3 presents the dynamic job search model, including a discussion of the calibration and the results. In section 4, I describe the two-period model from Baily (1978), and solve for the welfare derivative and optimal UI equation; section 5 contains the analytical results based upon these equations. Numerical results from the statistical extrapolation method follow in section 6. Section 7 provides a conclusion.

2 Optimal UI Literature and Fiscal Externalities

In general, optimal UI analysis focusses on balancing consumption-smoothing and risk-sharing benefits with moral hazard costs. UI insures workers against the risk of unemployment, weakening liquidity constraints and allowing consumption to be better smoothed across employed and unemployed states, but also subsidizes leisure and thus reduces the incentive to search for jobs, leading to an increase in unemployment. The typical optimal UI paper balances these marginal benefits and costs to maximize expected utility, and any activities of government unrelated to UI are ignored.

The two main approaches in the literature have attacked this problem in different ways.

---

2 A number of papers also study issues of the design of the UI system, including the optimal time path of benefits and the use of re-employment taxes or job-creation subsidies to complement UI. Hopenhayn and Nicolini (1997) and Fredriksson and Holmlund (2001) find that benefits should decline with time spent unemployed, while Coles and Masters (2006) and Coles (2008) advocate job-creation subsidies, and Hopenhayn and Nicolini (2009) supports taxes which increase and UI benefits which decrease with previous unemployment durations.

3 Usually, it is implicitly assumed that the government has access only to a UI payroll tax and pays only for UI benefits; the tax is often assumed to be lump-sum conditional upon employment, which is not, in the context of UI alone, an unreasonable assumption, as Chetty (2006) describes how, in the US, the UI payroll tax applies only to a small base of income, making it an inframarginal tax for most workers. Some studies also allow for job creation subsidies, or allow taxes to vary with past durations of unemployment, but even these papers continue to ignore the vast majority of government activity.
An economist studying the optimal UI problem from the sufficient statistics approach solves their model for an equation for the derivative of social welfare with respect to benefits, and finds a way to write this equation in terms of empirical quantities, or statistics which are sufficient for welfare analysis; these statistics encode all the relevant information from any underlying structural parameters, making the welfare equations robust to many modifications and modelling decisions. The macro approach, meanwhile, requires the specification of a dynamic job search model, for which functional forms and parameter values must be selected, allowing policy experiments to be performed numerically within the model.

While some studies concentrate entirely on theoretical contributions, many papers in the optimal UI literature do provide numerical results, and Table 1 summarizes estimates of the optimal UI replacement rate in the literature. The estimates vary significantly, but as stated before, the majority of the estimates indicate that benefits should be more generous, compared to the real UI system in the United States, which typically features benefits lasting 26 weeks and a statutory replacement rate of 0.5, which amounts to an effective mean replacement rate of about 0.46, due to maximum benefit limits.

The abstraction from government activities other than UI is, however, a universal feature, and so the literature is united in its erroneous minimization of the importance of fiscal externalities. While I emphasize revenue effects, from increased durations of unemployment and possibly higher subsequent wages, there is an additional externality which operates

---

4 This approach has been referred to as the “reduced-form” method upon occasion, but Chetty (2009) suggests the name “sufficient statistics,” reserving “reduced-form” for “program evaluation” techniques.

5 For instance, Chetty (2006) shows that the final equation from Baily (1978) is applicable to a far more general class of models; Baily’s simple two-period model depends only on the drop in consumption upon unemployment, the coefficient of relative risk-aversion, and the elasticity of durations of unemployment with respect to UI benefits.

6 For example, Chetty (2006) and Shimer and Werning (2007) provide sufficient statistic equations for optimal UI but do not present their own numerical estimates, and Kocherlakota (2004) and Hopenhayn and Nicolini (2009) focus on theoretical results for various aspects of optimal UI.

7 Further detail on the concept of fiscal externalities can be found in Lawson (2012b). Another interpretation of fiscal externalities is that they represent the effect of a pre-existing distortion: the income tax is assumed to be a pre-existing and fixed distortion, and thus, in the spirit of the Theory of the Second-Best, we may want to introduce or alter distortions in other areas of the economy to improve efficiency. This does, of course, require the existence of some non-policy-generated inefficiencies restricting the available policy choice set: I assume that there is some amount of required government spending, and that the government is unable to use lump-sum taxation to raise the required revenue. Neither of these assumptions are unusual.

8 This latter possibility has rarely been recognized in the optimal UI literature; Brown and Kaufold (1988) come closest with the recognition that, if UI affects how much human capital workers accumulate, it will have an effect on the tax base, though I will ignore this potential effect in the current paper. Baily (1978)
Table 1: Numerical Results from Studies of Optimal UI

<table>
<thead>
<tr>
<th>Paper</th>
<th>Optimal Replacement Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sufficient Statistics</strong></td>
<td></td>
</tr>
<tr>
<td>Baily (1978)</td>
<td>$0 - 0.34^*$</td>
</tr>
<tr>
<td>Gruber (1997)</td>
<td>$0 - 0.43^*$</td>
</tr>
<tr>
<td>Chetty (2008)</td>
<td>$&gt; 0.5^{**}$</td>
</tr>
<tr>
<td><strong>Structural Models</strong></td>
<td></td>
</tr>
<tr>
<td>Hansen and İmrohoroğlu (1992)</td>
<td>$0.15^*$ (with moral hazard)/0.65 (without)</td>
</tr>
<tr>
<td>Davidson and Woodbury (1997)</td>
<td>$0.66^*/1.30^{**}$</td>
</tr>
<tr>
<td>Hopenhayn and Nicolini (1997)</td>
<td>$&gt; 0.94^*$ (with optimal tax)</td>
</tr>
<tr>
<td>Acemoglu and Shimer (2000)</td>
<td>$&gt; 0.4^{**}$</td>
</tr>
<tr>
<td>Fredriksson and Holmlund (2001)</td>
<td>$0.38 - 0.42^*$</td>
</tr>
<tr>
<td>Wang and Williamson (2002)</td>
<td>$0.24^*/0.56^{**}$</td>
</tr>
<tr>
<td>Coles and Masters (2006)</td>
<td>$0.76^*$</td>
</tr>
<tr>
<td>Lentz (2009)</td>
<td>$0.43 - 0.82^*$</td>
</tr>
</tbody>
</table>

* corresponds to infinite-duration UI, ** to finite-duration (typically 26 weeks)

through the government budget constraint: if UI increases durations of unemployment, this also increases the total amount of UI benefits paid, which raises the required tax rate. The magnitude of this effect, however, does not depend on the size of the income tax, and thus has been fully accounted for in previous work; I take it into account as well, but the emphasis is placed upon externalities operating through changes in taxable income because they are dramatically affected by size of government, and I attempt to discover how much of an impact such revenue effects have on optimal UI calculations.9

This discussion indicates one important feature of the analysis: I study the effect of fiscal externalities on optimal UI calculations. I am not approaching the problem as one of changing a primitive parameter of the model, i.e. I am not considering the “true impact” on optimal UI of increasing the size of government. Rather, I argue that governments’ fiscal activities have always been much more extensive than just UI, but that this has been ignored actually derives an intermediate step in his calculations which accounts for a possible wage effect, but then assumes it to be zero on the grounds that it will be very small given the small size of the UI tax. More recently, Chetty (2006) and Chetty (2008) both conclude that wage or matching effects do not matter for optimal UI calculations, findings which are entirely due to their assumption of a lump-sum UI tax.

9This question is partly motivated by a paragraph in Chetty (2006), where he acknowledges that the previous literature has ignored fiscal externalities, though he refers to the effect of taxes on capital gains or dividends and not the general personal income tax which I will study, and goes on to state that “It would be useful to determine the magnitude of such fiscal externalities [to] assess whether they affect the calculation of the optimal benefit rate significantly.”
by the optimal UI literature, and I want to know how much the calculated level of optimal UI changes if the large amount of non-UI government spending is no longer ignored.\textsuperscript{10} This point should simply be kept in mind as we proceed, as it causes me to perform the analysis in ways which may otherwise seem somewhat unusual.

The two approaches used in this paper were chosen partly as substitute approaches which demonstrate the generality of the results, as mentioned earlier,\textsuperscript{11} but seen from within the public finance literature, the two parts of this paper are also highly complementary: while Chetty (2006) demonstrates that a wide variety of dynamic search models can be made compatible with the sufficient statistics approach and provide identical local results, Chetty (2009) acknowledges that the statistical extrapolations needed to make out-of-sample predictions in the sufficient statistics method may be seen as ad-hoc, and states that an alternative is to use the same sufficient statistics to calibrate a structural model. This is exactly what I do, as the moments used in the structural setting are derived directly from the sufficient statistics used later in the paper.

\section{Dynamic Job Search Model}

The analysis now begins with the dynamic job search model from Lentz (2009). The first subsection contains a description of the model, while the second explains how I calibrate the model; I then present the numerical results, and discuss the effects of fiscal externalities on the estimated optimal UI benefit level.

\textsuperscript{10}This, for instance, is why I calibrate the structural model twice, to match the real-world moments in each case: doing so indicates how my results compare to those estimated by researchers who try to match their models to the real world but ignore all government activities other than UI.

\textsuperscript{11}Both models, however, ignore any general equilibrium effects of UI on the job search process; Welch (1977) argues that more generous UI may lead firms to reduce wages to compensate for higher tax payments due to experience rating, whereas Acemoglu and Shimer (2000) suggest that UI may encourage firms to create higher-productivity jobs. Also, my use of a one-sided search specification in both models, while very common in the literature, means that I abstract from firms and assume that any changes in wages resulting from changes in UI represent actual productivity gains or losses. This could be inappropriate, for example if all matches are identical in productivity and wages are determined by Nash bargaining; however, the difficulties of evaluating match quality effects in any other way provide little alternative, and my assumption is not unusual in the literature, as seen in the discussion of wage effects in Chetty (2008), and the way in which Shimer and Werning (2008) explicitly equate wages and output, despite their search environment.
3.1 Model Setup

Lentz (2009) is a standard dynamic one-sided job search model, which is both conceptually simple to simulate and incorporates important features such as choice of search intensity and private savings; the only modifications I make are the introduction of non-UI government spending and a simplified specification of the search cost function.

The model features a representative infinitely-lived agent who makes stochastic transitions between states of employment and unemployment. When unemployed, the agent receives an after-tax UI benefit equal to \( b \), with infinite potential duration,12 and chooses search intensity \( s_t \) subject to a convex search cost function, where \( s_t \) is defined to be the probability of receiving a job offer.13 All jobs have an identical wage \( w \),14 and jobs exogenously end at a constant rate of \( \delta \) per period. Finally, agents cannot borrow, but can make savings which earn interest at a rate \( i \) per period, while the annual discount rate is \( \rho \); I define each period to be equal to a week, so the per-period discount factor is \( \beta = \left( \frac{1}{1+\rho} \right)^{\frac{1}{52}} \).

Thus, in all periods and states, agents decide on their level of consumption (or equivalently saving), while unemployed agents also choose a level of search intensity. The decision problem can be written as:

\[
\max_{\{C_t, s_t\}} E \sum_{t=0}^{\infty} \beta^t [u(C_t) - e(s_t)] \\
\text{s.t. : } k_{t+1} = (1+i)k_t + n_tw(1-\tau) + (1-n_t)b - C_t \\
C_t, k_t, s_t \geq 0 \\
Pr(n_{t+1} = 1|n_t = 1) = 1 - \delta, Pr(n_{t+1} = 0|n_t = 1) = \delta \\
Pr(n_{t+1} = 1|n_t = 0) = s_t, Pr(n_{t+1} = 0|n_t = 0) = 1 - s_t
\]

where \( C_t \) is consumption, \( \tau \) is the tax rate, \( k_t \) represents assets at time \( t \), and \( n_t \) is an indicator for employment where \( n = 1 \) indicates that the agent is employed.

For the purposes of simulating the model, it will be more convenient to write this recursively; let \( V_e(k) \) represent the maximum present value of being employed with assets equal

---

12By assuming that benefits are constant and never expire, I keep the analysis simple and focus on only one dimension of the optimal UI problem, namely the optimal level of benefits.
13This is a simplified specification with a closed-form solution for the individual’s search decision.
14Lentz assumes that each individual has a fixed wage, but that different individuals may face different wages, allowing him to match the wage distribution he observes in his data. I simply assume that all individuals and all jobs have the same wage.
to \( k \), while \( V_u(k) \) will be the analogous value of unemployment, and let \( k' \) represent next period’s assets:

\[
V_e(k) = \max_{k' \in \Gamma_{z}(k)} \left[ u((1 + i)k + w(1 - \tau) - k') + \beta[(1 - \delta)V_e(k') + \delta V_u(k')] \right]
\]

\[
V_u(k) = \max_{k' \in \Gamma_{a}(k), h \geq 0} \left[ u((1 + i)k + b - k') - c(s) + \beta(s V_e(k') + (1 - s) V_u(k')) \right]
\]

where \( \Gamma_{z}(k) = (k' \in \mathbb{R} | 0 \leq k' \leq (1 + r)k + z) \) is the set of permissible asset values. As in Lentz (2009), the numerical solutions always appear to yield concave value functions by asset level. I can now solve numerically for the economy-wide steady-state, and in that steady-state, the government budget constraint is given by:  

\[
(1 - u)w\tau = ub + G
\]

where \( u \) is the unemployment rate and \( G \) is the level of exogenous non-UI government spending. My goal will be to compare the results from cases when I use my best estimate of \( G \) with those where I assume that \( G = 0 \).

### 3.2 Calibration of Model

I now have to choose functional forms and parameter values, in order to be able to simulate the model. I assume a constant relative risk-aversion utility function with risk-aversion parameter \( R \), so \( u(C) = \frac{C^{1-R}}{1-R} \). The search cost function will be \( e(s) = \frac{(\theta s)^{1+\kappa}}{1+\kappa} \), until \( s = \bar{s} \), beyond which the marginal cost is infinite, making \( \bar{s} \) the maximum feasible search intensity. Finally, the way I specify \( b \) must be explained. To begin with, if the replacement rate is \( r \) and the baseline tax rate is \( \tau_0 \), the after-tax value of benefits for a recipient is \( r(1 - \tau_0)w \). However, the elasticities and other quantities used later are defined for the entire population that is eligible for UI, so I also multiply by the take-up rate of benefits, which Ebenstein and Stange (2010) find to be about 0.8 from 1990-2005. Additionally, because real-world
benefits have a finite duration, whereas the model assumes an infinite duration for simplicity, I multiply the value of benefits by \( \frac{15.8}{24.3} \), which is the ratio of mean compensated unemployment duration to mean total duration in the Mathematica sample of Chetty (2008). The final equation for \( b \) therefore is \( b = r(1 - \tau_0)w(0.8)\left(\frac{15.8}{24.3}\right) \); thus, I express \( b \) as an approximate present-value-equivalent infinite-duration benefit for an individual who becomes unemployed and is eligible for UI.

The selected parameter values are summarized in Table 2. The job separation rate is set to \( \delta = \frac{1}{213} \) to correspond with a median job duration of 4.1 years measured by the Bureau of Labor Statistics in January 2008. For the interest rate \( \hat{i} \), I follow the example of Hansen and İmrohoroğlu (1992) and Chetty (2008) in setting it to zero. The wage \( w \) is normalized to one. \( R = 2 \) is a standard value for the coefficient of relative risk-aversion in studies of UI, but Chetty (2008) states that his results imply a value of \( R = 5 \) in the context of unemployment, so I use both values. For the upper limit of search intensity, I use a value of \( \bar{s} = 0.5 \), which satisfies the intuition that it should not be possible for an individual to guarantee finding a job immediately.

Finally, I select initial values of \( r = 0.46 \) and \( \tau_0 = 0.23 \); the former is the mean replacement rate over 1988-2010 reported by the U.S. Department of Labor, while the latter incorporates the 15% federal rate of the typical UI recipient, 5% for a typical state income tax, and 3% for the Medicare tax. Therefore, the baseline value of \( b \) is 0.184, and these estimates imply a value of \( G \) approximately equal to 0.208. In cases where I assume that

\[ \text{Chetty (2008) finds that unemployed individuals tend to have little in the way of long-term savings, and Hansen and İmrohoroğlu (1992) argues that previous findings of near-zero average real returns on “highly liquid short-term debt” justify the assumption of a non-interest-bearing asset.} \]

\[ \text{For example, Chetty and Saez (2010) use } R = 2, \text{ and Lentz (2009) estimates a value of 2.21.} \]

\[ \text{Chetty (2006) states that “empirical studies that have identified large income effects on labor supply for the unemployed” are inconsistent with low values of } R, \text{ and argues that such a parameter must be chosen to be consistent with the context in which it is being considered.} \]

\[ \text{This upper limit is binding in very few instances.} \]

\[ \text{The DOL also reports another number, a mean replacement rate of 35%; this number represents average weekly benefit payments divided by average weekly earnings among workers covered by UI. However, workers who receive UI tend to have incomes that are somewhat lower than the average of all covered workers.} \]

\[ \text{The employee and employer shares of the Medicare tax add up to 2.9%. FICA taxes are not applicable to UI benefits, so I may be slightly understating the after-tax value of UI; however, I ignore the Social Security part of FICA taxes on the grounds that it is more of a pension contribution than a tax, but if a portion does correspond to a tax, I also tend to underestimate the value of } G, \text{ so there are offsetting biases on the relative sizes of UI and } G. \text{ Liebman, Luttmer, and Seif (2009) estimate a mean effective Social Security tax of 3.8% in their sample of 52- to 80-year-olds, though for a population of relatively modest-income and generally younger individuals, it would likely be lower.} \]
$G = 0$, meanwhile, I assume that the UI payroll tax is paid only by employees, so $\tau_0 = 0$ for the purpose of calculating UI benefits.

### Table 2: Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>job separation rate</td>
<td>$\frac{1}{213}$</td>
</tr>
<tr>
<td>$i$</td>
<td>real interest rate</td>
<td>0</td>
</tr>
<tr>
<td>$w$</td>
<td>per-period wage</td>
<td>1</td>
</tr>
<tr>
<td>$r_0$</td>
<td>baseline replacement rate</td>
<td>0.46</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>baseline tax rate</td>
<td>0.23</td>
</tr>
<tr>
<td>$R$</td>
<td>coefficient of relative risk-aversion</td>
<td>${2, 5}$</td>
</tr>
<tr>
<td>$\bar{s}$</td>
<td>maximum search intensity</td>
<td>0.5</td>
</tr>
<tr>
<td>$\theta$</td>
<td>search cost parameter</td>
<td>TBD</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>search cost parameter</td>
<td>TBD</td>
</tr>
<tr>
<td>$\rho$</td>
<td>annual discount rate</td>
<td>TBD</td>
</tr>
</tbody>
</table>

The remaining parameters, $\theta$, $\kappa$ and $\rho$, are set to make the model match a set of moments from the real world. There are a total of four cases to consider, specifically the interactions of $R \in \{2, 5\}$ and $G \in \{0, 0.208\}$; for each value of $R$, I calibrate the model twice, once for the true $G$ and once for $G = 0$, so as to match the real-world moments in each case. The moments I use are the unemployment rate $u$, the percentage gap between average consumption while employed and while unemployed $\frac{E(C_e) - E(C_u)}{E(C_e)}$, and the elasticity of unemployment durations with respect to benefits, which I denote as $E_u^b = \frac{b u}{du/db}$.

The unemployment rate is often used to calibrate job search models, and while all three moments clearly jointly determine the values of $\theta$, $\kappa$ and $\rho$, $u$ may be thought of as a natural choice for pinning down the level of the search cost function, which is primarily determined by $\theta$; $E_u^b$, meanwhile, is highly informative about the curvature parameter $\kappa$. My decision to calibrate the discount rate to match a moment, however, is unusual, as the standard approach is simply to choose a value for $\rho$ that seems reasonable; however, there does not appear to be a definite consensus on the right “reasonable value,”$^{24}$ and Lentz (2009) finds that his results are very sensitive to the gap between $i$ and $\rho$. Since the consumption gap is closely related to workers’ ability to maintain a buffer stock of assets, $\frac{E(C_e) - E(C_u)}{E(C_e)}$ is very useful for identifying the discount rate. The three moments I use here are also among the

---

$^{24}$A number around 5% seems most typical, but to take opposite extremes, Acemoglu and Shimer (2000) use an annual discount rate of about 10%, whereas Coles (2008) produces results for a zero discount rate.
sufficient statistics that will be used later in the paper.\footnote{I am therefore following the advice of Chetty (2009) in considering the use of such statistics to calibrate a structural model as an alternative to statistical extrapolation.}

The specific values used for the moments are as follows. The average unemployment rate among high-school graduates was 5.4\% over 1992-2000, and this group seems like a reasonable approximation to the population likely to receive UI, so $u = 0.054$. Gruber (1997) estimates a relationship of $\frac{E(C_e) - E(C_u)}{E(C_e)} = 0.222 - 0.265r$,\footnote{Gruber’s data is on food consumption from the PSID; obtaining good quality data on consumption across states of employment and unemployment has proven to be difficult.} which implies a value of 0.1001 at baseline.\footnote{In this model, it is very difficult to generate a consumption gap which declines with $b$, because in considering the steady-state, the “initial conditions” of the model are allowed to change with $b$, and thus agents accumulate fewer assets as $b$ increases. Whether this is taken as evidence against the empirical methods used to estimate the relationship of consumption to benefits or against the current model is beyond the scope of this paper, but since my later sufficient statistics analysis uses Gruber’s estimated relationship directly in the welfare derivative, it adds further credibility to the claim that I am considering strongly differing approaches to estimating optimal UI.} Finally, Chetty (2008) estimates an elasticity of unemployment durations with respect to benefits of 0.53,\footnote{This estimate is close to the middle of the typical range of estimates in the literature: Chetty (2008) describes the usual range of estimates as 0.4 to 0.8, while Fredriksson and Holmlund (2001) claim that their own calibrated value of 0.5 is “in the middle range of the available estimates.”} though this estimate is based on a sample of UI recipients, whereas the consumption estimates in Gruber (1997) were based on a sample of unemployed workers who were estimated to be eligible for UI, regardless of whether they actually took up benefits; therefore, I follow Gruber’s recommendation and multiply the elasticity by 0.48, which is the derivative of benefit receipt to benefit eligibility in his sample, giving a value of $E_u^b = 0.2544$.\footnote{The government can only control the benefits for which an individual is eligible, not what they actually receive. 0.48 is also very close to my value of $0.8 \times \frac{15.8}{24.3} = 0.52$ for the percentage of the time eligible unemployed individuals receive benefits.} These values are summarized in Table 3.

### Table 3: Moments

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u$</td>
<td>unemployment rate</td>
<td>0.054</td>
</tr>
<tr>
<td>$\frac{E(C_e) - E(C_u)}{E(C_e)}$</td>
<td>consumption gap between employment and unemployment</td>
<td>$0.222 - 0.265r = 0.1001$</td>
</tr>
<tr>
<td>$E_u^b = \frac{b du}{db}$</td>
<td>elasticity of $u$ wrt $b$</td>
<td>$0.48 \times 0.53 = 0.2544$</td>
</tr>
</tbody>
</table>

To numerically solve the model for any given set of parameters and a value of $b$, I begin with value function iteration: an initial guess is chosen for the value functions, and the maximization problem is solved for a range of asset values, which then provides a new guess
for the value functions; this process is repeated until the value functions have converged. I only evaluate the maximization problem and value function for a subset of the asset value grid on each iteration, and then use cubic spline interpolation to fill in intermediate values, as also done by Lentz. Next, the transition process of agents between states is iterated to calculate the steady-state distribution. I then evaluate the government budget surplus, and then re-set the tax rate and repeat the above steps until the budget is balanced, except in the baseline, where I know the starting tax rate $\tau_0 = 0.23$.

In order to calculate $E^{u}_b$, the model must be solved at baseline, and then solved again for a different level of $b$; since the numerical procedure involves discretizing the asset distributions, the results are slightly “lumpy” at high magnification, so I choose a replacement rate of $r = 0.56$, and compute the resulting arc elasticity. With both sets of numerical results in hand, I can then estimate the moments of interest in the simulated data and compare them to their real-world counterparts. Due to the “lumpiness” of the results, a precise numerical search for the minimum-distance parameters is not feasible; instead, I find values for the parameters that make the moments as close to the desired values as is practical. Finally, in each case, once the parameters have been calibrated, a grid search over $r$ can be performed to find the optimal level.

### 3.3 Numerical Results

We can now proceed to an evaluation of the results; appendix A contains the parameters used in each case, the resulting moments, and a further discussion of the methods used. The results for the optimal replacement rates and estimated welfare gains from moving to the optimum can be found in Table 4; welfare gains are expressed as percentages of initial spending on UI, and the final line of the table, labelled “Diff.”, displays the welfare gain from moving from the replacement rate believed to be optimal when $G$ is assumed equal to zero to the “true” optimum. Figures 1, 2, 3 and 4, meanwhile, display the unemployment rates and tax rates as a function of $r$ in each of the cases that I consider; as can be seen,

---

30. This variation is comparable to that studied by empirical work in this area; for example, Addison and Blackburn (2000) estimate a mean replacement rate of 0.44 with a standard deviation of 0.12.

31. I estimate the welfare gain per worker per year, convert this to dollars using a base of mean consumption while unemployed, and divide by annual spending on UI.
the relationship of \( u \) and \( \tau \) with respect to \( r \) is very similar in each case (except that \( \tau \) is considerably higher when \( G = 0.208 \)).

<table>
<thead>
<tr>
<th>( G )</th>
<th>( R = 2 )</th>
<th>( R = 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( r )</td>
<td>Welf. Gain*</td>
</tr>
<tr>
<td>0.208</td>
<td>0.00</td>
<td>12.69</td>
</tr>
<tr>
<td>0</td>
<td>0.36</td>
<td>0.54</td>
</tr>
<tr>
<td>Diff.**</td>
<td>-</td>
<td>6.92</td>
</tr>
</tbody>
</table>

Figure 1: Unemployment Rates and Tax Rates for \( R = 2 \) and \( G = 0.208 \)

We can see that a non-zero value of \( G \) significantly reduces the optimal replacement rate in both cases, but especially in the more standard case of \( R = 2 \), where the fiscal externalities I am studying reduce the optimum to zero, with an accompanying welfare gain of nearly 13% of current spending on UI. The welfare gains of moving between from the \( G = 0 \) optimum to the \( G = 0.208 \) optimum amount to about 7% of UI spending, or $3 billion on an economy-wide basis.

An extension of the job search model to a context with a wage distribution is in progress; when completed, it will represent the first attempt that I am aware of to use a simulated job search model to study optimal UI in the context of a non-degenerate wage distribution.\(^{32}\)

The results are clear-cut: the apparently minor modification of a standard search model to include non-UI spending substantially affects optimal UI calculations. But it might rea-

\(^{32}\)Acemoglu and Shimer (2000) allow for a wage distribution, but in their simulations this distribution consists of only two wages.
sonably be asked: what mechanisms are at work here? I have already explained the intuition behind fiscal externalities, but a model like the one I have been studying has a certain “black box” character to it. A simpler model will allow us to get inside the black box and make a more detailed, step-by-step analysis of fiscal externalities in the context of UI, and for this reason I now switch my attention to the original seminal paper in optimal UI, Baily (1978).

33Shimer and Werning (2007) support this view when they state that structural models “rely heavily on the entire structure of the model and its calibration, which sometimes obscures the economic mechanisms at work and their empirical validity.”
4 Baily (1978) Model

In this section, I analyze the two-period model of unemployment from Baily (1978). The first and second subsections present the model and derive a general version of the optimal benefit equation, while the following subsection explores this equation in further detail and provides the equations needed to perform the numerical analysis.

4.1 Model Setup

The only modification that I make to Baily’s model is to add $G$ to the government budget constraint. The model is very simple, but the analysis of Chetty (2006) and Lawson (2012b) indicate that the results from such a simple model can apply in much more general circumstances, and its simplicity is what makes it well suited to an exposition of the potential effects of fiscal externalities.

Time is finite and consists of two periods,\footnote{Baily states that his model represents a two-year time horizon, but we are not required to take the model so literally; it could stand for a world with a longer time horizon divided into two halves. We will see later that this interpretation could have a significant effect on the final numerical results.} with the interest and discount rates both set to zero. In the first period, the representative worker is employed at an exogenous wage $y$,\footnote{In an extension in appendix E.4, I consider the effect of allowing choice over initial labour supply.} and between periods they face a risk of unemployment: with exogenous probability $\alpha \in (0, 1)$, the worker keeps their initial job at the same wage for the entire second period, whereas they lose their job and become unemployed with probability $(1 - \alpha)$. If the worker becomes
unemployed, they choose search effort \( c \) (normalized into income units) and a desired wage \( y_n \), and then spend a fraction \((1 - \eta)\) of the second period unemployed and the remaining \( \eta \in (0, 1) \) at a new job at wage \( y_n \),\(^{36}\) where \( \eta \) is a deterministic function of \( c \) and \( y_n \):\(^{37}\)

\[
\eta = \eta(c, y_n), \quad \frac{\partial \eta}{\partial c} > 0, \quad \frac{\partial \eta}{\partial y_n} < 0.
\]

Individuals receive utility from consumption in each period according to the continuous function \( U(C) \), where \( U' > 0 \) and \( U'' < 0 \). \( k \) represents first-period savings, and so overall expected utility is given by:

\[
V = U(C_e^1) + \alpha U(C_e^2) + (1 - \alpha) U(C_u) \tag{1}
\]

where \( C_e^1 = y(1 - \tau) - k \), \( C_e^2 = y(1 - \tau) + k \) and \( C_u = (1 - \eta)(b - c) + \eta y_n(1 - \tau) + k \), and where once again \( \tau \) is the income tax rate and \( b \) the after-tax UI benefit.\(^{38}\)

The government budget constraint is written in the same way as before:

\[
[(1 + \alpha)y + (1 - \alpha)\eta y_n] \tau = (1 - \alpha)(1 - \eta)b + G \tag{2}
\]

where \( G \) again represents additional exogenous government expenditures. The source of the fiscal externalities can be seen in equations (1) and (2): post-unemployment income \( \eta y_n \) is subject to income tax, and thus part of the returns to the individual’s choice of \( c \) and \( y_n \) go to agents other than the individual worker. The important question, as we shall see, will be how UI benefits \( b \) affect this post-unemployment income.

### 4.2 Calculation of Welfare Derivative

The next step is to evaluate the government’s social welfare derivative. Looking at (1), we see that the worker’s lifetime expected utility can be written generally as \( V = V(k, c, y_n, b, \tau) \); the government sets the values of \( b \) and \( \tau \), and then the worker chooses \( \{k, c, y_n\} \) to maximize \( V \) taking \( \{b, \tau\} \) as given. The government planner has equally-weighted utilitarian preferences,

---

\(^{36}\)\( y_n \) is assumed to be deterministic, as if a worker defines the type of job (i.e. wage level) that they will search for, and will eventually find such a job, with it taking longer to find high-wage jobs. \( y_n \) is thus both a reservation wage and the wage upon re-employment.

\(^{37}\)In appendix E.1, I examine the consequences of a stochastic \( \eta \).

\(^{38}\)The assumption that, if the worker loses their job, utility in the second period is defined over total consumption implies no credit constraints within a period: the worker can borrow as much as necessary to smooth consumption during the second period. I consider a relaxation of this assumption in appendix E.2.
so they want to maximize $V$ at the individual’s optimum, choosing $b$ and $\tau$ so that the government budget constraint is satisfied in equilibrium. Since the individual maximizes subject to $\{b, \tau\}$, $\frac{\partial V}{\partial b} = \frac{\partial V}{\partial \tau} = \frac{\partial V}{\partial y_n} = 0$; therefore, for a small change in $b$, behavioural changes have no first-order effect on welfare. The envelope theorem then tells us that the welfare derivative can be written as a function of the two partial derivatives $\frac{\partial V}{\partial b}$ and $\frac{\partial V}{\partial \tau}$ and the derivative of the government budget constraint:

$$\frac{dV}{db} = \frac{\partial V}{\partial b} + \frac{\partial V}{\partial \tau} \frac{d\tau}{db}. \quad (3)$$

In a manner of speaking, $\frac{\partial V}{\partial b}$ represents the marginal benefit of increased UI, which is equivalent in utility terms to a marginal increase in consumption while unemployed, whereas the second term represents the marginal cost in the form of higher taxes, with $\frac{d\tau}{db}$ identifying the size of the tax increase needed to pay for higher benefits and $\frac{\partial V}{\partial \tau}$ the welfare cost of higher taxes in terms of lost consumption. This can be seen from the partial welfare derivatives:

$$\frac{\partial V}{\partial b} = (1 - \alpha)U'(C_u) \frac{\partial C_u}{\partial b}$$
$$= (1 - \alpha)(1 - \eta)U'(C_u) \quad (4)$$

$$\frac{\partial V}{\partial \tau} = U'(C_e^1) \frac{\partial C_e^1}{\partial \tau} + \alpha U'(C_e^2) \frac{\partial C_e^2}{\partial \tau} + (1 - \alpha)U'(C_u) \frac{\partial C_u}{\partial \tau}$$
$$= -yU'(C_e^1) - \alpha yU'(C_e^2) - (1 - \alpha)\eta y_n U'(C_u). \quad (5)$$

Therefore, the welfare derivative is:

$$\frac{dV}{db} = (1 - \alpha)(1 - \eta)U'(C_u) - [yU'(C_e^1) + \alpha yU'(C_e^2) + (1 - \alpha)\eta y_n U'(C_u)] \frac{d\tau}{db}. \quad (6)$$

I will leave $\frac{d\tau}{db}$ as it is for the time being, and return to it later. To simplify (6), I can replace $U'(C_e^1)$ and $U'(C_e^2)$ using the individual’s first-order condition for saving $k$ and a first-order Taylor series expansion of first-period marginal utility $U'(C_e^1)$ around $U'(C_u)$, specifically $U'(C_e^1) = U'(C_u) + \Delta C U''(\theta)$, where $\Delta C = C_e^1 - C_u$ and $\theta$ is in between $C_u$ and $C_e^1$; details are provided in appendix B.1. I then make two assumptions that are also made by Baily; specifically, I assume that $y_n = y$ in equilibrium, and that $C_e^1 U''(\theta) = C_u U''(C_u)$.

Further discussion can be found in the appendix, where it is explained that these assumptions contain biases in opposite directions. These assumptions further simplify the expression, and dividing by $U'(C_u)$ to put welfare into a dollar equivalent $\frac{dW}{db}$, I arrive at the results summarized in the following proposition.
Proposition 1. Using the assumptions that \( y_n = y \) and \( C_u U'(\theta) = C_u U''(C_u) \), the marginal value of increased benefits is given by:

\[
d W db = \frac{dV db}{U'(C_u)} = 2y \Delta C \frac{d\tau}{db} R - 2(1-u) \frac{d\tau}{db} - \omega
\]

where \( R = \frac{-C_u U''(C_u)}{U'(C_u)} \) is the coefficient of relative risk-aversion, \( u = \frac{(1-\alpha)(1-\eta)}{2} \) remains the unemployment rate, and \( \omega = \frac{(1-\alpha)(1-\eta)}{(1+\alpha)y+(1-\alpha)\eta y_n} \). The equation for the optimal value of \( b \) is thus given by:

\[
\Delta C \frac{d\tau}{db} R = (1-u) \frac{d\tau}{db} - \omega.
\]

Proof. The proof of this result can be found in appendix B.1.

I will calculate the numerical results later using elasticities, so I will also use the following equivalent results.

Corollary 1. The marginal value of increased benefits is also equal to:

\[
d W db = \frac{2u}{(1-u)\psi} \left[ \Delta C \frac{d\tau}{db} R E_0^\tau - (1-u) (E_0^\tau - \psi) \right]
\]

where \( E_0^\tau = \frac{b d\tau}{db} \) is the elasticity of \( \tau \) with respect to \( b \), and \( \psi = \frac{\omega b}{r} = \frac{2u b}{2u b + G} \) is the fraction of total government expenditures allocated to UI. Thus the equation for the optimum is:

\[
\Delta C \frac{d\tau}{db} R = (1-u) \frac{E_0^\tau - \psi}{E_0^\tau}.
\]

The meaning of these results will be interpreted in the next two subsections.\(^{39}\)

4.3 Analysis of Optimal Benefit Equation

Let us start with (7) and (8); in order to understand these equations, I need to evaluate \( \frac{d\tau}{db} \).

Total differentiation of the government budget constraint (2) gives:

\[
\frac{d\tau}{db} = \frac{(1-\alpha)(1-\eta) - (1-\alpha)b \frac{dn db}{db} - (1-\alpha)\tau y_n \frac{dn db}{db} - (1-\alpha)\eta \tau \frac{dn db}{db}}{(1+\alpha)y + (1-\alpha)\eta y_n}.
\]

The four terms in the numerator represent four separate components of the response of taxes to benefits. I will call the first the “mechanical effect”; even if there is no behavioural

\(^{39}\)It is clear from (5) that \( \frac{dV db}{db} < 0 \), and the pair of Baily’s assumptions described in appendix B.1 imply that \( \frac{dV db}{db} = -2y U'(C_u) \left[ (1-u) - \Delta C \frac{R}{C_u} \right] \); thus, if those assumptions are accurate, it follows that \( \Delta C \frac{R}{C_u} < 1 - u \). This result will be of use in section 5.
response to UI, if $b$ increases, the tax rate must increase to compensate, and
\[
\frac{(1-\alpha)(1-\eta)}{(1+\alpha)y+(1-\alpha)y_n}
\]
represents the size of this increase. The second component will be referred to as the “moral hazard effect”; it captures the fact that, if higher benefits increase the duration of unemployment, this increases the total amount of benefits received over time, requiring a further tax increase. The third and fourth components are the two “fiscal externality effects”; the first of them shows how longer unemployment durations also reduce the amount of taxes paid on labour income, raising the required tax increase still further, while the final component captures the fact that, if higher UI increases $y_n$, this increases tax revenues and reduces the necessary tax increase.

If we refer back to the definition of $\omega$ in Proposition 1, we can notice that it is exactly equal to the mechanical effect; this means that the $\frac{d\tau}{db} - \omega \frac{d\tau}{db}$ term on the right-hand side of (8) is the fraction of the total response of taxes to benefits generated by the moral hazard and fiscal externality components. We can now understand (8) as an intuitive way of balancing the marginal benefits and costs of increased UI: the left-hand side represents the welfare gain from increased UI in the form of increased consumption-smoothing, which is increasing both in the size of the consumption shock upon unemployment and the degree of risk-aversion; meanwhile, the right-hand side represents the cost of increased UI in terms of behavioural effects on the government budget. The mechanical effect $\omega$ represents a lump-sum transfer of income between employed and unemployed states, and thus is not a cost to society, so $\frac{d\tau}{db} - \omega \frac{d\tau}{db}$ is the fraction of the tax increase caused by socially costly behavioural responses, and is weighted by $1 - u$, which identifies how much income exists to be taxed.

Now we can shift our attention to (9) and (10). The intuition of (10) is exactly the same as that of (8); the left-hand side is identical, and on the right-hand side, $E^*_b$, represents the percentage increase in the tax rate for a percentage increase in $b$, while $\psi$, as the percentage of government expenditures on UI, is the percentage increase in $\tau$ required for mechanical reasons. The final fraction on the right, therefore, remains the fraction of the response of

---

40Of course, one could also attach the name of “fiscal externality” to the second component, since it involves a cost imposed through the government budget constraint, while the third component is clearly caused by moral hazard just like the second. However, I prefer to distinguish them in this way, partly because it makes the discussion straightforward, but also because the moral hazard component has been fully recognized in previous work, whereas the relative magnitude of the parts I refer to as the fiscal externalities depends to a great extent on the value of $G$. 

taxes to benefits caused by behavioural responses. Using (11), $E^\tau_b$ can be written as:

$$E^\tau_b = \psi \left( \psi + \frac{u}{1-u} \right) E^u_b \left( 1 \right) - \frac{(1 - \alpha) \eta}{2(1 - u)} E^y_b$$  \hspace{1cm} (12)

where $E^u_b$ remains the elasticity of unemployment with respect to $b$, and $E^y_b = \frac{b}{y_n} \frac{d\psi}{db}$ is the elasticity of post-unemployment wages $y_n$ with respect to $b$. The four components of the tax response can easily be seen here as well: the first $\psi$ is the mechanical effect, the $\psi$ and $\frac{2 - \gamma}{\gamma}$ multiplying $E^u_b$ represent the moral hazard effect and the first fiscal externality effect respectively, and the final term is the second fiscal externality.

Therefore, the equation for the optimum becomes:

$$\frac{\Delta C}{C^1_c} R = \left( 1 - u \right) \left( \psi + \frac{u}{1-u} \right) E^u_b \left( 1 \right) - \frac{(1 - \alpha) \eta}{2(1 - u)} E^y_b$$  \hspace{1cm} (13)

This is the equation which I will use to solve for the optimal value of $b$. However, even without a numerical analysis, we can make certain observations about this result. The standard implicit assumption in the literature that $G = 0$ would mean that $\psi = 1$; $\frac{u}{1-u}$, the ratio of unemployed to employed, is likely to be a relatively small number, and so in that case, the mechanical and moral hazard effects will be large compared to the fiscal externality effect, because the taxes that induce the fiscal externality are so small. However, if $G$ is large, $\psi$ will be small, and perhaps close to zero, meaning that the mechanical and moral hazard components will be small and the fiscal externality component will dominate.

Baily (1978) assumes $E^y_b = 0$, and joins the rest of the literature in making the assumption that $G = 0$ and thus $\psi = 1$, but in order to get his result he makes one additional assumption, which is that the $E^\tau_b$ in the denominator of the right-hand side of (10) is equal to one, or that $\psi + \left( \psi + \frac{u}{1-u} \right) E^u_b = 1$ in the denominator of (13).\(^{41}\) In this case, (13) collapse to exactly the result in Baily (1978), found also in Chetty (2006):

$$\frac{\Delta C}{C^1_c} R = \left( 1 - u \right) \left( 1 + \frac{u}{1-u} \right) E^u_b = E^u_b$$

As mentioned before, these are the three original sufficient statistics: $\frac{\Delta C}{C^1_c}$, $R$ and $E^u_b$.

\(^{41}\)Baily claims that this assumption leans towards overestimating the optimal $b$, to offset the conservative assumption that $C^1_c U''(\theta) \simeq C_u U''(C_u)$. However, this is incorrect; the former assumption is in fact conservative, and the latter depends on the value of the risk-aversion parameter, among other things.
5 Analytical Results

Before continuing to the numerical analysis based on (13), in the current section, I will present a series of results about my equations for $\frac{dW}{db}$ and for the optimal level of UI benefits. I will discuss $\frac{dW}{db}(b; G)$, the estimated welfare derivative at a particular value of $b$ given an estimated value of $G$, and let the estimated optimal value of $b$ for a given value of $G$ be denoted as $b^*(G)$. The analysis will consider how the results change when estimated quantities like $G$, and later $E^y_b$, are changed. As described earlier, it should be emphasized that this is not a change of a primitive parameter of the model; I are considering what happens in an estimated equation for the optimum when the sufficient statistics are not changed, since they reflect the unchanged real world, but the estimated value of $G$ plugged into the equation is significantly altered.\footnote{A helpful thought experiment is that of the “two researchers”: one who believes that the true value of $G$ is zero, and another who has estimated a positive value of $G$ from some real-world data. Our two researchers agree on all other sufficient statistics necessary to calculate the optimum; which one will estimate a larger optimal $b$, and by how much?}

My goal is to provide a general sense of how recognizing the true value of $G$ will tend to change any of the results estimated in the literature.

Throughout this section, I maintain two technical assumptions; the first is that $\frac{\Delta C}{C^1} R < 1 - u$, and the second is that $W$ (the integral of $\frac{dW}{db}$) is strictly quasi-concave in $b$. Footnote 3 points out that the first assumption follows immediately from $\frac{\partial V}{\partial \tau} < 0$ if Baily’s assumptions that $y_n = y$ and $C_1 U''(\theta) = C_u U''(C_u)$ are accurate; the latter assumption simply ensures that $\frac{dW}{db} = 0$ at only one value of $b$, so there is a unique maximum and $\frac{dW}{db}$ is positive for values of $b$ below the maximum and negative above.

I begin with an analysis of how the results change when I alter the selected value of $G$. The first result concerns the value of the welfare derivative at a given value of $b$, and is described in the proposition below.

**Proposition 2.** For $G_1 > 0$, $\frac{dW}{db}(b; G_1) - \frac{dW}{db}(b; 0)$ has the same sign as $\eta E^y_b - (1 - \eta) E^u_b$, or equivalently the same sign as $\frac{d(y_n)}{db}$.

**Proof.** The proof of this result can be found in appendix B.2.\qed

Therefore, if two researchers use (13) to estimate the baseline welfare derivative, one using
$G = 0$ and the other $G = G_1$, the latter will find a larger welfare gain from increasing $b$ if and only if $\frac{d(\eta y_n)}{db}$ is positive. The intuition is straightforward: ignoring $G$ greatly understates the revenue effects of changing $b$, and if higher UI benefits raise total post-unemployment earnings $\eta y_n$ (which is the only non-exogenous component of total earnings in the model), this revenue effect is welfare-increasing. Therefore, using a positive value of $G$, which implies higher taxes, amplifies the revenue effect and increases the welfare gain from raising benefits. If $\frac{d(\eta y_n)}{db}$ is negative, the reverse holds.

An immediate corollary arising from quasi-concavity is that, if the baseline estimated welfare derivative is zero for $G = 0$, and thus the current level of $b$ is estimated to be optimal for that case, then the optimum for the true $G$ will be larger or smaller according to the sign of $\frac{d(\eta y_n)}{db}$; if $\frac{d(\eta y_n)}{db} > 0$, $\frac{dW}{db}(b; G_1) > 0$ and quasi-concavity means that the optimum must be found at a higher $b$. A similar logic applies if $\frac{dW}{db}(b; G_1) = 0$; once we know that one of the welfare derivatives is zero, we only need to know the other to make a comparison. This result is summarized by the following proposition.

**Corollary 2.** If, for the current value of $b$, $\frac{dW}{db}(b; 0) = 0$ or $\frac{dW}{db}(b; G_1) = 0$, $b^*(G_1) > b^*(0)$ if and only if $\eta E^y_b - (1 - \eta) E^{u_b} > 0$, or equivalently if and only if $\frac{d(\eta y_n)}{db} > 0$.

Furthermore, if the welfare derivative is of opposite signs for $G = 0$ and $G = G_1$, then a comparison of the estimated optimal values of $b$ is simple; if, for example, $\frac{dW}{db}(b; 0) > 0$ and $\frac{dW}{db}(b; G_1) < 0$, it is clear that $b^*(0) > b^*(G_1)$. For a more general result, however, we need to be able to go beyond the local welfare derivative and make out-of-sample assumptions; as an illustration, consider Figure 5, which displays graphically how knowledge of a local welfare derivative doesn’t permit unambiguous conclusions about the optimum. Chetty (2009) recommends the method of statistical extrapolation that has been used by Baily (1978) and Gruber (1997): for each “sufficient statistic” in the optimal benefit equation, the best available data is used to select a functional form and parametrization of that statistic with respect to $b$, allowing us to extrapolate $\frac{dW}{db}$ out of sample and find the optimum. For this purpose, I define $\chi = \{\frac{\Delta C}{C_i}, R, \eta, E^u_b, E^y_b\}$ as the vector of underlying quantities in (13) which are not exogenously fixed, and let $\chi(b)$ denote a particular vector of extrapolated values of these quantities.\(^{43}\) This leads to the following corollary.

\(^{43}\)Strict quasi-concavity of $W$, when the latter is estimated out of sample using statistical extrapolations,
Corollary 3. For statistical extrapolations that do not depend on the estimated value of $G$, i.e. $\chi(b; G) = \chi(b)$, $b^*(G_1) > b^*(0)$ if and only if $\eta E^u_b - (1 - \eta) E^w_b > 0$, or equivalently if and only if $\frac{d(\eta y_n)}{db} > 0$, in between $b^*(0)$ and $b^*(G_1)$.

Proof. If a statistical extrapolation is used to find $b^*(0)$, and the same statistical extrapolation is used for the case of $G = G_1$, then $\frac{dW}{db}(b^*(0); G_1)$ takes the same sign as $\frac{d(\eta y_n)}{db}$. If that sign is positive, then by strict quasi-concavity $b^*(G_1) > b^*(0)$, and $\frac{d(\eta y_n)}{db}$ will continue to be positive at least until $b^*(G_1)$. If the sign is negative, the opposite is true. 

Therefore, if two researchers using $G = 0$ and $G = G_1$ use the same statistical extrapolations of the sufficient statistics, then the second researcher’s estimated optimal value $b^*(G_1)$ will be the larger of the two if and only if $\frac{d(\eta y_n)}{db} > 0$ in between the optimal values of $b$; the proof explains why the sign of $\frac{d(\eta y_n)}{db}$ will not change in that region.\footnote{This is not, however, a restrictive assumption relying on quasi-concavity; indeed, if everything is continuous, then a marginal increase in the estimated value of $G$ will lead to a marginal change in the optimal $b$ according to the sign of $\frac{d(\eta y_n)}{db}$. Supposing that $\frac{d(\eta y_n)}{db} > 0$, $b$ will only increase as long as it stays in a range where $\frac{d(\eta y_n)}{db} > 0$, so it can never increase out of this range, and thus a marginal change in $b$ cannot reverse this sign; therefore, changes in the estimated $G$ can never move the estimated optimal $b$ enough to change the sign of $\frac{d(\eta y_n)}{db}$. If there is a value of $b$ for which $\frac{d(\eta y_n)}{db}$ takes the opposite sign, there must be no value of $G$ such that this $b$ would be optimal.} This is arguably the most important result in this section, and explains the numerical results from the structural analysis earlier, as well as predicting additional numerical results to come: if UI benefits increase total earnings, this welfare-increasing fiscal externality will appear larger when we account for larger taxes, and we will want to increase benefits; if the effect of UI on wages is zero, however, as has commonly been assumed, the only behavioural effect of benefits will be to increase unemployment, reducing total earnings, and the fiscal externality will be implicitly places some restrictions on the extrapolations allowed.
negative. As we saw in the structural results earlier, the reduction in the optimal benefit level in this case can be quite substantial.

Next, I present results on the role of $E_y^b$, since I claim that the value of this parameter could be quite important; these results are relatively straightforward, and begin with the following proposition.

**Proposition 3.** For $E_y^{y2} > E_y^{y1}$, $\frac{dW}{db}(b; G, E_y^{y2}) > \frac{dW}{db}(b; G, E_y^{y1})$.

**Proof.** The proof of this result can be found in appendix B.3.

A higher value of $E_y^b$ means higher $b$ has a more positive effect on wages, meaning a smaller tax increase to pay for benefits, and thus a larger welfare gain from higher UI. The two following corollaries follow the pattern of the previous two, and pose no additional difficulties of interpretation.

**Corollary 4.** For the current value of $b$, if $\frac{dW}{db}(b; G, E_y^{y1}) = 0$ or $\frac{dW}{db}(b; G, E_y^{y2}) = 0$, or if $\frac{dW}{db}(b; G, E_y^{y1}) < 0$ and $\frac{dW}{db}(b; G, E_y^{y2}) > 0$, $b^*(G, E_y^{y2}) > b^*(G, E_y^{y1})$.

**Corollary 5.** For statistical extrapolations of $\chi_1 = \{\Delta C, R, \eta, E_u^b\}$ that do not depend on the estimated value of $E_y^b$, i.e. $\chi_1(b; E_y^b) = \chi_1(b)$, $b^*(G, E_y^{y2}) > b^*(G, E_y^{y1})$.

If we are at the estimated optimal value of $b$ for one of the values of $E_y^b$ under consideration, or if the signs of $\frac{dW}{db}$ are opposite, then we can make an unambiguous statement; more generally, once we define a statistical extrapolation that does not depend on $E_y^b$, we can state that a researcher choosing a larger value of $E_y^b$ will always find a larger optimal $b$, exactly as one would expect.

I have now shown that higher $E_y^b$ increases the optimal value of $b$, and found the conditions under which a higher value of $G$ increases or decreases the optimal $b$; the final analytical results consider the interaction of $G$ and $E_y^b$. As already mentioned, Baily (1978) is among several papers that acknowledge the fact that a parameter like $E_y^b$ enters into optimal UI calculations, but this effect is ignored on the grounds that, as the UI payroll tax is quite small, it will have little effect. However, my contention is that $E_y^b$ is much more important when $G$ is large, both to welfare and to the calculation of the optimal value of $b$, and I will now attempt to prove this, starting with the next proposition.
**Proposition 4.** For $G_1 > 0$ and $E_b^{y_2} > E_b^{y_1}$, $\frac{dW}{db}(b; G_1, E_b^{y_2}) - \frac{dW}{db}(b; G_1, E_b^{y_1}) > \frac{dW}{db}(b; 0, E_b^{y_2}) - \frac{dW}{db}(b; 0, E_b^{y_1})$.

*Proof.* The proof of this result can be found in appendix B.4.

This result can be summarized as saying that the effect of $E_b^{y}$ on the welfare derivative is increasing in $G$; it may be true that a researcher that ignores $G$ will find that the value of $E_b^{y}$ is relatively unimportant to their calculations, but when $G$ is potentially quite large, the tax rate will also be large, and $E_b^{y}$ will matter far more to the value of the welfare derivative. Alternatively, Proposition 4 could be interpreted as saying that the importance of $G$ to the welfare derivative is increasing in $E_b^{y}$.

There is one question left to answer: is $E_b^{y}$ more important, in some sense, to determining the value of the optimal $b$ when $G$ is large? For instance, is $b^*(G_1; E_b^{y_2}) - b^*(0; E_b^{y_2})$ increasing in $E_b^{y_2}$? This would fit with intuition; we already know that $b^*$ is increasing in $E_b^{y}$, and since the difference in the welfare derivative for different values of $E_b^{y}$ is increasing in $G$, it would be natural to assume that the increase in $b^*$ caused by $E_b^{y}$ would be larger when $G$ is larger. This, however, is not easy to prove without unusual and unintuitive assumptions; I can, however, prove a somewhat weaker result, as summarized below.

**Proposition 5.** For statistical extrapolations that do not depend on the estimated values of $G$ and $E_b^{y}$, if $\frac{dE_u}{db} \geq 0$, $\frac{dR}{db} = 0$, $E_u > -1$, and $\frac{d}{db} \left( \frac{\Delta C}{C} \right) > 0$, the following is true:

- $b^*(G_1, E_b^{y_2}) > b^*(0, E_b^{y_2})$ and $b^*(G_1, E_b^{y_1}) < b^*(0, E_b^{y_1})$ for $E_b^{y_2} > E_b^{y*} > E_b^{y_1}$, where $E_b^{y*}$ is defined such that $b^*(G_1, E_b^{y*}) = b^*(0, E_b^{y*})$, if such an $E_b^{y*}$ exists.

*Proof.* The proof of this result can be found in appendix B.4.

The appendix discusses the sufficient conditions in further detail, and explains why they are plausible and that they hold for nearly all of the non-zero numerical results in the next section. This proposition tells us that, although we can’t prove the stronger condition that $b^*(G_1; E_b^{y_2}) - b^*(0; E_b^{y_2})$ is increasing in $E_b^{y_2}$, we can state that for small values of $E_b^{y}$, $b^*(G_1, E_b^{y_2}) < b^*(0, E_b^{y_2})$, and vice-versa for sufficiently large values of $E_b^{y}$.

\[45\text{In the unlikely case that an increase in } E_b^{y} \text{ causes an increase in the optimal } b \text{ which makes } \eta \text{ decrease sufficiently quickly, the critical value } E_b^{y*} \text{ may not exist; in that case, } b^*(G_1) - b^*(0) \text{ is always negative, assuming that } E_b^{u} > 0.\]
summarized by the diagram in Figure 6. Therefore, at least locally around $E_b^{y^*}$, the stronger result is true; furthermore, in my numerical results, we will see that the stronger property does describe the behaviour of $b^*$ for the functional forms and parametrization that I use.

Figure 6: Consequences of Proposition 5

This series of results will now be demonstrated in the next section, as I proceed to perform the numerical evaluation of (13).

6 Numerical Results from Baily (1978) Model

My discussion of the numerical analysis of the Baily model begins with a brief subsection describing the method of statistical extrapolation; I then proceed to select the parameter values in the second subsection, and to evaluate the optimal benefit levels in the third subsection. The final two subsections contain a comparison of the results from this method with those obtained earlier from the dynamic job search model, and a discussion of the results.
6.1 Summary of Numerical Procedure

My goal in this section is to numerically evaluate (13) to find the optimal benefit level, and I begin by describing the procedure of statistical extrapolation used for this purpose. First of all, the optimal UI literature overwhelmingly solves for an optimal replacement rate rather than a dollar value of UI, so I will do the same here, as I did earlier in the structural analysis; as before, I define the replacement rate as 

\[ r = \frac{b}{(0.8)(\frac{15.8}{24.3})y(1-\tau_0)} \]

where \( \tau_0 \) is the true baseline real-world tax rate, which is therefore exogenously given, 0.8 is the take-up rate, and \( \frac{15.8}{24.3} \) is the ratio of mean compensated unemployment duration to mean total duration, and so \( r \) is simply a constant times \( b \).\(^{46}\) The steps in the procedure used to solve (13) for the optimal replacement rate are as follows, and in square brackets I provide an example of one of the sets of values and statistical extrapolations used later:

- select an equation for \( \frac{\Delta C}{C_s} \) as a (continuous) function of \( r \) \( \left[ \frac{\Delta C}{C_s} = 0.222 - 0.265r \right] \)
- select fixed values of \( E_u^b \) [0.2544], \( E_y^b \) [0.048] and \( R \) [2]
- select current values of \( r \) [0.46], \( u \) [0.054] and \( \eta \) [0.8]
- use \( 2u = (1 - \alpha)(1 - \eta) \) to solve for the fixed value of \( \alpha \)
- use the fixed value of \( E_u^b \) to define a functional form for \( \eta \) with respect to \( r \): \( (1 - \eta) = \phi r^{E_u^b} \), and use the current values of \( \eta \) and \( r \) to solve for \( \phi \) [0.2437]
- select the current value of \( \psi \) [0.0457], and specify the relationship of \( \psi \) to \( r \)

\[ \left[ \psi = \frac{0.7712.64}{0.2324.3} \frac{ur}{1-u} \right] \]

- solve the resulting non-linear equation in \( r \)

Discussion of the specific values chosen for the various quantities will be postponed to the next subsection; this example is meant simply to illustrate the method. If we proceed to use the values listed above to solve for the optimal \( r \), we obtain a value of approximately 0.393, or a replacement rate of 39.3%. On the other hand, if we assume that \( G = 0 \), this affects only the second-last step: \( \psi = 1 \) for all values of \( r \), and the optimal \( r \) would be

\(^{46}\)My goal is simply to normalize the benefit level to the units typically used in the literature.
approximately 0.472. However, now suppose we had initially chosen a larger value of $E_y^b$, such as 0.096; now the optimal values of $r$ are 0.594 when we use the estimate of $G$ and 0.484 when we assume $G = 0$. A higher value of $E_y^b$ leads to a higher optimal $b$ for each value of $G$; meanwhile, in the first case, with a relatively small value of $E_y^b$, using a larger estimate of $G$ reduces the optimal replacement rate, but with a larger value of $E_y^b$, the opposite happens, exactly as Proposition 5 suggested.

I can now use this procedure to provide numerical results for a range of plausible values of the sufficient statistics, to illustrate the analytical results and to further demonstrate how large an effect fiscal externalities can have on optimal UI calculations. The next two subsections will deal with these issues.

### 6.2 Sufficient Statistics and Extrapolations

I choose the values and functional forms according to the procedure described above. Many of these quantities have already been used earlier in the paper, so the discussion of most parameters will be kept brief.

Starting with the functional form of $\frac{\Delta C}{C^1}$, as before I use Gruber (1997), who estimates $\frac{\Delta C}{C^1} = 0.222 - 0.265r$. I also continue to use $E_y^b = 0.48 \times 0.53 = 0.2544$ from Chetty (2008) and Gruber (1997), and $R \in \{2, 5\}$.

For the elasticity of post-unemployment wages with respect to benefits, the empirical literature is fairly sparse and reports a wide range of results, which are summarized in Table 5; since I am focussing on the effect of higher benefit levels on wages, I list only papers addressing that specific question, though there are several papers which estimate the effect

---

47 None of the papers listed report coefficients in the form of an elasticity, so their coefficients (typically in log-level or level-level form) have been transformed into approximate elasticities using mean values of wages and/or benefit levels/replacement rates. Classen (1977) and Holen (1977) do not provide summary statistics for their dataset, so I use mean values from Burgess and Kingston (1976), which uses a smaller version of the same dataset used by Holen (1977).

48 Not listed are Blau and Robins (1986), who find a moderately large but not significant effect of UI benefits on the wage offer distribution, plus a positive effect of UI on reservation wages; Fitzenberger and Wilke (2007), who perform a Box-Cox quantile regression and do not arrive at a single estimate; and McCall and Chi (2008), which allows the effect of UI on wages to change over the spell of unemployment, and whose findings correspond to an initial elasticity of 0.10 which declines over time. Additionally, the estimate listed for Meyer (1989) is from one of 10 individual regressions which could be used, ranging from negative and marginally significant to positive and statistically insignificant (though economically significant); the author does not provide an indication of which is his preferred estimate, so the basic difference-in-differences estimate is used.
of longer benefit durations on wages, or the effect of some dimension of benefit generosity on other measures of job quality.\footnote{Among these, for instance, are Gaure, Roed, and Westlie (2008), which finds a positive effect of benefit duration on wages, and Lalove (2007), which does not; Centeno (2004), Centeno and Novo (2006), and Tatsiramos (2009), which find that a more generous UI system leads to greater subsequent job duration, and Portugal and Addison (2008), which does not.} In particular, Chetty (2008) focusses on two recent papers: Card, Chetty, and Weber (2007) and van Ours and Vodopivec (2008), which use natural-experiment methodologies to test for an effect of the potential duration of UI benefits on wages, using European data (from Austria and Slovenia respectively), and which find no significant effects. While the findings of these papers are at least suggestive, they may not be definitive in a North American context, given the different labour market structures and institutions found in Europe, such as higher union coverage, as acknowledged by Card, Chetty, and Weber (2007).

<table>
<thead>
<tr>
<th>Paper</th>
<th>Approx. Elasticity</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ehrenberg and Oaxaca (1976)</td>
<td>0.27 for older men</td>
<td>(0.12, 0.43)</td>
</tr>
<tr>
<td></td>
<td>0.06 for older women</td>
<td>(0.03, 0.09)</td>
</tr>
<tr>
<td></td>
<td>0.04 for young men</td>
<td>(-0.04, 0.12)</td>
</tr>
<tr>
<td></td>
<td>0.02 for young women</td>
<td>(-0.06, 0.10)</td>
</tr>
<tr>
<td>Burgess and Kingston (1976)</td>
<td>0.45</td>
<td>(0.26, 0.64)</td>
</tr>
<tr>
<td>Classen (1977)</td>
<td>0.03</td>
<td>(-0.16, 0.21)</td>
</tr>
<tr>
<td>Holen (1977)</td>
<td>0.64</td>
<td>(0.55, 0.72)</td>
</tr>
<tr>
<td>Meyer (1989)</td>
<td>-0.17</td>
<td>(-1.03, 0.69)</td>
</tr>
<tr>
<td>Maani (1993)*</td>
<td>0.11</td>
<td>(0.02, 0.20)</td>
</tr>
<tr>
<td>Addison and Blackburn (2000)</td>
<td>-0.05</td>
<td>(-0.14, 0.05)</td>
</tr>
</tbody>
</table>

*Maani (1993) uses data from New Zealand; all other papers in this table use American data.

More recent studies do seem to suggest smaller values of $E^y_b$, possibly around zero, but since the literature covers a wide range of values, and the analytical results in section 5 suggest that $E^y_b$ can have an important effect on the optimal benefit level, I use a set of possible values for $E^y_b$ covering the range found in Table 5,\footnote{Since estimates of $E^y_b$ are based on samples of UI recipients, I once again multiply by 0.48.} specifically $E^y_b = 0.48 \times \{-0.17, 0, 0.1, 0.2, 0.4, 0.64\}$.\footnote{Although it is often assumed that higher UI benefits will lead to higher wages, the elasticity could be negative; Welch (1977) predicts a reduction of wages by firms to compensate for higher benefits which they must finance due to experience rating.}

The baseline value of $r$ is set to 0.46 as before, and once again I use an initial unemploy-
ment rate of \( u = 0.054 \). The value used for \( \eta \), meanwhile, depends on the way the structure of the model is interpreted. If the model’s two-period structure is taken literally as representing two years, then I can use the finding of Chetty (2008) that the mean unemployment duration of individuals in his sample is 18.3 weeks to derive an estimate of \( \eta = \frac{52 - 18.3}{52} = 0.648 \). If, however, I interpret the model as representing a larger portion of an individual’s working life, perhaps its entirety, I would expect to find a larger value for \( \eta \); in this case, an approximation for \( \alpha \) can be obtained from the fact that Farber (1999) finds that 20.9\% of workers aged 45-64 had at least 20 years of tenure in 1996, so \( \eta = 1 - \frac{2u}{1-\alpha} = 0.864 \). To cover this range of possibilities, I use the set of values given by \( \eta = \{0.648, 0.725, 0.8, 0.864\} \).

The values for \( \alpha \) and \( \phi \) for each value of \( \eta \) are now easy to calculate, so I will not discuss them further here. Finally, I need an estimate for \( \psi \). At baseline values, \( \psi = \frac{ub}{(1-u)\tau y} = \frac{u}{1-u} \frac{1-\tau}{\tau} (0.8 \left(\frac{15.8}{24.3}\right) r, and the baseline tax rate once again is \( \tau = 0.23 \), so because the tax rate does not change much with \( r \), I use \( \psi = \frac{0.77 \cdot 12.64 \cdot ur}{0.23 \cdot 24.3 \cdot 1-u} \) , which at baseline is equal to 0.0457.

The parameter values are summarized in Table 6. I am now prepared to solve the non-linear first-order condition (13) defining the optimal replacement rate, an exercise which I pursue in the following subsection.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value/Extrapolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{\Delta C}{C} )</td>
<td>0.222 – 0.265r</td>
</tr>
<tr>
<td>( E_b^u )</td>
<td>0.2544</td>
</tr>
<tr>
<td>( E_{by} )</td>
<td>0.48 × {−0.17, 0, 0.1, 0.2, 0.4, 0.64}</td>
</tr>
<tr>
<td>( R )</td>
<td>{2, 5}</td>
</tr>
<tr>
<td>( r )</td>
<td>0.46</td>
</tr>
<tr>
<td>( u )</td>
<td>0.054</td>
</tr>
<tr>
<td>( \eta )</td>
<td>{0.648, 0.725, 0.8, 0.864}</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>{0.6932, 0.6073, 0.4600, 0.2059}</td>
</tr>
<tr>
<td>( \phi )</td>
<td>{0.4289, 0.3351, 0.2437, 0.1657}</td>
</tr>
<tr>
<td>( \psi )</td>
<td>( \frac{0.77 \cdot 12.64 \cdot ur}{0.23 \cdot 24.3 \cdot 1-u} = 0.0457 )</td>
</tr>
</tbody>
</table>

### 6.3 Optimal Replacement Rates

This subsection presents the results of the numerical analysis, in the form of optimal replacement rates for each of the cases under consideration. I solve for the optimal value of
and I report a numerical check on the second-order conditions in appendix C. Tables 7 and 8 below present the optimal $r$ for my parameter values, as well as the results when I set $G = 0$. Of course, $G = 0$ does not perfectly reproduce Baily’s results; to do so, I would also need to make the extra assumptions made by Baily, in which case the results vary only with $R$, giving us $r = 0.3577$ for $R = 2$ and $r = 0.6457$ for $R = 5$.

Table 7: Optimal Replacement Rates Calculated from (13) for $R = 2$

<table>
<thead>
<tr>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Value of $\eta$</td>
<td>$\eta$</td>
</tr>
<tr>
<td>$-0.0816$</td>
<td>$0.4500$</td>
</tr>
<tr>
<td>$0$</td>
<td>$0.4595$</td>
</tr>
<tr>
<td>$E_b$</td>
<td>$0.048$</td>
</tr>
<tr>
<td>$0.096$</td>
<td>$0.4707$</td>
</tr>
<tr>
<td>$0.192$</td>
<td>$0.4821$</td>
</tr>
<tr>
<td>$0.3072$</td>
<td>$0.4959$</td>
</tr>
</tbody>
</table>

Table 8: Optimal Replacement Rates Calculated from (13) for $R = 5$

<table>
<thead>
<tr>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Value of $\eta$</td>
<td>$\eta$</td>
</tr>
<tr>
<td>$-0.0816$</td>
<td>$0.6831$</td>
</tr>
<tr>
<td>$0$</td>
<td>$0.6866$</td>
</tr>
<tr>
<td>$E_b$</td>
<td>$0.048$</td>
</tr>
<tr>
<td>$0.096$</td>
<td>$0.6909$</td>
</tr>
<tr>
<td>$0.192$</td>
<td>$0.6951$</td>
</tr>
<tr>
<td>$0.3072$</td>
<td>$0.7004$</td>
</tr>
</tbody>
</table>

The difference between the $G = 0$ and $G = 0.208$ case is especially large for $R = 2$; for low values of $E_b^y$, allowing for fiscal externalities from the income tax causes the optimal replacement rate to drop to zero, whereas for higher values of $\eta$ and especially $E_b^y$, the replacement rate increases significantly, perhaps even above one. The numerical results are less extreme for $R = 5$, but the same pattern of findings is present there as well: when $G > 0$,

\footnote{I assume that the government is not interested in extracting payments from unemployed workers, and so any zeros reported in the tables are corner solutions. The choice of $r = 2$ as an upper limit is not a binding constraint except in the case of the extension in appendix E.3, where the squared consumption drop term presents the possibility, for some parameter values, of marginal utility seeming to increase without bound as $r$ gets large; this is due to the use of the local quadratic approximation for marginal utility, which becomes increasingly inaccurate as $C_u$ diverges from $C_1^*$.}
the optimal replacement rates spread out noticeably, becoming more sensitive to both \( E_b \) and \( \eta \).

For comparison with results from Chetty (2008), who only reports the baseline value of the welfare derivative, I also report the baseline values of \( \frac{dW}{db} \) in Tables 12 and 13 in appendix D, though the direct comparison to that paper is limited by the fact that the models are different, as well as the way marginal welfare is normalized into dollars (Chetty divides by marginal utility when re-employed, while I divide by \( U'(C_u) \)). The results are qualitatively similar to those in the tables above: for \( R = 2 \), values of \( \frac{dW}{db} \) cluster around zero when \( G = 0 \) but range from -0.08 to 0.18 when \( G > 0 \), and for \( R = 5 \), values around 0.045 when \( G = 0 \) spread out to cover the range from 0 to 0.15.

### 6.4 Comparison with Structural Results

Tables 6 and 7 display results for a wide variety of cases, but it is also of some interest to compare these results to those obtained from the structural analysis of the dynamic job search model earlier in the paper, for the specific subset of cases considered there. To do so, I use the specific values of the moments generated by the calibrated model, which can be found in Table 11, and input them into the statistical extrapolation method. This provides the results displayed in Table 9; the left side of the table contains the structural results and is simply a replication of Table 4, while the right side displays my new results, where the welfare gains once again represent a percentage of total UI spending.

**Table 9: Comparison of Results with \( E_b = 0 \)**

<table>
<thead>
<tr>
<th></th>
<th>Structural Model</th>
<th>Statistical Extrapolations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( R = 2 )</td>
<td>( R = 5 )</td>
</tr>
<tr>
<td></td>
<td>( G = 0.208 )</td>
<td>( G = 0 )</td>
</tr>
<tr>
<td></td>
<td>( 0.00 )</td>
<td>( 0.36 )</td>
</tr>
<tr>
<td>Welf. Gain</td>
<td>12.69</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>( 0.44 )</td>
<td>( 0.71 )</td>
</tr>
<tr>
<td>Welf. Gain</td>
<td>0.13</td>
<td>5.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td>6.92</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diff.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results for the optimal replacement rate are remarkably similar except for the case with \( R = 5 \) and an estimate of the real \( G \); in the latter case, the difference is generated partly by the fact that \( E_b \) increases with \( b \) in the dynamic job search model, whereas it is held fixed when I perform statistical extrapolations, but other differences in the models also
are likely responsible for part of the difference. Meanwhile, the estimates for the welfare gains tend to be larger than those from the simulations, again largely due to differences in models and assumptions about elasticities; in the $R = 2$ case, the estimated welfare gain of moving from $r = 0.46$ to $r = 0$ amounts to nearly 50% of initial UI spending, which is unrealistically large, and which is due to the assumption that unemployment goes to zero as $r \to 0$.\footnote{If, instead of allowing $\frac{dW}{db}$ to become very negative as $r \to 0$, I hold it constant at its $r = 0.33$ value (which is the peak of $\frac{dW}{db}$), which is likely a conservative assumption, the welfare gain drops to about 20%.

\section{Discussion of Numerical Results}

Let us now return to the results in Tables 7 and 8, and consider their meaning; first of all, the prediction of Proposition 5 that the optimal replacement rate should be lower for higher $G$ if $E^y_b$ is relatively small, but higher if $E^u_b$ is sufficiently positive, is strongly supported by the numerical results. Additionally, at least for $r > 0$, the results also support the stronger condition that $b^*(G_1) - b^*(0)$ is increasing in $E^u_b$. The prediction of Corollary 3 is also supported for all sets of parameter values: when $b^*(G_1) > b^*(0)$, $\eta E^u_b - (1 - \eta) E^u_b$ is found to be positive for all $b \in [b^*(0), b^*(G_1)]$, and vice-versa when $b^*(G_1) < b^*(0)$.

Practically speaking, however, how should we interpret these numerical results? For certain parameter values, a positive $G$ makes very little difference to the optimal replacement rate, because the two fiscal externality components of $E^\tau_b$ identified in (12) roughly cancel each other out. I would argue, however, that even if one’s prior fits one of these cases, it would be premature to conclude that we need not worry about $G$, until the literature has reached a greater consensus on the values of relevant parameters. This is because, in other cases, the difference in optimal replacement rates is quite large, and recent trends in the empirical literature suggest that some of these cases are quite plausible. The most typical values assumed for $R$ and $E^u_b$ in recent work would likely be 2 and 0, respectively, and my results in that case would indicate that fiscal externalities cause the optimal replacement rate to drop from about 0.46 to zero. On the other hand, if $E^u_b$ is as high as some papers have estimated, the optimal level of benefits could increase significantly; the situation is different again if $R = 5$. It is therefore clear that the values of certain parameters, particularly $E^u_b$,
are quite important to optimal UI calculations, and that even modest variation in the values of these parameters over a plausible range could be significant.

Some of the more extreme numbers, such as replacement rates of zero or over one, are unlikely to be completely realistic; the latter arise because large wage elasticities essentially transform UI from an insurance program into a job-match-upgrading program, and are dependent on my assumptions, such as the assumption that elasticities are constant out of sample, as well as the assumption of perfect credit markets and rational savings behaviour. But while the exact numbers may be questioned, but the central result is clear: full consideration of the fiscal externality effects of UI benefits can substantially alter the nature of the optimal UI problem and significantly change the numerical results.

Given the simplicity of the Baily model, I also perform a number of extensions, which can be found in appendix E: I allow for stochastic duration of unemployment, a restriction on borrowing during unemployment, the use of a second-order Taylor series expansion of marginal utility, and variable labour supply on the initial job. Although the numerical results do change somewhat (the first two extensions tend to move the optimal replacement rate closer to one, while the latter two reduce it), the results are still quite similar, and the qualitative conclusions are the same; the pairwise comparisons of optimal replacement rates given $G = 0$ versus $G = 0.208$ are nearly identical in each case.

7 Conclusion

The optimal UI literature has explored many of the aspects of the design and generosity of unemployment insurance systems, but there has not yet been a serious effort to account for the full role of fiscal externalities; it is this gap which I have attempted to fill with the current paper. My results have demonstrated how substantial an impact fiscal externalities resulting from income taxes may have on optimal UI calculations, and have indicated the previously unrecognized importance of parameters such as the elasticity of post-unemployment wages with respect to UI benefits; if this elasticity is significantly positive, recognition of greater fiscal externalities will cause us to adjust our estimates of the optimal replacement rate upwards, whereas we may prefer a significantly lower benefit level if the wage elasticity is small or even negative. My specific numerical findings indicate that a typical set of
parameters including a zero wage elasticity will tend to cause a large decrease in the optimal replacement rate, from around 0.4 to zero. The welfare gains from moving to the new optimum are estimated to be on the order of 10% of current government spending on UI.

These conclusions suggest the need for further empirical research on the effects of UI benefits on job matching and/or wages. At a minimum, it is hoped that this paper will cause future research to examine more closely the second-best nature of the design of labour market programs like UI. My insights can also be used in other areas of government policy, and work is currently underway in Lawson (2012a) to apply a similar analysis to post-secondary education.

A Technical Appendix for Structural Calibration and Simulation

Table 10 contains the parameters used in each case; Table 11 displays the moments calculated. In each case, the upper limit of the asset distribution was chosen so as to not be binding for all relevant cases, while the number of knots in the cubic spline was chosen so that increasing it further made no difference to the results. The spacing of the asset distribution was set at 0.005. Tests were made of all convergence parameters to ensure that further tightening had no non-negligible effect on results.

<table>
<thead>
<tr>
<th>Table 10: Calibrated Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = 2$</td>
</tr>
<tr>
<td>$G = 0.208$</td>
</tr>
<tr>
<td>$\rho$</td>
</tr>
<tr>
<td>$\theta$</td>
</tr>
<tr>
<td>$\kappa$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 11: Calculated Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = 2$</td>
</tr>
<tr>
<td>$G = 0.208$</td>
</tr>
<tr>
<td>$u$</td>
</tr>
<tr>
<td>$E_u$</td>
</tr>
<tr>
<td>$C_{e} - C_u$</td>
</tr>
</tbody>
</table>

B Algebra and Proofs

B.1 Proof of Proposition 1

The individual’s first-order condition for saving is given by:

$$\frac{\partial V}{\partial k} = -U'(C_e) + \alpha U'(C_e^2) + (1 - \alpha)U'(C_u) = 0$$
and I also need a first-order Taylor series expansion of $U'(C^1_e)$ around $U'(C_u)$:

$$U'(C^1_e) = U'(C_u) + \Delta C U''(\theta)$$

where $\theta$ is between $C_u$ and $C^1_e$, and $\Delta C = C^1_e - C_u$. Using both of these allows me to rewrite (6) as:

$$\frac{dW}{db} = 2y\Delta C U''(\theta)\frac{d\tau}{db} - [(1 + \alpha)y + (1 - \alpha)y_n]U'(C_u)\left[\frac{d\tau}{db} - \omega\right]$$

where $\omega = \frac{(1-\alpha)(1-\eta)}{(1+\alpha)y + (1-\alpha)y_n}$.

Next, I make two assumptions that are also found in Baily (1978). The first is that, in equilibrium, $y_n = y$, so that I can write all wages as $y$; the second is that $C^1_u U''(\theta) = C_u U''(C_u)$, which permits the second derivative of utility term to be incorporated into a coefficient of relative risk-aversion. The first assumption is generally conservative, because it overweights the income lost from unemployment; the second is described by Baily as conservative, but this is not always true, and for the values of the risk-aversion parameter that I use, it will tend to be a liberal assumption if utility is CRRA because it overstates the consumption-smoothing benefit implied by $U''(\theta)$. Combining these two assumptions, and dividing by $U'(C_u)$ to put the welfare derivative in dollar terms:

$$\frac{dW}{db} \equiv \frac{dV}{db} = 2y\Delta C C^1_e \frac{d\tau}{db} - 2(1 - u)y\left[\frac{d\tau}{db} - \omega\right]$$

where $R = \frac{-C_u U''(C_u)}{U'(C_u)}$ is the coefficient of relative risk-aversion, and $u = \frac{(1-\alpha)(1-\eta)}{2}$ is the unemployment rate. At the optimum, $\frac{dW}{db} = 0$, and this will be a unique optimum if $W$ is strictly quasi-concave; thus, the expression for the optimum is:

$$\frac{\Delta C}{C^1_e} R = (1 - u) \frac{d\tau}{db} - \omega.$$ 

### B.2 Proof of Proposition 2

Starting from (9), I can write:

$$\frac{dW}{db}(b; G) = \frac{2u}{1 - u} \left[ \frac{\Delta C}{C^1_e} R \frac{E^u}{\psi} - (1 - u) \left( \frac{E^u}{\psi} - 1 \right) \right]$$

and therefore the difference in welfare derivatives can be written as:

$$\frac{dW}{db}(b; G_1) - \frac{dW}{db}(b; 0) = \frac{2u}{1 - u} \left[ \frac{\Delta C}{C^1_e} R - (1 - u) \right] \left[ \left( \frac{E^u}{\psi} \right)_{G=G_1} - \left( \frac{E^u}{\psi} \right)_{G=0} \right].$$

Using (12) and the definition of $\psi$:

$$\frac{E^u}{\psi} = 1 + E^u + \frac{2ub + G}{2ub} \left[ \frac{(1 - \alpha)(1 - \eta)}{2(1 - u)} E^u - \frac{(1 - \alpha)\eta}{2(1 - u)} E^u \right]$$

and thus the welfare derivative difference becomes:

$$\frac{dW}{db}(b; G_1) - \frac{dW}{db}(b; 0) = \frac{(1 - \alpha)}{2(1 - u)^2 b} \left[ \frac{\Delta C}{C^1_e} R - (1 - u) \right] \left[ (1 - \eta) E^u - \eta E^u \right] G_1.$$

As I have assumed that $\frac{\Delta C}{C^1_e} R < 1 - u$, this will be positive if and only if $\eta E^u - (1 - \eta) E^u$ is positive. This latter expression can also be written as:

$$\eta E^u - (1 - \eta) E^u = \eta \frac{dy_n}{db} + b \frac{dy_n}{db} = \frac{b}{y_n} \frac{d(\eta y_n)}{db}.$$

and therefore $\frac{dW}{db}(b; G_1) - \frac{dW}{db}(b; 0)$ has the same sign as $\frac{d(\eta y_n)}{db}$.
B.3 Proof of Proposition 3

Once again, I start from (9); combined with (12), I immediately get:

\[
\frac{dW}{db}(b; G, E^2_b) - \frac{dW}{db}(b; G, E^1_b) = \frac{2u}{(1-u)\psi} \left[ \frac{\Delta C}{C_e} R - (1-u) \right] \left[ \frac{-(1-\alpha)\eta}{2(1-u)} \right] \left[ E^2_b - E^1_b \right].
\]

(14)

Given that \( E^2_b > E^1_b \), and that the middle two terms are both negative, this expression is always positive.

B.4 Proof of Propositions 4 and 5

I start with (14), and from the definition of \( \psi \):

\[
\left[ \frac{dW}{db}(b; G_1, E^2_b) - \frac{dW}{db}(b; G_1, E^1_b) \right] - \left[ \frac{dW}{db}(b; 0, E^2_b) - \frac{dW}{db}(b; 0, E^1_b) \right] = \frac{-(1-\alpha)\eta}{2(1-u)^2b} \left[ \frac{\Delta C}{C_e} R - (1-u) \right] \left[ E^2_b - E^1_b \right] G_1.
\]

This equation is clearly positive.

To prove Proposition 5, my task is somewhat more complicated. I begin with the fact that, at the optimum, \( \frac{\Delta C}{C_e} R E^*_b = (1-u)(E^*_b - \psi) \); using (12) and rearranging, this becomes:

\[
\frac{\Delta C}{C_e} R \psi(1 + E^*_b) - (1-u)\psi E^*_b = \frac{1-\alpha}{2(1-u)} [\eta E^*_b - (1-\eta) E^*_b] \left[ \frac{\Delta C}{C_e} R - (1-u) \right].
\]

We can observe that, because \( \frac{\Delta C}{C_e} R < 1-u, \eta E^*_b - (1-\eta) E^*_b > 0 \) if and only if \( \frac{\Delta C}{C_e} R(1+E^*_b)-(1-u)E^*_b < 0 \). I wish to show that \( b^*(G_1, E^2_b) > b^*(0, E^2_b) \) and \( b^*(G_1, E^1_b) < b^*(0, E^1_b) \) for \( E^2_b > E^y > E^1_b \), so Proposition 2 says that \( \eta E^*_b - (1-\eta) E^*_b \) must be positive for \( E^2_b \) and negative for \( E^1_b \). Then, if I define \( X(b) = \frac{\Delta C}{C_e} R(1+E^*_b)-(1-u)E^*_b \), I want \( X < 0 \) for \( E^2_b \) and \( X > 0 \) for \( E^1_b \); given that I am considering continuous statistical extrapolations, a sufficient and necessary condition is that \( \frac{dX}{db} < 0 \) at \( X = 0 \). The derivative is:

\[
\frac{dX}{db} = R(1+E^*_b) \frac{d(\frac{\Delta C}{C_e})}{db} + \frac{\Delta C}{C_e}(1+E^*_b) \frac{dR}{db} + \frac{\Delta C}{C_e} R \frac{dE^*_b}{db} - E^*_b \frac{d(1-u)}{db} - (1-u) \frac{dE^*_b}{db}
\]

and at \( X = 0 \), \( \frac{\Delta C}{C_e} R(1+E^*_b) = (1-u)E^*_b \), and thus:

\[
\frac{dX}{db} \bigg|_{X=0} = \frac{\Delta C}{C_e}(1+E^*_b) \frac{dR}{db} \bigg[ \frac{\Delta C}{C_e} R - (1-u) \bigg] \frac{dE^*_b}{db} + \frac{R(1+E^*_b)}{(1-u)} \left\{ \frac{1-u}{\frac{d}{db} \left( \frac{\Delta C}{C_e} \right)} - \frac{\Delta C}{C_e} \frac{d(1-u)}{db} \right\}.
\]

Sufficient conditions for this to be negative are that \( \frac{dR}{db} = 0, \frac{dE^*_b}{db} \geq 0, 1+E^*_b > 0 \) and \( 1-u \frac{d(\frac{\Delta C}{C_e})}{db} < \frac{\Delta C}{C_e} \frac{d(1-u)}{db} \). The first two assumptions are standard, and I will make them later in the numerical analysis; \( \frac{dR}{db} \) is commonly assumed to equal zero, as it would with a CRRA utility function, and Chetty (2006) states that estimates of \( \frac{dE^*_b}{db} \) “are broadly similar across studies with different levels of benefit generosity.” The third assumption is merely a formality, as a nearly universal finding of the empirical literature is that \( E^*_b \) is positive. The final assumption requires a bit more explanation; it is easiest to understand when written as

\[
\frac{d}{db} \left( \frac{\Delta C}{C_e} \right) < 0,
\]

where a bit of thought makes clear that this is likely to be satisfied in nearly every case that we might be interested in: the consumption gap \( \frac{\Delta C}{C_e} \) is likely to be much smaller than \( 1-u \), and to decline faster, as the former is always less than one and could reach or even drop below zero, whereas \( 1-u \) is always at least as large as \( \frac{1-\alpha}{2} \). In practice, this assumption is also strongly supported by the numerical results in section 5 in all cases in which the optimal replacement rate is above zero; my functional form assumptions cause \( \frac{d(1-u)}{db} \) to become unboundedly large and negative as benefits approach zero.
C Second-Order Conditions

Throughout the analytical results in section 5, I assume strict quasi-concavity, which ensures that the first-order condition (13) identifies the unique maximum. In the numerical results, I can test this assumption by plotting the estimated value of \( \frac{dW}{db} \) at intervals of 0.01 for \( r \in [0.01, 2] \), for each set of potential parameter values, and for the initial model as well as all extensions; I can then see if any failures of quasi-concavity appear over that range (outside that range, failures of quasi-concavity might be expected on the grounds that my assumptions become especially poor approximations).

For the baseline model, quasi-concavity fails in several cases when \( G = 0.208 \): for \( R = 5 \) and \( E_y \geq 0.048 \), there often appears to be a local minimum at low values of \( r \) (always less than 0.08). It should not be surprising, however, that these violations of strict quasi-concavity occur at low values of \( r \) when \( R \) is large, as those are exactly the values at which \( \Delta C \) may fail to hold, and my assumptions will tend to be most inaccurate, and at which the estimated \( E_y^b \) could turn negative. Over the vast majority of the range of \( r \) that I consider, however, \( \frac{dW}{db} \) seems to behave normally and in a way consistent with quasi-concavity.

Further failures of quasi-concavity are observed in each of the extensions in appendix E; in the first, second, and fourth extensions, local minima are also observed for \( R = 5 \) and \( E_y \geq 0.048 \). The first extension also presents cases for \( R = 2, E_y^b = 0 \) and \( \eta = \{0.648, 0.725\} \) where local maxima are observed at positive values of \( r \) but the global maximum is at \( r = 0 \), and similar cases are observed thrice in the second extension at high values of \( E_y^b \) and \( \eta \), as mentioned in appendix E.2, though these are anomalous as the optimal replacement rate should logically be close to one. A second local maximum also occurs in the second extension for \( R = 2, E_y^b = 0.048 \) and \( \eta = 0.648 \). In the third extension, local maxima are also observed at \( R = 5 \) and \( E_y^b = 0 \) for each \( \eta \), but once again the global maxima are at zero; this extension also presents the possibility of extreme behaviour for high values of \( r \), where the squared consumption drop term becomes unboundedly larger, but no such problems are observed for \( r \in [0.01, 2] \).

All plots are available upon request.

D Baseline Values of \( \frac{dW}{db} \)

Equation (9), when combined with (12), provides a way of evaluating \( \frac{dW}{db} \), and I present in Tables 12 and 13 the values of this derivative at the baseline value of \( r = 0.46 \). As discussed in subsection 6.3, the results are conceptually similar to those in Tables 6 and 7, in that a positive value of \( G \) causes the values in the table to “spread out.”

Table 12: Baseline Values of \( \frac{dW}{db} \) Calculated from (9) and (12) for \( R = 2 \)

<table>
<thead>
<tr>
<th>Initial Value of ( \eta )</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.864</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-0.0816)</td>
<td>-0.0008</td>
<td>-0.0011</td>
<td>-0.0016</td>
<td>-0.0026</td>
</tr>
<tr>
<td>(E_y^b)</td>
<td>0.048</td>
<td>0.0004</td>
<td>0.0006</td>
<td>0.0009</td>
</tr>
<tr>
<td>(0.096)</td>
<td>0.0012</td>
<td>0.0018</td>
<td>0.0029</td>
<td></td>
</tr>
<tr>
<td>(0.192)</td>
<td>0.0017</td>
<td>0.0024</td>
<td>0.0037</td>
<td>0.0059</td>
</tr>
<tr>
<td>(0.3072)</td>
<td>0.0027</td>
<td>0.0039</td>
<td>0.0059</td>
<td>0.0094</td>
</tr>
</tbody>
</table>

Table 13: Baseline Values of \( \frac{dW}{db} \) Calculated from (9) and (12) for \( R = 2 \)

<table>
<thead>
<tr>
<th>Initial Value of ( \eta )</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.864</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-0.0816)</td>
<td>-0.0418</td>
<td>-0.0487</td>
<td>-0.0605</td>
<td>-0.0809</td>
</tr>
<tr>
<td>(E_y^b)</td>
<td>0.048</td>
<td>-0.0165</td>
<td>-0.0124</td>
<td>-0.0054</td>
</tr>
<tr>
<td>(0.096)</td>
<td>-0.0071</td>
<td>0.0011</td>
<td>0.0150</td>
<td>0.0390</td>
</tr>
<tr>
<td>(0.192)</td>
<td>0.0117</td>
<td>0.0280</td>
<td>0.0558</td>
<td>0.1038</td>
</tr>
<tr>
<td>(0.3072)</td>
<td>0.0343</td>
<td>0.602</td>
<td>0.1048</td>
<td>0.1816</td>
</tr>
</tbody>
</table>

E Extensions

This appendix will analyze a variety of extensions to the Baily model, including stochastic duration of unemployment, within-period borrowing constraints, use of a second-order Taylor series expansion of marginal utility, and variable labour supply on the initial job. I will demonstrate that, although the formulas change
Table 13: Baseline Values of $\frac{dW}{db}$ Calculated from (9) and (12) for $R = 5$

<table>
<thead>
<tr>
<th>Initial Value of $\eta$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.864</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.0816$</td>
<td>0.0430</td>
<td>0.0428</td>
<td>0.0425</td>
<td>0.0420</td>
</tr>
<tr>
<td>$0$</td>
<td>0.0435</td>
<td>0.0435</td>
<td>0.0435</td>
<td>0.0435</td>
</tr>
<tr>
<td>$E^n_b$</td>
<td>0.048</td>
<td>0.0437</td>
<td>0.0438</td>
<td>0.0440</td>
</tr>
<tr>
<td>$E^n_b$</td>
<td>0.096</td>
<td>0.0440</td>
<td>0.0442</td>
<td>0.0446</td>
</tr>
<tr>
<td>$E^n_b$</td>
<td>0.192</td>
<td>0.0445</td>
<td>0.0449</td>
<td>0.0457</td>
</tr>
<tr>
<td>$E^n_b$</td>
<td>0.3072</td>
<td>0.0451</td>
<td>0.0458</td>
<td>0.0470</td>
</tr>
</tbody>
</table>

somewhat in each case, as do the specific numerical results, the qualitative effects of fiscal externalities change very little.

E.1 Stochastic Duration of Unemployment

I first consider the effect of allowing the duration of unemployment to be stochastic.\(^{54}\) Since Baily (1978) also analyzed this extension, I will follow his approach of defining the actual duration of unemployment $(1 - \tilde{\eta})$ as:

$$(1 - \tilde{\eta}) = [1 - \eta(c, y_n)] + v,$$

where $\eta$ is deterministic, and $v$ is a stochastic term with mean zero, and which is uncorrelated with $\eta$.\(^{55}\)

If I now denote second-period consumption if the worker loses their job as $\tilde{C}_u$, then:

$$\tilde{C}_u = (1 - \tilde{\eta})(b - c) + \tilde{\eta}y_n(1 - \tau) + k$$

$$= C_u - v\Delta y$$

where $C_u$ is defined as before, and $\Delta y = y_n(1 - \tau) - (b - c)$. I can now write utility as:

$$V = U[y(1 - \tau) - k] + \alpha U[y(1 - \tau) + k] + (1 - \alpha)E_v[U(\tilde{C}_u)].$$

The values of (4) and (5) now have to be replaced by:

$$\frac{\partial V}{\partial b} = (1 - \alpha)E_v[U'(\tilde{C}_u)(1 - \eta + v)]$$

$$\frac{\partial V}{\partial \tau} = -yU'(C^1_u) - \alpha yU'(C^2_u) - (1 - \alpha)y_nE_v[U'(\tilde{C}_u)(\eta - v)].$$

A first-order Taylor series expansion of $U'(C^1_e)$ gives $U'(C^1_u) = U'(C_u) + \Delta C U''(\theta)$ as before, and I perform a similar expansion of $U'(\tilde{C}_u)$:

$$U'(\tilde{C}_u) = U'(C_u) + U''(\delta)(\tilde{C}_u - C_u)$$

$$= U'(C_u) - v\Delta yU''(\delta)$$

where $\delta$ is somewhere between $C_u$ and $\tilde{C}_u$. Upon reaching this point in the calculations, Baily (1978) implicitly makes the assumption that $U''(\delta)$ is uncorrelated with $v$ and $v^2$, which is not generally true, but captures an intuition that the average first and second derivatives shouldn’t be too far from the respective derivatives at the average $C_u$, as well as greatly simplifying the algebra. I make the same assumption, and therefore:

$$E_v[U'(\tilde{C}_u)] = U'(C_u)$$

54 In these circumstances, it might be useful to see how finite potential duration of UI benefits would alter the results, but the model is not well suited to such an exploration, so as in Baily’s analysis, my results remain applicable only to the case of benefits that do not expire.

55 As noted by Baily, this can only be an approximation given that $(\eta - v)$ is constrained to lie in $(0, 1)$. 39
\[ E_v[U'(\tilde{C}_u)v] = -\Delta y E_v[U''(\delta)] Var(v). \]

As a result, the individual’s first-order condition for savings can now be written as:

\[
\frac{\partial V}{\partial k} = -U'(C'_e) + \alpha U'(C''_e) + (1 - \alpha) E_v[U'(\tilde{C}_u)] = 0
\]

\[
\alpha U'(C''_e) = U'(C'_e) - (1 - \alpha) U'(C_u).
\]

All of these results can be combined to give:

\[
\frac{\partial V}{\partial b} = (1 - \alpha)(1 - \eta)U'(C_u) - (1 - \alpha)\Delta y E_v[U''(\delta)] Var(v)
\]

As before, I next make the assumptions that \( y_n = y \) and \( C'_e U''(\theta) = C_u U''(C_u) \), and to this I add the assumption that \( E[U''(\delta)] = U''(\theta) \), which will tend towards underestimating the welfare gain from raising \( b \); then I can write the welfare derivative as:

\[
\frac{dW}{db} = 2y \frac{\Delta C}{C'_e} R \frac{d\tau}{db} + (1 - \alpha) \frac{\Delta y}{C'_e} Var(C) - (1 - \alpha) y_n \Delta y E_v[U''(\delta)] Var(v) = 2y \frac{d\tau}{db} - 2(1 - u)y \left[ \frac{d\tau}{db} - \omega \right].
\]

The budget constraint takes an expectation over all workers, and so is exactly the same as before. A bit of rearranging gives:

\[
\frac{dW}{db} = 2y \left[ \frac{\Delta C}{C'_e} R + \frac{\Delta y}{C'_e} Var(v) \right] \frac{d\tau}{db} - 2y(1 - u) \left[ 1 + \frac{\Delta y}{C'_e} Var(v) \right] \frac{d\tau}{db} - \omega,
\]

and therefore the equation for the optimum is:

\[
\frac{\Delta C}{C'_e} R + \frac{\Delta y}{C'_e} Var(v) = (1 - u) \left[ 1 + \frac{\Delta y}{C'_e} Var(v) \right] \frac{d\tau}{db} - \omega,
\]

or equivalently:

\[
\frac{\Delta C}{C'_e} R + \frac{\Delta y}{C'_e} Var(v) = (1 - u) \left[ 1 + \frac{\Delta y}{C'_e} Var(v) \right] \frac{E_y^0 - \psi}{E_y^0}. \tag{15}
\]

If I make the same assumptions as Baily, then this formula collapses to that used in his extension to stochastic unemployment durations. Most of the terms in (15) have exactly the same interpretation as before, or, as in the case of \( u \) and \( \eta \), still work as averages or expectations, and can be evaluated in the same way. However, there are two new terms to consider: \( \frac{\Delta y}{C'_e} \) and \( Var(v) \). The latter is also the variance of the duration of unemployment \( (1 - \eta) \), and to evaluate this parameter, I turn to Chetty (2008), who estimated a mean duration of unemployment of 18.3 weeks, and a standard deviation of 14.2, so I normalize the standard deviation by the mean and write \( \text{std} (v) = \frac{14.2}{18.3} (1 - \eta) \), and therefore \( Var(v) = \left( \frac{14.2}{18.3} \right)^2 (1 - \eta)^2 \). Meanwhile, in the absence of any better evidence, I will use Baily’s assumption that \( \frac{\Delta y}{C'_e} = 1 - r \). Evaluation of (15) then gives the results displayed below in Tables 14 and 15.

As can be seen, allowing for uncertain duration of unemployment tends to make the optimal rate closer to one, since the desire to provide full insurance is made greater by the uncertainty; this means a decrease in cases where the optimal rate was above one, as it is no longer as desirable to “over-insure” when unemployed individuals face uncertainty about duration. The qualitative conclusion remains the same, however, regarding the effect of the fiscal externality from income taxes: the optimal replacement rate decreases for lower values of \( \eta \) and especially \( E_y^0 \), whereas it increases for higher values; and the pairwise comparisons between the side-by-side tables are nearly identical, in the sense that, if the optimal replacement rate is higher for \( G = 0 \) than for \( G = 0.208 \) in Table 7 or 8, the same is true in Table 14 or 15, and vice-versa.

\[^{56}\text{There are potential offsetting biases in these calculations; Chetty (2008) uses a sample in which the duration of unemployment is truncated at 50 weeks, suggesting I may be underestimating } Var(v), \text{ but on the other hand, Chetty’s is an unconditional variance, some of which may be explained by individual characteristics, which means } Var(v) \text{ may be an overestimate.} \]
Table 14: Optimal Replacement Rates Calculated from (15) for $R = 2$

<table>
<thead>
<tr>
<th>Initial Value of $\eta$</th>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.048 0.725 0.8 0.864</td>
<td>0.048 0.725 0.8 0.864</td>
</tr>
<tr>
<td>-0.0816</td>
<td>0.629 0.6208 0.5812 0.5358</td>
<td>-0.0816 0 0 0 0</td>
</tr>
<tr>
<td>0</td>
<td>0.6590 0.6303 0.5969 0.5624</td>
<td>0.096 0.6047 0.6229 0.6773 0.7889</td>
</tr>
<tr>
<td>0.048</td>
<td>0.6626 0.6360 0.6063 0.5786</td>
<td>0.192 0.6995 0.7578 0.8098 1.0625</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6662 0.6417 0.6157 0.5951</td>
<td>0.3072 0.7924 0.8868 1.0522 1.3240</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6735 0.6532 0.6351 0.6293</td>
<td>0.3072 0.7924 0.8868 1.0522 1.3240</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.6824 0.6673 0.6590 0.6726</td>
<td>0.3072 0.7924 0.8868 1.0522 1.3240</td>
</tr>
</tbody>
</table>

Table 15: Optimal Replacement Rates Calculated from (15) for $R = 5$

<table>
<thead>
<tr>
<th>Initial Value of $\eta$</th>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.048 0.725 0.8 0.864</td>
<td>0.048 0.725 0.8 0.864</td>
</tr>
<tr>
<td>-0.0816</td>
<td>0.7691 0.7779 0.7560 0.7318</td>
<td>-0.0816 0.7175 0.6797 0.6244 0.5440</td>
</tr>
<tr>
<td>0</td>
<td>0.7986 0.7819 0.7626 0.7429</td>
<td>0 0.7453 0.7241 0.6994 0.6743</td>
</tr>
<tr>
<td>0.048</td>
<td>0.8002 0.7843 0.7665 0.7496</td>
<td>0.096 0.7779 0.7763 0.7880 0.8298</td>
</tr>
<tr>
<td>0.096</td>
<td>0.8017 0.7867 0.7705 0.7565</td>
<td>0.192 0.8104 0.8283 0.8760 0.9824</td>
</tr>
<tr>
<td>0.192</td>
<td>0.8048 0.7916 0.7786 0.7708</td>
<td>0.3072 0.8491 0.8901 0.9800 1.1599</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.8085 0.7975 0.7887 0.7888</td>
<td>0.3072 0.8491 0.8901 0.9800 1.1599</td>
</tr>
</tbody>
</table>

E.2 Within-Period Borrowing Constraints

Another unrealistic feature of the basic two-period model is the assumption that individuals can not only save or borrow as much as they want across periods, but can also perfectly smooth consumption within the second period. Recent work, however, in particular that of Chetty (2008), has emphasized the importance of liquidity constraints among the unemployed and the beneficial role of UI in loosening these constraints. I will therefore consider the case of no borrowing during unemployment; I assume that utility is additively time-separable within the second period, so that second period utility of a worker who loses their job is \((1 - \eta)U(C_u) + \eta U(C_n)\), where \(C_u\) is now per-period consumption while unemployed and \(C_n\) is per-period consumption when re-employed in a new job.\(^{57}\) I also assume that, if a worker loses their job, any savings from the first period are completely consumed while unemployed; none of those savings are kept for consumption when re-employed.\(^{58}\) I can therefore write total utility as:

\[
V = U[y(1 - \tau) - k] + \alpha U[y(1 - \tau) + k] + (1 - \alpha) \left[ (1 - \eta)U \left( (b - c) + \frac{k}{1 - \eta} \right) + \eta \left( y_n(1 - \tau) \right) \right]. \tag{16}
\]

(4) still holds, and (5) is now replaced by:

\[
\frac{\partial V}{\partial \tau} = -yU'(C_1^2) - \alpha yU'(C_2^2) - (1 - \alpha) \eta y_n U'(C_n)
\]

\(^{57}\)Chetty (2006) argues that, in his model, the nature of borrowing constraints does not change the optimal UI formula, as this effect will simply show up in the magnitude of the consumption drop. In a sense, this is correct in my analysis as well, but how I interpret borrowing constraints changes what I call the value of consumption during unemployment: I previously defined \(C_u\) as the total consumption in the second period if a worker experiences a spell of unemployment, whereas I now define \(C_u\) to be consumption while unemployed.

\(^{58}\)This assumption is made for simplicity, but given that unemployment durations are deterministic, and that workers can foresee at the beginning of time what \(c\) and \(y_n\) they would choose in period 2, then as long as \(y_n\) is not too far from \(y\) and \(b\) is not too large, there is no reason for them to save any more than they would want to consume in a spell of unemployment.
and I can replace \( U'(C_e^2) \) using the first-order condition for saving, as usual:

\[
\frac{\partial V}{\partial \tau} = -2yU'(C_1^1) + (1 - \alpha)yU'(C_u) - (1 - \alpha)\eta y_n U'(C_n).
\]

Next, I assume that \( C_n = C_1^1 \), which is consistent with the finding in Gruber (1997) that workers who lose their job in one year but are re-employed in the following year see their consumption return within to 4% of their pre-unemployment consumption at baseline; this suggests that \( C_n \), while a bit smaller than \( C_1^1 \), is quite close to the latter. I then get:

\[
\frac{\partial V}{\partial \tau} = -(2y + (1 - \alpha)\eta y_n |U'(C_1^1) + (1 - \alpha)yU'(C_u)
\]

and the usual Taylor series expansion of \( U'(C_1^1) \) gives:

\[
\frac{\partial V}{\partial \tau} = -[(1 + \alpha)y + (1 - \alpha)\eta y_n |U'(C_1^1) - [2y + (1 - \alpha)\eta y_n] \Delta C U''(\theta).
\]

Putting this together with (4):

\[
\frac{dW}{db} = \left[ 2 + (1 - \alpha)\eta \right] y \Delta C C_1^1 R \frac{d\tau}{db} - 2(1 - u)y \left( \frac{d\tau}{db} - \omega \right)
\]

and therefore the equation for the optimum is:

\[
\frac{\Delta C}{C_1^1} R = \frac{2(1 - u)}{2 + (1 - \alpha)\eta} \frac{E^\gamma - \psi}{E^\gamma}. \tag{17}
\]

\( E^\gamma \) is the same as before, and so this equation is almost identical to (10), and it is easy to introduce the extra \((1 - \alpha)\eta\) term into the calculations; solving for the optimal replacement rate generates the results found in Tables 16 and 17.

### Table 16: Optimal Replacement Rates Calculated from (17) for \( R = 2 \)

<table>
<thead>
<tr>
<th>Initial Value of ( \eta )</th>
<th>Optimal ( r ) for ( G = 0 ):</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0.648 )</td>
<td>0.725</td>
</tr>
<tr>
<td>-0.0816</td>
<td>0.4859</td>
</tr>
<tr>
<td>0</td>
<td>0.4934</td>
</tr>
<tr>
<td>( E^\gamma_b )</td>
<td>0.048</td>
</tr>
<tr>
<td>0.096</td>
<td>0.5035</td>
</tr>
<tr>
<td>0.192</td>
<td>0.5137</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.5260</td>
</tr>
</tbody>
</table>

### Table 17: Optimal Replacement Rates Calculated from (17) for \( R = 5 \)

<table>
<thead>
<tr>
<th>Initial Value of ( \eta )</th>
<th>Optimal ( r ) for ( G = 0 ):</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0.648 )</td>
<td>0.725</td>
</tr>
<tr>
<td>-0.0816</td>
<td>0.6963</td>
</tr>
<tr>
<td>0</td>
<td>0.6995</td>
</tr>
<tr>
<td>( E^\gamma_b )</td>
<td>0.048</td>
</tr>
<tr>
<td>0.096</td>
<td>0.7034</td>
</tr>
<tr>
<td>0.192</td>
<td>0.7073</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7120</td>
</tr>
</tbody>
</table>

### Table 17: Optimal Replacement Rates Calculated from (17) for \( R = 5 \)

<table>
<thead>
<tr>
<th>Initial Value of ( \eta )</th>
<th>Optimal ( r ) for ( G = 0 ):</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0.648 )</td>
<td>0.725</td>
</tr>
<tr>
<td>-0.0816</td>
<td>0.5849</td>
</tr>
<tr>
<td>0</td>
<td>0.6248</td>
</tr>
<tr>
<td>( E^\gamma_b )</td>
<td>0.048</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6723</td>
</tr>
<tr>
<td>0.192</td>
<td>0.7198</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7763</td>
</tr>
</tbody>
</table>

*Global Maximum/Local Maximum
The pattern of the results changes a little, as the optimal replacement rates generally tend to move closer to one (or more precisely, closer to $0.222 \over 20.265$). In a few cases with $R = 5$ and high $E_0^u$ and $\eta$, anomalous results are produced wherein the local maximum obtained at a replacement rate near one is not estimated to be the global maximum, which appears to occur at zero; in these cases, the assumption that unemployment goes to zero as benefits go to zero is partly responsible, as is a significant failure of some of my assumptions.\textsuperscript{59} However, aside from these cases, the changes tend to be quite small, surprisingly so given the shift in the nature of borrowing constraints, as zeros still occur for $R = 2$ and low values of $E_0^u$, and the qualitative comparisons are similar to those from the basic model. One explanation for this is that I still allow for unrestricted savings in the first period, so workers will take into account the borrowing constraints in the second period when they make their savings decision, and the desire for within-period consumption smoothing may be fairly small; a more severe form of liquidity constraint may require that savings be partly ruled out as well, which cannot be easily done within the current framework. Additionally, using the same expression for $\frac{\Delta C}{C_e}$ when I have redefined $C_u$ to be consumption while unemployed will tend to shift the results downwards.

### E.3 Second-Order Taylor Series Expansion of Marginal Utility

Chetty (2006) argues that ignoring third and higher derivatives of the utility function may be a mistake; he reports that, for simulation exercises using a CRRA utility function, using a first-order expansion of marginal utility can sometimes lead to an underestimate of the true optimal replacement rate on the order of 30%, whereas a revised welfare equation using a second-order expansion reduces this error to less than 4%.\textsuperscript{60} The model used by Chetty (2006) is somewhat different from mine, and he writes all marginal utilities in terms of consumption while employed rather than $U'(C_u)$, so the results are not directly comparable.\textsuperscript{60} However, I will now explore how the results change when I use the second-order Taylor series expansion of marginal utility.

To do so, I follow the approach of Chetty (2006) and rewrite my expression in terms of $U'(C_e^1)$ rather than $U'(C_u)$. This will hinder the comparability of my results with those of the baseline case, but is the only practical option under the circumstances. I begin with (4) and (5), and the standard first-order condition for saving. Next, I use a new Taylor series expansion of $U'(C_u)$ around $U'(C_e^1)$:

$$U'(C_u) = U'(C_e^1) - \Delta C U''(C_e^1) + \Delta C^2 U'''(\theta) \frac{\Delta C'}{C_e^1}$$

where $\theta$ is not necessarily the same value as before, but is still between $C_e^1$ and $C_u$. Using this to replace $U'(C_u)$ in both (4) and (5):

$$\frac{dV}{db} = 2u \left[ U'(C_e^1) - \Delta C U''(C_e^1) + \Delta C^2 U'''(\theta) \frac{\Delta C'}{C_e^1} \right]$$

$$- \left[ [(1 + \alpha)y + (1 - \eta)y] U'(C_e^1) + (1 - \alpha)(y - \eta y) \left( \Delta C U''(C_e^1) - \Delta C^2 U'''(\theta) \frac{\Delta C'}{C_e^1} \right) \right] \frac{d\tau}{db}.$$

Making the usual assumption that $y_u = y$, and adding the modified assumption that $\theta = C_e^1$, I can write:

$$\frac{dW_1}{db} = \frac{dV}{U'(C_e^1)} = 2u \left[ 1 + \frac{\Delta C}{C_e^1} R + \frac{1}{2} \left( \frac{\Delta C}{C_e^1} \right)^2 RP \right] - \left[ 2(1 - u)y - 2uy \left( \frac{\Delta C}{C_e^1} R + \frac{1}{2} \left( \frac{\Delta C}{C_e^1} \right)^2 RP \right) \right] \frac{d\tau}{db}$$

\textsuperscript{59}In these cases, $\frac{\Delta C}{C_e^1} R > \frac{2(1-u)}{2 + (1 - \alpha)y}$, so my assumptions imply $\frac{dV}{\tau} > 0$, whereas I also estimate that $\frac{d\tau}{db} < 0$.

\textsuperscript{60}The first-order Taylor series used in my paper is in fact an exact equality, not an approximation, given that I don’t specify the $\theta$; it is the assumption that $C_e^1 U''(\theta) = C_u U''(C_u)$ which generates the potential for error. As already discussed, that assumption tends to be a liberal one, but Chetty’s effective assumption that $U''(\theta)$ is equal to $U''$ at the average level of consumption while employed is a conservative one in his context, explaining why this leads to an underestimate in his paper. Baily’s assumption that the $E_0^u$ in the denominator of the right-hand side of (10) is equal to one is, in the context of his model, a significant reason for underestimation of the optimal $b$. 

43
where \( P = \frac{-C^*_1 U''(C_1)}{U''(C_2)} \) is the coefficient of relative prudence. A bit of rearranging finally gives:

\[
\frac{dW_1}{db} = 2y \left( \frac{\Delta C}{C_e} R + \frac{1}{2} \left( \frac{\Delta C}{C_e} \right)^2 RP \right) \frac{d\tau}{db} - 2(1 - u)y \left( 1 + \frac{\Delta C}{C_e} R + \frac{1}{2} \left( \frac{\Delta C}{C_e} \right)^2 RP \right) \left[ \frac{d\tau}{db} - \omega \right]
\]

and therefore the equation for the optimum is:

\[
\left( \frac{\Delta C}{C_e} R + \frac{1}{2} \left( \frac{\Delta C}{C_e} \right)^2 RP \right) = (1 - u) \left( 1 + \frac{\Delta C}{C_e} R + \frac{1}{2} \left( \frac{\Delta C}{C_e} \right)^2 RP \right) \frac{d\tau}{db} - \omega
\]

or equivalently:

\[
\left( \frac{\Delta C}{C_e} R + \frac{1}{2} \left( \frac{\Delta C}{C_e} \right)^2 RP \right) = (1 - u) \left( 1 + \frac{\Delta C}{C_e} R + \frac{1}{2} \left( \frac{\Delta C}{C_e} \right)^2 RP \right) \frac{E_b^L - \psi}{E_b^g}
\]

where \( \frac{d\tau}{db} \) and \( E_b^g \) are unchanged, and thus \( E_b^L \) is still given by (12).

I can use parameter values and functions as before, but there is one additional parameter to select: the coefficient of relative prudence. In a CRRA utility function \( U(C) = \frac{C^{1 - \eta}}{1 - \eta} \), \( P = \frac{-C^*_1 U''(C_1)}{U''(C_2)} = R + 1 \), so one possibility is to set \( P = R + 1 \). However, previous studies have tended to find low estimates of relative prudence; Merrigan and Normandin (1996) are on the high end of the results in the literature when they find estimates ranging from 1.78 to 2.33.\(^{61}\) I will therefore use a value of \( P = 2 \), and the results from evaluation of the optimal replacement rate are displayed in Tables 18 and 19.

### Table 18: Optimal Replacement Rates Calculated from (18) for \( R = 2 \)

<table>
<thead>
<tr>
<th>Initial Value of ( \eta )</th>
<th>Optimal ( r ) for ( G = 0 )</th>
<th>Optimal ( r ) for ( G = 0.208 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( -0.0816 )</td>
<td>0.648 0.725 0.8 0.864</td>
<td>0.648 0.725 0.8 0.864</td>
</tr>
<tr>
<td>( 0 )</td>
<td>0.4124 0.4124 0.4124 0.4124</td>
<td>0.0</td>
</tr>
<tr>
<td>( 0.048 )</td>
<td>0.4196 0.4227 0.4281 0.4373</td>
<td>0.048 0.0235 0.0914 0.3303 0.5041</td>
</tr>
<tr>
<td>( 0.096 )</td>
<td>0.4269 0.4331 0.4438 0.4624</td>
<td>0.096 0.3053 0.4385 0.5828 0.7553</td>
</tr>
<tr>
<td>( 0.192 )</td>
<td>0.4414 0.4539 0.4755 0.5133</td>
<td>0.192 0.5494 0.6758 0.8421 1.0644</td>
</tr>
<tr>
<td>( 0.3072 )</td>
<td>0.4588 0.4788 0.5138 0.5755</td>
<td>0.3072 0.7084 0.8545 1.0564 1.3382</td>
</tr>
</tbody>
</table>

### Table 19: Optimal Replacement Rates Calculated from (18) for \( R = 5 \)

<table>
<thead>
<tr>
<th>Initial Value of ( \eta )</th>
<th>Optimal ( r ) for ( G = 0 )</th>
<th>Optimal ( r ) for ( G = 0.208 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( -0.0816 )</td>
<td>0.648 0.725 0.8 0.864</td>
<td>0.648 0.725 0.8 0.864</td>
</tr>
<tr>
<td>( 0 )</td>
<td>0.6574 0.6574 0.6574 0.6574</td>
<td>0.0</td>
</tr>
<tr>
<td>( 0.048 )</td>
<td>0.6604 0.6618 0.6641 0.6682</td>
<td>0.048 0.5583 0.5833 0.6211 0.6770</td>
</tr>
<tr>
<td>( 0.096 )</td>
<td>0.6634 0.6662 0.6709 0.6792</td>
<td>0.096 0.6108 0.6503 0.7089 0.7936</td>
</tr>
<tr>
<td>( 0.192 )</td>
<td>0.6695 0.6750 0.6846 0.7012</td>
<td>0.192 0.6916 0.7519 0.8402 0.9668</td>
</tr>
<tr>
<td>( 0.3072 )</td>
<td>0.6768 0.6857 0.7011 0.7280</td>
<td>0.3072 0.7677 0.8471 0.9630 1.1300</td>
</tr>
</tbody>
</table>

The results this time are somewhat different quantitatively, in that I find optimal replacement rates of zero for low values of \( E_b^g \) even for \( R = 5 \). However, the qualitative comparison remain the same, right down to the specific pairwise comparisons: at low values of \( \eta \) and especially \( E_b^g \), the fiscal externality term considerably reduces the optimal replacement rate, whereas at higher values it considerably increases it.

\(^{61}\) Eisenhauer and Ventura (2003) are an exception in finding values of \( R \) and \( P \) in the 7 to 8 range, but they base their estimation on answers regarding willingness to pay for a security from a Bank of Italy survey of Italian households.
E.4 Variable Labour Supply

To this point, I have assumed that \( y \) is fixed; Chetty (2006) points out that, with a lump-sum tax on workers, the envelope condition means that whether or not individuals can change the amount of their labour supply while employed is irrelevant to the optimal UI calculation. However, with a proportional tax, labour supply effects on \( y \) matter though the government budget constraint. Let me begin by rewriting the utility function to allow for choice of \( y \) (assuming the worker must make the same choice of \( y \) in both the first and second periods if they retain their job); if disutility from work effort \( h(y) \) is separable from consumption, (1) becomes:

\[
V = U[y(1 - \tau) - k] + \alpha U[y(1 - \tau) + k] - (1 + \alpha)h(y) + (1 - \alpha)\eta(y(1 - \tau) + k) - (1 - \alpha)\eta_n(1 - \tau) + k.
\]

Because (4) and (5) are unchanged, both (7) and (9) remain valid; the only change is to the derivative of the government budget constraint. The latter now becomes:

\[
d\tau = \frac{(1 - \alpha)(1 - \eta) - (1 - \alpha)\eta \frac{dy}{db} - (1 - \alpha)\eta y_n \frac{dy}{db} - (1 - \alpha)\eta \tau \frac{dy}{db} - (1 + \alpha)\tau \frac{dy}{db}}{(1 + \alpha)y + (1 - \alpha)\eta y_n}.
\]

It is then simple to derive:

\[
E^*_b = \psi + \psi E^*_u + \frac{u}{1 - u} E^*_b - \frac{(1 - \alpha)\eta}{2(1 - u)} E^*_y = \frac{1 + \alpha}{2(1 - u)} \varepsilon^*_b
\]

where \( \varepsilon^*_b = \frac{b \cdot dy}{y \cdot db} \).

I now have to decide on a value for \( \varepsilon^*_b \); when \( b \) increases, \( \tau \) increases as well (unless \( E^*_b \) is so large as to actually lead to increased tax revenues, which cannot be the case in equilibrium), so some version of an elasticity of taxable income is required. Gruber and Saez (2002) find an elasticity of taxable income with respect to the net-of-tax rate of 0.4; using this, and assuming that the only effect of changes in \( b \) and \( \tau \) on \( y \) go through the channel of taxes, I can write:

\[
\varepsilon^*_b = \frac{dy}{db} = \frac{dy}{d\tau} \frac{d\tau}{db} = -0.4 \frac{b}{1 - \tau} \frac{d\tau}{db}.
\]

Since I cannot calculate \( \frac{d\tau}{db} \) without knowing \( \varepsilon^*_b \), I replace it with the partial derivative:

\[
\varepsilon^*_b \approx -0.4 \frac{b}{1 - \tau} \left[ \frac{(1 - \alpha)(1 - \eta)}{(1 + \alpha)y + (1 - \alpha)\eta y_n} \right] \approx -0.4 \frac{\tau}{1 - \tau} \psi.
\]

Finally, using a tax rate of \( \tau = 0.23 \) as before, I can write my estimate of the elasticity as \( \varepsilon^*_b = -\frac{0.092}{0.77} \psi \); I do not need to multiply this by 0.48, as this estimate is meant to apply to the entire universe of workers. The ensuing numerical results are displayed in Tables 20 and 21.

Table 20: Optimal Replacement Rates Calculated from (10) and (19) for \( R = 2 \)

<table>
<thead>
<tr>
<th>Initial Value of ( \eta )</th>
<th>0.048</th>
<th>0.725</th>
<th>0.8</th>
<th>0.864</th>
</tr>
</thead>
<tbody>
<tr>
<td>( -0.0816 )</td>
<td>0.3472</td>
<td>0.3429</td>
<td>0.3566</td>
<td>0.3595</td>
</tr>
<tr>
<td>( 0 )</td>
<td>0.3496</td>
<td>0.3548</td>
<td>0.3638</td>
<td>0.3795</td>
</tr>
<tr>
<td>( E^*_b )</td>
<td>0.048</td>
<td>0.3546</td>
<td>0.3619</td>
<td>0.3746</td>
</tr>
<tr>
<td>( 0.096 )</td>
<td>0.3596</td>
<td>0.3691</td>
<td>0.3856</td>
<td>0.4151</td>
</tr>
<tr>
<td>( 0.192 )</td>
<td>0.3697</td>
<td>0.3835</td>
<td>0.4080</td>
<td>0.4523</td>
</tr>
<tr>
<td>( 0.3072 )</td>
<td>0.3818</td>
<td>0.4012</td>
<td>0.4356</td>
<td>0.4993</td>
</tr>
<tr>
<td>Optimal ( r ) for ( G = 0 ):</td>
<td>0.048</td>
<td>0.0262</td>
<td>0.1405</td>
<td>0.3114</td>
</tr>
<tr>
<td>Initial Value of ( \eta )</td>
<td>0.048</td>
<td>0.0262</td>
<td>0.1405</td>
<td>0.3114</td>
</tr>
<tr>
<td>( -0.0816 )</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>( 0 )</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>( E^*_b )</td>
<td>0.048</td>
<td>0.0262</td>
<td>0.1405</td>
<td>0.3114</td>
</tr>
<tr>
<td>( 0.096 )</td>
<td>0.096</td>
<td>0.239</td>
<td>0.3887</td>
<td>0.5178</td>
</tr>
<tr>
<td>( 0.192 )</td>
<td>0.192</td>
<td>0.4807</td>
<td>0.5966</td>
<td>0.7597</td>
</tr>
<tr>
<td>( 0.3072 )</td>
<td>0.3072</td>
<td>0.6249</td>
<td>0.7660</td>
<td>0.9719</td>
</tr>
</tbody>
</table>

The inclusion of the labour supply elasticity tends to lower the optimal replacement rate, though in many cases not by much; in the \( R = 5 \) case, the effect is often almost negligible, whereas in the \( R = 2 \) case the effect can be somewhat larger for moderate values of \( E^*_b \). However, the essential point remains that consideration of the fiscal externality term can greatly change the results, either in a positive or negative direction (the pairwise comparisons are identical to the baseline); the magnitude of this effect is largely dependent on the parameter values chosen for \( \eta \) and \( E^*_b \).
Table 21: Optimal Replacement Rates Calculated from (10) (19) for \( R = 5\)

<table>
<thead>
<tr>
<th>( \eta )</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.864</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.6399</td>
<td>0.6496</td>
<td>0.6417</td>
<td>0.6436</td>
</tr>
<tr>
<td>0</td>
<td>0.6430</td>
<td>0.6450</td>
<td>0.6486</td>
<td>0.6548</td>
</tr>
<tr>
<td>( E_{B}^{y} )</td>
<td>0.048</td>
<td>0.6448</td>
<td>0.6477</td>
<td>0.6527</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6466</td>
<td>0.6504</td>
<td>0.6569</td>
<td>0.6687</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6503</td>
<td>0.6558</td>
<td>0.6656</td>
<td>0.6832</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.6548</td>
<td>0.6625</td>
<td>0.6763</td>
<td>0.7017</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \eta )</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.864</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.5211</td>
<td>0.5631</td>
<td>0.4723</td>
<td>0.4199</td>
</tr>
<tr>
<td>0</td>
<td>0.5653</td>
<td>0.5670</td>
<td>0.5700</td>
<td>0.5752</td>
</tr>
<tr>
<td>( E_{B}^{y} )</td>
<td>0.048</td>
<td>0.5910</td>
<td>0.6042</td>
<td>0.6273</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6163</td>
<td>0.6410</td>
<td>0.6838</td>
<td>0.7585</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6661</td>
<td>0.7131</td>
<td>0.7939</td>
<td>0.9334</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7241</td>
<td>0.7968</td>
<td>0.9207</td>
<td>1.1318</td>
</tr>
</tbody>
</table>

E.5 Summary of Extensions

I have now considered four possible extensions to the Baily model, to see how each might affect the results. In each case, altering the model generally does change the numerical results; allowing for stochastic duration of unemployment or second-period borrowing constraints tends to move the optimal replacement rates closer to one, whereas allowing for variable \( y \) or using the third derivative of marginal utility tends to reduce the estimated optimal replacement rate. These results, however, are all remarkable similar in terms of what they tell us about the importance of fiscal externalities; even the pairwise comparisons of optimal replacement rates given \( G = 0 \) versus \( G = 0.208 \) are nearly identical in each case. The numerical results may still be implausibly high or low in certain cases, and so earlier comments about potential limitations of the sufficient statistics approach still apply.

References


