Fiscal Externalities and Optimal Unemployment Insurance

Nicholas Lawson*
Princeton University

nlawson@princeton.edu

May 10, 2013

Abstract

In analyses of the optimal design of government programs, an important consideration is the extent to which a program can impact labour market outcomes. Typically, such analyses consider one government program in isolation, abstracting away from other roles of government, and thereby ignoring "fiscal externalities," or the effects of the program's labour market impacts on income tax revenues. This paper assesses the importance of fiscal externalities in the context of optimal unemployment insurance. I calibrate and simulate a standard dynamic job search model, and perform a theoretical and numerical analysis of the seminal Baily (1978) two-period model of unemployment, in order to answer the question: how much does the abstraction from other government activities matter for optimal UI calculations? Numerical results indicate that it matters a lot: with moderate risk-aversion and no effect of UI benefits on subsequent wages, the optimal replacement rate drops from around 40% to zero. Higher risk-aversion moderates this effect, and a large positive effect of UI on wages could significantly increase the optimal benefit level. The theoretical analysis indicates that accounting for fiscal externalities increases the estimated optimal benefit level if and only if higher benefits increase total post-unemployment earnings.

*I am indebted to David Lee and Richard Rogerson for many helpful comments and suggestions, and I am also grateful to Orley Ashenfelter for helpful discussions. I also wish to thank the participants of the Industrial Relations Section Graduate Lunch, the Public Finance Working Group, EconCon 2012, and the Labor Economics Seminar at Princeton University, as well as seminars at the University of Ottawa, the Max Planck Institute for Research on Collective Goods, the University of Mannheim, the Aix-Marseille School of Economics, the University of Miami, and New York University Abu Dhabi. Any errors or omissions are the responsibilty of the author.

1 Introduction

In an analysis of the optimal design of a government program, an important consideration is the program's impact on labour market outcomes, since this can affect both the direct cost of the program and the availability of tax revenues to finance it. This fact has motivated a considerable amount of empirical research analyzing the effects of social insurance programs on labour market outcomes, from wages to durations of unemployment to labour force attachment. For example, many studies find that more generous unemployment insurance significantly increases unemployment durations (e.g. Ehrenberg and Oaxaca (1976), Meyer (1990), and Chetty (2008)), leading to an increase in benefit payments and a reduction in revenues from the UI payroll tax. Similarly, employment disincentive effects of disability insurance have been documented in a number of papers (e.g. Bound (1989), Gruber (2000), and Chen and van der Klaauw (2008)), and a number of ways in which old age security benefits influence labour supply decisions among older workers have been identified (e.g. Blau (1994), Rust and Phelan (1997), and Coile and Gruber (2007)). In addition, tax revenue effects can be present even for programs which are not directly tied to labour force status: government policies regarding health insurance have been shown to affect incentives for retirement, job transitions, and entrepreneurship, as documented by Gruber and Madrian (1994), Boyle and Lahey (2010), and Fairlie, Kapur, and Gates (2011).[1]

Importantly, however, welfare analyses of each of these programs typically study it in isolation, abstracting away from other roles of government (for example, Feldstein (1985) and İmrohoroğlu, İmrohoroğlu, and Joines (1995) on social security, Golosov and Tsyvinski (2006) on disability insurance, and the analysis of public health insurance in Chetty and Saez (2010)).[2] The standard assumption is that the only actions of government are to levy a payroll tax and use the proceeds to fund the program in question. In reality, however,

[1] Additionally, revenue effects are a consideration for some programs beyond social insurance. Many studies find that financial aid to university students increases enrollment (see Kane (1994) and Dynarski (2003)), presumably leading to an increase in human capital and thus an increase in wages and tax revenues. This issue is studied in more detail in Lawson (2013).

[2] Parry and Oates (2000) recognize the importance of interactions between environmental policies and the tax system, and argue that this will apply to other programs and institutions that raise the cost of living, but they do not consider any of the programs mentioned above, restricting their discussion to areas of trade, agriculture, occupational licensing and monopolies. Some studies of the programs I discuss do take income taxes that pay for other spending into account, including İmrohoroğlu, İmrohoroğlu, and Joines (2003) and Laitner and Silverman (2012) in the area of social security, and Gruber (1996) and Bound, Cullen, Nichols, and Schmidt (2004) on disability insurance, but even in these cases, there is no mention of the importance of this component or the fact that including it represents an important departure from the rest of the literature.

many individual government programs account for only a small fraction of overall spending; this is particularly true in the case of a program like UI, which typically accounts for less than 1% of total government spending in the U.S. Therefore, the revenue needed to fund any particular program is a small fraction of the revenue from income taxes.

A critical implication of this is that an analysis of a program in isolation will not fully capture the welfare consequences of its labour market impacts, because those impacts affect not only the tax revenues used to pay for that program, but the general income tax revenues that pay for all government programs. I call this phenomenon a "fiscal externality": each government program that affects labour market outcomes has "external" effects on the government's ability to fund other programs, requiring a change in the tax rate or in expenditure on other programs that will have an additional effect on welfare.[3] As described in section 2, this phenomenon can also be thought of as an application of the Theory of the Second Best, in which optimal government policy must be considered in the context of a labour market that is already distorted by income taxes.

The important unanswered question is: how much do these fiscal externalities matter for welfare analysis? I will provide an answer in the specific context of unemployment insurance, a program that is the subject of a large optimal policy literature. Papers in this literature (see, for example, Baily (1978), Hansen and İmrohoroğlu (1992), Hopenhayn and Nicolini (1997), and Chetty (2008)) seek to balance the consumption-smoothing and risk-sharing benefits of UI with the moral hazard costs, i.e. the empirical regularity that more generous UI leads to longer durations of unemployment by inducing unemployed individuals to put less effort into searching for a job. However, this literature features a universal abstraction away from any role of government in the economy other than the UI system, thereby ignoring effects on income tax revenues.[4] Therefore, the tax revenue cost of longer unemployment durations has been understated, while on the other hand, higher UI benefits could increase income tax revenues by facilitating better matches and higher wages. Whichever of these

[3] In a situation in which the same government planner chooses the policy design and collects tax revenues, this impact is not external in the traditional sense, but the usual analysis behaves as if it is, by ignoring the tax collection activities. Section 2 describes the phenomenon in greater detail.

[4] Some studies allow for more complicated UI programs including job creation subsidies or taxes which vary with past durations of unemployment (e.g. Hopenhayn and Nicolini (1997) and Coles and Masters (2006)), but even these papers continue to ignore the vast majority of government activity. The abstraction from other activities of government is noted by Chetty (2006), who states 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."

effects dominates, the important point is that the labour market impacts of UI will impact the government's ability to pay for non-UI expenditures. This paper, therefore, asks the question: how much does the abstraction from government activities other than UI matter for optimal UI calculations? In particular, what effect does this omission have on the optimal level of UI benefits?

To answer these questions, I use both of the two main approaches from the optimal UI literature: a macro-based approach that involves calibrating and simulating a dynamic job search model, and a "sufficient statistics" approach that relies on reduced-from elasticities.[5] In the first approach, I use a standard structural job search model from Lentz (2009), which I calibrate to match a number of real-world moments under two scenarios: the first uses an estimate of the real value of total government spending, while the second ignores all government activities other than UI. I then simulate the model and solve numerically for the optimal replacement rate in each case.

For the second approach, I adopt the simple two-period model from the seminal work of Baily (1978), the first sufficient statistics paper in the area of UI. I solve this model for an analytical expression for the optimal level of benefits, which can be written in terms of a few reduced-form empirical values, or sufficient statistics for welfare.[6] I provide a series of analytical results about the importance of fiscal externalities, and I use a method of statistical extrapolation to calculate numerical results for the optimal benefit level. Using both approaches demonstrates the robustness of the results,[7] and while the structural approach is more natural for extrapolating out of sample, the simplicity of the Baily model allows for a more transparent demonstration of the central results.

The results from both of these approaches indicate that fiscal externalities are indeed important components of optimal UI analysis. In my baseline case, assuming a moderate degree of risk-aversion and no effect of UI benefits on post-unemployment wages, the optimal replacement rate drops from around 40% to zero once fiscal externalities are accounted for. A higher value of risk-aversion can moderate this decrease, but the optimal replacement

[5]Both models ignore any general equilibrium effects of UI on the job search process; I use a one-sided search specification that abstracts from firms, as is common in the literature.

[6]These quantities encode all relevant information from underlying structural parameters, making the welfare equations robust to many modifications and modelling decisions; Chetty (2006) shows that the results from Baily (1978) are applicable to a far more general class of models.

[7]The moments that I use to calibrate the structural model are derived from the sufficient statistics used later in the paper; Chetty (2009) proposes using the sufficient statistics both as the basis for statistical extrapolation and to calibrate a structural model.

rate still drops from about 70% to either 46% or 60%. However, if the elasticity of post-unemployment wages with respect to benefits is positive and sufficiently large, then the optimal replacement rate will increase, and can even approach 100%, making the magnitude of the wage elasticity a crucial quantity. Recent evidence suggests that this elasticity is unlikely to be large, but the empirical literature remains sparse, and therefore my findings indicate a need for further work on this question. The estimated welfare gains from moving to the new optimum are on the order of 10% of baseline UI spending. Finally, theoretical results from the Baily model show that accounting for fiscal externalities increases the implied optimal benefit level if and only if higher benefits increase total post-unemployment earnings, and the analysis indicates some broader lessons that could be applied in other areas of government policy.

The rest of the paper is organized as follows. Section 2 further explains the concept of fiscal externalities in the context of government policy. Section 3 presents the dynamic job search model, and includes 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 analytical results based on these equations. Numerical results from the statistical extrapolation method follow in section 6, and section 7 concludes.

2 What Are Fiscal Externalities?

There are a number of equivalent ways of describing the phenomenon which I have labelled a "fiscal externality" and which arises when one accounts for other government spending and taxation in the analysis of a government program.

In the introduction, a fiscal externality was defined as the "external" effect of one government program on the revenues available to fund all government activities, through that program's effects on labour market outcomes and thus on income tax revenues. This definition is similar to the way the term has been used in the literature on tax competition, as discussed by Wilson (1999). That literature analyzes how one government's actions can have fiscal effects on other governments, either in other jurisdictions or at separate levels of a federation. In the current paper, the fiscal effects of one government's actions are assumed

to be experienced by the same government, so these effects are not external to the planner.[8] However, by abstracting away from other activities of government, previous policy analysis has implicitly assumed that the fiscal effects are external to the relevant planner. Therefore, my analysis may therefore be thought of as "internalizing" the effects of UI on tax revenues.

An alternative definition focusses on the externality at the level of the individual worker: if a change in an individual's actions, such as how hard they search for a job, affects that individual's income and taxes paid, then the tax rate or amount of public goods provided must change to balance the budget. This imposes a cost or benefit on all workers, and therefore represents an external effect on other workers operating through the government budget constraint. Government programs that affect labour market outcomes interact with these externalities: a program that encourages income-increasing activities offsets the effect of taxes, internalizing the externality and increasing welfare, whereas programs that discourage effort and income aggravate the externality.

This explanation also indicates that a fiscal externality is an application of the Theory of the Second Best. If the government must raise a large amount of revenue to fund various programs, and can only raise the required revenue using an income tax, then the labour market is distorted prior to the implementation of a UI system, and in particular the search decision is distorted. Unemployment insurance may move search outcomes closer to or away from the first-best, but the important point is that, in a distorted labour market, the optimal second-best policy will generally not correspond to the policy that would have been preferred without distortions.

Each of these explanations of fiscal externalities highlight different aspects of the phenomenon, but it should be emphasized that the explanations are equivalent. That is, the existence of a large amount of government spending and taxation means that labour markets are distorted, and therefore that part of the returns to individual actions go to other people due to the effects of those actions on the government budget constraint, which will show up as positive or negative effects on tax revenues. The rest of the paper will be devoted to analyzing the effects of these fiscal externalities on optimal UI calculations.

[8] In reality, unemployment insurance policy may be set by a different level of government than that which collects most of the tax revenues; this is the case in the United States, where states set UI benefits. In this case, the phenomenon may in fact be interpreted as an externality between two levels of government, and my optimal policy analysis corresponds to federal intervention to solve this fiscal externality.

3 Dynamic Job Search Model

A substantial literature studies optimal UI using a structural model, and the numerical results of this literature are summarized in Table 1. Most of the studies listed find that UI should be more generous, compared to the typical U.S. system which features a replacement rate of about 50% and benefits that usually expire after 26 weeks. My analysis will be based on the model from Lentz (2009), which is a typical workhorse single-agent model, incorporating endogenous search intensity and private asset accumulation, as well as being intuitive and straightforward to simulate. My only modifications are the introduction of government spending outside of UI and a simplified functional form for the effort cost of job search.

Table 1: Numerical Results from Structural Studies of Optimal UI

Paper	Optimal Replacement Rate
Hansen and İmrohoroğlu (1992)	0.15^* (with moral hazard)/0.65 (without)
Davidson and Woodbury (1997)	$0.66^*/1.30^{**}$
Hopenhayn and Nicolini (1997)	$> 0.94^*$ (with optimal tax)
Acemoglu and Shimer (2000)	$> 0.4^{**}$
Fredriksson and Holmlund (2001)	$0.38 - 0.42^*$
Wang and Williamson (2002)	$0.24^*/0.56^{**}$
Coles and Masters (2006)	0.76^*
Lentz (2009)	$0.43 - 0.82^*$

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

The first subsection contains a description of the model, while the second explains the calibration; I then present the numerical results, and discuss the effects of fiscal externalities on the estimated optimal benefit level.

3.1 Model Setup

The model features a representative infinitely-lived risk-averse 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,[9] and chooses search intensity s_t subject to a convex search cost function $e(s_t)$, where s_t is the probability of receiving a job offer.[10] All jobs have an identical wage y, and jobs end exogenously at

[9] By 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.

[10] This simplified specification allows for a closed-form solution for the individual's search decision.

a constant rate of δ per period. Agents cannot borrow, but can make savings which earn interest at a rate of i per period; the annual discount rate is ρ, and I define a period to be equal to a week, so the per-period discount factor is $\left(\frac{1}{1+\rho}\right)^{\frac{1}{52}}$.

In all periods and states, agents decide on their level of consumption, while unemployed agents also choose how hard to search for a job. The decision problem can be written as:

$$\max_{\{c_t, s_t, k_{t+1}\}} E \sum_{t=0}^{\infty} \left(\frac{1}{1+\rho}\right)^{\frac{t}{52}} [U(c_t) - e(s_t)]$$
$$s.t.: k_{t+1} = (1+i)k_t + n_t y(1-\tau) + (1-n_t)b - c_t$$
$$c_t, k_{t+1} \geq 0, s_t \in [0,1]$$
$$Pr(n_{t+1} = 1 | n_t = 1) = 1 - \delta, Pr(n_{t+1} = 0 | n_t = 1) = \delta$$
$$Pr(n_{t+1} = 1 | n_t = 0) = s_t, Pr(n_{t+1} = 0 | n_t = 0) = 1 - s_t$$

where c_t is consumption, τ 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 is more convenient to write this recursively; let $V_e(k)$ represent the maximum present value of being employed with assets equal to k, while $V_u(k)$ will be the analogous value of unemployment, and let k' represent next period's assets:

$$V_e(k) = \max_{k' \in \Gamma_{y(1-\tau)}(k)} \left[U((1+i)k + y(1-\tau) - k') + \left(\frac{1}{1+\rho}\right)^{\frac{1}{52}} [(1-\delta)V_e(k') + \delta V_u(k')] \right]$$
$$V_u(k) = \max_{k' \in \Gamma_b(k), s \geq 0} \left[U((1+i)k + b - k') - e(s) + \left(\frac{1}{1+\rho}\right)^{\frac{1}{52}} [sV_e(k') + (1-s)V_u(k')] \right]$$

where $\Gamma_z(k) = (k' \in \mathbb{R} | 0 \leq k' \leq (1+i)k + z)$ is the set of permissable asset values.[11]

The agent is representative of a continuum of identical agents, and therefore I can consider the economy-wide steady-state, in which the government budget constraint is:[12]

$$(1-u)y\tau = ub + G$$

where u is the unemployment rate and G is the level of exogenous non-UI government spending.[13] The government chooses b and τ subject to this constraint to maximize steady-

[11] As in Lentz (2009), the numerical solutions always appear to yield concave value functions by asset level.

[12] Assuming a single proportional tax keeps the government budget constraint simple; a more complex tax system would not change the essential point that taxes paid are an increasing function of income.

[13] In principle, G could be made endogenous, but this would add considerable complexity while providing little new insight into the main point; if exogenous, G does not need to be accounted for in the individual's utility function.

state expected utility,[14] and my goal will be to compare the results from cases when the best estimate of G is used with the results obtained when I assume $G = 0$.

3.2 Calibration of the Model

Calibration requires choosing functional forms and parameter values, in order to be able to simulate the model. For functional forms, I assume constant relative risk-aversion utility with risk-aversion parameter R, so $U(c) = \frac{c^{1-R}}{1-R}$, and 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.[15]

Additionally, the way UI benefits are specified requires some explanation. If the replacement rate is r and the baseline tax rate is τ_0, then a natural value for after-tax benefits would be $r(1-\tau_0)y$.[16] However, although the model features universal unemployment benefits to make computation manageable, real-world benefits are of finite duration and not received by all unemployed individuals.[17] I therefore follow the general approach of Fredriksson and Holmlund (2001) in normalizing benefits in my model to be equal in expectation to real-world finite-duration benefits. I adjust for the finite duration of benefits by multiplying $r(1-\tau_0)y$ 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). I also multiply by the benefit take-up rate, which Ebenstein and Stange (2010) find to be about 0.8 from 1990-2005, to account for the fact that the empirical quantities used later are defined for the entire population that is eligible for UI, regardless of whether they take up benefits.[18] Therefore, the final equation is $b = r(1-\tau_0)y(0.8)\left(\frac{15.8}{24.3}\right)$; the replacement rate r represents the real-world replacement rate, and b is the corresponding value of an infinite-duration benefit that is equivalent in average dollar terms for an individual who becomes unemployed and is eligible for UI at that replacement rate.

[14] In keeping with the usual approach in the literature, I do not account for transitional dynamics; Lentz (2009), however, finds that such dynamics can have non-negligible effects on results.

[15] Thus, $\bar{s}$ is the maximum feasible search intensity; in the simulations, this upper limit will be binding in very few instances.

[16] The tax rate will not change much with benefits, so I assume that if b changes, benefits continue to be taxed at the baseline rate.

[17] I cannot ignore this issue, as is frequently done in the literature, because the magnitude of government spending on UI relative to other programs defines the size of fiscal externalities.

[18] Gruber (1997) argues that this is in fact the policy-relevant population, because government can control benefit eligibility but not benefit receipt. Fredriksson and Holmlund (2001) does the adjustment slightly differently, finding the average replacement rate for all unemployed individuals whether eligible for UI or not, across both UI and social assistance, accounting for UI benefit exhaustion.

The selected parameter values are summarized in Table 2. The job separation rate is set to $\delta = \frac{1}{260}$ to correspond with a median job duration of 5 years measured by the Bureau of Labor Statistics in January 2006 for high-school graduates, a group which is a reasonable proxy for UI recipients. For the interest rate i, I follow the example of Hansen and İmrohoroğlu (1992) and Chetty (2008) in setting it to zero.[19] The wage y is normalized to one. $R = 2$ is a standard value for the coefficient of relative risk-aversion in studies of UI,[20] but Chetty (2008) states that his results imply a value of about 5 in the context of unemployment,[21] so I use both values. For the upper limit of search intensity, I use $\bar{s} = 0.5$, which means that it is not possible for an individual to guarantee finding a job immediately.

Finally, I select initial values of $r_0 = 0.46$ 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 a 15% federal rate of the typical UI recipient, 5% for a typical state income tax, and 3% for the Medicare tax.[22] In cases where $G = 0$, meanwhile, the tax rate is that which pays for UI benefits, and I assume that the payroll tax is paid only by employees, so benefits are untaxed.

The remaining parameters, θ, κ and ρ, are set to make the model match a set of moments from the real world. The moments used are the unemployment rate u, the percentage gap between average consumption when employed and unemployed $\frac{E(c_e)-E(c_u)}{E(c_e)}$, and the elasticity of the unemployment rate with respect to benefits, which I denote as $E_b^u = \frac{b}{u}\frac{du}{db}$.

The unemployment rate is often used to calibrate job search models, and while all three moments jointly determine the parameter values, u is especially informative about the level of the search cost function, which is primarily determined by θ. E_b^u, meanwhile, is informative about the curvature parameter κ. Finally, the consumption gap is closely related to workers'

[19] 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.

[20] For example, Chetty and Saez (2010) use $R = 2$, and Lentz (2009) estimates a value of 2.21.

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

[22] The employee and employer shares of the Medicare tax add up to 2.9%. This is meant to represent a best approximation of a single tax rate applying to earned income and UI. In terms of the tax rate applying to UI, FICA taxes are not applicable to UI benefits, and some state taxes also do not apply to UI benefits, whereas I also ignore the possibility of some individuals being in higher tax brackets, due to their own income or that of a spouse. A marginal earned income tax rate of $\tau_0 = 0.23$, meanwhile, likely represents a conservative estimate, as I ignore the Social Security tax on the grounds that it is more of a pension contribution than a tax, and I also ignore the fact that some UI recipients may be on the downward-sloping part of the EITC, which would significantly increase the marginal tax on earned income.

Table 2: Parameters

Parameter	Definition	Value(s)
δ	job separation rate	$\frac{1}{260}$
i	real interest rate	0
y	per-period wage	1
r_0	baseline replacement rate	0.46
τ_0	baseline tax rate	0.23
R	coefficient of relative risk-aversion	$\{2, 5\}$
$\bar{s}$	maximum search intensity	0.5
θ	search cost parameter	TBD
κ	search cost parameter	TBD
ρ	annual discount rate	TBD

ability to maintain a buffer stock of assets, so $\frac{E(c_e)-E(c_u)}{E(c_e)}$ primarily identifies the discount rate.[23] These three moments are also derived from the sufficient statistics that will be used later in the paper.

The specific values used for the moments are summarized in Table 3, and are as follows. The unemployment rate u is set to 0.054 to match the average unemployment rate among high-school graduates during 1992-2010; combined with r_0 and τ_0, this implicitly defines $G = 0.208$. Gruber (1997) estimates a relationship of $\frac{E(c_e)-E(c_u)}{E(c_e)} = 0.222 - 0.265r$,[24] which implies a value of 0.1001 at baseline.[25] Finally, Chetty (2008) estimates an elasticity of unemployment durations with respect to benefits of 0.53,[26] though this estimate is based on a sample of UI recipients, whereas the consumption estimates in Gruber (1997) are from a

[23] My decision to set the discount rate to match a moment is unusual, as the standard approach is simply to choose a "reasonable" value for ρ; however, there is no definite consensus on the right "reasonable" value. A number around 4-5% is typical, but to take opposite extremes, Acemoglu and Shimer (2000) use an annual discount rate of nearly 11%, whereas Coles (2008) produces results for a zero discount rate. Lentz (2009) uses a value of 5.1%, but finds that his results are very sensitive to the gap between the interest and discount rates, motivating my attempt to use real-world data to pin down this parameter.

[24] 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. Gruber estimates the year-to-year drop in consumption for individuals who were employed in year $t-1$ and unemployed in t, which is a reasonable approximation to the consumption gap between average unemployed and employed individuals.

[25] In this model, it is difficult to generate a consumption gap which declines with b, because I consider the steady-state, and in steady-state agents accumulate fewer assets as b increases. The variances of $U(c)$ and $U'(c)$ do, however, decrease with b. Since my later sufficient statistics analysis uses Gruber's estimated relationship directly in the welfare derivative, including the strongly negative relationship between benefits and the consumption gap, this strengthens my claim to be testing the robustness of my conclusions to different assumptions.

[26] 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 value of 0.5 is "in the middle range of the available estimates."

sample of unemployed workers who were initially eligible for UI, regardless of whether they were actually receiving benefits. Therefore, I follow Gruber's recommendation and multiply the elasticity by 0.48, the derivative of benefit receipt to benefit eligibility in his sample.[27] If the average duration of unemployment is D, this gives $E_b^D = \frac{b}{D}\frac{dD}{db} = 0.2544$, and the fact that $u = \frac{D}{D+\frac{1}{\delta}}$ means that $E_b^u = (1-u)E_b^D = 0.946 \times 0.2544 = 0.2407$.

Table 3: Moments

Parameter	Definition	Value
u	unemployment rate	0.054
$\frac{E(c_e)-E(c_u)}{E(c_e)}$	consumption gap between employment and unemployment	$0.222 - 0.265r_0 = 0.1001$
$E_b^u = \frac{b}{u}\frac{du}{db}$	elasticity of u wrt b	$0.946 \times 0.48 \times 0.53 = 0.2407$

The model will be calibrated separately four times: for each value of $R \in \{2, 5\}$, I calibrate the model twice, once for $G = 0$ and once for the true $G = 0.208$. In this way, given each set of starting assumptions about R and the size of government, I find the parameters that match the real-world moments, and then find the level of UI that is optimal in each case.[28]

To numerically solve the model for any given set of parameters and a value of b, I begin by making a guess for the tax rate and doing 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 converge, and I only evaluate the maximization problem for a subset of the asset value grid on each iteration, and then use cubic spline interpolation to fill in intermediate points of the value functions, as 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 tax rate is $\tau_0 = 0.23$.

In order to calculate E_b^u, the model must be solved at baseline, and again for a different level of b; since the numerical procedure involves discretizing the asset distribution, the results are slightly "lumpy" at high magnification, so I use a replacement rate of $r = 0.56$,

[27] 0.48 is also very close to my value of $0.8 \times \frac{15.8}{24.3} \simeq 0.52$ for the percentage of time that initially eligible unemployed individuals receive benefits, which is closely related.

[28] This is not a comparative statics exercise; I am studying the effect of fiscal externalities on optimal UI *calculations*, not the impact on optimal UI of increasing the size of government. Governments' fiscal activities have always been much more extensive than just UI, but this has been ignored by the optimal UI literature, and I want to know how much the estimated optimum changes if this fact is no longer ignored.

and compute the resulting arc elasticity.[29] With both sets of numerical results in hand, I 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 match the moments as closely as is practical. Finally, in each case, once the parameters have been calibrated, I perform a grid search over r to find the optimal level.

3.3 Numerical Results

This subsection presents the numerical results; appendix A contains the parameters used in each case, the resulting moments, and a further discussion of the methods used.[30] The results for the optimal replacement rates and estimated welfare gains can be found in Table 4; the columns labelled "Welf. Gain" express the gain from moving from $r = 0.46$ to the optimum as percentage points of initial spending on UI, whereas the next columns, labelled "Diff.", display the welfare gain from moving between the replacement rate believed to be optimal when $G = 0$ and the "true" optimum.[31]

Table 4: Optimal Replacement Rates & Welfare Gains

	$R = 2$			$R = 5$		
	r	Welf. Gain	Diff.	r	Welf. Gain	Diff.
$G = 0.208$	0.00	11.81	6.80	0.46	0.00	6.08
$G = 0$	0.37	0.45	6.33	0.72	5.53	5.53

A non-zero value of G significantly reduces the optimal replacement rate in both cases, but especially when making the standard assumption of $R = 2$. In the latter case, fiscal externalities reduce the optimal benefit level to zero,[32] with an accompanying welfare gain

[29] This variation is comparable to that studied in the empirical literature; for example, Addison and Blackburn (2000) estimate a mean replacement rate of 0.44 with a standard deviation of 0.12 in their data.

[30] An over-identifying moment can be generated from the asset distribution, by comparing my simulated steady-state distribution to the asset distribution in the SIPP data of Chetty (2008). My model does not contain motives for saving other than self-insurance against unemployment, and I also restrict assets to be non-negative, so I cannot match the long left and right tails of a real-world distribution, but I can consider the median of my distribution. In the four cases I consider, the median level of assets generated by the model amounts to between 40% and 62% of a year's pre-tax labour income. If I assume that half of housing equity can be counted as liquid wealth, I find that the median liquid wealth in Chetty's sample is about 40% of mean annual income; if all of housing equity is counted, then median wealth is about 70% of mean income. Therefore, the centre of the asset distribution is on the right order of magnitude.

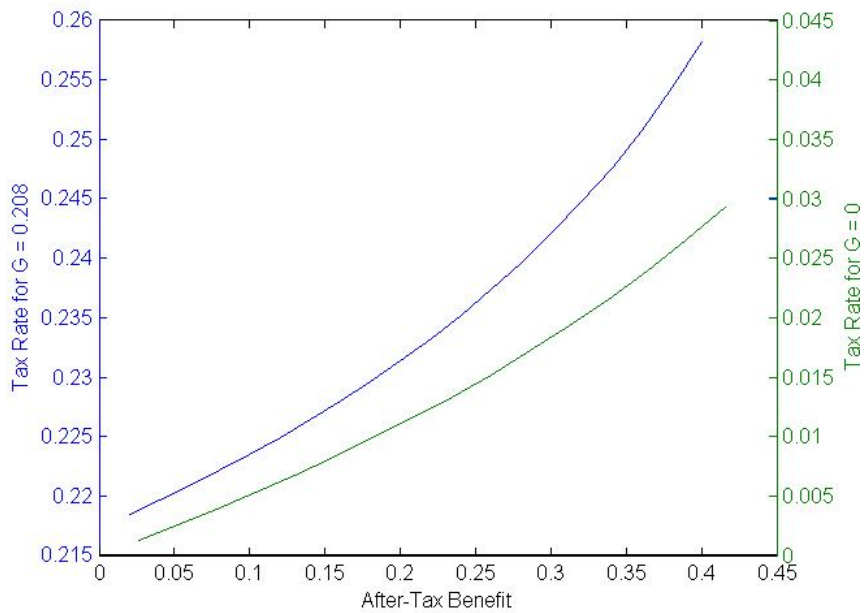
[31] I estimate the welfare gain per worker per year, convert this to dollars using a base of mean consumption while unemployed, divide by annual spending on UI and multiply by 100.

[32] More accurately, the optimum is very close to zero, as close as computationally feasible; an actual zero

of nearly 12% of current spending on UI, or about \$3.5 billion per year out of a typical \$30 billion in spending in normal economic times. The welfare gains of moving from the $G = 0$ optimum to the $G = 0.208$ optimum amount to about 6-7% of UI spending.

Figures 1 and 2, meanwhile, illustrate how much larger the revenue effects of UI's labour market impacts are when I account for income taxes. I plot the budget-balancing income tax rate as a function of the after-tax benefit b for both values of G, with the results for $R = 2$ in Figure 1 and those for $R = 5$ in Figure 2. The two figures are nearly identical, and show that taxes rise much more rapidly with b in the $G = 0.208$ case than when I abstract from G; each dollar given to an unemployed person costs more to the government when G is large.[33]

Figure 1: Tax Rates for $R = 2$

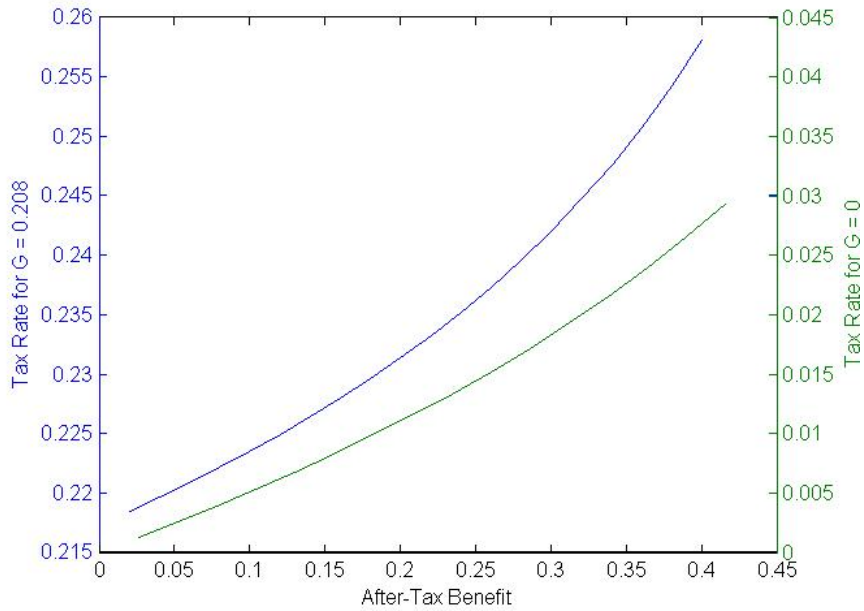


I have performed an extensive series of sensitivity analyses, with results that are displayed in appendix B.1. I consider different values of a wide variety of parameters and moments used in the baseline analysis, and the results are strongly robust to these modifications, with

benefit level, along with a zero interest rate and an upper limit to search intensity below one, would mean a probability of zero consumption in a period that is bounded away from zero.

[33]Since b is the after-tax benefit, a given value of b corresponds to a different replacement rate in the $G = 0.208$ and $G = 0$ cases, because an analysis that assumes $G = 0$ ignores income taxes. In this sense, the results in Table 4 understate the difference in after-tax UI dollars; for example, in the $R = 5$ case, the optimal replacement rate when $G = 0.208$ is about 63.9% of the replacement rate when $G = 0$, whereas the corresponding ratio in after-tax benefits b is about 49.2%.

Figure 2: Tax Rates for $R = 5$



the partial exception of the case with a positive interest rate, which significantly reduces the optimal benefit level, all the way to near zero if $R = 2$ even in the absence of fiscal externalities.

I have also performed two important extensions of the model: I allow for a non-degenerate wage distribution, and I model benefits that expire after 6 months, with results presented in appendix B.2. To summarize briefly, with a wage distribution, UI benefits can raise reservation wages and therefore the average wage in the economy. This leads to higher optimal replacement rates, especially in the case where $G > 0$; therefore, the effect of fiscal externalities on optimal UI may actually be positive due to a positive effect of unemployment benefits on tax revenues. When I model finite-duration benefits, on the other hand, the results appear very similar to the baseline, indicating that my conclusions hold with a more realistic (but less simple) approach to modelling UI benefits.

The results of the analysis are simple and striking: the minor modification of a standard search model to include non-UI spending can substantially affect optimal UI calculations. This analysis, however, does not present a clear demonstration of the mechanisms at work,[34]

[34]Shimer and Werning (2007) 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."

and so I will now switch my focus to a a simpler and more reduced-form model of unemployment in the sufficient statistics tradition. Specifically, I will focus on Martin Baily's (1978) original seminal paper in optimal UI, which will permit a more detailed, step-by-step analysis of fiscal externalities in the context of UI, and will generate an analytical solution as an intuitive function of observable empirical quantities.

4 Baily (1978) Model

Baily (1978) represents the first application of the sufficient statistics approach to optimal UI; his formula is used by Gruber (1997), and Chetty (2006) demonstrates that it applies to a wide range of job search models. Therefore, while the sufficient statistics approach has also been used in somewhat different ways by Chetty (2008) and Shimer and Werning (2007), I will return to Baily's original paper and analyze his two-period model of unemployment. 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 notation from Baily's paper is also altered to make it more compatible with the notation used earlier. The model is more reduced-form than the structural model earlier, but captures many of the same features, and the simplicity of the model makes it well suited to an exposition of the effects of fiscal externalities.

In Baily's model, time is finite and consists of two periods,[35] with the interest and discount rates both set to zero. In the first period, the representative worker is employed at an exogenous wage y,[36] and between periods they face a risk of unemployment: with exogenous probability δ the worker loses their job and becomes unemployed, whereas they keep their initial job at the same wage for the entire second period with probability $(1-\delta)$. If the worker becomes unemployed, they choose search effort e (normalized into income units) and a desired wage y_n. They will then spend a fraction $(1-s)$ of the second period unemployed

[35] Baily meant this to represent a two-year time horizon, but the model could also stand for a world with a longer time horizon divided into two halves.

[36] In an extension in appendix G.4, I consider the effect of allowing choice over initial labour supply.

and the remaining $s \in (0,1)$ at a new job at wage y_n,[37] where s is a deterministic function of e and y_n:[38]

$$s = s(e, y_n), \quad \frac{\partial s}{\partial e} > 0, \quad \frac{\partial s}{\partial y_n} < 0.$$

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, so overall expected utility is given by:

$$V = U(c_e^1) + (1-\delta)U(c_e^2) + \delta U(c_u) \tag{1}$$

where $c_e^1 = y(1-\tau) - k$, $c_e^2 = y(1-\tau) + k$ and $c_u = (1-s)(b-e) + sy_n(1-\tau) + k$,[39] and τ remains the income tax rate and b the after-tax UI benefit.

The government budget constraint over the two periods is:

$$[(2-\delta)y + \delta s y_n]\tau = \delta(1-s)b + 2G \tag{2}$$

where G again represents the new per-period exogenous government expenditures.

4.2 Calculation of Welfare Derivative

The next step is to evaluate the derivative of social welfare with respect to UI benefits. From (1), the worker's lifetime expected utility can be written generally as $V = V(e, y_n, k; b, \tau)$; the government sets the values of b and τ, and the worker chooses $\{e, y_n, k\}$ to maximize V taking $\{b, \tau\}$ as given. The government planner has equally-weighted utilitarian preferences, and therefore wants to maximize V at the individual's optimum, choosing b and τ so that the government budget constraint is satisfied in equilibrium. Since the individual chooses $\{e, y_n, k\}$ to maximize V, the partial derivatives with respect to these individual choices are zero: $\frac{\partial V}{\partial e} = \frac{\partial V}{\partial y_n} = \frac{\partial V}{\partial k} = 0$. In other words, for a small change in b, behavioural responses have no first-order effect on welfare, and the envelope theorem implies 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:

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

[37] y_n is assumed to be deterministic, such that 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.

[38] In appendix G.1, I examine the consequences of a stochastic s.

[39] 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 G.2.

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. The second term, meanwhile, 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 derivatives:

$$\begin{aligned}\frac{\partial V}{\partial b} &= \delta U'(c_u)\frac{\partial c_u}{\partial b} \\ &= \delta(1-s)U'(c_u) \qquad (4)\\ \frac{\partial V}{\partial \tau} &= U'(c_e^1)\frac{\partial c_e^1}{\partial \tau} + (1-\delta)U'(c_e^2)\frac{\partial c_e^2}{\partial \tau} + \delta U'(c_u)\frac{\partial c_u}{\partial \tau} \\ &= -yU'(c_e^1) - (1-\delta)yU'(c_e^2) - \delta s y_n U'(c_u). \qquad (5)\end{aligned}$$

Therefore, the welfare derivative is:

$$\frac{dV}{db} = \delta(1-s)U'(c_u) - \left[yU'(c_e^1) + (1-\delta)yU'(c_e^2) + \delta s y_n U'(c_u)\right]\frac{d\tau}{db}. \qquad (6)$$

The goal now is to rewrite (6) in a form that can be expressed in terms of observable empirical quantities, so it can be used numerically. I will leave $\frac{d\tau}{db}$ as it is for the time being, and return to it later. To simplify (6), I replace $U'(c_e^1)$ and $U'(c_e^2)$ using the individual's first-order condition for saving 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 θ is between c_u and c_e^1. This allows $\frac{dV}{db}$ to be written in terms of $U'(c_u)$. I then make two technical assumptions that are also made by Baily: I assume that $y_n = y$ and $c_e^1 U''(\theta) = c_u U''(c_u)$ in equilibrium; further discussion can be found in appendix C.1. These assumptions further simplify the expression, and I divide by $U'(c_u)$ to put welfare into a dollar equivalent and solve for $\frac{dW}{db} \equiv \frac{\frac{dV}{db}}{U'(c_u)}$, arriving at the results summarized in the following proposition.

Proposition 1. *Under the assumptions that $y_n = y$ and $c_e^1 U''(\theta) = c_u U''(c_u)$ in equilbrium, the marginal value of increased benefits is given by:*

$$\frac{dW}{db} = 2y\frac{\Delta c}{c_e^1}R\frac{d\tau}{db} - 2(1-u)y\left[\frac{d\tau}{db} - \omega\right] \qquad (7)$$

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

$$\frac{\Delta c}{c_e^1}R = (1-u)\frac{\frac{d\tau}{db} - \omega}{\frac{d\tau}{db}}. \qquad (8)$$

Proof. The proof of this result can be found in appendix C.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:*

$$\frac{dW}{db} = \frac{2u}{(1-u)\psi}\left[\frac{\Delta c}{c_e^1}RE_b^\tau - (1-u)\left(E_b^\tau - \psi\right)\right] \tag{9}$$

where $E_b^\tau = \frac{b}{\tau}\frac{d\tau}{db}$ *is the elasticity of* τ *with respect to* b*, and* $\psi = \frac{\omega b}{\tau} = \frac{ub}{ub+G}$ *is the fraction of total government expenditures allocated to UI. Thus the equation for the optimum is:*

$$\frac{\Delta c}{c_e^1}R = (1-u)\frac{E_b^\tau - \psi}{E_b^\tau}. \tag{10}$$

The meaning of these results will be interpreted in the next two subsections.[40]

4.3 Analysis of Optimal Benefit Equation

In order to be able to use the equations derived above, I need to evaluate the response of taxes to benefits. I begin with (7) and (8), for which I must evaluate $\frac{d\tau}{db}$. Total differentiation of the government budget constraint (2) gives:

$$\frac{d\tau}{db} = \frac{\delta(1-s) - \delta b\frac{ds}{db} - \delta\tau y_n\frac{ds}{db} - \delta s\tau\frac{dy_n}{db}}{(2-\delta)y + \delta s y_n}. \tag{11}$$

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 response to UI, if b increases, the tax rate must increase to compensate, and $\frac{\delta(1-s)}{(2-\delta)y+\delta s y_n}$ represents the size of this increase. The second component will be referred to as the "duration effect," and 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 "revenue effects": the first shows how longer unemployment durations also reduce the amount of taxes paid on labour income, raising the required tax increase still further, while the second captures the gain in tax revenues if higher UI increases y_n, reducing the necessary tax increase. Notice that while the magnitude of the

[40] It is clear from (5) that $\frac{\partial V}{\partial \tau} < 0$, and the pair of Baily's assumptions described in appendix C.1 imply that $\frac{\partial V}{\partial \tau} = -2yU'(c_u)\left[(1-u) - \frac{\Delta c}{c_e^1}R\right]$; thus, if those assumptions are accurate, it follows that $\frac{\Delta c}{c_e^1}R < 1-u$. This result will be used in section 5.

duration effect doesn't depend on the size of government, the revenue effects are multiplied by τ; this highlights the importance of the standard assumption that τ is a small payroll tax, rather than a large income tax.

It can now be seen that the ω defined in Proposition 1 is exactly equal to the mechanical effect, meaning that the $\frac{\frac{d\tau}{db}-\omega}{\frac{d\tau}{db}}$ term on the right-hand side of (8) is the fraction of the total response of taxes to benefits generated by the duration and revenue effects. (8) can now be understood 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 consumption-smoothing, which is increasing both in the magnitude of risk-aversion and the consumption shock upon unemployment, while the right-hand side represents the cost of increased UI in terms of behavioural effects on the government budget. The mechanical effect ω represents a lump-sum transfer of income between employed and unemployed states, and thus is not a cost to society, so $\frac{\frac{d\tau}{db}-\omega}{\frac{d\tau}{db}}$ 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.

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^τ represents the percentage increase in the tax rate for a percentage increase in b, while ψ, the share of government expenditures on UI, is the percentage increase in τ required for mechanical reasons. As in (8), therefore, the final fraction on the right is the fraction of the response of taxes to benefits caused by behavioural responses. Using (11), E_b^τ can be written as:

$$E_b^\tau = \psi + \left(\psi + \frac{u}{1-u}\right) E_b^D - \frac{\delta s}{2(1-u)} E_b^y \tag{12}$$

where $E_b^D = \frac{b}{1-s}\frac{d(1-s)}{db}$ is the elasticity of unemployment durations with respect to b, and $E_b^y = \frac{b}{y_n}\frac{dy_n}{db}$ is the elasticity of post-unemployment wages y_n with respect to b. The four components of the tax response are apparent here as well: the first ψ is the mechanical effect, the ψ and $\frac{u}{1-u}$ multiplying E_b^D represent the duration effect and the first revenue effect respectively, and the final term is the second revenue effect.

Therefore, the equation for the optimum is:

$$\frac{\Delta c}{c_e^1} R = (1-u) \frac{\left(\psi + \frac{u}{1-u}\right) E_b^D - \frac{\delta s}{2(1-u)} E_b^y}{\psi + \left(\psi + \frac{u}{1-u}\right) E_b^D - \frac{\delta s}{2(1-u)} E_b^y}. \tag{13}$$

This is the equation which I will use to solve for the optimal value of b. However, even without a numerical analysis, I can make certain observations about this result. The standard

assumption that $G = 0$ would mean that $\psi = 1$; $\frac{u}{1-u}$ and $\frac{\delta s}{2(1-u)}$ are both likely to be relatively small numbers, so the mechanical and duration effects will be large compared to the revenue effects, because the taxes that induce the revenue effects are so small. However, if G is large, ψ will be small, meaning that the revenue components will be at least on the same order of magnitude as the mechanical and duration components.

Baily (1978) assumes $E_b^y = 0$, and joins the rest of the literature in assuming that $G = 0$ and thus $\psi = 1$. However, in order to get his result, Baily makes an additional assumption which is equivalent to assuming that the E_b^τ in the denominator of the right-hand side of (10) is equal to one, or that $\psi + \left(\psi + \frac{u}{1-u}\right) E_b^D = 1$ in the denominator of (13).[41] In this case, (13) collapses to exactly the result in Baily (1978), also found in Chetty (2006):

$$\frac{\Delta c}{c_e^1} R = (1-u)\left(1 + \frac{u}{1-u}\right) E_b^D = E_b^D.$$

As mentioned before, $\frac{\Delta c}{c_e^1}$, R and E_b^D are the three original sufficient statistics for optimal UI. As in my results, the optimal UI benefit level balances off the consumption-smoothing benefits of UI with the cost from behavioural distortions.

5 Analytical Results

In this section, I will further analyze the equations derived above, and specifically I will present a series of analytical results about the 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 $b^*(G)$, the estimated optimal value of b for a given value of G. I consider how the results change when estimated quantities like G and E_b^y are changed.

It should be emphasized that this is not a comparative statics exercise, as I am not considering a change to a primitive parameter of the model; rather, I consider how the numerical results should be expected to change when the estimated value of G used in the calculations is altered. This represents a change in assumptions about the model, not a change in parameters, and so the values of the sufficient statistics are unaltered, since they reflect the unchanged real world to which the model is calibrated. A helpful thought

[41] Baily claims that this assumption leans towards overestimating the optimal b, to offset the conservative assumption that $c_e^1 U''(\theta) \simeq c_u U''(c_u)$. However, this is incorrect; the former assumption is in fact conservative, and the latter depends on the value of the risk-aversion parameter.

experiment is that of the "two researchers": one who assumes 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; my analysis answers the question: who will estimate a larger optimal b, and by how much?

Throughout this section, I maintain two technical assumptions; the first is that $\frac{\Delta c}{c_e^1}R < 1-u$, and the second is that W (the integral of $\frac{dW}{db}$) is strictly quasi-concave in b. Footnote 40 points out that the first assumption follows immediately from $\frac{\partial V}{\partial \tau} < 0$ if Baily's assumptions are accurate. The latter assumption, meanwhile, ensures that W is single-peaked, 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 $sE_b^y - (1-s)E_b^D$, or equivalently the same sign as $\frac{d(sy_n)}{db}$.*

Proof. The proof of this result can be found in appendix C.2. □

Therefore, if two researchers use (9) 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 $\frac{d(sy_n)}{db}$ is positive. Ignoring G greatly understates the revenue effects of changing b, and while higher UI is expected to increase durations of unemployment, it may also increase wages. If this wage effect is so large as to lead to an increase in total post-unemployment earnings sy_n, which is the only non-exogenous component of total earnings in the model, the overall revenue effect is positive and welfare-increasing. Therefore, using a positive value of G, which implies higher taxes, amplifies this positive revenue effect and increases the welfare gain from raising benefits. If $\frac{d(sy_n)}{db}$ is negative, the reverse holds.

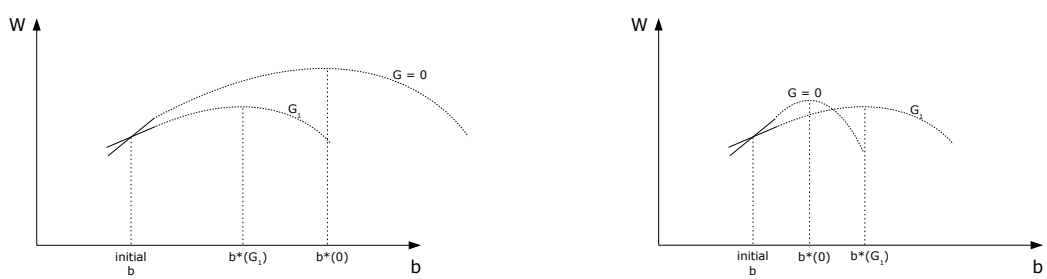
An immediate corollary arising from quasi-concavity is that, if the baseline welfare derivative is zero for $G = 0$, and thus the current level of b is estimated to be optimal in that case, then the optimum for the true G will be larger or smaller according to the sign of $\frac{d(sy_n)}{db}$. For example, if $\frac{d(sy_n)}{db} > 0$, $\frac{dW}{db}(b; G_1) > 0$ and quasi-concavity means that the optimum must be found at a higher b. A similar logic applies if $\frac{dW}{db}(b; G_1) = 0$; if one of the welfare derivatives

is zero, I only need to know the other to make a comparison. This result is summarized by the following corollary.

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 $sE_b^y - (1-s)E_b^D > 0$, or equivalently if and only if $\frac{d(sy_n)}{db} > 0$.*

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 $\frac{dW}{db}(b;G_1) < 0$, then clearly $b^*(0) > b^*(G_1)$. For a more general result, however, I need to go beyond the local welfare derivative and make out-of-sample assumptions; as an illustration, consider Figure 3, which displays graphically how knowledge of a local welfare derivative doesn't permit unambiguous conclusions about the optimum.[42] 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 available data is used to select the best functional form of that statistic with respect to b, allowing for an extrapolation of $\frac{dW}{db}$ out of sample to find the optimum. For this purpose, I define $\chi = \{\frac{\Delta c}{c_e^1}, R, s, E_b^D, E_b^y\}$ as the vector of sufficient statistics, the 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.

Figure 3: Two Possible Welfare Functions



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 $sE_b^y - (1-s)E_b^D > 0$, or equivalently if and only if $\frac{d(sy_n)}{db} > 0$, in between $b^*(0)$ and $b^*(G_1)$.*

[42] The values of W in the diagram are normalized to be equal at the initial b, but while it is clear that $\frac{dW}{db}(b;0) > \frac{dW}{db}(b;G_1)$ in the diagram, the dotted lines further to the right are meant to indicate that the sufficient statistics alone give no definite answer about the shape of these curves.

[43] Strict quasi-concavity of W, when the latter is estimated out of sample using statistical extrapolations, implicitly places some restrictions on the extrapolations allowed.

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 $\frac{d(sy_n)}{db}$ at $b^*(0)$. If that sign is positive, then by strict quasi-concavity $b^*(G_1) > b^*(0)$, and $\frac{d(sy_n)}{db}$ 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 $\frac{d(sy_n)}{db} > 0$ in between the optimal values of b; the proof explains why the sign of $\frac{d(sy_n)}{db}$ will not change in that region.[44] This is arguably the most important result in this section, and provides general intuition about fiscal externalities, as well as explaining my numerical results. If UI benefits increase total earnings, this welfare-increasing fiscal externality will appear larger when I account for larger taxes, and the optimal benefit level will increase. If, on the other hand, the effect of UI on wages is zero, 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 negative. As demonstrated in the structural results earlier, the reduction in the optimal benefit level in this case can be substantial.

Next, I present results on the role of E_b^y, to demonstrate that the value of this parameter could be important;[45] these results are straightforward, and begin with the following proposition.

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

Proof. The proof of this result can be found in appendix C.3. □

A higher value of E_b^y means that 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.

[44] This is not, however, a restrictive assumption relying on quasi-concavity. 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 $\frac{d(sy_n)}{db}$. Supposing that $\frac{d(sy_n)}{db} > 0$, b^* will only increase with G as long as it stays in a range where $\frac{d(sy_n)}{db} > 0$, so it can never increase out of this range, and thus a change in the estimated G can never move the estimated optimal b enough to change the sign of $\frac{d(sy_n)}{db}$. If there is a value of b for which $\frac{d(sy_n)}{db}$ takes the opposite sign, there must be no value of G such that this b would be optimal.

[45] The notation is now slightly altered to allow E_b^y to enter $\frac{dW}{db}$ and b^* as an argument, as I now consider how results vary when different values of E_b^y are selected.

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

Corollary 5. *For statistical extrapolations of $\chi_1 = \{\frac{\Delta c}{c_e^1}, R, s, E_b^D\}$ that do not depend on the estimated value of E_b^y, i.e. $\chi_1(b; E_b^y) = \chi_1(b)$, $b^*(G, E_b^{y2}) > b^*(G, E_b^{y1})$.*

If the current b is the estimated optimal value for one of the values of E_b^y under consideration, or if the signs of $\frac{dW}{db}$ are opposite, then I can make an unambiguous statement. More generally, once I define a statistical extrapolation that does not depend on E_b^y, I can state that a researcher choosing a larger value of E_b^y will always find a larger optimal b.

I have now shown that higher E_b^y 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 concern the interaction of G and E_b^y. As already mentioned, Baily (1978) is among the few papers that acknowledge the fact that a parameter like E_b^y could enter into optimal UI calculations, but he ultimately drops this parameter from his equation on the grounds that, since the UI payroll tax is quite small, it will have little effect on the results. However, E_b^y is more important when G is large, both to social welfare and to the calculation of the optimal value of b; a demonstration of this begins with the following proposition.

Proposition 4. *For $G_1 > 0$ and $E_b^{y2} > E_b^{y1}$, $\frac{dW}{db}(b; G_1, E_b^{y2}) - \frac{dW}{db}(b; G_1, E_b^{y1}) > \frac{dW}{db}(b; 0, E_b^{y2}) - \frac{dW}{db}(b; 0, E_b^{y1})$.*

Proof. The proof of this result can be found in appendix C.4. □

This proposition says that the effect of E_b^y on the welfare derivative is increasing in G. Thus, it may be true that a researcher who ignores G will find that the value of E_b^y is relatively unimportant to their calculations, but when G is large, the tax rate will also be large, and E_b^y will matter far more to the value of the welfare derivative. Proposition 4 can also be interpreted as saying that the importance of G to the welfare derivative is increasing in E_b^y.

The final analytical result addresses the question of whether E_b^y is more important in determining the value of the optimal b when G is large. The results so far make it logical to suspect that $b^*(G_1; E_b^y) - b^*(0; E_b^y)$ is increasing in E_b^y, i.e. that the increase in b^* caused by G is more positive when E_b^y is larger; after all, I have proved that b^* is increasing in E_b^y, and that the difference in the welfare derivative for different values of E_b^y is increasing in

G. This, however, cannot be proved without unusual and unintuitive assumptions; I can, however, prove a somewhat weaker result, as summarized below.

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

- *if $\exists$ an E_b^{y*} such that $b^*(G_1, E_b^{y*}) = b^*(0, E_b^{y*})$, then $b^*(G_1, E_b^{y2}) > b^*(0, E_b^{y2})$ for $E_b^{y2} > E_b^{y*}$ and $b^*(G_1, E_b^{y1}) < b^*(0, E_b^{y1})$ for $E_b^{y1} < E_b^{y*}$.*

Proof. The proof of this result can be found in appendix C.4. □

The appendix discusses the sufficient conditions in further detail, and explains why they are plausible and that they hold for all of the non-zero numerical results in the next section. This proposition says that, although I cannot prove the stronger condition that $b^*(G_1; E_b^y) - b^*(0; E_b^y)$ is increasing in E_b^y, I can state that for small values of E_b^y, $b^*(G_1, E_b^y) < b^*(0, E_b^y)$, and vice-versa for sufficiently large values of E_b^y;[46] this result is summarized by the diagram in Figure 4. Therefore, at least locally around E_b^{y*}, the stronger condition will hold, and I will show in my numerical results that the stronger condition does describe the general behaviour of b^* for the parameters and functional forms that I use.

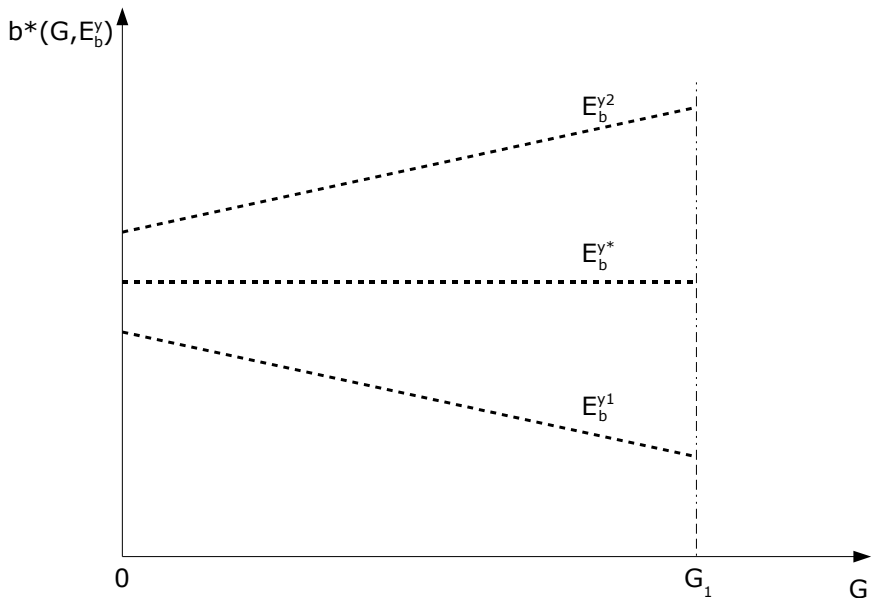
The results derived in this section apply in particular to UI, but similar results will also apply in the context of other government programs with impacts on the labour market. The idea that the direction of the change in optimal policy caused by fiscal externalities depends only on the direction of the program's effect on total taxable income is intuitive and more general than the current context, as is the result that effects of a program on wages are more important when the full size of government is taken into account.

6 Numerical Results from the Baily Model

My numerical analysis of the Baily model begins with a brief subsection describing the method of statistical extrapolation; parameter values are then discussed in the second subsection, and the optimal benefit levels are calculated in the third subsection. The final two subsections contain a comparison of the results with those obtained earlier from the dynamic job search model, and a general discussion of the results.

[46] In the unlikely case that an increase in E_b^y causes an increase in the optimal b which makes s decrease sufficiently quickly, the critical value E_b^{y*} may not exist; in that case, $b^*(G_1) - b^*(0)$ is always negative as long as $E_b^u > 0$.

Figure 4: Consequences of Proposition 5



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 as in the structural analysis earlier, I will do the same. As before, I define the replacement rate as $r = \frac{b}{(0.8)\left(\frac{15.8}{24.3}\right)y(1-\tau_0)}$, where τ_0 is the baseline real-world tax rate, 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. 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 $\frac{\Delta c}{c_e^1}$ as a function of r [$\frac{\Delta c}{c_e^1} = 0.222 - 0.265r$]
- select fixed values of E_b^D [0.2544], E_b^y [0.048] and R [2]
- select current values of r [0.46], u [0.054] and s [0.8]
- use $2u = \delta(1-s)$ to solve for the fixed value of δ [0.54]

- use the fixed value of E_b^D to define a functional form for s with respect to r: $(1-s) = \phi r^{E_b^D}$, and use the current values of s and r to solve for ϕ [0.2437]

- select the current value of ψ [0.0457], and specify the relationship of ψ to r $\left[\psi = 0.77\frac{12.64}{24.3}\frac{ur}{(1-u)\tau}\right]$

- 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 simply an illustration of the method. The values listed above imply an optimal $r = 0.393$, or a replacement rate of 39.3%; meanwhile, if $G = 0$, this affects only the second-last step, as $\psi = 1$ for all values of r, and the optimal r would be approximately 0.472. However, now suppose I had initially chosen a larger value of E_b^y, such as 0.096; now the optimal values of r are 0.594 using my estimate of G and 0.484 when $G = 0$. As predicted earlier, a higher value of E_b^y leads to a higher optimal b for each value of G. Meanwhile, in the first case, with a relatively small value of E_b^y, a larger estimate of G reduces the optimal replacement rate, whereas with a larger value of E_b^y, the opposite happens, exactly as Proposition 5 predicted.

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 of 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

Many of the quantities for the sufficient statistics have already been used earlier in the paper, so the discussion will be kept brief. Starting with the functional form of $\frac{\Delta c}{c_e^1}$, I use the estimate of $\frac{\Delta c}{c_e^1} = 0.222 - 0.265r$ from Gruber (1997). I also continue to use $E_b^D = 0.48 \times 0.53 = 0.2544$ 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.[47] Since I am focussing on the effect of higher benefit levels on wages, I list only papers

[47] None of the papers listed report coefficients in the form of an elasticity, so their coefficients have been transformed into approximate elasticities using mean values of wages and benefit levels. Classen (1977) and Holen (1977) do not provide summary statistics, so I use mean values from Burgess and Kingston (1976), who use a smaller version of the dataset used by Holen (1977).

addressing that specific question,[48] though numerous papers estimate the effect of longer benefit durations on wages, or of some dimension of benefit generosity on other measures of job quality.[49]

Table 5: Results of Empirical Literature on Benefit Elasticity of Wages

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

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

Chetty (2008), in particular, 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.[50] This corresponds with most recent studies in suggesting smaller or zero effects of UI on wages, but since the literature covers a wide range of values, and the analytical results in section 5 suggest that E_b^y can have an important effect on the optimal benefit level, I use a set of possible values for E_b^y covering the range found in Table 5,[51] specifically $E_b^y = 0.48 \times \{-0.17, 0, 0.1, 0.2, 0.4, 0.64\}$.[52]

[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), whose findings correspond to an initial elasticity of 0.10 which declines over the spell of unemployment. Additionally, the estimate listed for Meyer (1989) is from one of 10 individual regressions; the author does not designate a preferred estimate, so the basic difference-in-differences is used.

[49]Among these are Gaure, Roed, and Westlie (2008), who find a positive effect of benefit duration on wages, and Lalive (2007) and Schmieder, von Wachter, and Bender (2012), who do not (the latter paper finds a negative effect); Centeno (2004), Centeno and Novo (2006), and Tatsiramos (2009), who find that more generous UI leads to greater subsequent job duration, and Portugal and Addison (2008), who do not.

[50]These findings are at least suggestive, but 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).

[51]This elasticity could conceivably be negative; Welch (1977) predicts a reduction of wages by firms to compensate for higher benefits which they must finance due to experience rating.

[52]Since estimates of E_b^y are based on samples of UI recipients, I once again multiply by 0.48.

The baseline value of r is set to 0.46 as before, and I again use an initial unemployment rate of $u_0 = 0.054$. The value for s_0, meanwhile, depends on the way the structure of the model is interpreted. If the two periods are taken literally to represent two years, then the finding of Chetty (2008) that the mean unemployment duration in his sample is 18.3 weeks implies an estimate of $s_0 = \frac{52-18.3}{52} = 0.648$. If, however, the model represents a larger portion of an individual's working life, perhaps its entirety, then the fact that Farber (1999) finds that 20.9% of workers aged 45-64 had at least 20 years of tenure in 1996 can be interpreted to mean that $\delta = 0.791$, so $s_0 = 1 - \frac{2u_0}{\delta} = 0.863$. To cover this range of possibilities, I use the set of values given by $s_0 = \{0.648, 0.725, 0.8, 0.863\}$.

The values for δ and ϕ for each value of s_0 are easy to calculate, so I will not discuss them further here. Finally, at baseline values, $\psi = \frac{ub}{(1-u)\tau y} = \frac{u}{1-u}\frac{1-\tau_0}{\tau}(0.8)\left(\frac{15.8}{24.3}\right)r$, and the baseline tax rate is $\tau_0 = 0.23$, so $\psi = 0.77\frac{12.64}{24.3}\frac{ur}{(1-u)\tau}$, which equals 0.0457 at baseline.

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 is pursued in the following subsection.

Table 6: Sufficient Statistics & Parameters

Statistic	Value/Extrapolation
$\frac{\Delta c}{c_e^1}$	$0.222 - 0.265r$
E_b^D	0.2544
E_b^y	$0.48 \times \{-0.17, 0, 0.1, 0.2, 0.4, 0.64\}$
R	$\{2, 5\}$
r_0	0.46
u_0	0.054
s_0	$\{0.648, 0.725, 0.8, 0.863\}$
δ	$\{0.3068, 0.3927, 0.5400, 0.7883\}$
ϕ	$\{0.4289, 0.3351, 0.2437, 0.1669\}$
ψ	$0.77\frac{12.64}{24.3}\frac{ur}{(1-u)\tau}$

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 $r \in [0, 2]$,[53] and I report a numerical check on the second-order conditions in appendix D.

[53] I assume that the government is not interested in extracting payments from unemployed workers, so any zeros reported in the tables are corner solutions.

Tables 7 and 8 below present the optimal r for my parameter values, as well as the results when I set $G = 0$. The latter case does not perfectly reproduce Baily's results; to do so, I also need to make the extra assumptions made by Baily, in which case the results vary only with R, giving $r = 0.3577$ for $R = 2$ and $r = 0.6457$ for $R = 5$.

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

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.4500	0.4459	0.4390	0.4275
	0	0.4595	0.4595	0.4595	0.4595
E_b^y	0.048	0.4651	0.4675	0.4717	0.4789
	0.096	0.4707	0.4756	0.4842	0.4987
	0.192	0.4821	0.4921	0.5095	0.5399
	0.3072	0.4959	0.5122	0.5410	0.5922

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0	0	0	0
	0	0	0	0	0
E_b^y	0.048	0.0301	0.2633	0.3933	0.5237
	0.096	0.3737	0.4711	0.5959	0.7591
	0.192	0.5650	0.6835	0.8514	1.0841
	0.3072	0.7152	0.8638	1.0789	1.3788

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

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.6831	0.6815	0.6787	0.6741
	0	0.6866	0.6866	0.6866	0.6866
E_b^y	0.048	0.6887	0.6897	0.6914	0.6942
	0.096	0.6909	0.6928	0.6962	0.7020
	0.192	0.6951	0.6991	0.7061	0.7183
	0.3072	0.7004	0.7069	0.7184	0.7390

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.5518	0.5307	0.4953	0.4387
	0	0.5996	0.5996	0.5996	0.5996
E_b^y	0.048	0.6276	0.6402	0.6620	0.6987
	0.096	0.6554	0.6806	0.7240	0.7970
	0.192	0.7102	0.7601	0.8452	0.9853
	0.3072	0.7741	0.8523	0.9831	1.1914

The difference between the $G = 0$ and $G = 0.208$ cases is especially large for $R = 2$; for low values of E_b^y, fiscal externalities from the income tax cause the optimal replacement rate to drop to zero, whereas for higher values of s_0 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 $G > 0$, the optimal replacement rates spread out noticeably, becoming more sensitive to both E_b^y and s.

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 22 and 23 in appendix E. A direct comparison to that paper, however, is limited by the fact that the models are different, as well as by the different ways 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

To examine the robustness of my numerical results, I compare the results from the sufficient statistics method to those obtained from the structural analysis of the dynamic job search model earlier in the paper. For the subset of cases considered in the structural analysis, I use the specific values of the moments generated by the calibrated model, which can be found in Table 14, and input them into the statistical extrapolation. This produces the results displayed in Tables 9 and 10; the left side of each table contains the structural results and simply replicates Table 4, while the right side displays my new results, where the welfare gains once again represent percentage points of baseline UI spending.

Table 9: Comparison of Results with $E_b^y = 0$ and $R = 2$

Structural Model

	r	Welf. Gain	Diff.
$G = 0.208$	0.00	11.81	6.80
$G = 0$	0.37	0.45	6.33

Statistical Extrapolations

	r	Welf. Gain	Diff.
$G = 0.208$	0.00	46.37	46.27
$G = 0$	0.46	0.00	11.21

Table 10: Comparison of Results with $E_b^y = 0$ and $R = 5$

Structural Model

	r	Welf. Gain	Diff.
$G = 0.208$	0.46	0.00	6.08
$G = 0$	0.72	5.53	5.53

Statistical Extrapolations

	r	Welf. Gain	Diff.
$G = 0.208$	0.60	4.02	1.65
$G = 0$	0.69	10.25	1.58

The results for the optimal replacement rate are remarkably similar except for a modest difference in the case with $R = 5$ and $G = 0.208$.[54] Meanwhile, the estimated welfare gains from the sufficient statistics method tend to be larger than those from the structural model, largely due to differences in models and assumptions about elasticities. In particular, in the

[54] In the latter case, the difference is generated partly by the fact that E_b^D increases with b in the dynamic job search model, whereas it is held fixed during statistical extrapolations, but also because of other differences in the models, such as the fact that the steady-state asset distribution changes with b in the structural model.

$R = 2$ case, the estimated welfare gain of moving from $r = 0.46$ to $r = 0$ of nearly 50% of initial UI spending should not be taken seriously, as this is due to the assumption that unemployment goes to zero as r approaches zero.[55]

A similar comparison can be done for the case with a non-degenerate wage distribution, using the moments from the structural model in appendix B.2, and assuming $s_0 = 0.8$ for the statistical extrapolations; these results can be found in Tables 11 and 12. Here the results are somewhat different; allowing for a wage distribution significantly increases optimal replacement rates in all cases, but by much less in the statistical extrapolation method than when using the structural model. It is hard to compare this result to previous studies allowing for a wage distribution, simply because there have been so few such studies, with Acemoglu and Shimer (2000) the only optimal UI paper that I am aware of that considers a non-degenerate wage distribution.[56] However, the effect of fiscal externalities on the optimal benefit level is very similar across the different approaches.

Table 11: Comparison of Results with $E_b^y \simeq 0.0876$ and $R = 2$

Structural Model

	r	Welf. Gain	Diff.
$G = 0.208$	0.71	4.27	1.65
$G = 0$	0.55	0.48	1.64

Statistical Extrapolations

	r	Welf. Gain	Diff.
$G = 0.208$	0.57	1.24	0.77
$G = 0$	0.48	0.04	0.57

Table 12: Comparison of Results with $E_b^y \simeq 0.0876$ and $R = 5$

Structural Model

	r	Welf. Gain	Diff.
$G = 0.208$	1.08	51.98	0.06
$G = 0$	1.07	41.05	0.29

Statistical Extrapolations

	r	Welf. Gain	Diff.
$G = 0.208$	0.71	13.38	0.06
$G = 0$	0.70	11.04	0.06

6.5 Discussion of Numerical Results

The results in Tables 7 and 8 strongly support the predictions of section 5; in particular, the optimal replacement rate is lower for $G > 0$ if E_b^y is relatively small, but higher if E_b^y is sufficiently positive, as predicted by Proposition 5. In fact, at least for $r > 0$, the results

[55] If, instead of allowing $\frac{dW}{db}$ to become very negative as $r \to 0$, I hold it constant at its $r = 0.288$ value (which is the peak of $\frac{dW}{db}$), which is likely a conservative assumption, the welfare gain drops to about 20%.

[56] The wage distribution in Acemoglu and Shimer (2000) only features two mass points.

also support the stronger condition that $b^*(G_1) - b^*(0)$ is increasing in E_b^y. The prediction of Corollary 3 is also supported for all sets of parameter values: when $b^*(G_1) > b^*(0)$, $sE_b^y - (1-s)E_b^D$ is found to be positive for all $b \in [b^*(0), b^*(G_1)]$, and vice-versa when $b^*(G_1) < b^*(0)$.

The results from both approaches demonstrate that fiscal externalities can alter the nature of the optimal UI problem and significantly change the numerical results. In terms of practical policy recommendations, however, there is a certain amount of ambiguity. The most typical values of R and E_b^y in recent work are 2 and 0, respectively, in which case the optimal replacement rate is zero; however, given the sparsity of the literature estimating the effect of UI on wages, as well as uncertainty about the appropriate level of risk-aversion in the context of unemployment, it would be premature to advocate abolishing unemployment insurance. A value of E_b^y that is small enough to fit within the 95% confidence interval of nearly all empirical studies could partially or completely offset the negative fiscal effects of increased durations of unemployment and leave the optimal level of UI unchanged. Meanwhile, even larger effects of UI on wages, or a high degree of risk-aversion, could indicate that workers are currently underinsured and that UI benefits should be made more generous.

Given how variation in the values of these parameters over a plausible range could alter the results, one robust conclusion is that we need to know more, in particular on the empirical side when it comes to parameters such as R and E_b^y. All the same, given the evidence provided in recent years that indicates minimal effects of UI on wages, there is a case to be made that the efficiency costs of UI are larger than previously recognized.

I perform a number of sensitivity analyses, which can be found in appendix F. First, I try alternative values of E_b^D, specifically, $0.48 \times \{0.3, 0.8\}$, and unsurprisingly the optimal replacement rates move up in the former case and down in the latter; the effects of fiscal externalities remain significant in both cases.[57] I also try three other sensitivity analyses: I consider complete take-up of benefits, I use a larger value of the initial unemployment rate, specifically $u_0 = 0.064$, and I allow for different tax rates applying to UI benefits and earned income, specifically 0.15 for the former and 0.262 for the latter. The first two of these changes reduce the size of the difference between optimal replacement rates with $G = 0$ and $G > 0$, but I still obtain zeros for $R = 2$ and $E_b^y = 0$, and the analysis with alternative tax

[57] $E_b^D = 0.144$ is small enough to generate a positive local maximum for $R = 2$ and $E_b^y = 0$; to obtain a positive global maximum, I need E_b^D to drop below 0.096.

rates results in slightly increased effects of fiscal externalities.

Given the simplicity of the Baily model, I also perform a number of extensions, which can be found in appendix G. I allow for stochastic duration of unemployment, and restrictions on borrowing during unemployment, which both tend to move the optimal replacement rate closer to one, and I use a second-order Taylor series expansion of marginal utility, and allow for variable labour supply on the initial job, which both reduce the optimal benefit level. However, although the numerical results do change somewhat, the results are still quite similar, and the qualitative conclusions are unchanged: the pairwise comparisons of optimal replacement rates given the two values of G under consideration 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 any effort to account for the role of fiscal externalities, and I have demonstrated in this paper that this is an important omission. My results demonstrate how substantial an impact fiscal externalities resulting from income taxes can have on optimal UI calculations, and indicate 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, fiscal externalities will increase the estimated optimal replacement rate, whereas we may prefer a significantly lower benefit level if the wage elasticity is small or zero.

I present results from both of the main approaches in the existing optimal UI literature, specifically the macro-based structural approach and the sufficient statistics method, and I consider a wide range of parameter values. My baseline results, using the most typical set of parameters including a zero elasticity of wages with respect to benefits, feature an optimal replacement rate of zero. Given the remaining uncertainty about the appropriate parameter values, specifically the effect of UI on wages and the coefficient of relative risk-aversion, it would be premature to conclude that UI should be abolished, but my results do indicate that the efficiency costs of UI are likely to be more severe than has previously been recognized.

Since we need to know more about certain parameters and empirical values, future empirical work is needed, particularly on the effects of UI benefits on job matching and wages.

This paper also raises a number of new questions, about how past work on UI policy over the business cycle may be affected by fiscal externalities, and about the role of active labour market programs in reducing durations of unemployment and facilitating better matches. One lesson of this paper is that relatively small improvements in labour market efficiency can provide significant benefits when the labour market is already highly distorted, suggesting that the benefits provided by labour market programs considered in Card, Kluve, and Weber (2010) might be larger than previously realized.

Finally, the insights in this paper can also be generalized into other areas of government policy. In Lawson (2013), I apply a similar analysis to post-secondary education, and work is underway in considering optimal social insurance in a more general framework of fiscal interactions.

A Technical Appendix for Structural Calibration

Table 13 contains the parameters used in each case; Table 14 displays the moments calculated from the simulated data. 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; if average UI benefits are about \$300 per week, then a 46% replacement rate implies weekly wages of about \$650, and the asset distribution spacing corresponds to about \$3.25. Tests were made of all convergence parameters to ensure that further tightening had no non-negligible effect on results.

Table 13: Calibrated Parameters

	$R = 2$		$R = 5$	
	$G = 0.208$	$G = 0$	$G = 0.208$	$G = 0$
ρ	0.0109	0.01076	0.0485	0.0473
θ	32.0	26.3	92.0	40.3
κ	0.527	0.56	0.395	0.465

Table 14: Calculated Moments

	$R = 2$		$R = 5$	
	$G = 0.208$	$G = 0$	$G = 0.208$	$G = 0$
u	0.0539	0.0542	0.0538	0.0541
E_b^u	0.2405	0.2405	0.2408	0.2407
$\frac{E(c_e)-E(c_u)}{E(c_e)}$	0.1001	0.0998	0.1001	0.0999

B Sensitivity Analyses and Extensions with Structural Model

B.1 Sensitivity Analyses

I have performed a wide range of sensitivity analyses, to examine how the results change when the parameters or moments used in calibration are altered. The results for each of these cases can be found in Table 15, and the resulting complete sets of parameters and moments are available upon request.

Table 15: Optimal Replacement Rates & Welfare Gains

		$R = 2$			$R = 5$		
		r	Welf. Gain	Diff.	r	Welf. Gain	Diff.
$(1+i)^{52} = 0.03$	$G = 0.208$	0.01	41.59	0.00	0.00	11.42	7.64
	$G = 0$	0.01	24.17	0.00	0.36	0.26	1.36
$\delta = \frac{1}{364}$	$G = 0.208$	0.00	11.90	7.27	0.46	0.00	6.72
	$G = 0$	0.38	0.44	6.44	0.73	5.54	5.54
$t_b = 0.15, t_y = 0.262$	$G = 0.237$	0.00	12.76	7.49	0.44	0.11	8.47
	$G = 0$	0.37	0.45	6.43	0.72	5.53	6.28
$E_b^u = 0.1362$	$G = 0.208$	0.12	4.60	3.98	0.67	2.06	3.91
	$G = 0$	0.46	0.02	4.97	0.89	10.31	3.23
$E_b^u = 0.3633$	$G = 0.208$	0.00	19.90	8.96	0.30	2.55	9.07
	$G = 0$	0.32	1.64	6.01	0.59	1.82	7.11
perfect take-up	$G = 0.205$	0.04	9.09	3.75	0.48	0.03	3.32
	$G = 0$	0.34	0.80	4.37	0.66	3.43	2.84
$u = 0.07$	$G = 0.201$	0.01	11.70	6.78	0.42	0.19	6.71
	$G = 0$	0.37	0.54	6.16	0.70	4.41	5.74
consumption drop	$G = 0.208$	0.30	2.25	6.30	0.70	4.65	6.50
	$G = 0$	0.59	1.13	5.01	0.90	18.05	4.83
utility from leisure	$G = 0.208$	0.05	9.91	5.96	0.48	0.09	2.80
	$G = 0$	0.39	0.35	6.20	0.65	3.57	2.90
		$R = 2.5$			$R = 3$		
different risk-aversion	$G = 0.208$	0.09	7.82	6.98	0.18	4.62	7.00
	$G = 0$	0.45	0.01	6.13	0.52	0.24	5.95

It is striking how robust the baseline results are to the changes considered. A larger job separation rate and a larger unemployment rate both leave the results essentially unchanged, as does a modification of the tax system to allow for a tax rate of 0.15 applying to UI benefits and a rate of 0.262 for earned income.[58] Allowing for perfect take-up or utility from leisure when unemployed modestly reduces the effect of fiscal externalities, as does matching 0.1001 to the drop in consumption at the moment in job loss rather than the average consumption gap, though the latter also leads to significantly higher optimal benefit levels; however, in each of these cases the effect of fiscal externalities remains substantial.

The only case in which the results are dramatically altered is when a positive interest rate is used, as this reduces optimal replacement rates, perhaps so much as to eliminate the gap in optimal policy in the $R = 2$ case, but even then the welfare gain is much larger when fiscal externalities are considered, suggesting that the zero lower bound is "more binding" in that case. Allowing for lower and higher values of E_b^u also shifts the results, unsurprisingly leading to higher and lower optimal replacement rates respectively, but the

[58] Since Medicare benefits and some state taxes do not apply to UI benefits, I use a conservative estimate of 0.15 as the tax rate on UI. Meanwhile, Cushing (2005) uses the Social Security Administration's projections of future mortality rates to compute estimated marginal OASDI tax rates, and finds a rate of about 3.2% for 37-year-olds, which is the mean age of individuals in the SIPP sample of Chetty (2008), implying a total tax rate of 0.262 on earned income.

effect of fiscal externalities remains strong. Finally, intermediate values of R lead to results in between those for $R = 2$ and $R = 5$, and it appears that a risk-aversion coefficient just above two is sufficient to eliminate the zero-optimal-UI result.

B.2 Extensions

The first extension of the dynamic job search model considers a non-degenerate wage distribution, and represents the first attempt that I am aware of to use a simulated job search model to study optimal UI in a context in which workers face a distribution covering more than two wages. Job offers now contain a wage y drawn from a distribution $F(y)$, and an unemployed worker receiving such an offer decides whether to accept or to remain unemployed. Denoting the reservation wage by $\bar{y}$, the individual's recursive decision problem is:

$$V_e(k,y) = \max_{k' \in \Gamma_{y(1-\tau)}(k)} \left[U((1+i)k + y(1-\tau) - k') + \left(\frac{1}{1+\rho}\right)^{\frac{1}{52}} [(1-\delta)V_e(k',y) + \delta V_u(k')] \right]$$

$$V_u(k) = \max_{k' \in \Gamma_b(k), \bar{y}, s \geq 0} \left[U((1+i)k + b - k') - e(s) + \left(\frac{1}{1+\rho}\right)^{\frac{1}{52}} [s\tilde{V}_e(k',\bar{y}) + (1 - s(1-F(\bar{y})))V_u(k')] \right]$$

where $\tilde{V}_e(k',\bar{y}) = \int_{y \geq \bar{y}} V_e(k,y)dF(y)$.

The calibration of the model now allows for a constant in the search cost function, $e(s) = d + \frac{(\theta s)^{1+\kappa}}{1+\kappa}$, where d can be thought of as direct disutility from being unemployed; this is necessary in order to obtain the desired order of magnitude for the effect of UI on wages. Meanwhile, the wage is defined as $y = \underline{y} + y_{LN}$, where $\underline{y}$ is a constant and $y_{LN} \sim \ln N(\mu, \sigma^2)$; for the purpose of simulations, a discretized approximation is used with intervals of 0.002 over a central portion of the distribution and mass points at each end containing the remainder of the mass, at the mean value for said mass. The parameters are set to match the previous moments, as well as a mean wage of 1 at baseline and a wage elasticity $\frac{d\ln(E(y))}{d\ln(b)} = 0.02$. In the context and notation of the Baily model discussed later, if $s_0 = 0.8$, the latter corresponds to a wage elasticity of 0.0876, or a value of 0.1825 in the form usually evaluated in the empirical literature. The parameters and moments can be found in Tables 16 and 17, and the numerical results are in Table 18.

Table 16: Calibrated Parameters with Wage Distribution

	$R = 2$		$R = 5$	
	$G = 0.208$	$G = 0$	$G = 0.208$	$G = 0$
ρ	0.01148	0.01147	0.0525	0.0534
θ	9.94	9.335	14.54	11.60
κ	2.94	2.96	2.31	2.04
d	0.72	0.575	0.98	0.39
$\underline{y}$	0.8041	0.80335	0.5475	0.500
μ	-2.344	-2.337	-1.101	-0.968
σ	0.776	0.774	0.3823	0.3532

Allowing for this positive effect on wages leads to dramatically different results from those observed in Table 4; while the optimal replacement rates are higher in all cases, this is especially the case for $G = 0.208$, and fiscal externalities actually lead to an increase in the optimal benefit level (even if only slightly in the $R = 5$ case). If UI benefits increase reservation wages and thus subsequent wages, the resulting positive effect on income tax revenues can overturn the usual conclusion of lower optimal UI.

In my second extension, I attempt to model more realistically the finite duration of UI benefits; specifically, a period now represents a month rather than a week, and benefits expire after 6 months. I assume that individuals receive outside income of 0.1 per period, to ensure that consumption never reaches zero in the uninsured state. I ignore the question of take-up and simply define the benefit level for an individual

Table 17: Calculated Moments with Wage Distribution

	$R=2$		$R=5$	
	$G=0.208$	$G=0$	$G=0.208$	$G=0$
u	0.0541	0.0540	0.0540	0.0542
E_b^u	0.2392	0.2397	0.2418	0.2398
$\frac{E(c_e)-E(c_u)}{E(c_e)}$	0.1001	0.1001	0.1002	0.1002
E_b^y	0.0199	0.0200	0.0201	0.0196
$E(w)$	1.000	1.001	1.002	1.004

Table 18: Optimal Replacement Rates & Welfare Gains

	$R=2$			$R=5$		
	r	Welf. Gain	Diff.	r	Welf. Gain	Diff.
$G=0.208$	0.71	4.27	1.65	1.08	51.98	0.06
$G=0$	0.55	0.48	1.64	1.07	41.05	0.29

receiving benefits as $b = r(1-\tau_0)$; this will tend to bias downwards the importance of fiscal externalities, and means that my results may be more comparable to those with perfect takeup in Table 15. I also use a different functional form for the effort cost of search, specifically $e_t(s) = d - \frac{\theta_t}{1-\kappa}(1-s)^{1-\kappa} - \theta_t\left(s - \frac{1}{1-\kappa}\right)$. To capture duration dependency and ensure that a reasonable proportion of unemployment spells are of long duration and involve exhaustion of benefits, I allow θ_t to increase with time t spent out of work according to experimental results in Kroft, Lange, and Notowidigdo (2012);[59] for the same reason, I also set $\delta = \frac{1}{84}$, corresponding to an expected job duration of 7 years rather than 5. Parameters and moments are in Tables 19 and 20, and the numerical results are in Table 21.

Table 19: Calibrated Parameters with Finite-Duration Benefits

	$R=2$		$R=5$	
	$G=0.203$	$G=0$	$G=?$	$G=0$
ρ	0.0092	0.00915	0.0951	0.0918
θ_1	10.45	8.90	0.789	0.3705
κ	1.1	1.1	3.7	3.7
d	-0.708	-0.576	-2.266	-0.94346

This more realistically modelling choice actually makes relatively little difference to the results. The results with $R=2$ are quite similar to those from the baseline, except that the optimal replacement rate in the $G=0$ case is lower, because perfect take-up is assumed here (a 0.28 drop in optimal UI due to fiscal externalities compares very closely to the 0.30 drop in the perfect take-up case in Table 15). Meanwhile, the results with $R=5$ are even more dramatic, primarily because I allow for outside income of 0.1, so that the consumption-smoothing effects at lower levels of UI are less dramatic. Thus, I conclude that my conclusions are quite robust to alternative specifications of UI benefits.

[59] Kroft, Lange, and Notowidigdo (2012) find that the interview-finding rate drops from about 7% to about 4% over the first six months of an unemployment spell, then remains roughly constant. Given a θ_1 and a target search intensity s_1, I therefore find the θ_t for $t \in \{2,3,4,5,6,7\}$ (where 7 represents all periods of benefit exhaustion) that generates the same effort cost for $s_t = s_1 - (t-1)s_1/14$.

Table 20: Calculated Moments with Finite-Duration Benefits

	$R=2$		$R=5$	
	$G=0.203$	$G=0$	$G=?$	$G=0$
u	0.0538	0.0540	0.0539	0.0541
E_b^u	0.2410	0.2403	0.2411	0.2407
$\frac{E(c_e)-E(c_u)}{E(c_e)}$	0.1002	0.1002	0.0999	0.1000

Table 21: Optimal Replacement Rates & Welfare Gains

	$R=2$			$R=5$		
	r	Welf. Gain	Diff.	r	Welf. Gain	Diff.
$G=0.203/0.207$	0.00	10.62	3.84	0.25	1.09	9.41
$G=0$	0.28	1.39	2.69	0.72	3.11	7.58

C Algebra and Proofs

C.1 Proof of Proposition 1

The individual's first-order condition for saving is:

$$\frac{\partial V}{\partial k} = -U'(c_e^1) + (1-\delta)U'(c_e^2) + \delta U'(c_u) = 0$$

and I also need a first-order Taylor series expansion of $U'(c_e^1)$ around $U'(c_u)$:

$$U'(c_e^1) = U'(c_u) + \Delta c U''(\theta)$$

where θ is between c_u and c_e^1, and $\Delta c = c_e^1 - c_u$. Combining these allows me to rewrite (6) as:

$$\frac{dV}{db} = -2y\Delta c U''(\theta)\frac{d\tau}{db} - \left[(2-\delta)y + \delta s y_n\right] U'(c_u)\left[\frac{d\tau}{db} - \omega\right]$$

where $\omega = \frac{\delta(1-s)}{(2-\delta)y+\delta s y_n}$.

Next, I make two technical assumptions that are also found in Baily (1978). The first is that, in equilibrium, $y_n = y$, so that all wages can be written in terms of y; in other words, although the worker chooses y_n, and the effect of benefits on the individual's choice of y_n will be an important part of the analysis, I assume that the value of y_n that is chosen in equilibrium will not be very different from y. This assumption is generally conservative, because it overweights the income lost from unemployment. The second assumption is that $c_e^1 U''(\theta) = c_u U''(c_u)$, which permits the second derivative of utility to be incorporated into a coefficient of relative risk-aversion. Baily describes this assumption as conservative, but this is not always true, and for the risk-aversion coefficients that I use, it will 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, I find:

$$\frac{dW}{db} \equiv \frac{\frac{dV}{db}}{U'(c_u)} = 2y\frac{\Delta c}{c_e^1}R\frac{d\tau}{db} - 2(1-u)y\left[\frac{d\tau}{db} - \omega\right]$$

where $R = \frac{-c_u U''(c_u)}{U'(c_u)}$ is the coefficient of relative risk-aversion, and $u = \frac{\delta(1-s)}{2}$ 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:

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

C.2 Proof of Proposition 2

Starting from (9):

$$\frac{dW}{db}(b;G) = \frac{2u}{1-u}\left[\frac{\Delta c}{c_e^1}R\frac{E_b^\tau}{\psi} - (1-u)\left(\frac{E_b^\tau}{\psi} - 1\right)\right]$$

and therefore the difference in welfare derivatives is:

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

Using (12) and the definition of ψ:

$$\frac{E_b^\tau}{\psi} = 1 + E_b^D + \frac{ub+G}{ub}\left[\frac{\delta(1-s)}{2(1-u)}E_b^D - \frac{\delta s}{2(1-u)}E_b^y\right]$$

and thus the welfare derivative difference becomes:

$$\frac{dW}{db}(b;G_1) - \frac{dW}{db}(b;0) = \frac{\delta}{(1-u)^2 b}\left[\frac{\Delta c}{c_e^1}R - (1-u)\right]\left[(1-s)E_b^D - sE_b^y\right]G_1.$$

Since I assume that $\frac{\Delta c}{c_e^1}R < 1-u$, this right-hand side will be positive if and only if $sE_b^y - (1-s)E_b^D$ is positive. The latter expression can also be written as:

$$sE_b^y - (1-s)E_b^D = \frac{sb}{y_n}\frac{dy_n}{db} + b\frac{ds}{db} = \frac{b}{y_n}\frac{d(sy_n)}{db}.$$

and therefore $\frac{dW}{db}(b;G_1) - \frac{dW}{db}(b;0)$ has the same sign as $\frac{d(sy_n)}{db}$.

C.3 Proof of Proposition 3

Combining (9) and (12), I immediately get:

$$\frac{dW}{db}(b;G,E_b^{y2}) - \frac{dW}{db}(b;G,E_b^{y1}) = \frac{2u}{(1-u)\psi}\left[\frac{\Delta c}{c_e^1}R - (1-u)\right]\left[\frac{-\delta s}{2(1-u)}\right]\left[E_b^{y2} - E_b^{y1}\right]. \tag{14}$$

Given that $E_b^{y2} > E_b^{y1}$, and since the middle two terms are both negative, this expression is always positive.

C.4 Proof of Propositions 4 and 5

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

$$\left[\frac{dW}{db}(b;G_1,E_b^{y2}) - \frac{dW}{db}(b;G_1,E_b^{y1})\right] - \left[\frac{dW}{db}(b;0,E_b^{y2}) - \frac{dW}{db}(b;0,E_b^{y1})\right]$$

$$= \frac{-\delta s}{(1-u)^2 b}\left[\frac{\Delta c}{c_e^1}R - (1-u)\right]\left[E_b^{y2} - E_b^{y1}\right]G_1.$$

This equation is clearly positive, which concludes the proof of Proposition 4.

To prove Proposition 5, my task is somewhat more complicated. I begin with the fact that, at the optimum, $\frac{\Delta c}{c_e^1}RE_b^\tau = (1-u)(E_b^\tau - \psi)$; using (12) and rearranging, this becomes:

$$\frac{\Delta c}{c_e^1}R\psi(1+E_b^D) - (1-u)\psi E_b^D = \frac{\delta}{2(1-u)}\left[sE_b^y - (1-s)E_b^D\right]\left[\frac{\Delta c}{c_e^1}R - (1-u)\right].$$

Observe that, because $\frac{\Delta c}{c_e^1}R < 1-u$, $sE_b^y - (1-s)E_b^D > 0$ at the optimum if and only if $\frac{\Delta c}{c_e^1}R(1+E_b^D) - (1-u)E_b^D < 0$. I wish to show that $b^*(G_1,E_b^{y2}) > b^*(0,E_b^{y2})$ and $b^*(G_1,E_b^{y1}) < b^*(0,E_b^{y1})$ for

$E_b^{y2} > E_b^{y*} > E_b^{y1}$, so Proposition 2 says that $sE_b^y - (1-s)E_b^D$ must be positive for E_b^{y2} and negative for E_b^{y1}. Then, if I define $X(b) = \frac{\Delta c}{c_e^1}R(1+E_b^D) - (1-u)E_b^D$, I want $X < 0$ at the optimum for E_b^{y2} and $X > 0$ for E_b^{y1}; 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:

$$\frac{dX}{db} = R(1+E_b^D)\frac{d\left(\frac{\Delta c}{c_e^1}\right)}{db} + \frac{\Delta c}{c_e^1}(1+E_b^D)\frac{dR}{db} + \frac{\Delta c}{c_e^1}R\frac{dE_b^D}{db} - E_b^D\frac{d(1-u)}{db} - (1-u)\frac{dE_b^D}{db}$$

and at $X = 0$, $\frac{\Delta c}{c_e^1}R(1+E_b^D) = (1-u)E_b^D$, and thus:

$$\frac{dX}{db}|_{X=0} = \frac{\Delta c}{c_e^1}(1+E_b^D)\frac{dR}{db} + \left[\frac{\Delta c}{c_e^1}R - (1-u)\right]\frac{dE_b^D}{db} + \frac{R(1+E_b^D)}{(1-u)}\left[(1-u)\frac{d\left(\frac{\Delta c}{c_e^1}\right)}{db} - \frac{\Delta c}{c_e^1}\frac{d(1-u)}{db}\right].$$

Sufficient conditions for this to be negative are that $\frac{dR}{db} = 0$, $\frac{dE_b^D}{db} \geq 0$, $1 + E_b^D > 0$ and $(1-u)\frac{d\left(\frac{\Delta c}{c_e^1}\right)}{db} < \frac{\Delta c}{c_e^1}\frac{d(1-u)}{db}$. The first two assumptions are standard, and I make them in my 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 $\frac{dE_b^D}{db}$ "are broadly similar across studies with different levels of benefit generosity." The third assumption is a formality, as a nearly universal finding of the empirical literature is that E_b^D is positive. The final assumption requires a bit more explanation; it is easiest to understand when written as $\frac{d}{db}\left(\frac{\frac{\Delta c}{c_e^1}}{1-u}\right) < 0$. The consumption gap $\frac{\Delta c}{c_e^1}$ 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 $1 - \frac{\delta}{2}$. Therefore, this condition is likely to be satisfied in nearly every case of interest, and this assumption is strongly supported by the numerical results in section 6 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.

D 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 parameter values, and for the initial model as well as all sensitivity analyses and extensions. I can then see if any failures of quasi-concavity appear over that range; beyond $r = 2$, failures of quasi-concavity might be expected on the grounds that my assumptions become especially poor approximations. All plots are available upon request.

For the baseline model, quasi-concavity fails in several cases when $G = 0.208$: for $R = 5$ and $E_b^y \geq 0.048$, there appears to be a local minimum at low values of r (always less than 0.08). It is not surprising, however, that these violations of strict quasi-concavity occur at low values of r when R is large, as that is exactly when $\frac{\Delta c}{c_e^1}R < 1-u$ may fail to hold and my assumptions will tend to be most inaccurate, and at which the estimated E_b^τ could turn negative. Over the vast majority of the range of r that I consider, however, $\frac{dW}{db}$ behaves normally and consistent with quasi-concavity.

In each of the sensitivity analyses in appendix F, similar local minima are found at low r for $R = 5$ and $E_b^y \geq 0.048$. Additionally, in the extension with $E_b^D = 0.144$, for $R = 2$, $E_b^y = 0$ and all s_0, I observe local maxima at positive values of r with the global maximum at $r = 0$, as labelled in the results. Finally, in the perfect-take-up case, for $R = 2$, $E_b^y = 0.048$ and $s_0 = 0.648$, there is a second local maximum to the right of the global maximum of $r = 0.0393$.

Further failures of quasi-concavity are observed in each of the extensions in appendix G; in the first, second, and fourth extensions, local minima are observed for $R = 5$ and $E_b^y \geq 0.048$. The first extension presents cases for $R = 2$, $E_b^y = 0$ and $s_0 \leq 0.8$ where local maxima are observed at positive values of r but

the global maximum is at $r = 0$. Similar cases are observed thrice in the second extension at high values of E_b^y and s_0, though as discussed in appendix G.2, 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 $R = 2$, $E_b^y = 0.048$ and $s_0 = 0.648$. In the third extension, local maxima are also observed for $R = 5$, $E_b^y = 0$, and each s_0, but once again the global maximum is at zero.

E 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 22 and 23 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 7 and 8, in that a positive value of G causes the values in the table to "spread out."

Table 22: Baseline Values of $\frac{dW}{db}$ Calculated from (9) and (12) for $R = 2$

Baseline $\frac{dW}{db}$ for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	-0.0008	-0.0011	-0.0016	-0.0025
	0	-0.0000	-0.0000	-0.0000	-0.0000
E_b^y	0.048	0.0004	0.0006	0.0009	0.0014
	0.096	0.0008	0.0012	0.0018	0.0029
	0.192	0.0017	0.0024	0.0037	0.0058
	0.3072	0.0027	0.0039	0.0059	0.0094

Baseline $\frac{dW}{db}$ for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	-0.0418	-0.0487	-0.0605	-0.0805
	0	-0.0258	-0.0258	-0.0258	-0.0258
E_b^y	0.048	-0.0165	-0.0124	-0.0054	0.0063
	0.096	-0.0071	0.0011	0.0150	0.0384
	0.192	0.0117	0.0280	0.0558	0.1027
	0.3072	0.0343	0.0602	0.1048	0.1798

Table 23: Baseline Values of $\frac{dW}{db}$ Calculated from (9) and (12) for $R = 5$

Baseline $\frac{dW}{db}$ for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.0430	0.0428	0.0425	0.0420
	0	0.0435	0.0435	0.0435	0.0435
E_b^y	0.048	0.0437	0.0438	0.0440	0.0443
	0.096	0.0440	0.0442	0.0446	0.0452
	0.192	0.0445	0.0449	0.0457	0.0470
	0.3072	0.0451	0.0458	0.0470	0.0491

Baseline $\frac{dW}{db}$ for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.0185	0.0144	0.0073	-0.0046
	0	0.0280	0.0280	0.0280	0.0280
E_b^y	0.048	0.0337	0.0361	0.0402	0.0472
	0.096	0.0393	0.0441	0.0524	0.0664
	0.192	0.0505	0.0602	0.0768	0.1048
	0.3072	0.0640	0.0795	0.1061	0.1509

F Sensitivity Analyses in the Baily Model

In this appendix, I present results from a number of sensitivity analyses in the sufficient statistics approach. First, I use different values of E_b^D over a wide range; I try $E_b^D = 0.48 \times 0.3$, with results in Tables 24 and 25, and $E_b^D = 0.48 \times 0.8$, with results in Tables 26 and 27. Not surprisingly, the optimal replacement rates move up in the former case and down in the latter; the effects of fiscal externalities remain sizable in both cases. A value of E_b^D around 0.144 is where I begin to observe a positive local maximum when $R = 2$ and $E_b^y = 0$; for a positive global maximum, I need E_b^D to drop below 0.096.

Next, I ignore the question of take-up of benefits and only deflate benefits by the ratio of compensated to total unemployment duration; the ensuing results can be found in Tables 28 and 29. This tends to reduce the size of the difference between optimal replacement rates with $G = 0$ and $G > 0$, but I still observe zeros for $R = 2$ and $E_b^y = 0$.

Table 24: Optimal Replacement Rates for $R = 2$ and $E_b^D = 0.144$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.5908	0.5859	0.5776	0.5637
	0	0.6020	0.6020	0.6020	0.6020
	0.048	0.6087	0.6116	0.6167	0.6254
	0.096	0.6154	0.6213	0.6317	0.6494
	0.192	0.6289	0.6411	0.6624	0.6997
	0.3072	0.6454	0.6653	0.7007	0.7640

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0	0	0	0
	*0	0	0	0	0
	0.048	0.5311	0.5824	0.6585	0.7678
	0.096	0.6392	0.7181	0.8358	1.0062
	0.192	0.8028	0.9254	1.1095	1.3767
	0.3072	0.9598	1.1257	1.3739	1.7299

*Local Maximum of 0.3451 for this row

Table 25: Optimal Replacement Rates for $R = 5$ and $E_b^D = 0.144$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.7391	0.7371	0.7338	0.7283
	0	0.7435	0.7435	0.7435	0.7435
	0.048	0.7461	0.7473	0.7493	0.7528
	0.096	0.7487	0.7511	0.7552	0.7623
	0.192	0.7540	0.7589	0.7675	0.7824
	0.3072	0.7606	0.7686	0.7828	0.8081

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.6311	0.6066	0.5656	0.5005
	0	0.6874	0.6874	0.6874	0.6874
	0.048	0.7208	0.7358	0.7618	0.8058
	0.096	0.7543	0.7846	0.8368	0.9253
	0.192	0.8212	0.8819	0.9860	1.1588
	0.3072	0.9008	0.9968	1.1586	1.4177

Table 26: Optimal Replacement Rates for $R = 2$ and $E_b^D = 0.384$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.3129	0.3096	0.3039	0.2943
	0	0.3213	0.3213	0.3213	0.3213
	0.048	0.3262	0.3283	0.3317	0.3376
	0.096	0.3312	0.3352	0.3422	0.3541
	0.192	0.3411	0.3493	0.3636	0.3882
	0.3072	0.3531	0.3664	0.3897	0.4310

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0	0	0	0
	0	0	0	0	0
	0.048	0.0257	0.0544	0.1477	0.3031
	0.096	0.1503	0.2551	0.3873	0.5457
	0.192	0.3675	0.4800	0.6326	0.8357
	0.3072	0.5148	0.6460	0.8312	1.0832

Table 27: Optimal Replacement Rates for $R = 5$ and $E_b^D = 0.384$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.6292	0.6279	0.6256	0.6219
	0	0.6320	0.6320	0.6320	0.6320
	0.048	0.6337	0.6345	0.6359	0.6382
	0.096	0.6354	0.6370	0.6398	0.6445
	0.192	0.6389	0.6421	0.6478	0.6576
	0.3072	0.6430	0.6483	0.6576	0.6741

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.4764	0.4581	0.4273	0.3777
	0	0.5182	0.5182	0.5182	0.5182
	0.048	0.5422	0.5529	0.5713	0.6021
	0.096	0.5659	0.5870	0.6232	0.6836
	0.192	0.6117	0.6530	0.7228	0.8365
	0.3072	0.6642	0.7279	0.8337	1.0009

Table 28: Optimal Replacement Rates for $R = 2$ and Perfect Take-Up

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.4500	0.4459	0.4390	0.4275
	0	0.4595	0.4595	0.4595	0.4595
E_b^y	0.048	0.4651	0.4675	0.4717	0.4789
	0.096	0.4707	0.4756	0.4842	0.4987
	0.192	0.4821	0.4921	0.5095	0.5399
	0.3072	0.4959	0.5122	0.5410	0.5922

Optimal r for $G = 0.205$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0	0	0	0
	0	0	0	0	0
E_b^y	0.048	0.0393	0.3148	0.4086	0.5160
	0.096	0.3925	0.4719	0.5775	0.7184
	0.192	0.5516	0.6537	0.7996	1.0034
	0.3072	0.6821	0.8116	1.0001	1.2649

Table 29: Optimal Replacement Rates for $R = 5$ and Perfect Take-Up

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.6831	0.6815	0.6787	0.6741
	0	0.6866	0.6866	0.6866	0.6866
E_b^y	0.048	0.6887	0.6897	0.6914	0.6942
	0.096	0.6909	0.6928	0.6962	0.7020
	0.192	0.6951	0.6991	0.7061	0.7183
	0.3072	0.7004	0.7069	0.7184	0.7390

Optimal r for $G = 0.205$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.5786	0.5614	0.5322	0.4847
	0	0.6173	0.6173	0.6173	0.6173
E_b^y	0.048	0.6400	0.6503	0.6679	0.6978
	0.096	0.6626	0.6831	0.7184	0.7779
	0.192	0.7071	0.7479	0.8176	0.9332
	0.3072	0.7595	0.8237	0.9319	1.1064

I then try a larger value of the initial unemployment rate, specifically $u_0 = 0.064$. This leads to the results displayed in Tables 30 and 31. The optimal replacement rates spread out for $G = 0$, but the effects are more modest for $G > 0$, meaning a small reduction in the effect of fiscal externalities.

Finally, instead of using a tax rate of 0.23 to apply to both UI and earned income, I allow for one tax rate applied to UI benefits and another for earned income; as explained in appendix B.1, I use 0.15 as the tax rate on UI and 0.262 as the tax rate on earned income. The results are displayed in Tables 32 and 33, and the effects of fiscal externalities are slightly increased.

G Extensions to Baily Model

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

G.1 Stochastic Duration of Unemployment

I first consider the effect of allowing the duration of unemployment to be stochastic. I follow Baily's approach of defining the actual duration of unemployment $(1 - \tilde{s})$ as:

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

where s is deterministic, and v is a stochastic term with mean zero which is uncorrelated with s.[60]

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

$$\begin{aligned}\tilde{c_u} &= (1 - \tilde{s})(b - e) + \tilde{s} y_n (1 - \tau) + k \\ &= c_u - v\Delta y\end{aligned}$$

[60] As noted by Baily, this can only be an approximation given that $(s - v)$ is constrained to lie in $(0, 1)$.

Table 30: Optimal Replacement Rates for $R = 2$ and $u_0 = 0.064$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.4491	0.4443	0.4362	0.4227
	0	0.4603	0.4603	0.4603	0.4603
	0.048	0.4670	0.4698	0.4748	0.4833
	0.096	0.4736	0.4794	0.4895	0.5069
	0.192	0.4871	0.4989	0.5198	0.5562
	0.3072	0.5034	0.5228	0.5574	0.6195

Optimal r for $G = 0.203$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0	0	0	0
	0	0	0	0	0
	0.048	0.0306	0.2725	0.3978	0.5265
	0.096	0.3784	0.4743	0.5981	0.7610
	0.192	0.5672	0.6853	0.8532	1.0868
	0.3072	0.7170	0.8656	1.0813	1.3827

Table 31: Optimal Replacement Rates for $R = 5$ and $u_0 = 0.064$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.6828	0.6809	0.6777	0.6723
	0	0.6870	0.6870	0.6870	0.6870
	0.048	0.6895	0.6906	0.6926	0.6960
	0.096	0.6920	0.6943	0.6984	0.7053
	0.192	0.6971	0.7018	0.7101	0.7247
	0.3072	0.7033	0.7111	0.7249	0.7499

Optimal r for $G = 0.203$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.5554	0.5347	0.5003	0.4454
	0	0.6023	0.6023	0.6023	0.6023
	0.048	0.6299	0.6424	0.6639	0.7003
	0.096	0.6574	0.6824	0.7254	0.7981
	0.192	0.7116	0.7613	0.8462	0.9863
	0.3072	0.7753	0.8532	0.9840	1.1924

Table 32: Optimal Replacement Rates for $R = 2$ and Multiple Tax Rates

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.4500	0.4459	0.4390	0.4275
	0	0.4595	0.4595	0.4595	0.4595
	0.048	0.4651	0.4675	0.4717	0.4789
	0.096	0.4707	0.4756	0.4842	0.4987
	0.192	0.4821	0.4921	0.5095	0.5399
	0.3072	0.4959	0.5122	0.5410	0.5922

Optimal r for $G = 0.237$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0	0	0	0
	0	0	0	0	0
	0.048	0.0294	0.2543	0.3909	0.5249
	0.096	0.3708	0.4710	0.5987	0.7653
	0.192	0.5670	0.6880	0.8592	1.0963
	0.3072	0.7202	0.8717	1.0908	1.3960

Table 33: Optimal Replacement Rates for $R = 5$ and Multiple Tax Rates

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.6831	0.6815	0.6787	0.6741
	0	0.6866	0.6866	0.6866	0.6866
	0.048	0.6887	0.6897	0.6914	0.6942
	0.096	0.6909	0.6928	0.6962	0.7020
	0.192	0.6951	0.6991	0.7061	0.7183
	0.3072	0.7004	0.7069	0.7184	0.7390

Optimal r for $G = 0.237$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.5476	0.5259	0.4896	0.4317
	0	0.5968	0.5968	0.5968	0.5968
	0.048	0.6256	0.6387	0.6610	0.6989
	0.096	0.6543	0.6802	0.7249	0.8000
	0.192	0.7106	0.7620	0.8495	0.9935
	0.3072	0.7764	0.8567	0.9911	1.2046

where c_u is defined as before, and $\Delta y = y_n(1-\tau) - (b-e)$. Utility can now be written as:

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

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

$$\frac{\partial V}{\partial b} = \delta E_v[U'(\tilde{c_u})(1-s+v)]$$

$$\frac{\partial V}{\partial \tau} = -yU'(c_e^1) - (1-\delta)yU'(c_e^2) - \delta y_n E_v[U'(\tilde{c_u})(s-v)].$$

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

$$\begin{aligned} U'(\tilde{c_u}) &= U'(c_u) + U''(\gamma)(\tilde{c_u} - c_u) \\ &= U'(c_u) - v\Delta y U''(\gamma) \end{aligned}$$

where γ is somewhere between c_u and $\tilde{c_u}$. Upon reaching this point in the calculations, Baily (1978) implicitly makes an assumption that he does not state explicitly, which is that $U''(\gamma)$ is uncorrelated with v and v^2, capturing 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:

$$E_v[U'(\tilde{c_u})] = U'(c_u)$$

$$E_v[U'(\tilde{c_u})v] = -\Delta y E_v[U''(\gamma)]Var(v).$$

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

$$\frac{\partial V}{\partial k} = -U'(c_e^1) + (1-\delta)U'(c_e^2) + \delta U'(c_u) = 0$$

As before, I make the assumptions that $y_n = y$ and $c_e^1 U''(\theta) = c_u U''(c_u)$, and to this I add the assumption that $E[U''(\gamma)] = U''(\theta)$, which will tend towards underestimating the welfare gain from raising b. I then combine the results above and write the welfare derivative as:

$$\frac{dW}{db} = 2y\frac{\Delta c}{c_e^1}R\frac{d\tau}{db} + \delta\frac{\Delta y}{c_e^1}RVar(v) + \delta y\frac{\Delta y}{c_e^1}RVar(v)\frac{d\tau}{db} - 2(1-u)y\left[\frac{d\tau}{db} - \omega\right].$$

The budget constraint takes an expectation over all workers, and so is unchanged; therefore:

$$\frac{dW}{db} = \frac{2u}{(1-u)\psi}\left[\frac{\Delta c}{c_e^1}R + \frac{\Delta y}{c_e^1}\frac{RVar(v)}{1-s}\right]E_b^\tau - \frac{2u}{\psi}\left[1 + \frac{\Delta y}{c_e^1}\frac{RVar(v)}{1-s}\right][E_b^\tau - \psi]$$

and the equation for the optimum is:

$$\frac{\Delta c}{c_e^1}R + \frac{\Delta y}{c_e^1}\frac{RVar(v)}{1-s} = (1-u)\left[1 + \frac{\Delta y}{c_e^1}\frac{RVar(v)}{1-s}\right]\frac{E_b^\tau - \psi}{E_b^\tau}. \tag{15}$$

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 s, still work as averages or expectations, but there are two new terms to consider: $\frac{\Delta y}{c_e^1}$ and $Var(v)$. The latter is also the variance of the duration of unemployment $(1-s)$, and to evaluate this parameter, I turn to Chetty (2008), who estimates 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 $std(v) = \frac{14.2}{18.3}(1-s_0)$, and therefore $Var(v) = \left(\frac{14.2}{18.3}\right)^2(1-s_0)^2$.[61] Meanwhile, in the absence of any better evidence, I will use Baily's assumption that $\frac{\Delta y}{c_e^1} = 1-r$. Evaluation of (15) then gives the results displayed below in Tables 34 and 35.

Table 34: Optimal Replacement Rates Calculated from (15) for $R = 2$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.6529	0.6208	0.5812	0.5367
	0	0.6590	0.6303	0.5969	0.5630
E_b^y	0.048	0.6626	0.6360	0.6063	0.5790
	0.096	0.6662	0.6417	0.6157	0.5954
	0.192	0.6735	0.6532	0.6351	0.6292
	0.3072	0.6824	0.6673	0.6590	0.6720

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0	0	0	0
	0	0	0	0	0
E_b^y	0.048	0.5444	0.5334	0.5455	0.6008
	0.096	0.6034	0.6225	0.6788	0.7921
	0.192	0.7003	0.7611	0.8779	1.0739
	0.3072	0.7962	0.8945	1.0658	1.3354

Table 35: Optimal Replacement Rates Calculated from (15) for $R = 5$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.7961	0.7779	0.7560	0.7322
	0	0.7986	0.7819	0.7626	0.7432
E_b^y	0.048	0.8002	0.7843	0.7665	0.7499
	0.096	0.8017	0.7867	0.7705	0.7567
	0.192	0.8048	0.7916	0.7786	0.7708
	0.3072	0.8085	0.7975	0.7887	0.7887

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.7138	0.6756	0.6202	0.5423
	0	0.7427	0.7215	0.6969	0.6722
E_b^y	0.048	0.7597	0.7486	0.7428	0.7519
	0.096	0.7767	0.7758	0.7888	0.8317
	0.192	0.8107	0.8300	0.8803	0.9877
	0.3072	0.8510	0.8942	0.9868	1.1623

As can be seen, allowing for an 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 s and especially E_b^y, whereas it increases for higher values. Indeed, the pairwise comparisons between the side-by-side tables are 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 34 or 35, and vice-versa.

G.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 that they can also perfectly smooth consumption within the second period. Recent work, 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 $(1 - s)U(c_u) + sU(c_n)$, where c_u is now per-period consumption while unemployed and c_n is per-period consumption when re-employed in a new job.[62] 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

[61] 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 $Var(v)$, but on the other hand, Chetty's is an unconditional variance, some of which may be explained by individual characteristics, which means $Var(v)$ may be an overestimate.

[62] 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.

when re-employed.[63] I can therefore write total utility as:

$$V = U[y(1-\tau)-k] + (1-\delta)U[y(1-\tau)+k] + \delta\left[(1-s)U\left((b-e)+\frac{k}{1-s}\right) + sU\left(y_n(1-\tau)\right)\right].$$

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

$$\frac{\partial V}{\partial \tau} = -yU'(c_e^1) - (1-\delta)yU'(c_e^2) - \delta s y_n U'(c_n).$$

I replace $U'(c_e^2)$ using the first-order condition for saving, as before, and I assume that $c_n = c_e^1$, which is generally 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 to within 4% of their pre-unemployment consumption. Combining this with the usual Taylor series expansion of $U'(c_e^1)$:

$$\frac{\partial V}{\partial \tau} = -[(2-\delta)y + \delta s y_n]U'(c_u) - [2y + \delta s y_n]\Delta c U''(\theta).$$

Putting this together with (4):

$$\frac{dW}{db} = [2+\delta s]y\frac{\Delta c}{c_e^1}R\frac{d\tau}{db} - 2(1-u)y\left(\frac{d\tau}{db} - \omega\right)$$

and therefore the equation for the optimum is:

$$\frac{\Delta c}{c_e^1}R = \frac{2(1-u)}{2+\delta s}\frac{E_b^\tau - \psi}{E_b^\tau}. \tag{16}$$

E_b^τ is the same as before, so this equation is almost identical to (10), and it is easy to introduce the extra δs term into the calculations. Solving for the optimal replacement rate generates the results found in Tables 36 and 37.

Table 36: Optimal Replacement Rates Calculated from (16) for $R = 2$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.4850	0.4947	0.5098	0.5316
	0	0.4934	0.5063	0.5263	0.5551
E_b^y	0.048	0.4985	0.5133	0.5362	0.5694
	0.096	0.5035	0.5203	0.5463	0.5841
	0.192	0.5137	0.5344	0.5670	0.6146
	0.3072	0.5260	0.5518	0.5926	0.6534

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0	0	0	0
	0	0	0	0	0
E_b^y	0.048	0.2384	0.3621	0.4716	0.5905
	0.096	0.4156	0.5095	0.6274	0.7723
	0.192	0.5821	0.6956	0.8500	1.0509
	0.3072	0.7216	0.8621	1.0577	1.3160

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 $\frac{0.222}{0.265}$). In a few cases with $R = 5$ and high E_b^y and s_0, anomalous results occur in which 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 failure of an assumption that $\frac{\Delta c}{c_e^1}R < \frac{2(1-u)}{2+\delta s}$.[64] 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_b^y, 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 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. Additionally, using the same expression for $\frac{\Delta c}{c_e^1}$ when I have redefined c_u to be consumption while unemployed will tend to shift the results downwards, offsetting the tendency of optimal benefit levels to increase.

[63] Given that unemployment durations are deterministic, as long as y_n is not too far from y, there is no reason for a worker to save more than they would want to consume in a spell of unemployment.

[64] $\frac{\Delta c}{c_e^1}R > \frac{2(1-u)}{2+\delta s}$ implies $\frac{\partial V}{\partial \tau} > 0$, while I also estimate that $\frac{d\tau}{db} < 0$.

Table 37: Optimal Replacement Rates Calculated from (16) for $R = 5$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.6963	0.7002	0.7063	0.7151
	0	0.6995	0.7047	0.7128	0.7244
	0.048	0.7014	0.7074	0.7167	0.7301
	0.096	0.7034	0.7102	0.7207	0.7359
	0.192	0.7073	0.7157	0.7289	0.7481
	0.3072	0.7120	0.7225	0.7390	0.7636

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
E_b^y	-0.0816	0.5825	0.5775	0.5721	0.5697
	0	0.6229	0.6318	0.6456	0.6649
	0.048	0.6470	0.6651	0.6929	0.7324
	0.096	0.6713	0.6991	0.7422	0.8057
	0.192	0.7202	0.7681	0.8442	*0/0.9604
	0.3072	0.7786	0.8510	*0/0.9669	*0/1.1440

*Global Maximum/Local Maximum

G.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%. 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.[65] However, I will now explore how the results change when I use a second-order Taylor series expansion of marginal utility.

To do so, I must follow the approach of Chetty (2006) and rewrite my expression in terms of $U'(c_e^1)$ rather than $U'(c_u)$, although this will hinder the comparability of my results with those of the baseline case. 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 $U'(c_e^1)$:

$$U'(c_u) = U'(c_e^1) - \Delta c U''(c_e^1) + \Delta c^2 \frac{U'''(\theta)}{2}$$

where θ 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):

$$\frac{dV}{db} = 2u\left[U'(c_e^1) - \Delta c U''(c_e^1) + \Delta c^2 \frac{U'''(\theta)}{2}\right]$$
$$-\left[[(2-\delta)y + \delta s y_n]U'(c_e^1) + \delta(y - s y_n)\left(\Delta c U''(c_e^1) - \Delta c^2 \frac{U'''(\theta)}{2}\right)\right]\frac{d\tau}{db}.$$

I make the usual assumption that $y_n = y$, and add the modified assumption that $\theta = c_e^1$, and then a bit of rearranging gives:

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

where $P = \frac{-c_e^1 U'''(c_e^1)}{U''(c_e^1)}$ is the coefficient of relative prudence. Therefore, the equation for the optimum is:

$$\left(\frac{\Delta c}{c_e^1}R + \frac{1}{2}\left(\frac{\Delta c}{c_e^1}\right)^2 RP\right) = (1-u)\left(1 + \frac{\Delta c}{c_e^1}R + \frac{1}{2}\left(\frac{\Delta c}{c_e^1}\right)^2 RP\right)\frac{E_b^\tau - \psi}{E_b^\tau} \tag{17}$$

[65] The first-order Taylor series used in my paper is in fact an exact equality, not an approximation; 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, which explains why this leads to an underestimate in his paper. Baily's assumption that the E_b^τ 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.

where E_b^τ is unchanged thus and 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 $U(c) = \frac{c^{1-R}}{1-R}$, $P = \frac{-cU'''(c)}{U''(c)} = R + 1$, 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.[66] I will therefore use a value of $P = 2$, and the results from evaluation of the optimal replacement rate are displayed in Tables 38 and 39.

Table 38: Optimal Replacement Rates Calculated from (17) for $R = 2$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.4000	0.3947	0.3858	0.3707
	0	0.4124	0.4124	0.4124	0.4124
E_b^y	0.048	0.4196	0.4227	0.4281	0.4371
	0.096	0.4269	0.4331	0.4438	0.4620
	0.192	0.4414	0.4539	0.4755	0.5124
	0.3072	0.4588	0.4788	0.5138	0.5741

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0	0	0	0
	0	0	0	0	0
E_b^y	0.048	0.0238	0.0963	0.3317	0.5013
	0.096	0.3073	0.4384	0.5843	0.7581
	0.192	0.5502	0.6794	0.8513	1.0768
	0.3072	0.7128	0.8637	1.0720	1.3510

Table 39: Optimal Replacement Rates Calculated from (17) for $R = 5$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.6522	0.6499	0.6549	0.6392
	0	0.6574	0.6574	0.6574	0.6574
E_b^y	0.048	0.6604	0.6618	0.6641	0.6682
	0.096	0.6634	0.6662	0.6709	0.6790
	0.192	0.6695	0.6750	0.6846	0.7008
	0.3072	0.6768	0.6857	0.7011	0.7274

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0	0	0	0
	0	0	0	0	0
E_b^y	0.048	0.5565	0.5817	0.6201	0.6760
	0.096	0.6095	0.6499	0.7101	0.7955
	0.192	0.6922	0.7543	0.8452	0.9717
	0.3072	0.7704	0.8521	0.9700	1.1313

The results this time are somewhat different quantitatively, in that optimal replacement rates are zero for low values of E_b^y even for $R = 5$. However, the qualitative comparison remains the same, right down to nearly the exact same pairwise comparisons: at low values of s and especially E_b^y, the fiscal externality term considerably reduces the optimal replacement rate, whereas at higher values it considerably increases it.

G.4 Variable Labour Supply

To this point, I have assumed that y is fixed, and thus ignored any distortionary effects of taxes on labour supply once employed. 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, changes in y have an effect though the government budget constraint. Saez (2002) argues that much of the responsiveness of modest-income workers is on the extensive margin, which is already largely captured here by the decision about whether or not to seek work, but all the same I will examine how significant this effect could be. I begin by rewriting the utility function to allow for choice of y, assuming that 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, which I denote

[66] 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.

as $h(y)$, is separable from consumption, (1) becomes:

$$V = U[y(1-\tau)-k] + (1-\delta)U[y(1-\tau)+k] - (2-\delta)h(y) + \delta U[(1-s)(b-e) + sy_n(1-\tau)+k] - \delta h(sy_n).$$

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:

$$\frac{d\tau}{db} = \frac{\delta(1-s) - \delta b\frac{ds}{db} - \delta\tau y_n\frac{ds}{db} - \delta s\tau\frac{dy_n}{db} - (2-\delta)\tau\frac{dy}{db}}{(2-\delta)y + \delta s y_n}.$$

and rewritten in terms of elasticities, this is equivalent to:

$$E_b^\tau = \psi + \psi E_b^u + \frac{u}{1-u}E_b^D - \frac{\delta s}{2(1-u)}E_b^y - \frac{2-\delta}{2(1-u)}\varepsilon_b^y \tag{18}$$

where $\varepsilon_b^y = \frac{b}{y}\frac{dy}{db}$.

I now have to decide on a value for ε_b^y. When b increases, τ increases as well - unless E_b^y 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;[67] using this, and assuming that the only effect of changes in b and τ on y go through the channel of taxes:

$$\varepsilon_b^y = \frac{dy}{db}\frac{b}{y} = \frac{dy}{d\tau}\frac{d\tau}{db}\frac{b}{y} = -0.4\frac{b}{1-\tau}\frac{d\tau}{db}.$$

Since I cannot calculate $\frac{d\tau}{db}$ without knowing ε_b^y, I replace it with the partial derivative:

$$\varepsilon_b^y \simeq -0.4\frac{b}{1-\tau}\left[\frac{\delta(1-s)}{((2-\delta)y + \delta s y_n}\right] \simeq -0.4\frac{\tau}{1-\tau}\psi.$$

Finally, I use the baseline tax rate of $\tau = 0.23$, so my estimate of the elasticity is $\varepsilon_b^y = \frac{-0.092}{0.77}\psi$; 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 40 and 41.

Table 40: Optimal Replacement Rates Calculated from (10) and (18) for $R = 2$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.3412	0.3428	0.3456	0.3503
	0	0.3496	0.3548	0.3638	0.3791
E_b^y	0.048	0.3546	0.3619	0.3746	0.3966
	0.096	0.3596	0.3691	0.3856	0.4144
	0.192	0.3697	0.3835	0.4080	0.4513
	0.3072	0.3818	0.4012	0.4356	0.4978

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0	0	0	0
	0	0	0	0	0
E_b^y	0.048	0.0270	0.1511	0.3125	0.4579
	0.096	0.2854	0.3886	0.5185	0.6897
	0.192	0.4809	0.5989	0.7670	1.0039
	0.3072	0.6279	0.7733	0.9855	1.2867

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^y. However, the essential point remains that consideration of fiscal externalities can greatly change the results, either in a positive or negative direction; the pairwise comparisons are again nearly identical to the baseline.

[67] Their estimated elasticities for lower and moderate income individuals, who are more likely to end up on UI, are smaller, so this is likely to exaggerate the distortionary effects of taxation.

Table 41: Optimal Replacement Rates Calculated from (10) and (18) for $R = 5$

Optimal r for $G = 0$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.6399	0.6406	0.6417	0.6435
	0	0.6430	0.6450	0.6486	0.6547
E_b^y	0.048	0.6448	0.6477	0.6527	0.6615
	0.096	0.6466	0.6504	0.6569	0.6684
	0.192	0.6503	0.6558	0.6656	0.6828
	0.3072	0.6548	0.6625	0.6763	0.7011

Optimal r for $G = 0.208$:

		s_0			
		0.648	0.725	0.8	0.863
	-0.0816	0.5198	0.5020	0.4719	0.4229
	0	0.5639	0.5656	0.5685	0.5735
E_b^y	0.048	0.5898	0.6033	0.6266	0.6669
	0.096	0.6156	0.6408	0.6846	0.7598
	0.192	0.6666	0.7149	0.7983	0.9385
	0.3072	0.7263	0.8012	0.9282	1.1345

G.5 Summary of Extensions

In each of the four extensions considered, altering the model generally does change the numerical results; allowing for a 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.

References

Acemoglu, D., and R. Shimer (2000): "Productivity Gains from Unemployment Insurance," *European Economic Review*, 44(7), 1195–1224.

Addison, J. T., and M. L. Blackburn (2000): "The Effects of Unemployment Insurance on Postunemployment Earnings," *Labour Economics*, 7(1), 21–53.

Baily, M. N. (1978): "Some Aspects of Optimal Unemployment Insurance," *Journal of Public Economics*, 10(3), 379–402.

Blau, D. M. (1994): "Labor Force Dynamics of Older Men," *Econometrica*, 62(1), 117–156.

Blau, D. M., and P. K. Robins (1986): "Job Search, Wage Offers, and Unemployment Insurance," *Journal of Public Economics*, 29(2), 173–197.

Bound, J. (1989): "The Health and Earnings of Rejected Disability Insurance Applicants," *American Economic Review*, 79(3), 482–503.

Bound, J., J. B. Cullen, A. Nichols, and L. Schmidt (2004): "The Welfare Implications of Increasing Disability Insurance Benefit Generosity," *Journal of Public Economics*, 88(12), 2487–2514.

Boyle, M. A., and J. N. Lahey (2010): "Health Insurance and the Labor Supply Decisions of Older Workers: Evidence from a U.S. Department of Veterans Affairs Expansion," *Journal of Public Economics*, 94(7-8), 467–478.

Burgess, P. L., and J. L. Kingston (1976): "The Impact of Unemployment Insurance Benefits on Reemployment Success," *Industrial and Labor Relations Review*, 30(1), 25–31.

Card, D., R. Chetty, and A. Weber (2007): "Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," *Quarterly Journal of Economics*, 122(4), 1511–1560.

Card, D., J. Kluve, and A. Weber (2010): "Active Labour Market Policy Evaluations: A Meta-Analysis," *Economic Journal*, 120(548), F452–F477.

Centeno, M. (2004): "The Match Quality Gains from Unemployment Insurance," *Journal of Human Resources*, 39(3), 839–863.

Centeno, M., and A. A. Novo (2006): "The Impact of Unemployment Insurance Generosity on Match Quality Distribution," *Economics Letters*, 93(2), 235–241.

Chen, S., and W. van der Klaauw (2008): "The Work Disincentive Effects of the Disability Insurance Program in the 1990s," *Journal of Econometrics*, 142(2), 757–784.

Chetty, R. (2006): "A General Formula for the Optimal Level of Social Insurance," *Journal of Public Economics*, 90(10-11), 1879–1901.

——— (2008): "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 116(2), 173–234.

——— (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–488.

Chetty, R., and E. Saez (2010): "Optimal Taxation and Social Insurance with Endogenous Private Insurance," *American Economic Journal: Economic Policy*, 2(2), 85–114.

Classen, K. P. (1977): "The Effect of Unemployment Insurance on the Duration of Unemployment and Subsequent Earnings," *Industrial and Labor Relations Review*, 30(4), 438–444.

Coile, C., and J. Gruber (2007): "Future Social Security Entitlements and the Retirement Decision," *Review of Economics and Statistics*, 89(2), 234–246.

Coles, M. (2008): "Optimal Unemployment Policy in a Matching Equilibrium," *Labour Economics*, 15(4), 537–559.

Coles, M., and A. Masters (2006): "Optimal Unemployment Insurance in a Matching Equilibrium," *Journal of Labor Economics*, 24(1), 109–138.

Cushing, M. J. (2005): "Net Marginal Social Security Tax Rates over the Life Cycle," *National Tax Journal*, 58(2), 227–245.

Davidson, C., and S. A. Woodbury (1997): "Optimal Unemployment Insurance," *Journal of Public Economics*, 64(3), 359–387.

Dynarski, S. (2003): "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion," *American Economic Review*, 93(1), 279–288.

Ebenstein, A., and K. Stange (2010): "Does Inconvenience Explain Low Take-Up? Evidence from Unemployment Insurance," *Journal of Policy Analysis and Management*, 29(1), 111–136.

Ehrenberg, R. G., and R. L. Oaxaca (1976): "Unemployment Insurance, Duration of Unemployment, and Subsequent Wage Gain," *American Economics Review*, 66(5), 754–766.

Eisenhauer, J. G., and L. Ventura (2003): "Survey Measures of Risk Aversion and Prudence," *Applied Economics*, 35(13), 1477–1484.

Fairlie, R. W., K. Kapur, and S. Gates (2011): "Is Employer-Based Health Insurance a Barrier to Entrepreneurship?," *Journal of Health Economics*, 30(1), 146–162.

Farber, H. S. (1999): "Mobility and Stability: The Dynamics of Job Change in Labor Markets," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3b, Elsevier, pp. 2439–2483.

Feldstein, M. (1985): "The Optimal Level of Social Security Benefits," *Quarterly Journal of Economics*, 100(2), 303–320.

Fitzenberger, B., and R. A. Wilke (2007): "New Insights on Unemployment Duration and Post Unemployment Earnings in Germany: Censored Box-Cox Quantile Regression at Work," Discussion Paper No. 2609, IZA, Bonn, Germany.

Fredriksson, P., and B. Holmlund (2001): "Optimal Unemployment Insurance in Search Equilibrium," *Journal of Labor Economics*, 19(2), 370–399.

Gaure, S., K. Roed, and L. Westlie (2008): "The Impacts of Labor Market Policies on Job Search Behavior and Post-Unemployment Job Quality," Discussion Paper No. 3802, IZA.

Golosov, M., and A. Tsyvinski (2006): "Designing Optimal Disability Insurance: A Case for Asset Testing," *Journal of Political Economy*, 114(2), 257–279.

Gruber, J. (1996): "Disability Insurance Benefits and Labor Supply," Working Paper no. 5866, NBER, Cambridge, MA.

——— (1997): "The Consumption Smoothing Benefits of Unemployment Insurance," *American Economic Review*, 87(1), 192–205.

——— (2000): "Disability Insurance Benefits and Labor Supply," *Journal of Political Economy*, 108(6), 1162–1183.

Gruber, J., and B. C. Madrian (1994): "Health Insurance and Job Mobility: The Effects of Public Policy on Job-Lock," *Industrial and Labor Relations Review*, 48(1), 86–102.

Gruber, J., and E. Saez (2002): "The Elasticity of Taxable Income: Evidence and Implications," *Journal of Public Economics*, 84(1), 1–32.

Hansen, G. D., and A. İmrohoroğlu (1992): "The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard," *Journal of Political Economy*, 100(1), 118–142.

Holen, A. (1977): "Effects of Unemployment Insurance Entitlement on Duration and Job Search Outcome," *Industrial and Labor Relations Review*, 30(4), 445–450.

Hopenhayn, H. A., and J. P. Nicolini (1997): "Optimal Unemployment Insurance," *Journal of Political Economy*, 105(2), 412–438.

İmrohoroğlu, A., S. İmrohoroğlu, and D. H. Joines (1995): "A Life Cycle Analysis of Social Security," *Economic Theory*, 6(1), 83–114.

——— (2003): "Time-Inconsistent Preferences and Social Security," *Quarterly Journal of Economics*, 118(2), 745–784.

Kane, T. J. (1994): "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education," *Journal of Political Economy*, 102(5), 878–911.

Kroft, K., F. Lange, and M. J. Notowidigdo (2012): "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment," Working Paper no. 18387, NBER, Cambridge, MA.

Laitner, J., and D. Silverman (2012): "Consumption, Retirement and Social Security: Evaluating the Efficiency of Reform that Encourages Longer Careers," *Journal of Public Economics*, 96(7-8), 615–634.

Lalive, R. (2007): "Unemployment Benefits, Unemployment Duration, and Post-Unemployment Jobs: A Regression Discontinuity Approach," *American Economic Review: Papers and Proceedings*, 97(2), 108–112.

Lawson, N. (2013): "Fiscal Externalities, Liquidity Constraints and Grants to Post-Secondary Students," Unpublished Paper, Princeton University.

Lentz, R. (2009): "Optimal Unemployment Insurance in an Estimated Job Search Model with Savings," *Review of Economic Dynamics*, 12(1), 37–57.

Maani, S. A. (1993): "Post-Unemployment Wages, the Probability of Re-Employment, and the Unemployment Benefit," *New Zealand Economic Papers*, 27(1), 35–55.

McCall, B., and W. Chi (2008): "Unemployment Insurance, Unemployment Durations and Re-Employment Wages," *Economics Letters*, 99(1), 115–118.

Merrigan, P., and M. Normandin (1996): "Precautionary Saving Motives: An Assessment from UK Time Series of Cross-Sections," *Economic Journal*, 106(438), 1193–1208.

Meyer, B. D. (1989): "A Quasi-Experimental Approach to the Effects of Unemployment Insurance," Working Paper no. 3159, NBER, Cambridge, MA.

——— (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58(4), 757–782.

Parry, I. W., and W. E. Oates (2000): "Policy Analysis in the Presence of Distorting Taxes," *Journal of Policy Analysis and Management*, 19(4), 603–613.

Portugal, P., and J. T. Addison (2008): "Six Ways to Leave Unemployment," *Scottish Journal of Political Economy*, 5(4), 393–419.

Rust, J., and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, 65(4), 781–831.

Saez, E. (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117(3), 1039–1073.

Schmieder, J. F., T. von Wachter, and S. Bender (2012): "The Effect of Unemployment Insurance Extensions on Reemployment Wages," Unpublished Paper.

Shimer, R., and I. Werning (2007): "Reservation Wages and Unemployment Insurance," *Quarterly Journal of Economics*, 122(3), 1145–1185.

Tatsiramos, K. (2009): "Unemployment Insurance in Europe: Unemployment Duration and Subsequent Employment Stability," *Journal of the European Economic Association*, 7(6), 1225–1260.

van Ours, J. C., and M. Vodopivec (2008): "Does Reducing Unemployment Insurance Generosity Reduce Job Match Quality?," *Journal of Public Economics*, 92(3-4), 684–695.

Wang, C., and S. D. Williamson (2002): "Moral Hazard, Optimal Unemployment Insurance and Experience Rating," *Journal of Monetary Economics*, 49(7), 1337–1371.

Welch, F. (1977): "What Have We Learned from Empirical Studies of Unemployment Insurance?," *Industrial and Labor Relations Review*, 30(4), 451–461.

Wilson, J. D. (1999): "Theories of Tax Competition," *National Tax Journal*, 52(2), 269–304.