THREE ESSAYS ON THE ECONOMICS
OF FISCAL EXTERNALITIES

NICHOLAS PETER LAWSON

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF ECONOMICS

ADVISOR: DAVID S. LEE

SEPTEMBER 2013
Abstract

The three chapters of this dissertation investigate the effect of fiscal externalities on the welfare analysis of government social programs.

The first chapter introduces the subject of fiscal externalities, defined as the effects of a social program’s labour market impacts on income tax revenues, and assesses their importance in the context of optimal unemployment insurance. I calibrate and simulate a dynamic job search model, and perform a theoretical and numerical analysis of Baily’s (1978) two-period model of unemployment. The numerical results are significantly altered when fiscal externalities are taken into account: with moderate risk-aversion, the optimal replacement rate drops from around 40% to zero. However, higher risk-aversion moderates this effect, and a large positive effect of UI on wages could significantly increase the optimal benefit level.

In the second chapter, I evaluate the optimal tuition subsidy policy for post-secondary education while taking into account two of the most prominent justifications for government financial aid: liquidity constraints and fiscal externalities. I present a simple model and apply a sufficient statistics method to solve for an equation for the welfare gain from increasing aid. I then use statistical extrapolations and a calibration and simulation of my model to estimate optimal student grants. I find that financial aid should be more generous, and that the optimal amount of aid roughly corresponds to eliminating tuition at public universities. These results are largely unchanged if students face no liquidity constraints, whereas general equilibrium effects of tuition subsidies on wages can significantly affect the results.
In chapter 3, I add interactions between social programs to my study of fiscal externalities. I consider an example in which individuals can substitute between unemployment insurance and disability insurance, and demonstrate that the fiscal effects of substitution could significantly raise the optimal generosity of UI. I then present a general model which can be applied to any program of state-contingent transfers, and solve the model for a derivative of social welfare with respect to an individual program, with a simple and intuitive result that depends directly on the magnitude of fiscal externalities and program interaction effects.
Acknowledgements

I am deeply grateful to my advisor, David Lee, for his guidance, advice and encouragement; I am continually amazed at how much I have learned from him. I also wish to thank Richard Rogerson for his very helpful advice and for challenging me to consider alternative views on numerical research. I have also benefitted greatly from discussion and support from Orley Ashenfelter, and from my fellow students, particularly Yan Lau, Dean Spears, and Nikolaj Harmon.

I gratefully acknowledge financial support from Princeton University, the Richard A. Lester Fellowship for Industrial Relations, and the Fellowship of Woodrow Wilson Scholars. Chapter 1 of this dissertation was presented at EconCon 2012 at Princeton University, and I thank the participants there as well as seminar participants at Princeton and elsewhere for their feedback.

Finally, I would like to give special thanks to my family - my parents Monica and Scott Lawson, my sister Victoria and my brother Christopher - for their love and support throughout.
Contents

1 Fiscal Externalities and Optimal Unemployment Insurance ........................................ 1

1.1 Introduction .................................................................................................................. 1

1.2 What Are Fiscal Externalities? ...................................................................................... 6

1.3 Dynamic Job Search Model .......................................................................................... 8

1.3.1 Model Setup ............................................................................................................ 9

1.3.2 Calibration of the Model ......................................................................................... 10

1.3.3 Numerical Results ................................................................................................. 16

1.4 Baily (1978) Model ...................................................................................................... 19

1.4.1 Model Setup ......................................................................................................... 20

1.4.2 Calculation of Welfare Derivative .......................................................................... 21

1.4.3 Analysis of Optimal Benefit Equation .................................................................... 24

1.5 Analytical Results ........................................................................................................ 26

1.6 Numerical Results from the Baily Model ................................................................ 33

1.6.1 Summary of Numerical Procedure ....................................................................... 34

1.6.2 Sufficient Statistics and Extrapolations ................................................................. 36

1.6.3 Optimal Replacement Rates .................................................................................. 38

1.6.4 Comparison with Structural Results ....................................................................... 40

vi
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.6.5</td>
<td>Discussion of Numerical Results</td>
<td>42</td>
</tr>
<tr>
<td>1.7</td>
<td>Conclusion</td>
<td>44</td>
</tr>
<tr>
<td>1.8</td>
<td>Appendix</td>
<td>45</td>
</tr>
<tr>
<td>1.8.1</td>
<td>Technical Appendix for Structural Calibration</td>
<td>45</td>
</tr>
<tr>
<td>1.8.2</td>
<td>Sensitivity Analyses and Extensions with Structural Model</td>
<td>46</td>
</tr>
<tr>
<td>1.8.3</td>
<td>Algebra and Proofs</td>
<td>51</td>
</tr>
<tr>
<td>1.8.4</td>
<td>Second-Order Conditions</td>
<td>55</td>
</tr>
<tr>
<td>1.8.5</td>
<td>Baseline Values of $\frac{dW}{db}$</td>
<td>57</td>
</tr>
<tr>
<td>1.8.6</td>
<td>Sensitivity Analyses in the Baily Model</td>
<td>57</td>
</tr>
<tr>
<td>1.8.7</td>
<td>Extensions to Baily Model</td>
<td>58</td>
</tr>
<tr>
<td>2</td>
<td>Fiscal Externalities, Liquidity Constraints and Grants to Post-Secondary</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>73</td>
</tr>
<tr>
<td>2.2</td>
<td>The Role of Liquidity Constraints and Fiscal Externalities in Post-Secondary</td>
<td></td>
</tr>
<tr>
<td>2.2.1</td>
<td>Fiscal Externalities</td>
<td>77</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Liquidity Constraints</td>
<td>80</td>
</tr>
<tr>
<td>2.2.3</td>
<td>My Analysis of Grants to Students</td>
<td>83</td>
</tr>
<tr>
<td>2.3</td>
<td>A Simple Model of University Education</td>
<td>85</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Model Setup</td>
<td>85</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Welfare Calculations</td>
<td>89</td>
</tr>
<tr>
<td>2.4</td>
<td>Numerical Results</td>
<td>93</td>
</tr>
</tbody>
</table>
2.4.1 Sufficient Statistics Method ........................................ 93
2.4.2 Statistical Extrapolation ........................................... 96
2.4.3 Simulation of Structural Model ..................................... 98
2.5 Separating the Effects of Liquidity Constraints and Fiscal Externalities .... 102
  2.5.1 No Liquidity Constraints ........................................... 102
  2.5.2 No Fiscal Externalities ............................................ 104
  2.5.3 Relative Importance of Fiscal Externalities and Liquidity Constraints 105
2.6 Heterogeneity ............................................................ 106
  2.6.1 Heterogeneity in Liquidity Constraints and Two-Tier Grants ...... 106
  2.6.2 Heterogeneous Returns to Education ............................. 108
2.7 General Equilibrium Effects .......................................... 110
  2.7.1 GE Effects on PSE Wage Premium ............................... 111
  2.7.2 Wage Spillovers .................................................. 113
  2.7.3 Both GE Effects Combined ...................................... 115
2.8 Conclusion .............................................................. 116
2.9 Appendix ................................................................. 118
  2.9.1 Liquidity Term with Heterogeneous Constraints .................. 118
  2.9.2 Calculation of $\bar{\varepsilon}_b$ .................................. 120
  2.9.3 Sensitivity Analyses .............................................. 120

3 Interactions Between Government Social Programs: An Application and
  General Formula .......................................................... 126
3.1 Introduction .......................................................... 126

viii
3.2 Optimal UI With DI as Substitute Pathway . . . . . . . . . . . . . . . . 128
3.3 Understanding Program Interaction Effects & Fiscal Externalities . . . . . 135
3.4 Theoretical Analysis of General Model . . . . . . . . . . . . . . . . . . . 136
  3.4.1 General Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 136
  3.4.2 Calculation of Welfare Derivative . . . . . . . . . . . . . . . . . . . . 140
  3.4.3 Analysis of Welfare Derivative . . . . . . . . . . . . . . . . . . . . . . 144
3.5 Applications of the General Approach . . . . . . . . . . . . . . . . . . . . 149
  3.5.1 Social Insurance . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 149
  3.5.2 Human Capital Development . . . . . . . . . . . . . . . . . . . . . . . 151
3.6 Discussion and Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . 153
Chapter 1

Fiscal Externalities and Optimal Unemployment Insurance

1.1 Introduction

In an analysis of the optimal design of a government program, an important consideration is that 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 various 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).\textsuperscript{1}

Importantly, however, welfare analyses of one 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)).\textsuperscript{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, many government programs on their own account for only a small fraction of overall spending; this is particularly true in the case of a program like UI, which routinely 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

\textsuperscript{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 chapter 2.

\textsuperscript{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.
that pay for all government programs. I will refer to this phenomenon as 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 1.2, 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 of longer durations of unemployment. 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 effects dominates, the important point

---

\(^3\)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 1.2 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.”
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. 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. I provide a series of
analytical results about the importance of fiscal externalities, and I use a method of statis-
tical extrapolation to calculate numerical results for the optimal benefit level. Using both
approaches demonstrates the robustness of the results, and while the structural approach

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.
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, where I assume 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 rate still drops significantly. 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 1.2 further explains the concept of fiscal externalities in the context of government policy. Section 1.3 presents the dynamic job search model, and includes a discussion of the calibration and the results. In section 1.4, I describe the two-period model from Baily (1978), and solve for the welfare derivative and optimal UI equation; section 1.5 contains analytical results based on these equations. Numerical results from the statistical extrapolation method follow in section 1.6, and section
1.7 concludes.

1.2 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, and is closer to the concept described in the analysis of income tax policy in Buchanan (1966).9

---

8In 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.

9Buchanan (1966) complained that “theoretical welfare economics has been developed largely on the assumption that a public sector of the economy does not exist.” He argued that both sides of the fiscal situation should be modelled, and that therefore we will want to encourage activities that expand the tax base and discourage those that decrease it. Buchanan, however, was only concerned with a narrow conception of tax policy; the fact that individual responses to taxation impose externalities on other taxpayers was used
An alternative definition focuses 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 in either case, 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 highlights 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 to motivate an informal analysis of ways to better design the tax system to mitigate these externalities, rather than considering the consequences of such externalities for other policies.
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.

1.3 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 1.1. 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 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>
<thead>
<tr>
<th>Paper</th>
<th>Optimal Replacement Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hansen and Imrohoroglu (1992)</td>
<td>0.15* (with moral hazard)/0.65 (without)</td>
</tr>
<tr>
<td>Davidson and Woodbury (1997)</td>
<td>0.66*/1.30**</td>
</tr>
<tr>
<td>Hopenhayn and Nicolini (1997)</td>
<td>&gt; 0.94* (with optimal tax)</td>
</tr>
<tr>
<td>Acemoglu and Shimer (2000)</td>
<td>&gt; 0.4**</td>
</tr>
<tr>
<td>Fredriksson and Holmlund (2001)</td>
<td>0.38 – 0.42*</td>
</tr>
<tr>
<td>Wang and Williamson (2002)</td>
<td>0.24*/0.56**</td>
</tr>
<tr>
<td>Coles and Masters (2006)</td>
<td>0.76*</td>
</tr>
<tr>
<td>Lentz (2009)</td>
<td>0.43 – 0.82*</td>
</tr>
</tbody>
</table>

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

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.
1.3.1 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,\(^{10}\) 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.\(^ {11}\) All jobs have an identical wage $y$, and jobs end exogenously at a constant rate of $\delta$ per period. Agents cannot borrow, but can make savings which earn interest at a rate of $i$ per period; the annual discount rate is $\rho$, 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:

$$\max_{\{c_t, s_t, k_{t+1}\}} \sum_{t=0}^{\infty} \left(\frac{1}{1+\rho}\right)^{\frac{t}{52}} \left[U(c_t) - e(s_t)\right]$$

s.t.:

$$k_{t+1} = (1 + i)k_t + n_ty(1 - \tau) + (1 - n_t)b - c_t$$

$$c_t, k_{t+1} \geq 0, s_t \in [0, 1]$$

$$Pr(n_{t+1} = 1|n_t = 1) = 1 - \delta, Pr(n_{t+1} = 0|n_t = 1) = \delta$$

$$Pr(n_{t+1} = 1|n_t = 0) = s_t, Pr(n_{t+1} = 0|n_t = 0) = 1 - s_t$$

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

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

\(^{10}\)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.

\(^{11}\)This simplified specification allows for a closed-form solution for the individual’s search decision.
while \( V_u(k) \) will be the analogous value of unemployment, and let \( k' \) represent next period’s assets:

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

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

where \( \Gamma_z(k) = (k' \in \mathbb{R} | 0 \leq k' \leq (1 + i)k + z) \) is the set of permissable asset values.\(^{12}\)

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:\(^{13}\)

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

where \( u \) is the unemployment rate and \( G \) is the level of exogenous non-UI government spending.\(^{14}\) The government chooses \( b \) and \( \tau \) subject to this constraint to maximize steady-state expected utility,\(^{15}\) 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 \( G = 0 \).

### 1.3.2 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.\(^{16}\)

\(^{12}\) As in Lentz (2009), the numerical solutions always appear to yield concave value functions by asset level.

\(^{13}\) Assuming a single proportional tax keeps the government budget constraint simple; the basic result will hold with a more complex tax system if taxes paid remain an increasing function of income.

\(^{14}\) 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.

\(^{15}\) 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.

\(^{16}\) Thus, \( \bar{s} \) is the maximum feasible search intensity; in the simulations, this upper limit will be binding in very few instances.
Additionally, the way UI benefits are specified requires some explanation. If the replacement rate is $r$ and the baseline tax rate is $\tau_0$, then a natural value for after-tax benefits would be $r(1 - \tau_0)y$.\(^{17}\) 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.\(^{18}\) 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.\(^{19}\) 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.\(^{20}\)

The selected parameter values are summarized in Table 1.2. The job separation rate is set to $\delta = \frac{1}{260}$ to correspond with a median job duration of 5 years measured by the

---

\(^{17}\)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.

\(^{18}\)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.

\(^{19}\)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) make a slightly different adjustment, 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.

\(^{20}\)As discussed later, in appendix 1.8.2, I also model benefits that expire after 6 months. This is much less simple to implement numerically, but the results are similar.
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. $^21$ The wage $y$ is normalized to one. $R = 2$ is a standard value for the coefficient of relative risk-aversion in studies of UI, $^22$ but Chetty (2008) states that his results imply a value of about 5 in the context of unemployment, $^23$ 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. $^24$ 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, $\theta$, $\kappa$ and $\rho$, 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

---

$^21$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.


$^23$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 $R$.

$^24$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 1.2: Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>job separation rate</td>
<td>$\frac{1}{260}$</td>
</tr>
<tr>
<td>$i$</td>
<td>real interest rate</td>
<td>0</td>
</tr>
<tr>
<td>$y$</td>
<td>per-period wage</td>
<td>1</td>
</tr>
<tr>
<td>$r_0$</td>
<td>baseline replacement rate</td>
<td>0.46</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>baseline tax rate</td>
<td>0.23</td>
</tr>
<tr>
<td>$R$</td>
<td>coefficient of relative risk-aversion</td>
<td>${2, 5}$</td>
</tr>
<tr>
<td>$\bar{s}$</td>
<td>maximum search intensity</td>
<td>0.5</td>
</tr>
<tr>
<td>$\theta$</td>
<td>search cost parameter</td>
<td>TBD</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>search cost parameter</td>
<td>TBD</td>
</tr>
<tr>
<td>$\rho$</td>
<td>annual discount rate</td>
<td>TBD</td>
</tr>
</tbody>
</table>

of the unemployment rate with respect to benefits, which I denote as $E_b^u = \frac{b \, du}{du}$. 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 $\theta$. $E_b^u$, meanwhile, is informative about the curvature parameter $\kappa$. Finally, the consumption gap is closely related to workers’ ability to maintain a buffer stock of assets, so $\frac{E(c_e) - E(c_u)}{E(c_e)}$ primarily identifies the discount rate.\(^{25}\) 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 1.3, 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 $\tau_0$, this implicitly defines $G = 0.208$. Gruber (1997) estimates a relationship of $\frac{E(c_e) - E(c_u)}{E(c_e)} = 0.222 - 0.265r$,\(^{26}\) which

\(^{25}\)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 $\rho$; 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.

\(^{26}\)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
implies a value of 0.1001 at baseline. Finally, Chetty (2008) estimates an elasticity of unemployment durations with respect to benefits of 0.53, though this estimate is based on a sample of UI recipients, whereas the consumption estimates in Gruber (1997) are from a 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. If the average duration of unemployment is $D$, this gives $E^D_b = \frac{b}{D} \frac{dD}{db} = 0.2544$, and the fact that $u = \frac{D}{D + \frac{1}{3}}$ means that $E^u_b = (1 - u)E^D_b = 0.946 \times 0.2544 = 0.2407$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u$</td>
<td>unemployment rate</td>
<td>0.054</td>
</tr>
<tr>
<td>$\frac{E(c_e) - E(c_u)}{E(c_e)}$</td>
<td>consumption gap between employment and unemployment</td>
<td>$0.222 - 0.265r_0 = 0.1001$</td>
</tr>
<tr>
<td>$E^u_b = \frac{b}{u} \frac{du}{db}$</td>
<td>elasticity of $u$ wrt $b$</td>
<td>$0.946 \times 0.48 \times 0.53 = 0.2407$</td>
</tr>
</tbody>
</table>

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.

---

27In 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.

28This 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.”

290.48 is also very close to my value of $0.8 \times \frac{0.52}{2.43} \approx 0.52$ for the percentage of time that initially eligible unemployed individuals receive benefits, a closely related quantity.

30This 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
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 performing 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 \( E_u^m \), 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 \( r = 0.56 \), and compute the resulting arc elasticity.\(^{31}\) 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.

\(^{31}\)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.
1.3.3 Numerical Results

This subsection presents the numerical results; appendix 1.8.1 contains the parameters used in each case, the resulting moments, and a further discussion of the methods used. The results for the optimal replacement rates and estimated welfare gains can be found in Table 1.4; the column 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 column, labelled “Diff.”, display the welfare gain from moving between the replacement rate believed to be optimal when $G = 0$ and the “true” optimum.

Table 1.4: Optimal Replacement Rates & Welfare Gains

<table>
<thead>
<tr>
<th></th>
<th>$r$</th>
<th>Welf. Gain</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A. $R = 2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G = 0.208$</td>
<td>0.00</td>
<td>11.81</td>
<td>6.80</td>
</tr>
<tr>
<td>$G = 0$</td>
<td>0.37</td>
<td>0.45</td>
<td>6.33</td>
</tr>
<tr>
<td>B. $R = 5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G = 0.208$</td>
<td>0.46</td>
<td>0.00</td>
<td>6.08</td>
</tr>
<tr>
<td>$G = 0$</td>
<td>0.72</td>
<td>5.53</td>
<td>5.53</td>
</tr>
</tbody>
</table>

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, with an accompanying welfare gain

---

32 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.

33 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.

34 More accurately, the optimum is very close to zero, as close as computationally feasible; an actual zero 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.
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 1.1 and 1.2, 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 1.1 and those for $R = 5$ in Figure 1.2. 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.35

![Figure 1.1: Tax Rates for $R = 2$](image)

35Since $b$ 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 1.4 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%.
I have performed an extensive series of sensitivity analyses, with results that are displayed in appendix 1.8.2. 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 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 1.8.2. 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 overall 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,\(^{36}\) and so I will now switch my focus to 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.

### 1.4 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.

\(^{36}\)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.”
1.4.1 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,\(^{37}\) with the interest and discount rates both set to zero. In the first period, the representative worker is employed at an exogenous wage \( y \),\(^{38}\) and between periods they face a risk of unemployment: with exogenous probability \( \delta \) 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 and the remaining \( s \in (0, 1)\) at a new job at wage \( y_n \),\(^{39}\) where \( s \) is a deterministic function of \( e \) and \( y_n \):\(^{40}\)

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

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

\(^{37}\)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.

\(^{38}\)In an extension in appendix 1.8.7, I consider the effect of allowing choice over initial labour supply.

\(^{39}\)\( y_n \) 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.

\(^{40}\)In appendix 1.8.7, I examine the consequences of a stochastic \( s \).
expected utility is given by:

\[ V = U(c_e^1) + (1 - \delta)U(c_e^2) + \delta U(c_u) \]  

(1.1)

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

The government budget constraint over the two periods is:

\[ [(2 - \delta)y + \delta sy_n]\tau = \delta(1 - s)b + 2G \]  

(1.2)

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

### 1.4.2 Calculation of Welfare Derivative

The next step is to evaluate the derivative of social welfare with respect to UI benefits. From (1.1), 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 \( \tau \), 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 \( \tau \) 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:

\[ \frac{dV}{db} = \frac{\partial V}{\partial b} + \frac{\partial V}{\partial \tau} \frac{d\tau}{db}. \]  

(1.3)

\[41\] 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 1.8.7.
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:

\[
\frac{\partial V}{\partial b} = \delta U'(c_u) \frac{\partial c_u}{\partial b} = \delta (1 - s) U'(c_u) \tag{1.4}
\]

\[
\frac{\partial V}{\partial \tau} = U'(c_e^1) \frac{\partial c_e^1}{\partial \tau} + (1 - \delta) U'(c_e^2) \frac{\partial c_e^2}{\partial \tau} + \delta U'(c_u) \frac{\partial c_u}{\partial \tau} = -y U''(c_e^1) - (1 - \delta) y U''(c_e^2) - \delta s y_n U'(c_u). \tag{1.5}
\]

Therefore, the welfare derivative is:

\[
\frac{dV}{db} = \delta (1 - s) U'(c_u) - \left[ y U'(c_e^1) + (1 - \delta) y U'(c_e^2) + \delta s y_n U'(c_u) \right] \frac{d\tau}{db}. \tag{1.6}
\]

The goal now is to rewrite (1.6) in a form that can be expressed in terms of observable empirical quantities, so that 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 (1.6), 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 \( U'(c_e) \) around \( U'(c_u) \), specifically \( U'(c_e^1) = U'(c_u) + \Delta c U''(\theta) \), where \( \Delta c = c_e^1 - c_u \) and \( \theta \) is between \( c_u \) and \( c_e^1 \). This allows \( \frac{dV}{db} \) to be written in terms of \( U'(c_u) \) and \( U''(\theta) \). 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 1.8.3. These assumptions further simplify the expression, and I divide by \( U'(c_u) \) to put
welfare into a dollar equivalent and solve for \( \frac{dW}{db} \equiv \frac{dV}{U'(c_u)} \), arriving at the results summarized in the following proposition.

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

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

(1.7)

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 \( \omega = \frac{\delta(1-s)}{(2-\delta)y + \delta y_n} \). The equation for the optimal value of \( b \) is thus given by:

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

(1.8)

**Proof.** The proof of this result can be found in appendix 1.8.3.

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:

\[
\frac{dW}{db} = \frac{2u}{(1-u)\psi} \left[ \frac{\Delta c}{c_e^1} R E^*_b - (1-u) (E^*_b - \psi) \right]
\]

(1.9)

where \( E^*_b = \frac{b}{\tau} \frac{d\tau}{db} \) is the elasticity of \( \tau \) 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:

\[
\frac{\Delta c}{c_e^1} R = (1 - u) \frac{E^*_b - \psi}{E^*_b}.
\]

(1.10)

The meaning of these results will be interpreted in the next subsection.\(^{42}\)

\(^{42}\)It is clear from (1.5) that \( \frac{dV}{d\tau} < 0 \), and the pair of Baily’s assumptions described in appendix 1.8.3 imply that \( \frac{dV}{d\tau} = -2yU'(c_u) \left[ (1-u) - \frac{\Delta c}{c_e^1} R \right] \); 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 1.5.
1.4.3 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 (1.7) and (1.8), for which I must evaluate $\frac{d\tau}{db}$. Total differentiation of the government budget constraint (1.2) gives:

$$\frac{d\tau}{db} = \delta (1 - s) - \delta \delta y_n \frac{ds}{db} - \delta s \tau y_n \frac{dy}{db} + \delta s \tau y_n.$$  \hfill (1.11)

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 $\delta (1 - s) (2 - \delta) y + \delta s y_n$ 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 duration effect doesn’t depend on the size of government, the revenue effects are multiplied by $\tau$; this highlights the importance of the standard assumption that $\tau$ is a small payroll tax, rather than a large income tax.

It can now be seen that the $\omega$ defined in Proposition 1 is exactly equal to the mechanical effect, meaning that the $\frac{d\tau - \omega}{dt}$ term on the right-hand side of (1.8) is the fraction of the total response of taxes to benefits generated by the duration and revenue effects. (1.8) 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 \( \omega \) represents a lump-sum transfer of income between employed and unemployed states, and thus is not a cost to society, so \( \frac{d\tau}{d\tau} - \omega \) 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 (1.10) is exactly the same as that of (1.8): the left-hand side is identical, and on the right-hand side, \( E^\tau_b \) represents the percentage increase in the tax rate for a percentage increase in \( b \), while \( \psi \), the share of government expenditures on UI, is the percentage increase in \( \tau \) required for mechanical reasons. As in (1.8), therefore, the final fraction on the right is the fraction of the response of taxes to benefits caused by behavioural responses.

Using (1.11), \( E^\tau_b \) can be written as:

\[
E^\tau_b = \psi + \left( \psi + \frac{u}{1-u} \right) E^D_b - \frac{\delta s}{2(1-u)} E^y_b
\] (1.12)

where \( E^D_b = \frac{b}{1-s} \frac{d(1-s)}{db} \) is the elasticity of unemployment durations with respect to \( b \), and \( E^y_b = \frac{b}{y_n} \frac{dy_n}{db} \) 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 \( \psi \) is the mechanical effect, the \( \psi \) and \( \frac{u}{1-u} \) multiplying \( E^D_b \) 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:

\[
\frac{\Delta c}{c^1} R = (1 - u) \psi + \left( \psi + \frac{u}{1-u} \right) E^D_b - \frac{\delta s}{2(1-u)} E^y_b.
\] (1.13)
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 would be large compared to the revenue effects, because the taxes that induce the revenue effects are so small. However, if \( G \) is large, \( \psi \) 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 \( E_y^b = 0 \), 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 \( E_{\tau}^b \) in the denominator of the right-hand side of (1.10) is equal to one, or that \( \psi + \left( \psi + \frac{u}{1-u} \right) E^D_b = 1 \) in the denominator of (1.13).\(^{43}\) In this case, (1.13) collapses to exactly the result in Baily (1978), also found in Chetty (2006):

\[
\frac{\Delta c}{c_e} R = (1-u) \left( 1 + \frac{u}{1-u} \right) E^D_b = E^D_b.
\]

Therefore, \( \frac{\Delta c}{c_e} \), \( R \) and \( E^D_b \) 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.

### 1.5 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

---

\(^{43}\)Baily claims that this assumption leans towards overestimating the optimal \( b \), to offset the conservative assumption that \( c_1^U''(\theta) \simeq c_u U''(c_u) \). However, this is incorrect; the former assumption is in fact conservative, and the latter depends on the value of the risk-aversion parameter.
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 $E^y_b$ 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 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 $\frac{\Delta}{c^*_z} R < 1 - u$, and the second is that $W$ (the integral of $\frac{dW}{db}$) is strictly quasi-concave in $b$. Footnote 42 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 $sE^y_b - (1 - s)E^D_b$, or equivalently the same sign as $\frac{d(sy_n)}{db}$.

Proof. The proof of this result can be found in appendix 1.8.3.

Therefore, if two researchers use (1.9) 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 $sE^y_b - (1 - s)E^D_b > 0$, 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 1.3, which displays graphically how knowledge of a local welfare derivative doesn’t permit unambiguous conclusions about the optimum.\(^{44}\) 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 $\chi = \{ \Delta \frac{A c}{c}, R, s, E_b^D, E_b^y \}$ as the vector of sufficient statistics, the underlying quantities in (1.13) which are not exogenously fixed, and let $\chi(b)$ denote a particular vector of extrapolated values of these quantities.\(^{45}\) This leads to the following corollary.

**Figure 1.3: Two Possible Welfare Functions**

---

**Corollary 3.** For statistical extrapolations that do not depend on the estimated value of $G$,

\(^{44}\)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.

\(^{45}\)Strict quasi-concavity of $W$, when the latter is estimated out of sample using statistical extrapolations, implicitly places some restrictions on the extrapolations allowed.
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 \( b^*(G_1) \).

**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.\(^{46}\) 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 param-

\(^{46}\)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.
eter could be important; 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 1.8.3.

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.

**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 $\frac{dW}{db}(b; G, E_b^{y2}) > 0$, then $b^*(G, E_b^{y2}) > b^*(G, E_b^{y1})$.

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

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

---

\(^{47}\)The notation is now slightly altered to allow $E_b^{y}$ to enter into $\frac{dW}{db}$ and $b^*$ as an argument, as I now consider how results vary when different values of $E_b^{y}$ are selected.
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, $E_y^y$ 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 $E_y^{y_2} > E_y^{y_1}$, $\frac{dW}{db}(b; G_1, E_y^{y_2}) - \frac{dW}{db}(b; G_1, E_y^{y_1}) > \frac{dW}{db}(b; 0, E_y^{y_2}) - \frac{dW}{db}(b; 0, E_y^{y_1})$.

*Proof.* The proof of this result can be found in appendix 1.8.3. □

This proposition says that the effect of $E_y^y$ 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 $E_y^y$ is relatively unimportant to their calculations, but when $G$ is large, the tax rate will also be large, and $E_y^y$ 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 $E_y^y$.

The final analytical result addresses the question of whether $E_y^y$ 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 $b^*(G_1; E_y^y) - b^*(0; E_y^y)$ is increasing in $E_y^y$, i.e. that the increase in $b^*$ caused by $G$ is more positive when $E_y^y$ is larger; after all, I have proved that $b^*$ is increasing in $E_y^y$, and that the difference in the welfare derivative for different values of $E_y^y$ is increasing in $G$. This suspicion, however, cannot be turned into proof 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 $E_y^y$, if $\frac{dE_y^D}{db} \geq 0$, $\frac{dR}{db} = 0$, $E_y^D > -1$, and $\frac{d}{db} \left( \frac{\Delta \nu}{\nu^2} \right) < 0$, the following is true:
• if \( \exists E^y_b \) such that \( b^*(G_1, E^y_b) = b^*(0, E^y_b) \), then \( b^*(G_1, E^y_2) > b^*(0, E^y_2) \) for \( E^y_b > E^y_2 \) and \( b^*(G_1, E^y_1) < b^*(0, E^y_1) \) for \( E^y_1 < E^y_2 \).

Proof. The proof of this result can be found in appendix 1.8.3. \( \square \)

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 \( b^*(G_1; E^y_b) - b^*(0; E^y_b) \) is increasing in \( E^y_b \), I can state that for small values of \( E^y_b \), \( b^*(G_1, E^y_b) < b^*(0, E^y_b) \), and vice-versa for sufficiently large values of \( E^y_b \);\(^{48}\) this result is summarized by the diagram in Figure 1.4. Therefore, at least locally around \( E^y_0 \), 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.

1.6 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 sub-

\(^{48}\)In the unlikely case that an increase in \( E^y_b \) causes an increase in the optimal \( b \) which makes \( s \) decrease sufficiently quickly, the critical value \( E^y_b \) may not exist; in that case, \( b^*(G_1) - b^*(0) \) is always negative as long as \( E^y_b > 0 \).
section, 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.

1.6.1 Summary of Numerical Procedure

My goal in this section is to numerically evaluate (1.13) 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 $r = \frac{b}{(0.8) \left( \frac{\tau_0}{24.3} \right) y (1 - \tau_0)}$, where $\tau_0$ 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 (1.13) 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 $\Delta c\Delta e$ as a function of $r$ $[\Delta c\Delta e = 0.222 - 0.265r]$

• select fixed values of $E_D^b [0.2544]$, $E_y^b [0.048]$ 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 $\delta [0.54]$

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

• select the current value of $\psi [0.0457]$, and specify the relationship of $\psi$ to $r$

$$\left[\psi = 0.77^{12.64}_{24.3} \frac{ur}{(1-u)r}\right]$$

• 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 $E_y^b$, 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 $E_y^b$ leads to a higher optimal $b$ for each value of $G$. Meanwhile, in the first case, with a relatively small value of $E_y^b$, a larger estimate of $G$ reduces the optimal replacement rate, whereas with a larger value of $E_y^b$, 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.

1.6.2 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 $\frac{\Delta c}{c_t}$, I use the estimate of $\Delta C_{1e} = 0.222 - 0.265r$ from Gruber (1997). I also continue to use $E^D_b = 0.48 \times 0.53 = 0.2544$ from Chetty (2008) and Gruber (1997), and $R \in \{2, 5\}$.

For the elasticity of post-unemployment wages with respect to benefits, the empirical literature is fairly sparse and reports a wide range of results, which are summarized in Table 1.5.49 Since I am focusing on the effect of higher benefit levels on wages, I list only papers addressing that specific question,50 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.51

Chetty (2008), in particular, focuses on two recent papers, Card, Chetty, and Weber

---

49None 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).

50Not 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.

51Among 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.
Table 1.5: Results of Empirical Literature on Benefit Elasticity of Wages

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

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

(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. This corresponds with most recent studies in suggesting small or zero effects of UI on wages, but since the literature covers a wide range of values, and the analytical results in section 1.5 suggest that $E_{yb}$ can have an important effect on the optimal benefit level, I use a set of possible values for $E_{yb}$ covering the range found in Table 1.5, specifically $E_{yb} = 0.48 \times \{-0.17, 0, 0.1, 0.2, 0.4, 0.64\}$.54

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 \( \delta \) and \( \phi \) for each value of \( s_0 \) are easy to calculate, so I will not discuss them further here. Finally, at baseline values, \( \psi = \frac{ub}{(1-u)\tau y} = \frac{u}{1-u} \frac{1-\tau y}{\tau} (0.8) \left(\frac{15.8}{24.3}\right) r \), 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 1.6. I am now prepared to solve the non-linear first-order condition (1.13) defining the optimal replacement rate, an exercise which is pursued in the following subsection.

### Table 1.6: Sufficient Statistics & Parameters

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value/Extrapolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{4\pi}{c_1} )</td>
<td>0.222 – 0.265r</td>
</tr>
<tr>
<td>( E_b^D )</td>
<td>0.2544</td>
</tr>
<tr>
<td>( E_b^y )</td>
<td>0.48 \times {-0.17, 0, 0.1, 0.2, 0.4, 0.64}</td>
</tr>
<tr>
<td>( R )</td>
<td>{2, 5}</td>
</tr>
<tr>
<td>( r_0 )</td>
<td>0.46</td>
</tr>
<tr>
<td>( u_0 )</td>
<td>0.054</td>
</tr>
<tr>
<td>( s_0 )</td>
<td>{0.648, 0.725, 0.8, 0.863}</td>
</tr>
<tr>
<td>( \delta )</td>
<td>{0.3068, 0.3927, 0.5400, 0.7883}</td>
</tr>
<tr>
<td>( \phi )</td>
<td>{0.4289, 0.3351, 0.2437, 0.1669}</td>
</tr>
<tr>
<td>( \psi )</td>
<td>0.77 \frac{12.64}{24.3} \frac{ur}{(1-u)\tau}</td>
</tr>
</tbody>
</table>

#### 1.6.3 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
$r \in [0, 2]$, and I report a numerical check on the second-order conditions in appendix 1.8.4. Tables 1.7 and 1.8 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 1.7: Optimal Replacement Rates Calculated from (1.13) for $R = 2$

<table>
<thead>
<tr>
<th></th>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.0816</td>
<td>0.4590</td>
<td>0.4695</td>
<td>0.4790</td>
<td>0.4827</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.4595</td>
<td>0.4595</td>
<td>0.4595</td>
<td>0.4595</td>
</tr>
<tr>
<td>$E_b^y$</td>
<td>0.048</td>
<td>0.4651</td>
<td>0.4675</td>
<td>0.4717</td>
<td>0.4789</td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.4707</td>
<td>0.4756</td>
<td>0.4842</td>
<td>0.4987</td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>0.4821</td>
<td>0.4921</td>
<td>0.5095</td>
<td>0.5399</td>
</tr>
<tr>
<td></td>
<td>0.3072</td>
<td>0.4959</td>
<td>0.5122</td>
<td>0.5410</td>
<td>0.5922</td>
</tr>
</tbody>
</table>

Table 1.8: Optimal Replacement Rates Calculated from (1.13) for $R = 5$

<table>
<thead>
<tr>
<th></th>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.0816</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$E_b^y$</td>
<td>0.048</td>
<td>0.3001</td>
<td>0.2633</td>
<td>0.3933</td>
<td>0.5237</td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.3737</td>
<td>0.4711</td>
<td>0.5959</td>
<td>0.7591</td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>0.5650</td>
<td>0.6835</td>
<td>0.8514</td>
<td>1.0841</td>
</tr>
<tr>
<td></td>
<td>0.3072</td>
<td>0.7152</td>
<td>0.8638</td>
<td>1.0789</td>
<td>1.3788</td>
</tr>
</tbody>
</table>

Table 1.7: Optimal Replacement Rates Calculated from (1.13) for $R = 2$

<table>
<thead>
<tr>
<th></th>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.0816</td>
<td>0.5518</td>
<td>0.5307</td>
<td>0.4953</td>
<td>0.4387</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.5996</td>
<td>0.5996</td>
<td>0.5996</td>
<td>0.5996</td>
</tr>
<tr>
<td>$E_b^y$</td>
<td>0.048</td>
<td>0.6276</td>
<td>0.6402</td>
<td>0.6620</td>
<td>0.6987</td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.6554</td>
<td>0.6806</td>
<td>0.7240</td>
<td>0.7970</td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>0.7102</td>
<td>0.7601</td>
<td>0.8452</td>
<td>0.9853</td>
</tr>
<tr>
<td></td>
<td>0.3072</td>
<td>0.7741</td>
<td>0.8523</td>
<td>0.9831</td>
<td>1.1914</td>
</tr>
</tbody>
</table>

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

---

55I assume that the government is not interested in extracting payments from unemployed workers, so any zeros reported in the tables are corner solutions.
$R = 5$, 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 1.20 and 1.21 in appendix 1.8.5. 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.

1.6.4 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 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 1.12, and input them into the statistical extrapolation. This produces the results displayed in Table 1.9; the left side contains the structural results and simply replicates Table 1.4, while the right side displays my new results, where the welfare gains once again represent percentage points of baseline UI spending.

The results for the optimal replacement rate are remarkably similar except for a modest difference in the case with $R = 5$ and $G = 0.208$.\textsuperscript{56} Meanwhile, the estimated welfare gains

\textsuperscript{56}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,
Table 1.9: Comparison of Results with $E_b^y = 0$

<table>
<thead>
<tr>
<th>Structural Model</th>
<th></th>
<th></th>
<th>Statistical Extrapolations</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>Welf. Gain</td>
<td>Diff.</td>
<td></td>
<td>$r$</td>
</tr>
<tr>
<td></td>
<td>A.  $R = 2$</td>
<td></td>
<td></td>
<td>B.  $R = 5$</td>
<td></td>
</tr>
<tr>
<td>$G = 0.208$</td>
<td>0.00</td>
<td>11.81</td>
<td>6.80</td>
<td>$G = 0.208$</td>
<td>0.00</td>
</tr>
<tr>
<td>$G = 0$</td>
<td>0.37</td>
<td>0.45</td>
<td>6.33</td>
<td>$G = 0$</td>
<td>0.46</td>
</tr>
<tr>
<td>$G = 0.208$</td>
<td>0.46</td>
<td>0.00</td>
<td>6.08</td>
<td>$G = 0.208$</td>
<td>0.60</td>
</tr>
<tr>
<td>$G = 0$</td>
<td>0.72</td>
<td>5.53</td>
<td>5.53</td>
<td>$G = 0$</td>
<td>0.69</td>
</tr>
</tbody>
</table>

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 $R = 2$ 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.\(^{57}\)

A similar comparison can be done for the case with a non-degenerate wage distribution, using the moments from the structural model in appendix 1.8.2, and assuming $s_0 = 0.8$ for the statistical extrapolations; these results can be found in Table 1.10. 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.\(^{58}\) However, the effect of fiscal externalities on the optimal benefit level is very similar across the different approaches.

\(^{57}\)such as the fact that the steady-state asset distribution changes with $b$ in the structural model.

\(^{58}\)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
Table 1.10: Comparison of Results with $E_b^y \simeq 0.0876$ and $R = 2$

<table>
<thead>
<tr>
<th></th>
<th>Structural Model</th>
<th></th>
<th>Statistical Extrapolations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>Welf. Gain</td>
<td>Diff.</td>
</tr>
<tr>
<td></td>
<td>A.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$G = 0.208$</td>
<td>0.71</td>
<td>4.27</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.55</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>B.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$G = 0.208$</td>
<td>1.08</td>
<td>51.98</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>1.07</td>
<td>41.05</td>
</tr>
</tbody>
</table>

1.6.5 Discussion of Numerical Results

The results in Tables 1.7 and 1.8 strongly support the predictions of section 1.5; 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 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.

\(^{58}\) The wage distribution in Acemoglu and Shimer (2000) only features two mass points.
A value of $E^y_b$ 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 $E^y_b$. 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 1.8.6. First, I try alternative values of $E^D_b$, 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.\(^{59}\) 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 $E^y_b = 0$, and the analysis with alternative tax rates results in slightly increased effects of fiscal externalities.

Given the simplicity of the Baily model, I also perform a number of extensions to the\(^{59}\) $E^D_b = 0.144$ is small enough to generate a positive local maximum for $R = 2$ and $E^y_b = 0$; to obtain a positive global maximum, I need $E^D_b$ to drop below 0.096.
model, which can be found in appendix 1.8.7. 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.

1.7 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 the second chapter of this dissertation, I apply a similar analysis to post-secondary education, and future work will also consider optimal social insurance in a more general framework of fiscal interactions.

1.8 Appendix

1.8.1 Technical Appendix for Structural Calibration

Table 1.11 contains the parameters used in each case; Table 1.12 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 1.11: Calibrated Parameters

<table>
<thead>
<tr>
<th></th>
<th>$R = 2$</th>
<th>$G = 0.208$</th>
<th>$G = 0$</th>
<th>$R = 5$</th>
<th>$G = 0.208$</th>
<th>$G = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\rho)</td>
<td>0.0109</td>
<td>0.01076</td>
<td>0.0485</td>
<td>0.0473</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta)</td>
<td>32.0</td>
<td>26.3</td>
<td>92.0</td>
<td>40.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\kappa)</td>
<td>0.527</td>
<td>0.56</td>
<td>0.395</td>
<td>0.465</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.12: Calculated Moments

<table>
<thead>
<tr>
<th></th>
<th>$R = 2$</th>
<th>$G = 0.208$</th>
<th>$G = 0$</th>
<th>$R = 5$</th>
<th>$G = 0.208$</th>
<th>$G = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(u)</td>
<td>0.0539</td>
<td>0.0542</td>
<td>0.0538</td>
<td>0.0541</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(E_u)</td>
<td>0.2405</td>
<td>0.2405</td>
<td>0.2408</td>
<td>0.2407</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(E(c_e) - E(c_u))</td>
<td>0.1001</td>
<td>0.0998</td>
<td>0.1001</td>
<td>0.0999</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.8.2 Sensitivity Analyses and Extensions with Structural Model

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 1.13, and the complete sets of resulting parameters and moments are available upon request.

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.
Table 1.13: Optimal Replacement Rates & Welfare Gains

<table>
<thead>
<tr>
<th></th>
<th>$R = 2$</th>
<th></th>
<th>$R = 5$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>Welf. Gain</td>
<td>Diff.</td>
<td>$r$</td>
</tr>
<tr>
<td>(1 + $i$)$^{0.24}$ = 0.03</td>
<td>$G = 0.208$</td>
<td>0.01</td>
<td>41.59</td>
<td>0.00</td>
</tr>
<tr>
<td>$\delta = \frac{1}{364}$</td>
<td>$G = 0.208$</td>
<td>0.00</td>
<td>11.90</td>
<td>7.27</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.38</td>
<td>0.44</td>
<td>6.44</td>
</tr>
<tr>
<td>$t_b = 0.15, t_y = 0.262$</td>
<td>$G = 0.237$</td>
<td>0.00</td>
<td>12.76</td>
<td>7.49</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.37</td>
<td>0.45</td>
<td>6.43</td>
</tr>
<tr>
<td>$E_b^u = 0.1362$</td>
<td>$G = 0.208$</td>
<td>0.12</td>
<td>4.60</td>
<td>3.98</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.46</td>
<td>0.02</td>
<td>4.97</td>
</tr>
<tr>
<td>$E_b^u = 0.3633$</td>
<td>$G = 0.208$</td>
<td>0.00</td>
<td>19.90</td>
<td>8.96</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.32</td>
<td>1.64</td>
<td>6.01</td>
</tr>
<tr>
<td>perfect take-up</td>
<td>$G = 0.205$</td>
<td>0.04</td>
<td>9.09</td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.34</td>
<td>0.80</td>
<td>4.37</td>
</tr>
<tr>
<td>$u = 0.07$</td>
<td>$G = 0.201$</td>
<td>0.01</td>
<td>11.70</td>
<td>6.78</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.37</td>
<td>0.54</td>
<td>6.16</td>
</tr>
<tr>
<td>consumption drop</td>
<td>$G = 0.208$</td>
<td>0.30</td>
<td>2.25</td>
<td>6.30</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.59</td>
<td>1.13</td>
<td>5.01</td>
</tr>
<tr>
<td>utility from leisure</td>
<td>$G = 0.208$</td>
<td>0.05</td>
<td>9.91</td>
<td>5.96</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.39</td>
<td>0.35</td>
<td>6.20</td>
</tr>
<tr>
<td>different risk-aversion</td>
<td>$G = 0.208$</td>
<td>0.09</td>
<td>7.82</td>
<td>6.98</td>
</tr>
<tr>
<td></td>
<td>$G = 0$</td>
<td>0.45</td>
<td>0.01</td>
<td>6.13</td>
</tr>
</tbody>
</table>

and a rate of 0.262 for earned income.⁶⁰ 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 of 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

---

⁶⁰Since 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.
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 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.

**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:

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

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

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 1.14 and 1.15, and the numerical results are in Table 1.16.

Table 1.14: Calibrated Parameters with Wage Distribution

<table>
<thead>
<tr>
<th></th>
<th>( R = 2 )</th>
<th>( R = 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( G = 0.208 )</td>
<td>( G = 0.208 )</td>
</tr>
<tr>
<td></td>
<td>( G = 0 )</td>
<td>( G = 0 )</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.01148</td>
<td>0.01147</td>
</tr>
<tr>
<td>( \theta )</td>
<td>9.94</td>
<td>14.54</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>2.94</td>
<td>2.31</td>
</tr>
<tr>
<td>( d )</td>
<td>0.72</td>
<td>0.98</td>
</tr>
<tr>
<td>( y )</td>
<td>0.8041</td>
<td>0.80335</td>
</tr>
<tr>
<td>( \mu )</td>
<td>-2.344</td>
<td>-1.101</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.776</td>
<td>0.3823</td>
</tr>
</tbody>
</table>

Table 1.15: Calculated Moments with Wage Distribution

<table>
<thead>
<tr>
<th></th>
<th>( R = 2 )</th>
<th>( R = 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( G = 0.208 )</td>
<td>( G = 0.208 )</td>
</tr>
<tr>
<td></td>
<td>( G = 0 )</td>
<td>( G = 0 )</td>
</tr>
<tr>
<td>( u )</td>
<td>0.0541</td>
<td>0.0540</td>
</tr>
<tr>
<td>( E_0^{yu} )</td>
<td>0.2392</td>
<td>0.2418</td>
</tr>
<tr>
<td>( E(c_e) - E(c_u) )</td>
<td>0.1001</td>
<td>0.1002</td>
</tr>
<tr>
<td>( E_0^{yu} )</td>
<td>0.0199</td>
<td>0.0201</td>
</tr>
<tr>
<td>( E(w) )</td>
<td>1.000</td>
<td>1.002</td>
</tr>
</tbody>
</table>

Table 1.16: Optimal Replacement Rates & Welfare Gains

<table>
<thead>
<tr>
<th></th>
<th>( R = 2 )</th>
<th>( R = 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( G = 0.208 )</td>
<td>( G = 0.208 )</td>
</tr>
<tr>
<td></td>
<td>( G = 0 )</td>
<td>( G = 0 )</td>
</tr>
<tr>
<td>( r )</td>
<td>0.71</td>
<td>1.08</td>
</tr>
<tr>
<td>Welf. Gain</td>
<td>4.27</td>
<td>51.98</td>
</tr>
<tr>
<td>Diff.</td>
<td>1.65</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>0.55</td>
<td>1.07</td>
</tr>
<tr>
<td>Welf. Gain</td>
<td>0.48</td>
<td>41.05</td>
</tr>
<tr>
<td>Diff.</td>
<td>1.64</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Allowing for this positive effect on wages leads to dramatically different results from those observed in Table 1.4; 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). Therefore, 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 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 1.13. I also use a different functional form for the effort cost of search, specifically $e_t(s) = d - \theta_t (1 - s)^{1 - \kappa} - \theta_t (s - \frac{1}{1 - \pi})$.

To capture duration dependency and ensure that a reasonable proportion of unemployment spells are of long duration and involve exhaustion of benefits, I allow $\theta_t$ to increase with time $t$ spent out of work according to experimental results in Kroft, Lange, and Notowidigdo (2012); 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 1.17 and 1.18, and the numerical results are in Table 1.19.

This more realistic 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

---

$^{61}$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 $\theta_t$ and a target search intensity $s_1$, I therefore find the $\theta_t$ for $t \in \{2, 3, 4, 5, 6, 7\}$ (where 7 represents all periods of benefit exhaustion) that generates the same effort cost for $s_t = s_1 - (t - 1)s_1/14$. 

50
Table 1.17: Calibrated Parameters with Finite-Duration Benefits

<table>
<thead>
<tr>
<th></th>
<th>$R = 2$</th>
<th>$R = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$G = 0.203$</td>
<td>$G = 0$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.0092</td>
<td>0.00915</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>10.45</td>
<td>8.90</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>$d$</td>
<td>-0.708</td>
<td>-0.576</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$R = 2$</th>
<th>$R = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$G = ?$</td>
<td>$G = 0$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.0951</td>
<td>0.0918</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.789</td>
<td>0.3705</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>3.7</td>
<td>3.7</td>
</tr>
<tr>
<td>$d$</td>
<td>-2.266</td>
<td>-0.94346</td>
</tr>
</tbody>
</table>

Table 1.18: Calculated Moments with Finite-Duration Benefits

<table>
<thead>
<tr>
<th></th>
<th>$R = 2$</th>
<th>$R = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$G = 0.203$</td>
<td>$G = 0$</td>
</tr>
<tr>
<td>$u$</td>
<td>0.0538</td>
<td>0.0540</td>
</tr>
<tr>
<td>$E(u_t)$</td>
<td>0.2410</td>
<td>0.2403</td>
</tr>
<tr>
<td>$E(c_t^e - E(c_t^u))$</td>
<td>0.1002</td>
<td>0.1002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$R = 2$</th>
<th>$R = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$G = ?$</td>
<td>$G = 0$</td>
</tr>
<tr>
<td>$u$</td>
<td>0.0539</td>
<td>0.0541</td>
</tr>
<tr>
<td>$E(u_t)$</td>
<td>0.2411</td>
<td>0.2407</td>
</tr>
<tr>
<td>$E(c_t^e - E(c_t^u))$</td>
<td>0.0999</td>
<td>0.1000</td>
</tr>
</tbody>
</table>

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 1.13). 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.

1.8.3 Algebra and Proofs

Proof of Proposition 1

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

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

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

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

where $\theta$ is between $c_u$ and $c_t^1$, and $\Delta c = c_t^1 - c_u$. Combining these allows me to rewrite (1.6) as:

$$\frac{dV}{db} = -2y\Delta c U''(\theta)\frac{d\tau}{db} - [(2 - \delta)y + \delta \kappa] U'(c_u) \left[\frac{d\tau}{db} - \omega\right]$$
Table 1.19: Optimal Replacement Rates & Welfare Gains

<table>
<thead>
<tr>
<th></th>
<th>( R = 2 )</th>
<th></th>
<th>( R = 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( r )</td>
<td>Welf. Gain</td>
<td>Diff.</td>
</tr>
<tr>
<td>( G = 0.203/0.207 )</td>
<td>0.00</td>
<td>10.62</td>
<td>3.84</td>
</tr>
<tr>
<td>( G = 0 )</td>
<td>0.28</td>
<td>1.39</td>
<td>2.69</td>
</tr>
</tbody>
</table>

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

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 \( c_eU''(\theta) = c_uU''(c_u) \), 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:

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

where \( R = \frac{-c_uU''(c_u)}{U'(c_u)} \) 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:

\[
\frac{\Delta c}{c_e^1}R = (1-u)\frac{d\tau}{db} - \omega.
\]
Proof of Proposition 2

Starting from (1.9):
\[
\frac{dW}{db}(b; G) = 2u \left( \frac{\Delta c}{c_e^1} R \frac{E_v^f}{\psi} - (1 - u) \left( \frac{E_b^v}{\psi} - 1 \right) \right)
\]

and therefore the difference in welfare derivatives is:
\[
\frac{dW}{db}(b; G_1) - \frac{dW}{db}(b; 0) = 2u \frac{\Delta c}{c_e^1} R - (1 - u) \left[ \left( \frac{E_b^v}{\psi} \right)_{G=G_1} - \left( \frac{E_b^v}{\psi} \right)_{G=0} \right].
\]

Using (1.12) and the definition of $\psi$:
\[
\frac{E_b^v}{\psi} = 1 + E_b^D + \frac{ub + G}{ub} \left[ \frac{\delta(1 - s)}{2(1 - u)} E_b^D - \frac{\delta s}{2(1 - u)} E_y^y \right]
\]

and thus the welfare derivative difference becomes:
\[
\frac{dW}{db}(b; G_1) - \frac{dW}{db}(b; 0) = \frac{\delta}{(1 - u)^2 b} \left[ \frac{\Delta c}{c_e^1} R - (1 - u) \right] \left[ (1 - s) E_b^D - sE_y^y \right] G_1.
\]

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

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

Proof of Proposition 3

Combining (1.9) and (1.12), I immediately get:
\[
\frac{dW}{db}(b; G, E_b^{y2}) - \frac{dW}{db}(b; G, E_b^{y1}) = \frac{2u}{(1 - u)\psi} \left[ \frac{\Delta c}{c_e^1} R - (1 - u) \right] \left[ -\frac{\delta s}{2(1 - u)} \right] \left[ E_b^{y2} - E_b^{y1} \right].
\]

(1.14)

Given that $E_b^{y2} > E_b^{y1}$, and since the middle two terms are both negative, this expression is always positive.
Proof of Propositions 4 and 5

I start with (1.14), and from the definition of $\psi$:

$$
\left[\frac{dW}{db}(b; G_1, E_b^{y_2}) - \frac{dW}{db}(b; G_1, E_b^{y_1})\right] - \left[\frac{dW}{db}(b; 0, E_b^{y_2}) - \frac{dW}{db}(b; 0, E_b^{y_1})\right]
= \frac{-\delta s}{(1-u)^3b} \frac{\Delta c}{c^1_e} R - (1-u) \left[ E_b^{y_2} - E_b^{y_1} \right] G_1.
$$

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, $\frac{\Delta c}{c^1_e} RE_b^y = (1-u) (E_b^y - \psi)$; using (1.12) and rearranging, this becomes:

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

Observe that, because $\frac{\Delta c}{c^1_e} R < 1 - u$, $s E_b^y - (1-s) E_b^D > 0$ at the optimum if and only if $\frac{\Delta c}{c^1_e} R (1 + E_b^D) - (1-u) E_b^D < 0$. I wish to show that $b^*(G_1, E_b^{y_2}) > b^*(0, E_b^{y_2})$ and $b^*(G_1, E_b^{y_1}) < b^*(0, E_b^{y_1})$ for $E_b^{y_2} > E_b^{y_1} > E_b^y$, so Proposition 2 says that $s E_b^y - (1-s) E_b^D$ must be positive for $E_b^{y_2}$ and negative for $E_b^{y_1}$. Then, if I define $X(b) = \frac{\Delta c}{c^1_e} R (1 + E_b^D) - (1-u) E_b^D$, I want $X < 0$ at the optimum for $E_b^{y_2}$ and $X > 0$ for $E_b^{y_1}$; 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:

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

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

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

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 $(1-u) \frac{d}{db} \left( \frac{\Delta c}{c^1_e} \right) < \frac{\Delta c}{c^1_e} \frac{d(1-u)}{db}$. 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 $\frac{dE^D}{db}$ “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 $E^D_b$ is positive. The final assumption requires a bit more explanation; it is easiest to understand when written as $\frac{d}{db} \left( \frac{\Delta c}{c^*_1} \right) < 0$. The consumption gap $\frac{\Delta c}{c^*_1}$ 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 1.6 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.

1.8.4 Second-Order Conditions

Throughout the analytical results in section 1.5, I assume strict quasi-concavity, which ensures that the first-order condition (1.13) 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 $E^y_b \geq 0.048$, 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 $\Delta c / c^2 R < 1 - u$ may fail to hold and my assumptions will tend to be most inaccurate, and at which the estimated $E^r_c$ could turn negative. Over the vast majority of the range of $r$ that I consider, however, $dW/db$ behaves normally and consistent with quasi-concavity.

In each of the sensitivity analyses in appendix 1.8.6, similar local minima are found at low $r$ for $R = 5$ and $E^y_b \geq 0.048$. Additionally, in the extension with $E^D_b = 0.144$, for $R = 2$, $E^y_b = 0$ 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 $R = 2$, $E^y_b = 0.048$ 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 1.8.7; in the first, second, and fourth extensions, local minima are observed for $R = 5$ and $E^y_b \geq 0.048$. The first extension presents cases for $R = 2$, $E^y_b = 0$ 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 $E^y_b$ and $s_0$, though as discussed in appendix 1.8.7, these are anomalous as the optimal replacement rate should logically be close to one. A second local maximum also occurs in the second extension for $R = 2$, $E^y_b = 0.048$ and $s_0 = 0.648$. In the third extension, local maxima are also observed for $R = 5$, $E^y_b = 0$, and each $s_0$, but once again the global maximum is at zero.
1.8.5 Baseline Values of $\frac{dW}{db}$

Equation (1.9), when combined with (1.12), provides a way of evaluating $\frac{dW}{db}$, and I present in Tables 1.20 and 1.21 the values of this derivative at the baseline value of $r = 0.46$. As discussed in subsection 1.6.3, the results are conceptually similar to those in Tables 1.7 and 1.8, in that a positive value of $G$ causes the values in the table to “spread out.”

Table 1.20: Baseline Values of $\frac{dW}{db}$ Calculated from (1.9) and (1.12) for $R = 2$

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E^D_b$</td>
<td>-0.0816</td>
<td>-0.0008</td>
<td>-0.0011</td>
<td>-0.0016</td>
</tr>
<tr>
<td>0</td>
<td>-0.0000</td>
<td>-0.0000</td>
<td>-0.0000</td>
<td>-0.0000</td>
</tr>
<tr>
<td>0.048</td>
<td>0.0004</td>
<td>0.0006</td>
<td>0.0009</td>
<td>0.0014</td>
</tr>
<tr>
<td>0.096</td>
<td>0.0008</td>
<td>0.0012</td>
<td>0.0018</td>
<td>0.0029</td>
</tr>
<tr>
<td>0.192</td>
<td>0.0017</td>
<td>0.0024</td>
<td>0.0037</td>
<td>0.0058</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.0027</td>
<td>0.0039</td>
<td>0.0059</td>
<td>0.0094</td>
</tr>
</tbody>
</table>

Table 1.21: Baseline Values of $\frac{dW}{db}$ Calculated from (1.9) and (1.12) for $R = 5$

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E^D_b$</td>
<td>-0.0816</td>
<td>-0.0118</td>
<td>-0.0428</td>
<td>-0.0605</td>
</tr>
<tr>
<td>0</td>
<td>-0.0258</td>
<td>-0.0258</td>
<td>-0.0258</td>
<td>-0.0258</td>
</tr>
<tr>
<td>0.048</td>
<td>-0.0165</td>
<td>-0.0124</td>
<td>-0.0054</td>
<td>0.0063</td>
</tr>
<tr>
<td>0.096</td>
<td>-0.0071</td>
<td>0.0011</td>
<td>0.0150</td>
<td>0.0384</td>
</tr>
<tr>
<td>0.192</td>
<td>0.0117</td>
<td>0.0280</td>
<td>0.0558</td>
<td>0.1027</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.0343</td>
<td>0.0602</td>
<td>0.1048</td>
<td>0.1798</td>
</tr>
</tbody>
</table>

1.8.6 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 $E^D_b$ over a wide range; I try $E^D_b = 0.48 \times 0.3$, with results in Tables 1.22 and 1.23, and $E^D_b = 0.48 \times 0.8$, with results in Tables 1.24 and 1.25. 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 $E^D_b$
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 1.26 and 1.27. 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 $E_{b}^{y} = 0$.

I then try a larger value of the initial unemployment rate, specifically $u_0 = 0.064$. This leads to the results displayed in Tables 1.28 and 1.29. 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 1.8.2, 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 1.30 and 1.31, and the effects of fiscal externalities are slightly increased.

### 1.8.7 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.
Table 1.22: Optimal Replacement Rates for $R = 2$ and $E_D^b = 0.144$

<table>
<thead>
<tr>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_0$</td>
<td>0.648</td>
</tr>
<tr>
<td>0.0816</td>
<td>0.5908</td>
</tr>
<tr>
<td>0</td>
<td>0.6020</td>
</tr>
<tr>
<td>$E_D^b$</td>
<td>0.048</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6154</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6289</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.6454</td>
</tr>
<tr>
<td>*Local Maximum of 0.3451 for this row</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.23: Optimal Replacement Rates for $R = 5$ and $E_D^b = 0.144$

<table>
<thead>
<tr>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_0$</td>
<td>0.648</td>
</tr>
<tr>
<td>0.0816</td>
<td>0.7391</td>
</tr>
<tr>
<td>0</td>
<td>0.7435</td>
</tr>
<tr>
<td>$E_D^b$</td>
<td>0.048</td>
</tr>
<tr>
<td>0.096</td>
<td>0.7487</td>
</tr>
<tr>
<td>0.192</td>
<td>0.7540</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7606</td>
</tr>
</tbody>
</table>

Table 1.24: Optimal Replacement Rates for $R = 2$ and $E_D^b = 0.384$

<table>
<thead>
<tr>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_0$</td>
<td>0.648</td>
</tr>
<tr>
<td>0.0816</td>
<td>0.3129</td>
</tr>
<tr>
<td>0</td>
<td>0.3213</td>
</tr>
<tr>
<td>$E_D^b$</td>
<td>0.048</td>
</tr>
<tr>
<td>0.096</td>
<td>0.3312</td>
</tr>
<tr>
<td>0.192</td>
<td>0.3411</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.3531</td>
</tr>
</tbody>
</table>

Table 1.25: Optimal Replacement Rates for $R = 5$ and $E_D^b = 0.384$

<table>
<thead>
<tr>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_0$</td>
<td>0.648</td>
</tr>
<tr>
<td>0.0816</td>
<td>0.6292</td>
</tr>
<tr>
<td>0</td>
<td>0.6320</td>
</tr>
<tr>
<td>$E_D^b$</td>
<td>0.048</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6554</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6899</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.6430</td>
</tr>
</tbody>
</table>

59
Table 1.26: Optimal Replacement Rates for $R = 2$ and Perfect Take-Up

Optimal $r$ for $G = 0$:  

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.4500</td>
<td>0.4595</td>
<td>0.4390</td>
<td>0.4275</td>
</tr>
<tr>
<td>0</td>
<td>0.4595</td>
<td>0.4505</td>
<td>0.4595</td>
<td>0.4595</td>
</tr>
<tr>
<td>$E_b^v$</td>
<td>0.048</td>
<td>0.4651</td>
<td>0.4675</td>
<td>0.4717</td>
</tr>
<tr>
<td>0.096</td>
<td>0.4707</td>
<td>0.4756</td>
<td>0.4842</td>
<td>0.4987</td>
</tr>
<tr>
<td>0.192</td>
<td>0.4821</td>
<td>0.4921</td>
<td>0.5095</td>
<td>0.5399</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.4959</td>
<td>0.5122</td>
<td>0.5410</td>
<td>0.5922</td>
</tr>
</tbody>
</table>

Optimal $r$ for $G = 0.205$:  

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0.6831</td>
<td>0.6815</td>
<td>0.6787</td>
<td>0.6741</td>
</tr>
<tr>
<td>$E_b^v$</td>
<td>0.048</td>
<td>0.6887</td>
<td>0.6897</td>
<td>0.6914</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6909</td>
<td>0.6928</td>
<td>0.6962</td>
<td>0.7020</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6951</td>
<td>0.6991</td>
<td>0.7061</td>
<td>0.7183</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7004</td>
<td>0.7069</td>
<td>0.7184</td>
<td>0.7390</td>
</tr>
</tbody>
</table>

Table 1.27: Optimal Replacement Rates for $R = 5$ and Perfect Take-Up

Optimal $r$ for $G = 0$:  

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.6831</td>
<td>0.6815</td>
<td>0.6787</td>
<td>0.6741</td>
</tr>
<tr>
<td>0</td>
<td>0.6866</td>
<td>0.6866</td>
<td>0.6866</td>
<td>0.6866</td>
</tr>
<tr>
<td>$E_b^v$</td>
<td>0.048</td>
<td>0.6887</td>
<td>0.6897</td>
<td>0.6914</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6909</td>
<td>0.6928</td>
<td>0.6962</td>
<td>0.7020</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6951</td>
<td>0.6991</td>
<td>0.7061</td>
<td>0.7183</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7004</td>
<td>0.7069</td>
<td>0.7184</td>
<td>0.7390</td>
</tr>
</tbody>
</table>

Optimal $r$ for $G = 0.205$:  

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0.5786</td>
<td>0.5614</td>
<td>0.5222</td>
<td>0.4847</td>
</tr>
<tr>
<td>$E_b^v$</td>
<td>0.048</td>
<td>0.6400</td>
<td>0.6503</td>
<td>0.6679</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6626</td>
<td>0.6831</td>
<td>0.7184</td>
<td>0.7779</td>
</tr>
<tr>
<td>0.192</td>
<td>0.7071</td>
<td>0.7479</td>
<td>0.8176</td>
<td>0.9332</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7595</td>
<td>0.8237</td>
<td>0.9319</td>
<td>1.1064</td>
</tr>
</tbody>
</table>

Table 1.28: Optimal Replacement Rates for $R = 2$ and $u_0 = 0.064$

Optimal $r$ for $G = 0$:  

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.4491</td>
<td>0.4443</td>
<td>0.4362</td>
<td>0.4227</td>
</tr>
<tr>
<td>0</td>
<td>0.4603</td>
<td>0.4603</td>
<td>0.4603</td>
<td>0.4603</td>
</tr>
<tr>
<td>$E_b^v$</td>
<td>0.048</td>
<td>0.4670</td>
<td>0.4698</td>
<td>0.4748</td>
</tr>
<tr>
<td>0.096</td>
<td>0.4736</td>
<td>0.4794</td>
<td>0.4895</td>
<td>0.5069</td>
</tr>
<tr>
<td>0.192</td>
<td>0.4871</td>
<td>0.4989</td>
<td>0.5198</td>
<td>0.5562</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.5034</td>
<td>0.5228</td>
<td>0.5574</td>
<td>0.6195</td>
</tr>
</tbody>
</table>

Optimal $r$ for $G = 0.203$:  

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0.5786</td>
<td>0.5614</td>
<td>0.5222</td>
<td>0.4847</td>
</tr>
<tr>
<td>$E_b^v$</td>
<td>0.048</td>
<td>0.3006</td>
<td>0.2725</td>
<td>0.2578</td>
</tr>
<tr>
<td>0.096</td>
<td>0.3784</td>
<td>0.4743</td>
<td>0.5981</td>
<td>0.6173</td>
</tr>
<tr>
<td>0.192</td>
<td>0.5672</td>
<td>0.7479</td>
<td>0.8176</td>
<td>0.9332</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7170</td>
<td>0.8656</td>
<td>1.0813</td>
<td>1.3827</td>
</tr>
</tbody>
</table>

Table 1.29: Optimal Replacement Rates for $R = 5$ and $u_0 = 0.064$

Optimal $r$ for $G = 0$:  

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.6828</td>
<td>0.6809</td>
<td>0.6777</td>
<td>0.6723</td>
</tr>
<tr>
<td>0</td>
<td>0.6870</td>
<td>0.6870</td>
<td>0.6870</td>
<td>0.6870</td>
</tr>
<tr>
<td>$E_b^v$</td>
<td>0.048</td>
<td>0.6895</td>
<td>0.6906</td>
<td>0.6926</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6920</td>
<td>0.6943</td>
<td>0.6984</td>
<td>0.7053</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6971</td>
<td>0.7018</td>
<td>0.7101</td>
<td>0.7247</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7033</td>
<td>0.7111</td>
<td>0.7249</td>
<td>0.7499</td>
</tr>
</tbody>
</table>

Optimal $r$ for $G = 0.203$:  

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.5554</td>
<td>0.5347</td>
<td>0.5005</td>
<td>0.4154</td>
</tr>
<tr>
<td>0</td>
<td>0.6023</td>
<td>0.6023</td>
<td>0.6023</td>
<td>0.6023</td>
</tr>
<tr>
<td>$E_b^v$</td>
<td>0.048</td>
<td>0.6209</td>
<td>0.6242</td>
<td>0.6639</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6574</td>
<td>0.6824</td>
<td>0.7254</td>
<td>0.7981</td>
</tr>
<tr>
<td>0.192</td>
<td>0.7116</td>
<td>0.7613</td>
<td>0.8462</td>
<td>0.9863</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7753</td>
<td>0.8532</td>
<td>0.9840</td>
<td>1.1924</td>
</tr>
</tbody>
</table>
Table 1.30: Optimal Replacement Rates for $R = 2$ and Multiple Tax Rates

<table>
<thead>
<tr>
<th></th>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.0816$</td>
<td>0.4900</td>
<td>0.4459</td>
<td>0.4390</td>
<td>0.4275</td>
<td></td>
</tr>
<tr>
<td>$0.0000$</td>
<td>0.4595</td>
<td>0.4595</td>
<td>0.4595</td>
<td>0.4595</td>
<td></td>
</tr>
<tr>
<td>$E_b^y$</td>
<td>0.048</td>
<td>0.4651</td>
<td>0.4675</td>
<td>0.4717</td>
<td>0.4789</td>
</tr>
<tr>
<td>$E_b^y$</td>
<td>0.096</td>
<td>0.4707</td>
<td>0.4756</td>
<td>0.4842</td>
<td>0.4987</td>
</tr>
<tr>
<td>$0.192$</td>
<td>0.4821</td>
<td>0.4921</td>
<td>0.5095</td>
<td>0.5399</td>
<td></td>
</tr>
<tr>
<td>$0.3072$</td>
<td>0.4959</td>
<td>0.5122</td>
<td>0.5410</td>
<td>0.5922</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.31: Optimal Replacement Rates for $R = 5$ and Multiple Tax Rates

<table>
<thead>
<tr>
<th></th>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.0816$</td>
<td>0.4500</td>
<td>0.4459</td>
<td>0.4390</td>
<td>0.4275</td>
<td></td>
</tr>
<tr>
<td>$0.0000$</td>
<td>0.4595</td>
<td>0.4595</td>
<td>0.4595</td>
<td>0.4595</td>
<td></td>
</tr>
<tr>
<td>$E_b^y$</td>
<td>0.048</td>
<td>0.4614</td>
<td>0.4675</td>
<td>0.4717</td>
<td>0.4789</td>
</tr>
<tr>
<td>$E_b^y$</td>
<td>0.096</td>
<td>0.4707</td>
<td>0.4756</td>
<td>0.4842</td>
<td>0.4987</td>
</tr>
<tr>
<td>$0.192$</td>
<td>0.4821</td>
<td>0.4921</td>
<td>0.5095</td>
<td>0.5399</td>
<td></td>
</tr>
<tr>
<td>$0.3072$</td>
<td>0.4959</td>
<td>0.5122</td>
<td>0.5410</td>
<td>0.5922</td>
<td></td>
</tr>
</tbody>
</table>

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 $s$.\(^{62}\)

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

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

$$= c_u - v\Delta y$$

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

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

\(^{62}\)As noted by Baily, this can only be an approximation given that $(s - v)$ is constrained to lie in $(0, 1)$. 61
(1.4) and (1.5) now have to be replaced by:

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

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

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

\[
U'(\tilde{c}_u) = U'(c_u) + U''(\gamma)(\tilde{c}_u - c_u)
\]

\[
= U'(c_u) - v\Delta y U''(\gamma)
\]

where \(\gamma\) 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:

\[
\frac{\partial V}{\partial k} = -U'(c^1_e) + (1 - \delta)U'(c^2_e) + \delta U'(c_u) = 0
\]

As before, I make the assumptions that \(y_n = y\) and \(c^1_e U''(\theta) = c_u U''(c_u)\), 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:

\[
\frac{dW}{db} = 2y \frac{\Delta c}{c^1_e} R \frac{d\tau}{db} + \delta \frac{\Delta y}{c^1_e} RV ar(v) + \delta y \frac{\Delta y}{c^1_e} RV ar(v) \frac{d\tau}{db} - 2(1 - u)y \left[ \frac{d\tau}{db} - \omega \right].
\]
The budget constraint takes an expectation over all workers, and so is unchanged; therefore:

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

and the equation for the optimum is:

\[
\frac{\Delta c}{c_e} R + \frac{\Delta y RVar(v)}{c_e} (1-s) = (1-u) \left[ 1 + \frac{\Delta y RVar(v)}{c_e} \right] \left( E_b^r - \psi \right).
\] (1.15)

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 (1.15) 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: \( \frac{\Delta y}{c_e} \) 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 \( Var(v) = \left( \frac{14.2}{18.3} \right)^2 (1-s_0)^2 \). Meanwhile, in the absence of any better evidence, I will use Baily’s assumption that \( \frac{\Delta y}{c_e} = 1-r \). Evaluation of (1.15) then gives the results displayed below in Tables 1.32 and 1.33.

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

\[\text{There are potential offsetting biases in these calculations; Chetty (2008) uses a sample in which the duration of unemployment is truncated at 50 weeks, suggesting I may be underestimating Var(v), but on the other hand, Chetty’s is an unconditional variance, some of which may be explained by individual characteristics, which means Var(v) may be an overestimate.}\]
Table 1.32: Optimal Replacement Rates Calculated from (1.15) for \( R = 2 \)

Optimal \( r \) for \( G = 0 \):

<table>
<thead>
<tr>
<th></th>
<th>( s_0 )</th>
<th>0.148</th>
<th>0.175</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0.569</td>
<td>0.620</td>
<td>0.581</td>
<td>0.6367</td>
</tr>
<tr>
<td></td>
<td>0.0816</td>
<td>0.6629</td>
<td>0.6208</td>
<td>0.5812</td>
<td>0.5367</td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.6662</td>
<td>0.6303</td>
<td>0.5969</td>
<td>0.5630</td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>0.6735</td>
<td>0.6532</td>
<td>0.6351</td>
<td>0.6292</td>
</tr>
<tr>
<td></td>
<td>0.3072</td>
<td>0.6824</td>
<td>0.6673</td>
<td>0.6590</td>
<td>0.6720</td>
</tr>
</tbody>
</table>

Table 1.33: Optimal Replacement Rates Calculated from (1.15) for \( R = 5 \)

Optimal \( r \) for \( G = 0 \):

<table>
<thead>
<tr>
<th></th>
<th>( s_0 )</th>
<th>0.148</th>
<th>0.175</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0.796</td>
<td>0.781</td>
<td>0.756</td>
<td>0.7322</td>
</tr>
<tr>
<td></td>
<td>0.0816</td>
<td>0.8002</td>
<td>0.7843</td>
<td>0.7665</td>
<td>0.7499</td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.8017</td>
<td>0.7867</td>
<td>0.7705</td>
<td>0.7567</td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>0.8048</td>
<td>0.7916</td>
<td>0.7786</td>
<td>0.7708</td>
</tr>
<tr>
<td></td>
<td>0.3072</td>
<td>0.8085</td>
<td>0.7975</td>
<td>0.7887</td>
<td>0.7887</td>
</tr>
</tbody>
</table>

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 \( E_y^b \), 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 1.7 or 1.8, the same is true in Table 1.32 or 1.33, and vice-versa.

### 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.\(^{64}\) I also assume that, if a worker loses their job, any savings from the first period are completely consumed while unemployed; none of those savings are kept for consumption when re-employed.\(^{65}\) I can therefore write total utility as:

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

(1.4) still holds, and (1.5) is now replaced by:

\[\frac{\partial V}{\partial \tau} = -yU'(c_1^e) - (1 - \delta)yU'(c_2^e) - \delta s y_n U'(c_n).\]

I replace \(U'(c_2^e)\) using the first-order condition for saving, as before, and I assume that \(c_n = c_1^e\), 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 \(U'(c_1^e)\):

\[\frac{\partial V}{\partial \tau} = -[(2 - \delta)y + \delta sy_n]U'(c_u) - [2y + \delta sy_n] \Delta cU''(\theta).\]

Putting this together with (1.4):

\[\frac{dW}{db} = [2 + \delta s]y \frac{\Delta c}{c_1^e} R \frac{d\tau}{db} - 2(1 - u)y \left( \frac{d\tau}{db} - \omega \right)\]

\(^{64}\)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.

\(^{65}\)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.
and therefore the equation for the optimum is:

\[ \frac{\Delta c}{c_{1e}^t} R = \frac{2(1 - u) E^r_b}{2 + \delta s} - \psi. \] (1.16)

\( E^r_b \) is the same as before, so this equation is almost identical to (1.10), and it is easy to introduce the extra \( \delta s \) term into the calculations. Solving for the optimal replacement rate generates the results found in Tables 1.34 and 1.35.

**Table 1.34:** Optimal Replacement Rates Calculated from (1.16) for \( R = 2 \)

<table>
<thead>
<tr>
<th>( s_0 )</th>
<th>Optimal ( r ) for ( G = 0 ):</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.648</td>
<td>0.725</td>
</tr>
<tr>
<td>-0.0816</td>
<td>0.4850</td>
</tr>
<tr>
<td>0</td>
<td>0.4934</td>
</tr>
<tr>
<td>0.048</td>
<td>0.4985</td>
</tr>
<tr>
<td>0.096</td>
<td>0.5035</td>
</tr>
<tr>
<td>0.192</td>
<td>0.5137</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.5260</td>
</tr>
</tbody>
</table>

**Table 1.35:** Optimal Replacement Rates Calculated from (1.16) for \( R = 5 \)

<table>
<thead>
<tr>
<th>( s_0 )</th>
<th>Optimal ( r ) for ( G = 0 ):</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.648</td>
<td>0.725</td>
</tr>
<tr>
<td>-0.0816</td>
<td>0.6963</td>
</tr>
<tr>
<td>0</td>
<td>0.6995</td>
</tr>
<tr>
<td>0.048</td>
<td>0.7014</td>
</tr>
<tr>
<td>0.096</td>
<td>0.7034</td>
</tr>
<tr>
<td>0.192</td>
<td>0.7073</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7120</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( s_0 )</th>
<th>Optimal ( r ) for ( G = 0.208 ):</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.648</td>
<td>0.725</td>
</tr>
<tr>
<td>-0.0816</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.048</td>
<td>0.2384</td>
</tr>
<tr>
<td>0.096</td>
<td>0.4156</td>
</tr>
<tr>
<td>0.192</td>
<td>0.5821</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7216</td>
</tr>
</tbody>
</table>

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 \( E^y_b \) 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 \( \Delta c R < \frac{2(1-u)}{2+\delta s} \). However, aside from \( \Delta c R > \frac{2(1-u)}{2+\delta s} \) implies \( \frac{\partial V}{\partial \tau} > 0 \), while I also estimate that \( \frac{d\tau}{db} < 0 \).
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 $E_b^w$, 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 $\Delta c$ 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.

**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. However, I will now explore how the results change when I use a second-order Taylor series expansion of marginal utility.

---

67The first-order Taylor series used in my paper is in fact an exact equality, not an approximation; it is the assumption that $c_1^e U''(\theta) = c_u U''(c_u)$ 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 $E_b^w$ in the denominator of the right-hand side of (1.10) is equal to one is, in the context of his model, a significant reason for underestimation of the optimal $b$. 

---
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 (1.4) and (1.5), and the standard first-order condition for saving. Next, I use a new Taylor series expansion of \(U'(c_u)\) around \(U'(c_e^1)\):

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

where \(\theta\) is not necessarily the same value as before, but is still between \(c_e^1\) and \(c_u\). Using this to replace \(U'(c_u)\) in both (1.4) and (1.5):

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

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:

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

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

\[
\left( \frac{\Delta c}{c_e^1} R + \frac{1}{2} \left( \frac{\Delta c}{c_e^1} \right)^2 RP \right) = (1 - u) \left( 1 + \frac{\Delta c}{c_e^1} R + \frac{1}{2} \left( \frac{\Delta c}{c_e^1} \right)^2 RP \right) \frac{E_{b}^r - \psi}{E_{b}^r} \quad (1.17)
\]

where \(E_b^r\) is unchanged thus and still given by (1.12).

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. I will therefore use a value of $P = 2$, and the results from evaluation of the optimal replacement rate are displayed in Tables 1.36 and 1.37.

Table 1.36: Optimal Replacement Rates Calculated from (1.17) for $R = 2$

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.4000</td>
<td>0.3947</td>
<td>0.3858</td>
<td>0.3707</td>
</tr>
<tr>
<td>0</td>
<td>0.4124</td>
<td>0.4124</td>
<td>0.4124</td>
<td>0.4124</td>
</tr>
<tr>
<td>$E_y^p$</td>
<td>0.048</td>
<td>0.4196</td>
<td>0.4227</td>
<td>0.4281</td>
</tr>
<tr>
<td>0.096</td>
<td>0.4269</td>
<td>0.4331</td>
<td>0.4438</td>
<td>0.4620</td>
</tr>
<tr>
<td>0.192</td>
<td>0.4414</td>
<td>0.4539</td>
<td>0.4755</td>
<td>0.5124</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.4588</td>
<td>0.4788</td>
<td>0.5138</td>
<td>0.5741</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$E_y^p$</td>
<td>0.048</td>
<td>0.0238</td>
<td>0.0963</td>
<td>0.3317</td>
</tr>
<tr>
<td>0.096</td>
<td>0.3073</td>
<td>0.4384</td>
<td>0.5843</td>
<td>0.7581</td>
</tr>
<tr>
<td>0.192</td>
<td>0.5502</td>
<td>0.6794</td>
<td>0.8513</td>
<td>1.0768</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.7128</td>
<td>0.8637</td>
<td>1.0720</td>
<td>1.3510</td>
</tr>
</tbody>
</table>

Table 1.37: Optimal Replacement Rates Calculated from (1.17) for $R = 5$

<table>
<thead>
<tr>
<th>$s_0$</th>
<th>0.648</th>
<th>0.725</th>
<th>0.8</th>
<th>0.863</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.0816</td>
<td>0.6522</td>
<td>0.6499</td>
<td>0.6549</td>
<td>0.6392</td>
</tr>
<tr>
<td>0</td>
<td>0.6574</td>
<td>0.6574</td>
<td>0.6574</td>
<td>0.6574</td>
</tr>
<tr>
<td>$E_y^p$</td>
<td>0.048</td>
<td>0.6604</td>
<td>0.6618</td>
<td>0.6641</td>
</tr>
<tr>
<td>0.096</td>
<td>0.6634</td>
<td>0.6662</td>
<td>0.6709</td>
<td>0.6790</td>
</tr>
<tr>
<td>0.192</td>
<td>0.6695</td>
<td>0.6750</td>
<td>0.6846</td>
<td>0.7008</td>
</tr>
<tr>
<td>0.3072</td>
<td>0.6768</td>
<td>0.6857</td>
<td>0.7011</td>
<td>0.7274</td>
</tr>
</tbody>
</table>

The results this time are somewhat different quantitatively, in that optimal replacement rates are zero for low values of $E_y^p$ 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 $E_y^p$, the fiscal externality term considerably reduces the optimal replacement rate, whereas at higher values it considerably increases it.

---

68Eisenhauer 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.
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 as \( h(y) \), is separable from consumption, (1.1) 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 (1.4) and (1.5) are unchanged, both (1.7) and (1.9) remain valid; the only change is to the derivative of the government budget constraint. The latter now becomes:

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

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

\[
E^\tau_b = \psi + \psi E^u_b + \frac{u}{1 - u} E^D_b - \frac{\delta s}{2(1 - u)} E^y_b - \frac{2 - \delta}{2(1 - u)} \varepsilon^y_b
\]

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

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

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

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

Finally, I use the baseline tax rate of $\tau = 0.23$, so my estimate of the elasticity is $\varepsilon^y_{b} = -0.092 \frac{0.77}{\psi}$; 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 1.38 and 1.39.

Table 1.38: Optimal Replacement Rates Calculated from (1.10) and (1.18) for $R = 2$

<table>
<thead>
<tr>
<th>Optimal $r$ for $G = 0$:</th>
<th>Optimal $r$ for $G = 0.208$:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_0$ 0.648 0.725 0.8 0.863</td>
<td>$s_0$ 0.648 0.725 0.8 0.863</td>
</tr>
<tr>
<td>0.0816 0.3412 0.3125 0.3456 0.3503</td>
<td>0.0816 0.0270 0.1511 0.3125 0.4579</td>
</tr>
<tr>
<td>0 0.3496 0.3548 0.3638 0.3791</td>
<td>0 0.2854 0.3886 0.5185 0.6897</td>
</tr>
<tr>
<td>$E^y_{b}$ 0.048 0.3546 0.3619 0.3746 0.3966</td>
<td>$E^y_{b}$ 0.048 0.2070 0.1511 0.3125 0.4579</td>
</tr>
<tr>
<td>0.096 0.3596 0.3691 0.3856 0.4144</td>
<td>0.096 0.2854 0.3886 0.5185 0.6897</td>
</tr>
<tr>
<td>0.048 0.3596 0.3619 0.3746 0.3966</td>
<td>0.096 0.2854 0.3886 0.5185 0.6897</td>
</tr>
<tr>
<td>0.192 0.3697 0.3835 0.4080 0.4513</td>
<td>0.192 0.4809 0.5989 0.7670 1.0039</td>
</tr>
<tr>
<td>0.3072 0.3818 0.4012 0.4356 0.4978</td>
<td>0.3072 0.6279 0.7733 0.9855 1.2867</td>
</tr>
</tbody>
</table>

The inclusion of the labour supply elasticity tends to lower the optimal replacement rate, though in many cases not by much; in the $R = 5$ case, the effect is often almost negligible, whereas in the $R = 2$ case the effect can be somewhat larger for moderate values of $E^y_{b}$.

69 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 1.39: Optimal Replacement Rates Calculated from (1.10) and (1.18) for $R = 5$

<table>
<thead>
<tr>
<th></th>
<th>$s_0 = 0.648$</th>
<th>$s_0 = 0.725$</th>
<th>$r_0 = 0.8$</th>
<th>$r_0 = 0.863$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>0.648</td>
<td>0.6406</td>
<td>0.6417</td>
<td>0.6435</td>
</tr>
<tr>
<td>$s_0$</td>
<td>0.6430</td>
<td>0.6450</td>
<td>0.6486</td>
<td>0.6547</td>
</tr>
<tr>
<td>$E_y b_0$</td>
<td>0.048</td>
<td>0.6448</td>
<td>0.6477</td>
<td>0.6527</td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.6504</td>
<td>0.6569</td>
<td>0.6615</td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>0.6558</td>
<td>0.6656</td>
<td>0.6828</td>
</tr>
<tr>
<td></td>
<td>0.3072</td>
<td>0.6625</td>
<td>0.6763</td>
<td>0.7011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$s_0 = 0.648$</th>
<th>$s_0 = 0.725$</th>
<th>$r_0 = 0.8$</th>
<th>$r_0 = 0.863$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>-0.0816</td>
<td>0.5198</td>
<td>0.6209</td>
<td>0.4719</td>
</tr>
<tr>
<td>$s_0$</td>
<td>0</td>
<td>0.5639</td>
<td>0.5656</td>
<td>0.5685</td>
</tr>
<tr>
<td>$E_y b_0$</td>
<td>0.048</td>
<td>0.5898</td>
<td>0.6033</td>
<td>0.6266</td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.6156</td>
<td>0.6408</td>
<td>0.6846</td>
</tr>
<tr>
<td></td>
<td>0.192</td>
<td>0.6666</td>
<td>0.7149</td>
<td>0.7983</td>
</tr>
<tr>
<td></td>
<td>0.3072</td>
<td>0.7263</td>
<td>0.8012</td>
<td>0.9282</td>
</tr>
</tbody>
</table>

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.

**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.
Chapter 2

Fiscal Externalities, Liquidity Constraints and Grants to Post-Secondary Students

2.1 Introduction

Post-secondary education (PSE) has been the focus of a considerable amount of research by economists, especially in recent decades in North America as tuition and fee increases have motivated increased discussion and concern about the affordability of post-secondary study and appropriate government policy in this area; see, for example, the surveys in Kane (2006) and McPherson and Schapiro (2006).

In particular, a significant economic literature studies the existence and magnitude of student borrowing constraints, with varying conclusions as illustrated by Carneiro and Heckman (2002) and Belley and Lochner (2007). At the same time, a growing literature seeks to provide insights into how education subsidy policy can be used to offset tax distortions affecting individuals' education choices, a phenomenon which may be called a fiscal externality (see, for example, Bovenberg and Jacobs (2005)). Liquidity constraints and fiscal externalities are likely the most frequently discussed justifications for government financial aid to post-
secondary students; Bovenberg and Jacobs (2005), for instance, cite fiscal externalities and capital market imperfections (along with spillovers and redistributive motives) as two of the most prominent explanations for education subsidies.

However, in these literatures, one important practical question has received very little attention: in numerical terms, what is the optimal tuition subsidy policy? Many papers focus on documenting the size of fiscal externalities and borrowing constraints (see, for example, Trostel (1993) and Cameron and Taber (2004) respectively), or on theoretical analysis, as in Bovenberg and Jacobs (2005). However, a very small literature has attempted to provide numerical answers: Trostel (1996) and Akyol and Athreya (2005) find optimal subsidies of around 40% of the resources invested in education, while Bohacek and Kapicka (2008) find lower optimal subsidies that vary from 0 to 20%. All three papers focus only on fiscal externalities, ignoring the issue of liquidity constraints, and their calculations require further assumptions to translate the results into dollar amounts. Additionally, each paper uses a parametrization of a complex structural models, leaving open the question of how well their preferred parameters fit important moments of data from the education sector.

In this paper, I will provide a new answer to this practical question. I present a simple model of PSE in which both liquidity constraints and fiscal externalities are accounted for, and use the “sufficient statistics” method described by Chetty (2009) to solve the model for an equation for the marginal welfare gain from increasing grants to PSE students; this method involves writing the welfare derivative in terms of observable empirical quantities, generating results that are robust to different combinations of the underlying structural parameters. I then evaluate this welfare derivative and determine whether government grants should be

---

1 Bohacek and Kapicka (2008), however, ignore direct costs of post-secondary education.
more or less generous. Finally, I use two methods, statistical extrapolations of my welfare equation and a calibration and simulation of my simple model, to estimate the optimal dollar amount of student grants.

My main conclusion is that financial aid should be more generous, and my preferred estimate of the optimal policy corresponds to eliminating tuition at the typical public university.² My results are robust to a variety of alterations and extensions, and are largely unchanged if I assume that students face no liquidity constraints; however, the results are sensitive to the nature of general equilibrium effects of PSE graduate supply on relative and absolute wages.

The primary contribution of my paper is to provide credible numerical estimates of the welfare effects of tuition policy which can inform policy discussion. As well as providing an answer to the simple question of optimal tuition subsidy policy, my analysis helps us to understand what we do and don’t know, and in which areas we need to know more, by indicating whether the results are sensitive to particular factors. The finding that liquidity constraints do not greatly affect my results suggests that the empirical literature seeking evidence for or against such constraints may be of limited policy relevance. On the other hand, it is clear that we need to know more about the size and nature of general equilibrium effects. I also argue that further structural work that takes into account multiple dimensions of heterogeneity would be valuable, as it would provide answers to alternative policy questions about targeting financial aid at marginal or high-return students.

An additional contribution is to show how the sufficient statistics method can be adapted

²This result is similar to the finding in Saez (2002) that an Earned Income Tax Credit is optimal when low-income behavioural responses to taxation are concentrated on the extensive margin: the decision to attend PSE is an important extensive margin, and thus a large transfer to individuals who undertake that action may be efficient.
to the context of PSE, and to use this method to provide results which are robust to much of the variety of assumptions and modelling strategies found in the literature. My paper represents the first attempt that I am aware of to perform sufficient statistics welfare analysis of PSE.

The rest of the paper is organized as follows. Section 2.2 provides a brief analysis of how each of the liquidity constraint and fiscal externality motivations may justify government funding of students, and discusses the relevant literatures. Section 2.3 lays out my simple model, and solves it for a sufficient statistics expression for the derivative of social welfare with respect to student grants. Section 2.4 provides the main numerical results. Section 2.5 then performs the experiment of shutting down the liquidity constraint and fiscal externality motivations for financial aid one by one, to examine the robustness of the results to such modifications. Sections 2.6 and 2.7 extend the model to include heterogeneity and general equilibrium effects. Section 2.8 provides a conclusion.

2.2 The Role of Liquidity Constraints and Fiscal Externalities in Post-Secondary Education

Significant literatures study the existence and magnitude of fiscal externalities and liquidity constraints in post-secondary education. I therefore begin by discussing these literatures and their main findings and limitations, to provide motivation for the analysis that follows. Subsequently, I will explain how I combine the insights from these literatures into a novel framework for the analysis of tuition subsidy policy.
2.2.1 Fiscal Externalities

The term “fiscal externality” has been used to describe a variety of concepts in different literatures, but in this paper I use it to refer to effects of government programs upon labour market outcomes, which in turn affect government tax revenues and expenditures, influencing the government’s ability to fund its entire range of programs. In this way, one program has an “external” effect on the funding for other programs: if a program affects individuals’ decisions in such a way as to increase their total income, for example, this increases income tax revenues and allows for more spending in other areas, or lower tax rates, with added benefits to society.³

Alternatively, I can describe the phenomenon at the individual level: when choices of an individual affect their own taxable income, and thereby the tax revenues collected from that worker, those choices have external effects on other people through the tax system, and the magnitude of those effects are larger when government expenditures (and therefore taxes) are large. In the context of PSE, obtaining education is believed to augment workers’ skills and increase their future earnings, so the decision to attend university or college⁴ provides external benefits to all workers through an increase in future tax revenues. Programs that influence such individual choices, therefore, interact with this already-existing externality and affect the tax revenues available to society as a whole.

de Bartolome (1999) points out that a fiscal externality also indicates a distortion to

³ To the extent that the revenue increase is experienced by the same level of government that implements the program in question, the effect is not external in the usual sense. However, if different levels of government are involved, this phenomenon becomes very similar to that described in the tax competition literature, as summarized by Wilson (1999).

⁴ My analysis focusses primarily on 4-year university attendance, but I will use “PSE” throughout this paper to refer to post-secondary education, rather than “university” or “college,” as the latter term in particular has different meanings in different countries.
individual actions. That is, if an individual's action increases tax revenues, this implies that the individual isn't receiving the full returns to that action, indicating a distortion to incentives and an insufficient amount of that action being performed. This is the context in which fiscal externalities are usually described when it comes to education policy: the pre-existing tax distortion on the education decision can be offset by a tuition subsidy which encourages students to attend higher education and increase their skills and productivity, thus increasing efficiency. A fiscal externality, therefore, is also an application of the Theory of the Second Best: a pre-existing inefficiency in the form of some exogenously required government spending and a distortionary tax alters the efficient policy in another area.\footnote{It is necessary to assume the necessity both of this government expenditure and of a distortionary tax system; in other words, I require an “irreducible distortion,” or a distortion that is not itself generated by policy. Otherwise, the advice of Browning (1999) to remove the initial policy generating the distortion and thereby to obtain the first-best would be valid. The generality of Browning (1999)'s analysis is limited by the fact that it is based on a special case in which the policy generating the distortion and the policy instrument which could be used to offset it are identical.}

The idea of fiscal benefits from education is not a new one, and economists have been explicitly stating it as a motivation for subsidizing education at least as far back as Singer (1972), who discusses how education can generate external benefits through the taxes paid on increased income and a reduction in expenditures on the social safety net. Simulations in Trostel (1993) show that proportional income taxation could have a significant negative effect on investment in human capital. Subsequently, Trostel (2010) quantifies the fiscal benefits of PSE to government, estimating that net government spending per university degree is negative in the United States,\footnote{Similar findings for high school graduation are discussed in a New York Times editorial, Levin and Rouse (January 25, 2012), which points out that reducing undereducation in that context would pay for itself. Damon and Glewwe (2011) also undertake a related examination of the costs and benefits of state financing of public universities in Minnesota, finding that the fiscal benefits outweigh the costs; however, they ignore federal taxes and conclude that a large portion of the gains at the state level come from wage spillovers from educated workers, which I will consider in section 2.7.2.} as direct expenditures of about $71000 (in present-
value 2005 dollars) are more than offset by expenditure savings of $56000 and increased tax revenues amounting to $197000.\textsuperscript{7}

The use of public provision of education for fiscal reasons has been studied by numerous papers: after Guesnerie and Roberts (1984) showed that quantity controls can be welfare-enhancing in a Second Best world, Del Rey (2001) and Greco (2011) apply this insight to education, showing that if taxes distort the quantity of education downwards, public provision can reduce these tax-induced inefficiencies. Alternatively, the use of publicly-provided private goods such as education to weaken self-selection constraints in redistribution is studied by Boadway and Marchand (1995), among others.

A large and mostly theoretical literature, meanwhile, studies policies of the type I will be focussing on: financial incentives for post-secondary education. Bovenberg and Jacobs (2005) present a model in which it is optimal to use a subsidy to education to perfectly offset income taxes, returning human capital investment to the first-best amount. They describe this result as an extension of the production efficiency theorem of Diamond and Mirrlees. Numerous papers then follow which qualify this finding by seeking conditions under which education should be effectively subsidized or taxed, i.e. under which the quantity of education should be induced to move above or below the first-best amount. Richter (2009), Richter and Braun (2010), and Braun (2010) all find that education should be subsidized beyond the

\textsuperscript{7}Therefore, from the government’s perspective, a university degree is a money-generating machine, as each dollar spent at present is more than recovered later. However, this does not mean that a tuition subsidy is necessarily a money-generating machine; that is, it is not necessarily the case that an increase in subsidies would not require a tax increase and would therefore be “free” in fiscal terms. Increasing student grants does encourage more students to attend university, but it also requires increased payments to inframarginal students, so what matters is how sensitive enrollment is to tuition. In my baseline numerical case, I do find that a marginal increase in tuition subsidies from the current level is “free,” but as grants increase, this quickly ceases to be true. This means that taxes eventually have to increase, and since I assume a simple proportional tax, this will tend to generate negative redistributional consequences, making it even more striking that a utilitarian social planner would find significant welfare gains from raising grants.
first-best if the human capital accumulation function has an increasing elasticity with respect to education, which Braun (2010) argues is likely. Finally, the way in which the returns to education interact with other characteristics and decisions is analyzed by Maldonado (2008) and Jacobs and Bovenberg (2011). Both find that education should be effectively taxed if it is complementary with ability, while the latter paper also shows that education should be subsidized relative to first-best if it is complementary with labour effort, leaving as an empirical matter which of these effects dominates.

Only a few papers, however, attempt any numerical evaluation: Trostel (1996), Akyol and Athreya (2005), and Bohacek and Kapicka (2008) calculate an optimal percentage subsidy on investments in education of around 40% or lower, using approaches that involve parametrizing a complex structural model of the education decision; in each case, they also assume away borrowing constraints. There is, therefore, no prior work that calculates optimal dollar values of tuition subsidies using a sufficient statistics approach which is robust to much of the variety in assumptions and modelling strategies in the literature.

### 2.2.2 Liquidity Constraints

The possibility that borrowing constraints may prevent some young people from making efficient choices about their education is one of the most frequently mentioned justifications for government financial aid to students, especially students from low-income families; see, for example, Kane (1999). There is also a considerable literature studying the presence and magnitude of these constraints, often consisting of papers that look for causal effects of family income on PSE enrollment, on the grounds that students from higher-income families
will be less likely to face constraints. Some studies have also looked for effects of income on the frequency of delayed or part-time enrollment, as well as the relative sensitivity of enrollment to tuition versus expected returns to education.

However, the existence of liquidity constraints among students remains the subject of a persistent empirical controversy, as a number of papers have argued that constraints among PSE students are small or negligible. Cameron and Taber (2004) and Shea (2000), for example, find little or no effect of income on enrollment, and a series of papers by James Heckman and various co-authors, as summarized in Cunha, Heckman, Lochner, and Masterov (2006), find that the income-enrollment relationship in the 1979 NLSY is close to zero after controlling for various measures of skill and family background. They argue that this means that the income gap in enrollment is primarily caused by long-run family factors and not short-run liquidity constraints. Kane (2006), however, points out that this evidence cannot rule out liquidity constraints, because current income is not a perfect proxy for the availability of financial resources for education, and measures of family background may also serve as (imperfect) proxies for family resources. Therefore, a distinction between “long-run” and “short-run” factors in the data may not be particularly meaningful.

Other papers, meanwhile, do claim to find positive effects of income on enrollment,

---

8 However, Brown, Scholz, and Seshadri (2009) argue that by modelling the financial relationship between child and parents, it can be shown that constrained students need not necessarily be poor, and they provide suggestive evidence along these lines.

9 Delayed and part-time enrollment are inefficient in simple models of education choice, so although imperfect information about earning capabilities could be responsible, they are sometimes taken as evidence of constraints on education choices; see, for example, Ellwood and Kane (2000) and Kane (1996). Meanwhile, Ellwood and Kane (2000) and Kane (1999) cite greater sensitivity of enrollment to tuition than to measures of future (or expected) returns to education as suggestive evidence of liquidity constraints, although uncertainty about future earnings and returns to education could also be responsible.

10 Shea (2000) does find that income has a significant effect for children from the poverty subsample of the PSID and for children whose fathers have less than 12 years of schooling.
including Acemoglu and Pischke (2001) and Coelli (2011). Belley and Lochner (2007) repeat the analysis of Heckman and co-authors on data from the 1997 NLSY, and show that income has become a much more important determinant of enrollment since the 1979 NLSY.

Furthermore, any empirical analysis is complicated by the fact that an absence of income effects on enrollment is not a sufficient condition for a lack of liquidity constraints, as it is possible for students’ consumption or other choices (such as working while in school) to be constrained even if the enrollment decision is not significantly affected, as pointed out by Belley and Lochner (2007) and Lochner and Monge-Naranjo (2008). Additionally, Belley and Lochner (2007) find that an unconstrained student population should exhibit a negative causal effect of income on enrollment, a theoretical result which also holds in my analysis, providing further evidence against the idea that liquidity constraints imply positive effects of income on enrollment.

Therefore, my interpretation of the literature is that, although there is suggestive evidence on both sides, it is difficult to find a definitive test for liquidity constraints among PSE students, and so the empirical controversy is likely to persist. I will therefore consider a range of possibilities in my numerical analysis, and my results may shed some light on the relevance of this issue for optimal financial aid policy.

Discussions of optimal policy have been limited in papers that discuss liquidity constraints; papers that claim to find no evidence of constraints, such as Cameron and Taber (2004), are generally dismissive of the scope for welfare gains from financial aid, while else-

---

11 In fact, Coelli (2011) finds that parental job loss leads to a significant reduction in enrollment, and cannot conclusively disentangle income loss from other consequences of job loss.

12 Policy simulations in Keane and Wolpin (2001) predict exactly this type of response.

13 Cameron and Taber (2004) simply state that “additional subsidies aimed at improving credit access will have little net value and little impact on overall schooling attainment.”
where it is commonly argued that guaranteed loans can solve whatever liquidity constraints may exist.\textsuperscript{14} However, several papers that study structural models of the education decision, notably Keane and Wolpin (2001) and Johnson (2012), argue that raising borrowing limits in guaranteed loan programs will have very little effect on enrollment or graduation, because of a precautionary savings motive;\textsuperscript{15} students may not want to build up excessive amounts of debt early in life for fear of a negative income shock shortly after entering the labour market.\textsuperscript{16} Therefore, I focus my analysis on grants to students, as Johnson (2012) specifically argues that tuition subsidies will be much more effective than loans in raising university completion. I abstract away from changes in guaranteed loan policy, though I do implicitly model guaranteed loans since I allow borrowing limits to be determined by data on the magnitude of liquidity constraints, so if guaranteed loans are effective in reducing constraints, the borrowing limits I calculate should be quite loose.

\textbf{2.2.3 My Analysis of Grants to Students}

My discussion of the literatures on fiscal externalities and liquidity constraints in PSE has shown that, although a great deal of effort has been put into studying the existence and theoretical properties of these phenomena, very little work has been done on their practical policy implications. Therefore, the main contribution of this paper will be to calculate optimal tuition subsidy policy, taking into account both fiscal externalities and liquidity

\textsuperscript{14}Bovenberg and Jacobs (2005) argue that “loans rather than subsidies are the most direct way to address liquidity constraints.”

\textsuperscript{15}Debt-aversion could also have a similar effect. The baseline version of my model does not include any uncertainty and therefore cannot be used to model a precautionary savings motive, though an extension in appendix 2.9.3 does consider uncertainty about post-first-period earnings.

\textsuperscript{16}Upon entering the job market, the individual would no longer qualify for new low-interest loans, making a negative shock especially costly. Rothstein and Rouse (2011) show that student loans can affect early career choices in a way that is consistent with credit constraints.
To do so, I will use a simple and intuitive model of PSE that allows me to highlight the essential tradeoffs of financial aid policy. My analysis represents the first application of the “sufficient statistics” method described by Chetty (2009) to the context of post-secondary education; I use this method to derive an expression for the derivative of social welfare with respect to student grants. This approach uses analysis of a model to identify a set of empirical statistics which are sufficient for welfare analysis of the policy in question,\(^{17}\) and the main benefit of this approach is that the underlying primitive parameters and functional forms do not need to be specified; empirical measurement of the sufficient statistics is all that is needed to make welfare predictions. However, there are also costs: the welfare derivative, evaluated using current estimates of a set of sufficient statistics, is only valid locally. In order to make out-of-sample predictions and solve for the optimal policy, some variety of extrapolation is required, and Chetty (2009) suggests two options: statistical extrapolation of the sufficient statistics, and using the statistics to calibrate and simulate a structural model. I will implement both approaches, allowing me to demonstrate the robustness of my results to alternative assumptions.

The model and analysis is deliberately kept simple; rather than calibrating or estimating a structural model with multiple dimensions of heterogeneity, I focus on a basic model as a first-order approximation to reality, emphasizing the most important features.\(^{18}\) However, my results hold in a far more general analysis, as demonstrated by the analysis in the third

\(^{17}\)As illustrated by Shimer and Werning (2007) and Chetty (2008) in the area of unemployment insurance, the set of sufficient statistics is not necessarily unique.

\(^{18}\)A more comprehensive structural analysis would be complementary to the current analysis, in that it would provide answers to alternative questions about how to target financial aid at heterogeneous populations, but without the transparency of the current analysis.
chapter of this dissertation, where I produce an expression that simplifies to exactly the same form. My analysis is in keeping with the request in Lipsey (2007) for Second-Best policy analysis that aims for “piecemeal improvements in context-specific situations,” by providing clear numerical results from an intuitive model of PSE.

In discussing fiscal externalities and liquidity constraints, I abstract away from non-monetary motivations for government support of students. My intention is to be both simple and conservative, and so I ignore other potential positive externalities from PSE, such as social benefits from better-educated citizens, as mentioned by Kane (1999) and McPherson and Schapiro (2006) and discussed in detail by Lochner (2011).

### 2.3 A Simple Model of University Education

In this section, I will present my model of post-secondary education, followed by the calculations leading to an expression for the derivative of social welfare with respect to student grants.

#### 2.3.1 Model Setup

Time is finite, and consists of twelve periods, each corresponding to 4 years, representing a normal working life of 48 years (say, from age 18 to 65 inclusive).\(^\text{19}\) In the first period, each individual has a choice of attending university or working at wage \(Y_{01}\), and this choice is represented by \(s = \{0, 1\}\), where 1 indicates attendance. In periods \(t = 2, \ldots, 12\), the individual works at a wage \(Y_{st}\) that depends on the education choice in the first period,

---

\(^{19}\)This is equivalent to the usual two-period model of education, but in such a model the periods would be of drastically different lengths, making discounting and the comparison of incomes across the two periods more complicated. Since the income gains from PSE are central to my analysis, I consider that the analytical simplicity of a 2-period approach is offset by greater ease of interpretation and implementation with 12 periods.
where $Y_{1t} > Y_{0t}$. I therefore model only the attainment of a 4-year university degree, and I assume a single jurisdiction, with a single government planner, abstracting from issues of coordination of policy in a federation and from mobility of students across countries.

I assume that the real interest and discount rates are both equal to 3% per year, and since a period is equal to 4 years, I will use $r = 0.12$ for the interest rate and $\beta = \frac{1}{1.12} \approx 0.893$ for the discount rate. I also allow for real wage growth, calculated from the average net compensation series used by the Social Security Administration for the computation of the national average wage index, deflated using the CPI; the average real growth rate over 1991-2008 is almost exactly 1%, so I allow wages to grow at $g = 0.04$ per period.

The individual’s utility from consumption $c$ while in university is $u(c)$, whereas it is $v(c)$ while employed, allowing for direct utility or disutility from university attendance as well as different utility from consumption in the two states. Both utility functions obey the usual properties of $u'$, $v' > 0$ and $u''$, $v'' < 0$, and I denote individuals by $i$. If an individual chooses not to attend university, then since the interest and discount rates are equal, they will simply set consumption to a constant value $c_{0i}$ in each period, and receive lifetime utility of $U_{0i} = \sum_{t=1}^{12} \beta^{t-1} v(c_{0i})$. If they do attend university, they will set per-period post-schooling consumption $c_{1i}$ to some constant value, and choose some value $c_{ui}$ of consumption while in school, receiving lifetime utility of $U_{1i} + \eta_i$, where $U_{1i} = u(c_{ui}) + \sum_{t=2}^{12} \beta^{t-1} v(c_{1i})$ and where $\eta_i$ captures any idiosyncratic portion of the utility or disutility from schooling. I therefore consider a “representative-agent” setting in which there is heterogeneity in taste.

---

20 These wages are assumed to be exogenous, and there is no uncertainty. I allow for uncertainty in future incomes in an extension in appendix 2.9.3.

21 I thereby ignore questions of mismatch, or situations in which particular majors or subjects of study may have different returns and individuals may make socially suboptimal choices of major. This could be an important dimension of heterogeneity to consider in future work.

22 These issues are studied by Del Rey (2001) and Poutvaara (2008), among others.
for schooling but not income or returns to education; heterogeneity in the latter dimension will be considered in section 2.6.2.

The individual’s budget constraints (one for each value of \( s \)) can be written in the following way:

\[
\sum_{t=1}^{12} \left( \frac{1}{1 + r} \right)^{t-1} c_{t\|i}^0 = (1 - \tau) \sum_{t=1}^{12} \left( \frac{1}{1 + r} \right)^{t-1} Y_{0t}
\]

\[
c_{ui} + \sum_{t=2}^{12} \left( \frac{1}{1 + r} \right)^{t-1} c_{t\|i}^1 = (b - e) + (1 - \tau) \sum_{t=2}^{12} \left( \frac{1}{1 + r} \right)^{t-1} Y_{1t}
\]

where \( e \) is the direct cost of university to the individual, \( \tau \) is the marginal tax rate, and \( b \) is the government grant given to students. For simplicity, I restrict attention to a lump-sum grant for now, though I will consider a 2-tier grant scheme in section 2.6.1. To simplify the notation, let me define \( R_x = \sum_{t=x}^{12} \left( \frac{1}{1 + r} \right)^{t-1} \) and \( \gamma_x = \sum_{t=x}^{12} \left( \frac{1 + \theta}{1 + r} \right)^{t-1} \); then the budget constraints can be written as:

\[
R_1 c_{0\|i}^0 = (1 - \tau) \gamma_1 Y_{01}
\]

\[
c_{ui} + R_2 c_{1\|i}^1 = (b - e) + (1 - \tau) \gamma_2 Y_{11}.
\]

I also allow students to face a liquidity constraint, which will take the form of a limit \( A_i \) to...
the debt that a student may accumulate:\footnote{26} \[ c_{ui} - (b - e) \leq A_i. \]

Therefore, the individual’s maximization problem is to choose \( \{s_i, c_{vi}^0, c_{vi}^1, c_{ui}\} \) to maximize \( V_i = s_i(U_{1i} + \eta_i) + (1 - s_i)U_{0i}: \]

\[ V_i = s_i[u(c_{ui}) + R_2 v(c_{vi}^1) + \eta_i - \lambda_{1i}(c_{ui} + R_2 c_{vi}^1 - (b - e) - (1 - \tau)\gamma_2 Y_{11}) - \mu_i(c_{ui} - (b - e) - A_i)] + (1 - s_i)[R_1 v(c_{vi}^0) - \lambda_{0i}(R_1 c_{vi}^0 - (1 - \tau)\gamma_1 Y_{01})]. \]

The government chooses \( b \) and \( \tau \) subject to a budget constraint:

\[ Sb + G = \tau[S\gamma_2 Y_{11} + (1 - S)\gamma_1 Y_{01}] = \tau \bar{Y} \]

where \( S = E(s_i) \) is the mean of \( s_i \) across the population, or the fraction of the population attending university, \( G \) is the discounted total of other (exogenous) government spending over the 12 periods,\footnote{27} and \( \bar{Y} \) is mean total discounted lifetime income.\footnote{28} If \( V_i \) is total lifetime utility of individual \( i \), and social welfare \( V \) is utilitarian with equal weights, then \( V = E(V_i) \) and the social welfare gain from increasing \( b \) is:

\[ \frac{dV}{db} = \frac{\partial V}{\partial b} + \frac{\partial V}{\partial \tau} \frac{d\tau}{db} = E \left( \frac{\partial V_i}{\partial b} \right) + E \left( \frac{\partial V_i}{\partial \tau} \right) \frac{d\tau}{db}. \]  

\[ 26 \text{Given the lack of uncertainty when employed, this constraint will never bind on an individual after completing university, nor on an individual who chooses not to attend university.} \]

\[ 27 \text{Because } G \text{ is exogenous, I do not need to account for it in individual welfare.} \]

\[ 28 \text{The assumption that a lump-sum tax cannot be used, necessitating a proportional tax, is the first pre-existing distortion, as increases in income now translate into increased tax revenues that can be shared among all individuals. However, } G \text{ magnifies this distortion: if a large sum of government spending is necessary, } \tau \text{ will be large, and so will the distortion to individuals’ decisions.} \]
2.3.2 Welfare Calculations

I will now solve the model for an empirically-implementable version of (2.1). First, I evaluate the terms in (2.1), making use of the (unwritten) first-order conditions of the individual’s maximization problem:

\[
\frac{\partial V_i}{\partial b} = s_i (\lambda_{1i} + \mu_i) = s_i u'(c_{ui})
\]

\[
\frac{\partial V_i}{\partial \tau} = -s_i \lambda_{1i} \gamma_2 Y_{11} - (1 - s_i) \lambda_{0i} \gamma_1 Y_{01} = -s_i \gamma_2 Y_{11} v'(c_{vi}) - (1 - s_i) \gamma_1 Y_{01} v'(c_{vi})
\]

\[
\frac{dt}{db} = \frac{S}{\bar{Y}} \left[ 1 + \varepsilon_{Sb} - \left( 1 + \frac{G}{Sb} \right) \varepsilon_{Yb} \right]
\]

where \(\varepsilon_{ab}\) represents the (total derivative) elasticity of \(a\) with respect to \(b\). Defining \(E_0[\cdot]\) and \(E_1[\cdot]\) as expectations over individuals for which \(s_i = 0\) and \(s_i = 1\) respectively, the welfare derivative is:

\[
\frac{dV}{db} = SE_1[u'(c_{ui})] - [S \gamma_2 Y_{11} E_1[v'(c_{vi})] + (1 - S) \gamma_1 Y_{01} E_0[v'(c_{vi})]] \frac{S}{\bar{Y}} \left[ 1 + \varepsilon_{Sb} - \left( 1 + \frac{G}{Sb} \right) \varepsilon_{Yb} \right].
\]

Notice that \(\eta_i\) only affects the choice of \(s_i\), and that the debt limit has no effect on the consumption of those who do not attend university; therefore, \(c_{vi}^0 = c_v^0\) is constant across individuals and \(E_0[v'(c_v^0)] = v'(c_v^0)\). Next, for some quantity \(c^*\), I can write \(S \gamma_2 Y_{11} E_1[v'(c_{vi}^1)] + (1 - S) \gamma_1 Y_{01} v'(c_v^0) = \bar{Y} v'(c^*)\), where I expect \(c_{vi}^1 > c^* > c_v^0\) and therefore \(v'(c^*) < v'(c_v^0)\). The expression for \(\frac{dV}{db}\) thus becomes:

\[
\frac{dV}{db} = SE_1[u'(c_{ui})] - SV'(c^*) \left[ 1 + \varepsilon_{Sb} - \left( 1 + \frac{G}{Sb} \right) \varepsilon_{Yb} \right].
\]

Finally, to normalize the welfare gain into a dollar amount, define \(\frac{dW}{db} \equiv \frac{dV}{v'(c_v^0)}\); this expresses the welfare gain in terms of an equivalent amount of additional consumption among.
non-graduates. Therefore:

\[
\frac{dW}{db} = SE\left[ u'(c_{ui})v'(c^*) - v'(c_0)\right] - Sv'(c^*) - v'(c_0)\varepsilon_S + \left( 1 + \frac{G}{Sb} \right) \varepsilon_{Yb}.
\]

(2.3)

By making the assumption that \( v'(c^*) \simeq v'(c_0) \) above, I am overstating the relative importance of the fiscal effect; given that the optimum will occur where \( \frac{d\tau}{db} \) is positive, this will lead to an underestimate of the optimal \( b \).

(2.3) provides a simple and intuitive illustration of the welfare consequences of tuition subsidies, or indeed of any government income transfer program: the ratio of marginal utilities measures the welfare gain or loss from taking a dollar from one group and giving it to another, while the subsequent terms express the additional revenue cost or saving to the government from behavioural responses. A formula of this sort can generally be adapted to consider any transfer program, where the welfare gain from redistribution is measured against the resource cost of distortions. In the current context, there may be a welfare gain from redistributing towards students if they are borrowing-constrained, or there may be a welfare loss from a redistribution away from lower-income uneducated individuals towards people who will be higher-income graduates in the future. However, the revenue effect of tuition subsidies is likely to be positive, at least at baseline, so that there is a fiscal benefit of grants to students. Therefore, even if tuition subsidies redistribute in the “wrong” direction, they may increase total welfare because the gains to PSE graduates can be considerably larger than the losses to high school graduates.

\[29\] Since presumably \( v'(c^*_0) > E[v'(c_{vi})] \), this is less than the dollar amount I would get if I normalized by the mean marginal utility while employed.

\[30\] In the terminology of Okun (1975), this revenue effect measures how leaky the bucket is.
By design, this model has been conspicuous in its simplicity, but the result is very general; as mentioned earlier, a far more general analysis in chapter 3 produces an expression which simplifies to exactly this form. What the simplicity buys me is, on the one hand, ease of interpretation, and on the other, a starting point for the next step in my analysis: the replacement of the ratio of marginal utilities with some empirically observable quantity. Similar to the analysis of unemployment insurance in Chetty (2008), I will decompose the marginal utility term into “liquidity” and “substitution” effects.31

For simplicity, let me first assume that debt limits are the same for all individuals; the result is robust to a distribution of debt limits under certain assumptions, as I show in appendix 2.9.1, but the intuition is clearer in the simplest case. Thus, since the only heterogeneity enters in the form of $\eta_i$, consumption choices if schooling is undertaken are identical for all individuals, i.e. $c_{ui} = c_u, c_{vi} = c^1_v$ for all $i$. An individual chooses to attend university if $U_1 + \eta_i \geq U_0$, or:

$$\eta_i \geq R_1 v(c^0_v) - u(c_u) - R_2 v(c^1_v).$$

I assume that $\eta_i$ follows some continuously differentiable distribution $F(\eta)$, with a density given by $f(\eta)$. It follows that $S = 1 - F[R_1 v(c^0_v) - u(c_u) - R_2 v(c^1_v)]$, and therefore:

$$\frac{\partial S}{\partial b} = f(\eta^*)u'(c_u)$$

$$\frac{\partial S}{\partial a_1} = f(\eta^*)[u'(c_u) - v'(c^0_v)]$$

$$\frac{\partial S}{\partial w_1} = -f(\eta^*)v'(c^0_v)$$

31Chetty (2008) refers to the substitution effect as a “moral hazard” effect. Chetty (2006) instead proposes using a ratio of values of consumption, but a lack of suitable data on the consumption of students and its relation to government grants, as well as the implausibility of treating consumption while in school as directly comparable to that while employed, makes such an approach unsuitable here.
where $\eta^*$ is the critical value, and where $a_1$ and $w_1$ are artificial concepts representing dollars of additional income in the first period, with $a_1$ being an unconditional lump-sum of cash and $w_1$ an amount of additional after-tax employment income. It follows that I can rewrite (2.3) in the following way:

$$\frac{dW}{db} \simeq S \left[ L - \varepsilon_S b + \left( 1 + \frac{G}{S b} \right) \varepsilon_{Y b} \right]$$

where $L = \frac{\partial S}{\partial a_{1}} - \frac{\partial S}{\partial b} - \frac{\partial S}{\partial w_{1}}$.

The $\frac{\partial S}{\partial a_{1}}$ in the numerator of $L$ is the liquidity effect, whereas I call the $\frac{\partial S}{\partial b} - \frac{\partial S}{\partial a_{1}} = - \frac{\partial S}{\partial w_{1}}$ in the denominator the substitution effect, as it represents the effect on enrollment of changing relative prices without providing immediate income to students. $L$ therefore represents the ratio of the liquidity and substitution effects, and a higher value indicates a relatively larger liquidity or redistribution effect of student grants. (2.4) is the equation I will use in my sufficient statistics analysis in the next section, as it allows me to estimate the welfare gain from a marginal change in $b$, given values of the quantities which appear in the equation.

However, before beginning the numerical analysis, notice what must be true in my model if $\frac{\partial S}{\partial a_{1}} = 0$, i.e. if there is no causal effect of income on enrollment: this implies that $u'(c_u) = u'(c_u^0)$, and I expect that $v'(c_u^0) > v'(c_u^1)$, so therefore $u'(c_u) > v'(c_u)$. However, the absence of liquidity constraints requires $u'(c_u) = v'(c_u)$; therefore, a precisely-estimated zero effect of income on enrollment is in fact evidence in favour of liquidity constraints. If individuals were unconstrained, income should have a negative causal effect on enrollment, because a dollar of income would be more valuable to those who do not attend PSE; this is exactly the point discovered by Belley and Lochner (2007).\(^{32}\) As an illustration, when I study a

\(^{32}\)This conclusion is, however, dependent on the assumptions underlying the model; for instance, if the utility function for employed individuals was different across educated and uneducated individuals, it would not necessarily be true that $v'(c_u^0) > v'(c_u^1)$. 

92
case without liquidity constraints in section 2.5.1, I find that the model implies that each $1000 of initial assets should reduce enrollment by between 0.13 and 0.27 percentage points. Therefore, my results support the contention of Belley and Lochner (2007) that previous findings indicating that income has no causal effect on PSE enrollment should not necessarily be taken as evidence against liquidity constraints.

### 2.4 Numerical Results

In this section, I will focus on providing numerical results, starting with equation (2.4). First of all, using estimates of the current values of each of the sufficient statistics in (2.4), I calculate an estimated value of $\frac{dW}{db}$ and thereby determine if financial aid ought to be increased or decreased. To go beyond this local derivative, I must make additional assumptions, and I follow the advice of Chetty (2009) in trying two different approaches, which take up the subsequent two subsections: I perform statistical extrapolations of the quantities in (2.4), predicting their values out of sample, and I use the sufficient statistics to calibrate my model, permitting me to simulate the model to find the optimum.

#### 2.4.1 Sufficient Statistics Method

To evaluate (2.4), I must specify values for a number of quantities. To begin with, I use $S = 0.388$, which is the estimate of the enrollment rate of 18-24-year-olds in 2007 from NCES (2011).\(^{33}\) I also specify $b = 2$ (defining monetary amounts as thousands of dollars per year), using data on federal aid and state grants in 2007-08 from NCES (2011) and applying

---

\(^{33}\)I intend this value as a compromise. Given that the population of 18-24-year-olds includes individuals who have already completed or dropped out of PSE, this is an underestimate of the proportion of individuals ever enrolled in PSE (Lovenheim (2011) finds that 52% of his sample has completed more than 12 years of schooling). However, it is also an overestimate of the proportion actually completing a degree, which was 28.7% in 2007 according to NCES (2011).
the formula of Epple, Romano, and Sieg (2006) for turning loans and work-study into grant equivalents.\footnote{In 2007-08, 27.6\% of undergraduates received federal grants averaging $2800, 34.7\% received federal loans averaging $5100, 5.6\% received federal work-study averaging $2300, and 16.4\% received state grants averaging $2500; Epple, Romano, and Sieg (2006) suggest using a formula of $\text{aid} = \text{grants} + 0.25 \times \text{loans} + 0.5 \times \text{workstudy}$, which gives a per-person average of $1690. Lacking data on other government aid and tax credits, I round this total up to $2000.}

In choosing $b = 2$, I am assuming that the out-of-pocket tuition cost $e$ prior to government aid covers the marginal cost of PSE. At public institutions, this may not be accurate if universities' need for public funding increases with enrollment,\footnote{The finding in Trostel (2010) that the cost to government per degree is $71400 in present value does not prove this, as this is an average cost, not a marginal cost. Additionally, Trostel emphasizes that this is likely to be an overestimate.} but increased education should also lead to reductions in necessary spending on things like corrections, social insurance and social assistance. In fact, Trostel (2010) finds that additional expenditures on appropriations per degree are roughly offset by reductions in other government spending,\footnote{Trostel (2010) conservatively estimates the reduction in expenditures on such things as Medicaid, UI, and corrections per degree as $55800 in present value.} supporting my assumption that the marginal cost of PSE is well captured by $e$; in that case, increased education appropriations and reduced social spending cancel out of the government budget constraint and I can ignore both in my analysis. My conclusions, however, are not sensitive to this assumption; in appendix 2.9.3, I redo all the calculations using Trostel’s most pessimistic estimates, and the results are only slightly changed.

Deming and Dynarski (2009) summarize the literature on the price response of PSE attendance, and conclude that the general consensus is that a $1000 increase in price leads to a 4 percentage point decline in attendance, which implies an elasticity of $\varepsilon_{SB} \simeq 0.2$. Most estimates of the effect of price on both enrollment and degree completion are of this magnitude, so I will treat it as my baseline case. However, Dynarski (2008) estimates that
$2500 of financial aid leads to a 4 percentage point increase in degree completion from a base of 27%, which suggests a value closer to 0.1, so I will present results for this case as well.

As discussed earlier, numerous papers argue that income has no causal effect on enrollment, i.e. \( \frac{\partial S}{\partial a} = 0 = L \), which will be my preferred estimate. However, several papers do claim to find a positive income effect, the largest of which is Coelli (2011), whose results imply an effect on enrollment 25% as large as my preferred estimate of \( \frac{\partial S}{\partial b} \), suggesting \( L = \frac{1}{3} \); I will therefore present results for both values.

To calculate a value for \( \bar{\varepsilon}_{Yb} \), I assume that each year of schooling increases earnings by a constant 8%, and that the elasticity of taxable income is 0.4, as found by Gruber and Saez (2002). Utilizing these estimates, appendix 2.9.2 demonstrates that I can write the elasticity as

\[
\bar{\varepsilon}_{Yb} = \left[ \frac{(\gamma_2 (1.08)^4 - \gamma_1) S}{1 - (1.08)^4} - \frac{0.4\tau}{1 - \tau} \left( 1 + \frac{G}{Sb} \right)^{-1} \right] \frac{1 - \tau}{1 - 1.4\tau} \bar{\varepsilon}_{Sb} = \frac{0.4\tau}{1 - 1.4\tau} \left( 1 + \frac{G}{Sb} \right)^{-1}.
\]

Finally, for \( \frac{G}{Sb} \), I begin with my estimates of \( b = 2 \) and \( S = 0.388 \); I then need to estimate \( \tau \bar{Y} \) in order to be able to compute \( G \). I use a value of \( \tau = 0.23 \), which incorporates a 15% federal tax rate, a 5% state tax, and 3% for the Medicare tax. For \( \bar{Y} \), I turn to the CPS 2008 Annual Social and Economic Supplement, which estimates the mean earnings of a high school graduate in 2007 to be $33609, which I round up to \( Y_{01} = 34 \), meaning that \( Y_{11} = 34 (1.08)^4 = 46.26 \) and \( \bar{Y} = 301.661 \). Therefore, \( G = 68.606 \) and \( \frac{G}{Sb} = 88.410 \).

\[\text{37 This conclusion requires an assumption that the impact of parental job loss on enrollment found in Coelli (2011) comes entirely through the channel of lost income, making it a reasonable choice for an upper bound.}\]

\[\text{38 The range of estimates in Card (1999) are described in Sianesi and Van Reenen (2003) as being from 6% to 11%, but more recent estimates are higher: Dynarski (2008) summarizes several higher estimates of returns to university, and Heckman, Lochner, and Todd (2006) and Carneiro, Heckman, and Vytlacil (2010) estimate “policy relevant treatment effects” of tuition subsidies that range from 9% to 25%, making 8% a conservative estimate.}\]

\[\text{39 I adjust equilibrium earnings for changes in taxes according to this elasticity, but I do not attempt to model the labour supply decision; in this way, I will tend to produce an underestimate of the optimal } b, \text{ since I overstate the cost of tax increases by ignoring increases in leisure.}\]

\[\text{40 The federal and state rates are chosen to be appropriate for the typical high school graduate; they will be conservative for many graduates of university. In my analysis of heterogeneous returns to education in subsection 2.6.2, I attempt to model the tax system in more detail.}\]
Plugging in each of these values, I find values for $\frac{dW}{db}$ as displayed in Panel A of Table 2.1. They are positive and substantial, suggesting that a one dollar per year increase in student grants from the baseline of $2000 would provide a welfare increase equivalent to between 18 and 54 cents per year for 4 years. A 1% increase in $b$ to $2020$, therefore, would provide an average individual with an annual welfare gain of between $0.52$ and $1.57$ over their lifetime; aggregated up to an economy-wide level, this corresponds to a net welfare gain of $101$ to $305$ million per year for an increase in yearly grant spending of about $126$ million,\(^{41}\) indicating a very large return to public investments in education.

Table 2.1: Results from Sufficient Statistics and Extrapolation using (2.4)

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{sb}$</th>
<th>$L$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>A. Estimate of $\frac{dW}{db}$ at $b = 2$</td>
<td>0.1811</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Optimal Student Grants</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. Welfare Gains from Moving to Optimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
</tr>
</tbody>
</table>

### 2.4.2 Statistical Extrapolation

To make predictions out of sample, and thereby to produce an estimate of the optimal level of student grants, I need to make some functional form assumptions. In the current subsection, I make functional form assumptions about the sufficient statistics themselves; this approach is proposed in Chetty (2009), and has previously been used in sufficient statistic studies of unemployment insurance, including Baily (1978), Gruber (1997), and my first chapter.

\(^{41}\)I use the Census Bureau’s April 2010 estimate of the 18-64 population as 194,296,087.
First, let me now denote the baseline values of quantities using hats, i.e. \( \hat{b} = 2 \) and \( \hat{S} = 0.388 \). I assume a constant value of \( \varepsilon_{Sb} \) at either 0.1 or 0.2, implying that \( S = \phi b \varepsilon_{Sb} \),\(^\text{42}\) where \( \phi = \frac{\hat{S}}{b \varepsilon_{Sb}} \). I can then construct \( \frac{G}{Sb} = 88.41 \frac{\hat{b}}{Sb} \), and \( \varepsilon_{Yb} = \left[ \frac{\gamma_2 (1.08)^4 - \gamma_1 S}{\gamma_2 (1.08)^4 S + \gamma_1 (1-S)} - \frac{0.4\tau}{1-\tau} \left( 1 + \frac{G}{Sb} \right)^{-1} \right] \times \frac{1-\tau}{1-1.4\tau} \varepsilon_{Sb} - \frac{0.4\tau}{1-1.4\tau} \left( 1 + \frac{G}{Sb} \right)^{-1} \) as above, but holding \( \tau \) fixed at 0.23 for simplicity.

This leaves only the liquidity term to be constructed. I will follow the simple approach of Belley, Frenette, and Lochner (2011), who find a 16 percentage point gap in attendance at 4-year PSE institutions between the highest and lowest family income quartiles; I therefore assume that a 16 percentage point increase in enrollment is needed to eliminate the liquidity effect. I assume that once \( L \) reaches zero, or if it is initially zero, then it remains zero as \( b \) increases further; therefore, if the initial \( L \) is \( \frac{1}{3} \), I assume \( \frac{\partial S}{\partial a_1} = \frac{1}{4} \frac{0.16 - (S - \hat{S})}{0.16} \frac{\partial S}{\partial b} \), and thus \( L = \max\{ \frac{25}{16} \frac{(0.16 - (S - \hat{S}))}{0.16 - (S - \hat{S})}, 0 \} \).

Putting all of this together, I get the results displayed in Panels B and C of Table 2.1; for estimates of the welfare gain from moving to the optimum, I numerically integrate \( \frac{dW}{db} \) from \( b = 2 \) to the optimum. I then express the welfare gain in two ways: I multiply by 4 to get the dollar amount of an equivalent one-year per-person consumption increase, and I also divide by \( \hat{S} \hat{b} \) to express the gain as a percentage of the initial size of the student grant program; these latter values are shown in brackets in Panel C.

NCES (2011) estimates that median tuition at public 4-year universities was $5689 in 2007-08, so my results would suggest that, in the worst-case scenario, net tuition should be eliminated, and government appropriations increased accordingly.\(^\text{43}\) With a larger respon-\(^\text{42}\)This is generally consistent with the finding in Nielsen, Sørensen, and Taber (2010) of a considerably lower response of enrollment to price in Denmark, where financial support to students is much more generous.\(^\text{43}\)Although this is outside the scope of the current analysis, a policy of abolishing tuition is likely to be more effective and cost-effective than offsetting tuition with financial aid, for reasons of salience and reduced administration. Courant, McPherson, and Resch (2006) argue that the “old tradition of making public higher education ‘free’ has much to recommend it,” and claim that this policy might be efficient if enrollment is
siveness of enrollment to tuition or more serious borrowing constraints, the optimal policy would also include a yearly stipend which reaches as high as $8000;\textsuperscript{44} in the baseline case of $\varepsilon_{Sb} = 0.2$ and $\hat{L} = 0$, the optimal stipend is about $2400 per year. Meanwhile, the estimated welfare gains are substantial, particularly in comparison to the size of the policy change, and aggregating to an economy-wide level, they indicate annual welfare improvements of between $6.6 billion and $34.6 billion, or as much as 0.24% of GDP.\textsuperscript{45}

### 2.4.3 Simulation of Structural Model

A second option for out-of-sample prediction is to calibrate my simple structural model. I begin by assuming CRRA utility, so $u(c) = \frac{c^{1-\theta}}{1-\theta}$ and $v(c) = \frac{c^{1-\rho}}{1-\rho}$. I specify $Y_{01} = 34y$ and $Y_{11} = 34(1.08)^4y$, where $y = \alpha(1 - \tau)^{ETI}$ and $\alpha = \frac{1}{(1-\tau)^{ETI}}$, so $y = 1$ at baseline and shifts with $\tau$ to reflect effort responses to taxation. I also assume that $\eta$ follows a logistic distribution with mean $\mu$ and scale parameter $\sigma$.\textsuperscript{46}

I use $e = 5.7$ to represent public tuition, and $G = 68.606$ as described earlier. I assume that all individuals face the same debt limit $A$, so I have to solve for 5 parameters: $\{A, \theta, \rho, \mu, \sigma\}.\textsuperscript{47}$ However, I only have three sufficient statistic conditions: $\hat{S} = 0.388$, sufficiently sensitive to tuition, but they do not evaluate the welfare implications of the policy themselves.

\textsuperscript{44}The optimal value of $b$ is much more sensitive to the liquidity effect when $\varepsilon_{Sb}$ is small because my assumptions about how $L$ changes with $b$ mean that it takes much longer for the constraint to vanish in that case.

\textsuperscript{45}As a comparison, I can approximate the welfare gains from solving the educational mismatch found by Robst (2008), who evaluates wage penalties from mismatch between education and occupation on dimensions of both quantity and type of education. Robst finds little evidence of inefficient private decision-making, as the majority of mismatched individuals state supply-side reasons for their situation; in that case, the only welfare losses are from the fiscal externality, and an upper bound on the welfare gain can be computed under the assumption that mismatched individuals are indifferent between their current situation and an “appropriate” education-occupation match (so that there are no personal utility losses from being forced to take a different degree or job). In this case, the upper bound of the welfare gain is 0.36% of GDP, demonstrating that welfare gains from the source I am considering are of at least the same order of magnitude. All calculations are available upon request.

\textsuperscript{46}Unlike with a normal distribution, the logistic allows for calibration to be done analytically.

\textsuperscript{47}$\mu$ is not normalized to zero because $u(c)$ and $v(c)$ both have zero intercepts, so $\mu$ represents the difference in intercepts, the mean direct utility or disutility from schooling.
$\varepsilon_{sb} = \{0.1, 0.2\}$ and $\hat{L} = \{0, \frac{1}{3}\}$, so I need to incorporate additional data.

One piece of data I can use is some comparison of $\hat{c}_u$, $\hat{c}_0^v$ and $\hat{c}_1^v$; any ratio of two of these, along with the equation for the debt limit and the first-order conditions, will define all three. One possibility is to use consumption values from the Consumer Expenditure Survey, where I find that, on average, university graduates consumed 73.9\% of their pre-tax income and high school graduates consumed 83.4\% in 2007. The NBER’s TAXSIM calculator for 2007 allows me to transform these into percentages of after-tax income (ignoring state taxes and assuming a single-earner married couple), and if I then apply those values to my estimates of $Y_{01}$ and $Y_{11}$, I get $\hat{c}_v^1 = 1.2758 \hat{c}_v^0$. An alternative is to use results in Keane and Wolpin (2001) implying student consumption (not including room and board) of $8077 in 1987, plus the estimate from NCES (2011) of average room and board expenses in 1987-88 of $3037, compared to average per-equivalent-person consumption of $15816 in 1988 as estimated by Cutler and Katz (1991). I then find $\hat{c}_u = 0.7318\bar{c}$, where $\bar{c}$ is average consumption while employed, which implies $\hat{c}_v^1 = 1.2749 \hat{c}_v^0$. Given the similarities of these estimates, I will use $\hat{c}_v^1 = 1.275 \hat{c}_v^0$.

Finally, I can use an external estimate of relative risk-aversion to pin down one of $\theta$ and $\rho$. A CRRA parameter of 1, implying log utility, is typical; Gourinchas and Parker (2002), in particular, estimate a relative risk-aversion parameter during one’s working life of around unity. I therefore assume $\nu(c) = \ln(c)$, but I also try $\rho = 2$ in appendix 2.9.3.

My calibration method begins by using $\hat{c}_v^1 = 1.275 \hat{c}_v^0$ to solve for $A$; the condition $u'(c_u) = (\hat{L} + 1)u'(\hat{c}_0^v)$ then allows me to solve for $\theta$, and I can use these results and the conditions that $\hat{S} = 0.388$ and $\varepsilon_{sb} = \{0.1, 0.2\}$ to find $\mu$ and $\sigma$. I then simulate the model for various values of $b$ to find the optimum, and the results are displayed in Table 2.2.
Table 2.2: Results from Calibration and Simulation

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{sb}$</th>
<th>$L_0$</th>
<th>$\frac{1}{3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1259</td>
<td>0.2550</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4207</td>
<td>0.5502</td>
</tr>
</tbody>
</table>

A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{sb}$</th>
<th>Optimal Student Grants</th>
<th>Welfare Gains from Moving to Optimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>$3876$</td>
<td>$469$ (15.1%)</td>
</tr>
<tr>
<td>0.2</td>
<td>$5944$</td>
<td>$3327$ (107.2%)</td>
</tr>
</tbody>
</table>

B. Optimal Student Grants

C. Welfare Gains from Moving to Optimum

The results for the optimal level of $b$ are somewhat smaller than those in Table 2.1, typically by around $2000, with a much larger drop in the case where $\varepsilon_{sb} = 0.1$ and $\hat{L} = \frac{1}{3}$, in which my statistical extrapolations imply that the liquidity effect goes away very slowly. However, the qualitative conclusions are similar, in that eliminating tuition remains the optimal policy in all but one case. The estimated welfare gains, meanwhile, are actually larger when $\varepsilon_{sb} = 0.2$, reaching $23.3$ billion in the baseline case; overall, they vary from a low of $3.3$ billion to a high of $43.6$ billion, or 0.3% of GDP.

Figure 2.1 displays the values of $S$ over the relevant range, where it can be seen that in all cases, but especially those with $\varepsilon_{sb} = 0.2$, the optimal policy (indicated by the squares) involves inducing significant increases in the fraction of the population that attends PSE. Figure 2.2 displays the budget-balancing tax rates, and it is remarkable (though hard to see) that, in the cases with $\varepsilon_{sb} = 0.2$, a small increase in $b$ from the current level leads to a lower tax rate, because average income increases enough that the increased grants more than pay for themselves. This quickly ceases to be true as grants increase further, but if this standard estimate of the responsiveness of enrollment to tuition is correct, then at present
we are slightly on the wrong side of the “financial aid Laffer curve,” and thus there are Pareto improvements available from a small increase in tuition subsidies: taxes do not have to rise until the grant level reaches about $2550. Beyond that, the tax rate does rise, which means that there is redistribution away from high school graduates, which is socially costly, and yet the losses of high school graduates are more than offset by the considerable gains of PSE graduates until $b$ is well over $5000$.

Figure 2.1: Values of $S$

To further test the robustness of my results, I attempt a number of extensions and alterations to the model in appendix 2.9.3. A higher degree of risk-aversion slightly lowers optimal grants, while uncertainty about future income raises them; an alternative specification of government spending has varying effects depending on whether the statistical extrapolation or calibration method is used. The qualitative conclusions, however, are very
similar across all extensions.

### 2.5 Separating the Effects of Liquidity Constraints and Fiscal Externalities

In this section, as a sensitivity analysis, I perform the experiment of “switching off” the liquidity constraints and the fiscal externalities one at a time, to determine which contributes more to the argument for more generous financial aid.

#### 2.5.1 No Liquidity Constraints

I begin by assuming away liquidity constraints, and I focus on the structural approach as there is no transparent way to impose a zero-liquidity-constraint condition in the sufficient
statistics method.\footnote{As noted earlier, $L = 0$ does not correspond to no liquidity constraints, since an absence of liquidity constraints actually requires $v'(c_i^1) = u'(c_u)$, which means $\frac{\partial S}{\partial a_1} < 0$. The sufficient statistics method is difficult to use in this case because there is no way to impose the theoretical condition of no liquidity constraints without knowing what empirical value of $L$ this would correspond to. One way to calculate the implied $L$ is by simulating a structural model, and the results in this section suggest a value of $L = -0.2157$ in the absence of liquidity constraints.}

I begin by using $c_i^1 = 1.275 \hat{c}_i^0$ to solve for values of consumption, and then I use $v'(c_i^1) = u'(c_u)$, the no-liquidity-constraint condition, to solve for $\theta$. The rest of the calibration procedure continues as before, and the results are displayed in Table 2.3. The results are similar to before; the values of $\frac{dW}{db}$ are somewhat smaller than in Table 2.2, but in the baseline case the optimal benefit level is actually higher than in the $\hat{L} = 0$ case, and the welfare gain is nearly identical. It therefore appears that my general conclusion is not sensitive to the existence of liquidity constraints.

Table 2.3: Results from Calibration and Simulation with No Liquidity Constraints

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{Sb}$</th>
<th>A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$</th>
<th>B. Optimal Student Grants</th>
<th>C. Welfare Gains from Moving to Optimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.0430</td>
<td>$3263$</td>
<td>$110$ (3.6%)</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3374</td>
<td>$6630$</td>
<td>$3314$ (106.8%)</td>
</tr>
</tbody>
</table>


This is also the clearest demonstration in the paper of the “Theory of the Second Best” character of fiscal externalities: with a totally unconstrained population and no fiscal externalities, there would be no reason to interfere with the private decision about attending PSE. When taxes are needed to pay for other government spending, optimal policy in the area of education changes dramatically.
2.5.2 No Fiscal Externalities

Next, I instead shut off the fiscal externality, in the sense that I ignore $G$ and assume that $\tau_t$ is a lump-sum tax imposed on employed workers, growing at rate $g$ per period, so $\tau_t = (1 + g)^{t-1}\tau$. Re-doing my initial analysis in this context is straightforward, and the resulting equation for the welfare gain from increasing $b$ is:

$$\frac{dW}{db} \approx S \left( L - \frac{\gamma_1}{\gamma_1 - S} \varepsilon_{Sb} \right).$$

Implementing this formula, I get the results in Table 2.4. The values of $\frac{dW}{db}$ are much smaller, as are most of the optimal values of $b$; if $L = 0$ then there is no reason whatsoever to subsidize education (I set $b = 0$ as a lower bound). In the case of $\varepsilon_{Sb} = 0.1$ and $\hat{L} = \frac{1}{3}$, the estimated optimal $b$ is very large, but this is an anomaly resulting from the assumption that the fiscal externality is negative but small whereas the liquidity effect is significant and takes a long time to completely dissipate.

Table 2.4: Results from Sufficient Statistics and Extrapolation with no Fiscal Externalities

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{Sb}$</th>
<th>0</th>
<th>$\frac{1}{3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A. Estimate of $\frac{dW}{db}$ at $b = 2$</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>-0.0407</td>
<td>0.0886</td>
</tr>
<tr>
<td>0.2</td>
<td>-0.0814</td>
<td>0.0479</td>
</tr>
<tr>
<td></td>
<td>B. Optimal Student Grants</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$0$</td>
<td>$19218$</td>
</tr>
<tr>
<td>0.2</td>
<td>$0$</td>
<td>$3600$</td>
</tr>
<tr>
<td></td>
<td>C. Welfare Gains from Moving to Optimum</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$295$ (9.5%)</td>
<td>$1859$ (59.9%)</td>
</tr>
<tr>
<td>0.2</td>
<td>$539$ (17.4%)</td>
<td>$143$ (4.6%)</td>
</tr>
</tbody>
</table>

The structural approach also follows in the usual way. The results can be found in Table 2.5, and present conclusions that are similar to those in Table 2.4, with slightly larger values

---

This result also follows directly from (2.4) if I assume $G = 0$ and if I take $\varepsilon_{Sb}$ to be the elasticity of the tax base $\gamma_1 - S$. 

---

49 This result also follows directly from (2.4) if I assume $G = 0$ and if I take $\varepsilon_{Sb}$ to be the elasticity of the tax base $\gamma_1 - S$. 

104
for \( \frac{dW}{db} \) and optimal grants, with the exception of a much smaller optimal \( b \) in the \( \varepsilon_{sb} = 0.1 \), \( \hat{L} = \frac{1}{3} \) case. Therefore, using both approaches, I find that fiscal externalities are important to establishing beneficial effects of significantly increased grants to PSE students; severe liquidity constraints would otherwise be required in order to support significant grant increases, and in the baseline case, the optimal policy would involve reducing or even abolishing tuition subsidies in the absence of fiscal externalities.

Table 2.5: Results from Calibration and Simulation with No Fiscal Externalities

<table>
<thead>
<tr>
<th>Value of ( \varepsilon_{sb} )</th>
<th>( L = 0 )</th>
<th>( L = \frac{1}{3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Numerical Estimate of ( \frac{dW}{db} ) at ( b = 2 )</td>
<td>( -0.0080 )</td>
<td>( 0.1214 )</td>
</tr>
<tr>
<td>( -0.0457 )</td>
<td>( 0.0837 )</td>
<td></td>
</tr>
<tr>
<td>B. Optimal Student Grants</td>
<td>( $1769 )</td>
<td>( $6017 )</td>
</tr>
<tr>
<td>( $1087 )</td>
<td>( $3961 )</td>
<td></td>
</tr>
<tr>
<td>C. Welfare Gains from Moving to Optimum</td>
<td>( $4 \ (0.1%) )</td>
<td>( $976 \ (31.5%) )</td>
</tr>
<tr>
<td>( $82 \ (2.6%) )</td>
<td>( $333 \ (10.7%) )</td>
<td></td>
</tr>
</tbody>
</table>

2.5.3 Relative Importance of Fiscal Externalities and Liquidity Constraints

To summarize my results in this section, it appears that the liquidity term makes little difference to the optimal level of \( b \); eliminating all liquidity constraints does not significantly reduce the optimal \( b \) or change its qualitative implications, as can be seen from the comparison of Tables 2.2 and 2.3. Fiscal externalities, on the other hand, seem to be far more important. The logic is that, while liquidity constraints may well be a motivation for financial aid, fiscal externalities on their own can justify eliminating tuition and possibly providing stipends in most cases, by which point any liquidity constraints will have ceased to be a major concern. Thus, liquidity constraints appear to be of second-order importance.
when designing optimal financial aid policy for post-secondary students.

2.6 Heterogeneity

In this section, I undertake a structural analysis of heterogeneity along two dimensions. First, I consider heterogeneity in liquidity constraints, allowing me to calculate the optimal 2-tier grant scheme and compare it to my main results. Then I allow for a distribution of wage premiums from post-secondary education to assess how this alters my conclusions.

2.6.1 Heterogeneity in Liquidity Constraints and Two-Tier Grants

In appendix 2.9.1, I examine how robust the sufficient statistics condition in (2.4) is to a distribution of debt limits; an alternative examination of the robustness of the results to heterogeneous liquidity constraints can be performed using a structural approach. I allow for two groups, each representing half of the population,\(^5\) one of which is unconstrained while the other faces a debt limit \(A\). I calibrate the model for \(\{A, \theta, \mu, \sigma\}\) using the sufficient statistics as averages, and then solve for the optimal lump-sum student grant, with the results displayed in Table 2.6. The values of \(\frac{dW}{db}\) are slightly smaller than in Table 2.2, which is to be expected because the logistic distribution for \(\eta\) has an increasing hazard (see appendix 2.9.1), but the rest of the results are generally close to those from the baseline calculations; in most cases, including the baseline case, the optimal level of \(b\) is actually higher.

With this calibrated model in hand, I can go one step further and consider what policy the government would want to set if they could observe individuals’ debt limits; with two types of individuals, the government could introduce a two-tier grant system, with one grant

\(^5\)Brown, Scholz, and Seshadri (2009) find that approximately half of the children in their sample did not receive post-schooling cash transfers from their parents, which they claim as an indicator for student liquidity constraints.
Table 2.6: Results from Calibration and Simulation with Heterogeneous Liquidity Constraints

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{Sb}$</th>
<th>$L$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.1236</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4159</td>
</tr>
<tr>
<td>B. Optimal Student Grants</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$4227$</td>
</tr>
<tr>
<td>0.2</td>
<td>$6509$</td>
</tr>
<tr>
<td>C. Welfare Gains from Moving to Optimum</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$546$ (17.6%)</td>
</tr>
<tr>
<td>0.2</td>
<td>$3811$ (122.8%)</td>
</tr>
</tbody>
</table>

amount $b_1$ for the constrained group and another amount $b_2$ for unconstrained students.

It is straightforward to numerically maximize welfare (still measured as equally-weighted utilitarian social welfare) over the pair $(b_1, b_2)$, and the results for this exercise can be found in Table 2.7. Not surprisingly, it is always optimal to provide more generous aid to the constrained group, but substantial grants to the unconstrained group are still optimal with the standard estimate of $\varepsilon_{Sb} = 0.2$, as the fiscal externality motive remains strong; in the baseline case with $L = 0$, it remains optimal to abolish tuition, plus a stipend of just over $1000 for the constrained group. The welfare gains over and above those from the lump-sum policy in Table 2.6 are relatively small in most cases.

Table 2.7: Results from Calibration and Simulation with Two-Tier Grants

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{Sb}$</th>
<th>$L$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>A. Optimal Two-Tier Student Grants $(b_1/b_2)$</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$5261$/$1566$</td>
</tr>
<tr>
<td>0.2</td>
<td>$6925$/$5767$</td>
</tr>
<tr>
<td>B. Welfare Gains from Moving to Optimum</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$844$ (27.2%)</td>
</tr>
<tr>
<td>0.2</td>
<td>$3870$ (124.7%)</td>
</tr>
</tbody>
</table>
2.6.2 Heterogeneous Returns to Education

In this subsection, I investigate how sensitive the results are to allowing for heterogeneous returns to education.

I assume that the PSE wage premium $R$ (where $Y_{11} = RY_{01}$) follows some distribution $G(R)$, and to be precise I use a quadratic approximation to the marginal treatment effect distribution presented in Figure 4 of Carneiro, Heckman, and Vytlacil (2011). I divide the population into 100 equal masses denoted by $j = \{1, 2, ..., 100\}$, with wage premia equal to $\{G^{-1}(0.005), G^{-1}(0.015), ..., G^{-1}(0.995)\}$, and then I allow for a distribution of $\eta$ for each mini-population, where $\eta$ is allowed to be correlated with $R$. In particular, I let $\eta_{ij} = \bar{\eta}_j + \eta_i$, where $\bar{\eta}_j$ is deterministic for each $j$ and $\eta_i$ comes from a logistic distribution with mean 0 and scale parameter $\sigma$. I specify $\bar{\eta}_j = U_0 - U_{1j} + z - \mu_s \left( \frac{j-1}{J} \right)^{12}$, where $U_{1j} = u(c_{uj}) + R_2 v(c_{vj})$, as this generates a pattern of responsiveness to $b$ which is roughly consistent with that found in Carneiro, Heckman, and Vytlacil (2011).

Allowing for a distribution of wage premia makes it important to model the tax system more realistically; I assume that the state and Medicare tax rates do not vary with income, but I use an approximation to the US federal system in 2008, with a 15% marginal rate up to $41500 and a 25% rate beyond. To account for the personal exemption of $3500 and the standard deduction of $5450, as well as the fact that the first $8025 of taxable income is only taxed at a 10% rate, I provide a universal tax refund of $1743.75. To avoid discontinuities in the marginal tax schedule, I use a smoothed approximation to the tax rate between $39000 and $44000, specifically a sine connecting $\tau = 0.23$ at $39000$ to $\tau = 0.33$ at $44000$. I assume that the tax rate threshold moves up with wage growth, and that when taxes need
to adjust to balance the budget, the base tax rate is the one that moves.

When calibrating, I select values for \( \{A, \theta, \mu_s, \sigma, z\} \) in order to match five quantities, three of which are familiar: 
\[
E_1[u'(c_{ui})] = (\hat{L} + 1)v'(c_0^u), \quad \hat{S} = 0.388, \quad \text{and} \quad \epsilon_{Sb} = \{0.1, 0.2\},
\]
although in this case \( \epsilon_{Sb} \) is interpreted as a partial derivative. I also choose \( z \) to generate a probability of attendance of 95\% for the highest-return group, and I use the fact that university graduates consume 73.9\% of their pre-tax income and high school graduates consume 83.4\% to motivate setting 
\[
\frac{E_1(c_{1i}^u)}{E_1(Y_1)} = \frac{0.739}{0.834}.
\]

This leads to the results presented in Table 2.8. The striking finding is that the welfare derivative at baseline is significantly larger, because the average return to education among those induced to go to school is higher using the estimates from Carneiro, Heckman, and Vytlacil (2011). However, there are diminishing returns to inducing PSE attendance, because increasingly generous grants induce students with lower monetary returns to go to school; therefore, optimal grants are lower when \( \epsilon_{Sb} = 0.2 \), though they are larger when \( \epsilon_{Sb} = 0.1 \) because the returns to inducing PSE attendance do not decline as quickly in that case. In the baseline case, if heterogeneous returns of this magnitude do exist, it may no longer be optimal to completely eliminate tuition, but a significant increase in the generosity of grants is still indicated, and the welfare gains are significantly larger than before, amounting to $103.4 billion per year.

This analysis provides us with a sense of how heterogeneity in returns can affect the results, but naturally has more of a “black box” character than the baseline analysis. The current method is not as well suited to answering questions about how financial aid could be better targetted at students on the margin of attending PSE or from groups with high
Table 2.8: Results from Calibration and Simulation with Heterogeneous Returns to Education

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{Sb}$</th>
<th>$L$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$</td>
<td>1.1593</td>
</tr>
<tr>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>2.9816</td>
</tr>
<tr>
<td>B. Optimal Student Grants</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$5248$</td>
</tr>
<tr>
<td>0.2</td>
<td>$4574$</td>
</tr>
<tr>
<td>C. Welfare Gains from Moving to Optimum</td>
<td>$7036 (226.7%)$</td>
</tr>
<tr>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>$14767 (475.7%)$</td>
</tr>
</tbody>
</table>

returns, and therefore future work using structural models with observed and unobserved heterogeneity would be useful in providing answers to such questions.

2.7 General Equilibrium Effects

In this section, I explore the controversial issue of general equilibrium effects and demonstrate how sensitive my results are to their existence and magnitude. I begin by looking at how the wage premium may shift with the supply of PSE graduates, using two different estimates of the elasticity of substitution. Then I consider the possibility of spillovers, or positive externalities of PSE onto the wages of other workers; and finally I combine both of these factors into one model.

51 There is also the question of whether incentives should be aimed at particular fields of study. The idea that some individuals may make socially suboptimal choices of type of education is considered in the education mismatch literature, which has primarily focussed on the idea of overeducation, as studied by the meta-analysis of Groot and van den Brink (2000); Robst (2007) focussed attention on the qualitative match between field of study and occupation. However, Bauer (2002) and Tsai (2010) argue that much of the observed penalties from overeducation are in fact fixed effects from unobserved heterogeneity, and Leuven and Oosterbeek (2011) are pessimistic about causal interpretations of the empirical findings of the literature.
2.7.1 GE Effects on PSE Wage Premium

Analysis in papers such as Katz and Murphy (1992) suggests that changes in the supply of university graduates may have significant effects on relative wages. In Heckman, Lochner, and Taber (1998a) and Heckman, Lochner, and Taber (1998b) it is shown that this has consequences for the effectiveness of tuition subsidies: if increased PSE attendance lowers the PSE wage premium, then grants to students can only induce a small increase in attendance before declines in the wage premium completely offset the increased incentives to attend. Heckman, Lochner, and Taber (1998a) estimate an elasticity of substitution between high school and university graduates of 1.441, and show that this means that the effect of a tuition subsidy on PSE attendance in general equilibrium is about one-tenth the size of the partial equilibrium effect.

However, this conclusion is sensitive to assumptions about the usage of skill in the economy, as Lee (2005) finds general equilibrium effects of tuition subsidies that are more than 90% as large as the partial equilibrium values. Also, there is reason to believe that the short-run effects on relative wages of an increase in supply of PSE graduates may overstate the long-run effect if increased supply of skills leads to technological change to take advantage of those skills. Acemoglu (1998), Kiley (1999) and Acemoglu (2002) present models in which an increased supply of skilled workers leads to technological adjustment that creates more jobs designed for skilled workers, with the skill premium then increasing over time, possibly above the original level. The magnitude of these general equilibrium effects therefore remains an unanswered question, and one that is deserving of further study, but in the present

---

52 Additionally, Dupuy and Marey (2008) find that the elasticity of substitution has not been constant over time, rising significantly during the 1990s, and Walker and Zhu (2008) find that a significant increase in the supply of PSE graduates in the UK during 1994-2006 did not lead to a decline in the PSE premium.
analysis I will not take a position on what elasticity of substitution is correct; I will instead present results corresponding to both the Heckman, Lochner, and Taber (1998a) and Lee (2005) cases.

I begin by assuming a CES production function over high school and university graduates, specifically:

\[ Y_t = \zeta_t \left( a S_t^\kappa + (1 - a) S_t^{\kappa_0} \right)^{\frac{1}{\kappa}} \]

where \( S_t = S \) and \( S_t = 1 - S \). I assume that wages and the production function are specific to the generation in question, i.e. that vintage effects make the human capital of different cohorts perfectly non-substitutable, thereby producing an upper bound on general equilibrium effects.\(^{53}\) Therefore the wage of a PSE graduate is \( Y_{t1} = \frac{\partial Y_t}{S_t} \) and the wage of a high school graduate is \( Y_{t0} = \frac{\partial Y_t}{S_t} \), and \( a \) is chosen to make \( \frac{Y_{t1}}{Y_{t0}} = 1.08^4 \) at baseline. Calibration proceeds in the same way as before, since the only derivative used there is \( \frac{dS}{db} \), which I assume is evaluated at constant wages.

I produce results for two values of the elasticity of substitution, which can be written as \( \frac{1}{1-\kappa} \), namely 1.441 as in Heckman, Lochner, and Taber (1998a), and 350, which generates a ratio of general equilibrium to partial equilibrium effects that is comparable to Lee (2005).\(^{54}\) These results are displayed in Tables 2.9 and 2.10.

If high- and low-education workers are not good substitutes for each other, as argued by Heckman, Lochner, and Taber (1998a), then my findings confirm those of the latter paper in that the role of tuition subsidies in increasing PSE enrollment is minimal in the absence of

---

\(^{53}\) That is, I assume that the population share of PSE graduates adjusts immediately to that of the current generation, rather than allowing for an adjustment to a new long-run equilibrium. In this I follow the approach of Heckman, Lochner, and Taber (1998b), who state that short-run general equilibrium effects on enrollment with rational expectations are also very small.

\(^{54}\) The average of the ratio for men (1.05) and for women (1.52) in Lee (2005) is 0.9266; the average ratio across the four cases displayed in Table 2.10 is 0.9250.
liquidity constraints; in order to justify substantial increases in grants, significant liquidity constraints are required. However, with a much higher elasticity of substitution as in Lee (2005), the results are nearly identical to those from my baseline analysis. The magnitude of these general equilibrium effects, therefore, is of considerable importance, clearly demonstrating the importance of future work that can shed more light onto this phenomenon.

### 2.7.2 Wage Spillovers

A number of papers have sought evidence of positive wage spillovers from PSE education, i.e. a positive externality of education manifesting itself in higher wages for other workers,
resulting from off-the-job interactions or some form of social capital. Moretti (2004a) and Moretti (2004b) represent two prominent examples that do find significant effects, whereas Ciccone and Peri (2006) do not; Lange and Topel (2006) survey the literature and provide additional estimates. Damon and Glewwe (2011) evaluate the literature and conclude that the estimate produced by Lange and Topel (2006), which implies that a one percentage point increase in the population with a bachelor’s degree increases average wages by 0.2% within education group, represents a “very conservative” estimate of the effect, with other estimates often in the range of 1%. I will therefore proceed by using this estimate that each percentage point increase in PSE enrollment raises average wages by 0.2%.

This effect can easily be incorporated into the simulation to find a numerical estimate of \( \frac{dw}{db} \), but moving away from \( \hat{b} = 2 \), it does not seem plausible that this spillover would remain at the same level as \( S \) increases. Therefore, I have experimented with several assumptions about how marginal gains from spillovers decline with \( S \), and in Table 2.11 below I present results where the wage increase per percentage point of attendance is \( \frac{\delta}{S^2} \), where \( \delta = 0.002(0.388^2) \).

<table>
<thead>
<tr>
<th>Value of ( \varepsilon_{sb} )</th>
<th>( L )</th>
<th>( \frac{1}{3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Numerical Estimate of ( \frac{dw}{db} ) at ( b = 2 )</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>1.6318</td>
<td>1.8020</td>
</tr>
<tr>
<td>0.2</td>
<td>3.1094</td>
<td>3.3830</td>
</tr>
<tr>
<td><strong>B. Optimal Student Grants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$18291</td>
<td>$17179</td>
</tr>
<tr>
<td>0.2</td>
<td>$15819</td>
<td>$15027</td>
</tr>
<tr>
<td><strong>C. Welfare Gains from Moving to Optimum</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$42839 (1380.1%)</td>
<td>$45899 (1478.7%)</td>
</tr>
<tr>
<td>0.2</td>
<td>$72361 (2331.2%)</td>
<td>$79748 (2569.2%)</td>
</tr>
</tbody>
</table>
Even with this “very conservative” assumption about wage spillovers, the welfare gain from increasing grants to PSE students is now enormous; a 1% increase in $b$ to $2020$ generates an annual economy-wide gain of $1.74$ billion in the baseline case of $\varepsilon_{Sb} = 0.2$ and $\hat{L} = 0$. Furthermore, even with spillovers that diminish at the rate of $S^2$, the optimal grants are very large, higher than the median value of tuition, room and board at public universities of $13035$ in 2007-08, and the welfare gains are also very large, with a value of $506.6$ billion in the baseline case, or 3.6% of GDP. Also, notably, because of the spillovers to uneducated individuals, there is considerable scope for Pareto improvements; in all cases, Pareto gains can be obtained from marginal increases in $b$ up to at least $10000$, and at the optimum, both high school and PSE graduates are better off than when $b = 2$.

The results for the optimum can only be a rough approximation, given the lack of evidence on how spillovers would change with $S$, but the magnitude of the welfare derivative alone indicates that wage spillovers that might have been considered small in previous work are actually extremely important, which indicates a need for further work in this area.

### 2.7.3 Both GE Effects Combined

To give an indication of which of these general equilibrium effects may be expected to dominate if they occur at the same time, I can easily combine them. In Table 2.12, I present results with an elasticity of substitution of 1.441, as in Heckman, Lochner, and Taber (1998a), and wage spillovers of 0.2% per percentage point of PSE graduates, as in Damon and Glewwe (2011). The effects offset each other to a significant degree; optimal grants are now larger than in Table 2.2 if liquidity constraints are significant, and smaller with $\hat{L} = 0$ though still indicating increases in generosity. These results are only illustrations.
of the considerable sensitivity of optimal policy to the existence of general equilibrium effects; future work should focus on providing stronger evidence about their existence, magnitude and direction.

Table 2.12: Results from Calibration and Simulation with Elasticity of Substitution = 1.441 and Spillovers

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{Sb}$</th>
<th>$L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\frac{1}{3}$</td>
<td>0.2415</td>
</tr>
</tbody>
</table>

A. Numerical Estimate of $\frac{dw}{db}$ at $b = 2$

| 0.1 | 0.0472 | 0.2415 |
| 0.2 | 0.0537 | 0.2526 |

B. Optimal Student Grants

| 0.1 | $\$3200$ | $\$8239$ |
| 0.2 | $\$3359$ | $\$8475$ |

C. Welfare Gains from Moving to Optimum

| 0.1 | $\$112$ (3.6%) | $\$2777$ (89.5%) |
| 0.2 | $\$144$ (4.6%) | $\$3007$ (96.9%) |

### 2.8 Conclusion

In this paper, I have presented a simple model of post-secondary education, to allow for an analysis that takes seriously the Second-Best nature of the optimal tuition subsidy problem, and to permit me to derive plausible numerical results with clear policy implications. My results indicate that fiscal externalities provide justification for greater government support for students. The preferred estimates indicate as a first-order policy recommendation the elimination of tuition (at least at public schools, with perhaps a stipend of equal value for private institutions); some cases also recommend a partial stipend for living expenses. These results are robust to a number of sensitivity analyses and extensions of the basic model; an analysis with heterogenous liquidity constraints suggests it is generally optimal to provide more generous support to individuals who are borrowing-constrained, but at usual parame-
ter estimates substantial support should still be given to completely unconstrained groups. Furthermore, although the previous research on liquidity constraints has been enveloped in controversy about their magnitude and even their existence, my conclusions are driven by the fiscal externality component and are even largely robust to an elimination of liquidity constraints, suggesting that the latter are of second-order importance to policy.

The one factor which can alter these conclusions is the existence of significant general equilibrium effects of tuition subsidies on wages. If effects on relative wages are as severe as those estimated by Heckman, Lochner, and Taber (1998a), then the case for abolishing tuition rests entirely on the existence of significant liquidity constraints. On the other hand, even modest wage spillovers could make a case for large stipends on top of free tuition. Thus, further work that models and estimates wage formation in general equilibrium is called for.

Additionally, structural analysis of models with multiple dimensions of heterogeneity would be complementary to the current analysis in that it would generate answers to additional policy questions of interest. In particular, policies targeting financial aid more effectively at students close to the margin of attending PSE or with higher returns to education are likely to be more efficient; in the main analysis in this paper, I focussed on a simple case of a lump-sum student grant and a proportional tax, and found that even though such an arrangement will tend to involve redistribution away from lower-income high school graduates and towards higher-income PSE graduates, a utilitarian social planner would still prefer more generous grants. If these grants can be targeted at marginal groups, the revenue requirements will be lower and it will be more likely that Pareto gains can be obtained.

Therefore, some caution is called for in interpreting my results; there is more that we need to know. However, my analysis does suggest a baseline conclusion of eliminating public
tuition, and provides a clear guide to future research by highlighting the areas where we
need to know more. This paper also provides a methodological advance through a novel
application of the sufficient statistics method to the area of post-secondary education, and
demonstrates that empirical work on liquidity constraints among students may be of limited
policy relevance.

2.9 Appendix

2.9.1 Liquidity Term with Heterogeneous Constraints

To be as general as possible, let me allow for the possibility that \( \eta_i \) and \( A_i \) are jointly
distributed according to some bivariate distribution function \( F(\eta, A) \). Let me define \( S_A(A) \)
to be the probability of university attendance for an individual with debt limit \( A \); this can
be written as:

\[
S_A(A) = 1 - F_{\eta|A}[R_1 v(c_0^v) - u(c_u(A)) - R_2 v(c_1^v(A))|A]
\]

where \( F_{\eta|A} \) represents the conditional cdf. Then the overall probability of university atten-
dance is simply \( S = \int_A S_A(A) f_A(A)dA \), where \( f_A \) is the marginal density of \( A \).

Next, observe that:

\[
\frac{\partial S}{\partial b} = \int_A \frac{\partial S_A(A)}{\partial b} f_A(A)dA = \int_A f_{\eta|A}(\eta^*_A | A) f_A(A) u'(c_u(A))dA
\]

\[
\frac{\partial S}{\partial a_1} = \int_A \frac{\partial S_A(A)}{\partial a_1} f_A(A)dA = \int_A f_{\eta|A}(\eta^*_A | A) f_A(A) [u'(c_u(A)) - v'(c_0^v)]dA
\]

where \( \eta^*_A \) is the critical value for \( A_i = A \). Therefore, using the definition of \( L \) from the text:

\[
L = \frac{\int_A f_{\eta|A}(\eta^*_A | A) f_A(A) [u'(c_u(A)) - v'(c_0^v)]dA}{\int_A f_{\eta|A}(\eta^*_A | A) f_A(A) v'(c_0^v)dA}
\]

Meanwhile, the term I wish to replace is \( \frac{E_1[u'(c_{ui})] - v'(c_0^v)}{v'(c_0^v)} \); this is greater or less than \( L \) as:

\[
E_1[u'(c_{ui})] \geq \frac{\int_A f_{\eta|A}(\eta^*_A | A) f_A(A) u'(c_u(A))dA}{\int_A f_{\eta|A}(\eta^*_A | A) f_A(A)dA}
\]
\[
\frac{\int_A \left[ 1 - F_{\eta|A}(\eta^*_A|A) \right] f_A(A) u'(c_u(A)) dA}{\int_A \left[ 1 - F_{\eta|A}(\eta^*_A|A) \right] f_A(A) dA} \gtrless \frac{\int_A f_{\eta|A}(\eta_A^*|A) f_A(A) u'(c_u(A)) dA}{\int_A f_{\eta|A}(\eta_A^*|A) f_A(A) dA}.
\]

If the conditional hazard rate \( \frac{f_{\eta|A}(\eta^*_A|A)}{1 - F_{\eta|A}(\eta^*_A|A)} \) is constant, these two terms will be equal, and I can safely replace \( E_1[u'(c_u)] - \frac{u'(c_0)}{\nu'(c_0)} \) with \( L \) in (2.3). More generally, let me continue by substituting \( h(A) \) for the conditional hazard rate, and let me also write \( k(A) = \left[ 1 - F_{\eta|A}(\eta^*_A|A) \right] f_A(A) \) to represent the measure of enrollees at a particular value of \( A \); then the comparison becomes:

\[
\frac{\int_A k(A) u'(c_u(A)) dA}{\int_A k(A) dA} \gtrless \frac{\int_A h(A) u'(c_u(A)) dA}{\int_A h(A) dA}
\]

\[
\frac{\int_A k(A) h(A) dA}{\int_A k(A) dA} \gtrless \frac{\int_A k(A) h(A) u'(c_u(A)) dA}{\int_A k(A) dA}
\]

\[
E_1[h(A)] E_1[u'(c_u(A))] \gtrless E_1[h(A) u'(c_u(A))]
\]

\[
0 \gtrless \text{Cov}_1[h(A), u'(c_u(A))].
\]

Therefore, if the covariance of the hazard and the marginal utility among those attending university is close to zero, it will be a reasonable approximation to insert \( L \) into (2.3). Meanwhile, I will tend to underestimate the liquidity effect if the covariance is negative, which would follow, for instance, if \( h(A) \) is increasing in \( A \) (given that \( u'(c_u(A)) \) should be non-increasing in \( A \)).

It is hard to say if the hazard would be increasing; I expect \( 1 - F_{\eta|A}(\eta^*_A|A) \) to be higher for higher \( A \), which means that I also need \( f_{\eta|A}(\eta_A^*|A) \) to be increasing in \( A \). Dynarski (2002) argues that there is no consistent evidence of greater responsiveness of low-income students to price, which, given that I expect \( u'(c_u(A)) \) is higher for low-income students, would imply \( f_{\eta|A}(\eta_A^*|A) \) increasing in \( A \); but some studies do find evidence of such greater responsiveness (for instance, Kane (1994)), which provides less encouraging evidence.
Given that $\eta_A^*$ is decreasing in $A$, I would want the hazard to be decreasing in $\eta$, which would be the case for distributions such as the Pareto and the $\chi^2$ for degrees of freedom less than 2 (with 2 degrees of freedom, the hazard is constant). However, many other distributions, including the logistic that I use in my calibration, feature an increasing hazard, in which case my overestimate of $L$ would tend to offset the conservative assumptions elsewhere in the model.

2.9.2 Calculation of $\varepsilon_{\bar{Y}b}$

First, assuming that the only effects of $b$ on $\bar{Y}$ are from $b$'s effect on schooling and from the effect of the tax change on earnings $Y_{01}$ and $Y_{11}$, I can write:

$$\varepsilon_{\bar{Y}b} = \frac{b}{\bar{Y}} \frac{d\bar{Y}}{db} = \frac{b}{\bar{Y}} \left[ \frac{\partial \bar{Y}}{\partial S} \frac{dS}{db} + \frac{\partial \bar{Y}}{\partial \tau} \frac{d\tau}{db} \right].$$

It is clear that $\frac{\partial \bar{Y}}{\partial S} = \gamma_2 Y_{11} - \gamma_1 Y_{01} = [\gamma_2 (1.08)^4 - \gamma_1] Y_{01}$, and given that I assume that the elasticity of taxable income is 0.4, I have $\frac{\partial \bar{Y}}{\partial \tau} = -0.4 \bar{Y} \frac{1}{1-\tau}$. Using (2.2) for $\frac{d\tau}{db}$, the equation for $\varepsilon_{\bar{Y}b}$ becomes:

$$\varepsilon_{\bar{Y}b} = \left[ \gamma_2 (1.08)^4 - \gamma_1 \right] Y_{01} \frac{S}{\bar{Y}} \varepsilon_{Sb} - 0.4 \frac{Sb}{(1-\tau)\bar{Y}} \left[ 1 + \varepsilon_{Sb} - (1 + \frac{G}{Sb}) \varepsilon_{\bar{Y}b} \right]$$

and rearranging, I arrive at:

$$\varepsilon_{\bar{Y}b} = \left[ \frac{[\gamma_2 (1.08)^4 - \gamma_1] S}{\gamma_2 (1.08)^4 S + \gamma_1 (1-S)} - 0.4 \frac{1}{1-\tau} \left(1 + \frac{G}{Sb}\right)^{-1} \right] \frac{1-\tau}{1-1.4\tau} \varepsilon_{Sb} - \frac{0.4\tau}{1-1.4\tau} \left(1 + \frac{G}{Sb}\right)^{-1} \varepsilon_{\bar{Y}b}.$$

2.9.3 Sensitivity Analyses

This section will be devoted to an examination of the robustness of my results. I begin with an analysis of the sensitivity of my results to the coefficient of relative risk-aversion, and then I use the estimates of fiscal costs and benefits from Trostel (2010) to assess the impact on
my conclusions of how these fiscal effects are modelled. I also extend the model to consider uncertainty about future incomes. The quantitative results are only slightly altered in each case, and the qualitative conclusions remain very similar.

**Sensitivity of Results to Risk-Aversion**

My first sensitivity analysis considers how the results change when I specify a coefficient of relative risk-aversion of $\rho = 2$ for the employed state. Since I only need to specify this parameter when using the structural method, it will only affect my simulation results. Calibration proceeds as before, and simulation yields the results displayed in Table 2.13. The optimal values of $b$ and welfare effects are a bit smaller in most cases, but the conclusion of approximately abolishing tuition continues to hold in the baseline case.

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{sb}$</th>
<th>$L$</th>
<th>$\frac{1}{\delta}$</th>
<th>$\frac{dW}{db}$ at $b = 2$</th>
<th>$\hat{b}$</th>
<th>$\varepsilon$</th>
<th>$b$</th>
<th>Welfare Gains from Moving to Optimum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0.1505</td>
<td>0.2786</td>
<td>0.1</td>
<td>$3660$</td>
<td>$5051$</td>
<td>$489$ (15.8%) $1637$ (52.7%)</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.4185</td>
<td>0.5473</td>
<td>0.2</td>
<td>$5412$</td>
<td>$6506$</td>
<td>$2795$ (90.0%) $4835$ (155.8%)</td>
</tr>
</tbody>
</table>

**Evidence from Trostel (2010) on Fiscal Effects of Education**

In this subsection, I will test the robustness of my results to a different choice of $\hat{b}$; specifically, I perform my analysis again using the most pessimistic estimates from Trostel (2010), in which he concludes that each year of PSE costs the government $17850 and saves expenditures amounting to $13950 in present value. I therefore select $\hat{b} = 18$, increasing $e$ to 21.7
to correspond, and I assume that each year of schooling also saves expenditures amounting to \( p = 14 \).\footnote{If instead I set \( p = 16 \) to correspond to the baseline case in which I assume that government appropriations for education are exactly offset by reductions in other expenditures, all results are identical to those in section 2.4 except that the optimal grants and welfare gains are almost all larger using statistical extrapolations, due to functional form assumptions.} This changes the government budget constraint: I now divide \( G \) into two components, one exogenous component denoted by \( G_1 \), and one component \( G_2 = (1 - S)p \) representing the expenditures which can be eliminated with increased schooling. Therefore, the derivative of the budget constraint is:

\[
\frac{d\tau}{db} = \frac{S}{Y} \left[ 1 + \left( \frac{b - p}{b} \right) \varepsilon_{sb} - \left( 1 + \frac{G}{Sb} \right) \varepsilon_{yb} \right]
\]

and inserting this into \( \frac{dW}{db} \), I derive the following variant of (2.4):

\[
\frac{dW}{db} \approx S \left[ L - \left( \frac{b - p}{b} \right) \varepsilon_{sb} + \left( 1 + \frac{G}{Sb} \right) \varepsilon_{yb} \right]. \tag{2.5}
\]

Because the baseline value of \( b \) is 9 times larger, the earlier values of \( \varepsilon_{sb} = \{0.1, 0.2\} \) are now replaced by \( \varepsilon_{sb} = \{0.9, 1.8\} \). For the optimal grants, let me write them as \( \tilde{b} = b - 16 \) to make them comparable to earlier results; evaluating (2.5) and using the same statistical extrapolations as before leads to the results displayed in Table 2.14. The values of \( \frac{dW}{db} \) are smaller now, but the optimal grants are generally larger, as are the welfare gains at the optimum, due to the assumptions involved, particularly that of a constant value of \( \varepsilon_{sb} \). The baseline result involves an optimal stipend of over \$3600\) and a welfare gain amounting to \$36.6 billion.

Calibration and simulation follows the same procedure as before, and the results are found in Table 2.15. In every case, the welfare derivative at baseline is smaller, as are the optimal grants and the welfare gains from moving to the optimum; the optimal grants drop by less
Table 2.14: Results from Sufficient Statistics and Extrapolation using (2.5)

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{Sb}$</th>
<th>$\frac{1}{3}$</th>
<th>$\frac{1}{3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{L}$</td>
<td>$\hat{L}$</td>
</tr>
<tr>
<td>A. Estimate of $\frac{dW}{db}$ at $b = 18$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>0.1370</td>
<td>0.2664</td>
</tr>
<tr>
<td>1.8</td>
<td>0.3267</td>
<td>0.4560</td>
</tr>
<tr>
<td>B. Optimal Student Grants $b = b - 16$</td>
<td>$\hat{b}$</td>
<td>$\hat{b}$</td>
</tr>
<tr>
<td>0.9</td>
<td>$7795$</td>
<td>$8883$</td>
</tr>
<tr>
<td>1.8</td>
<td>$9343$</td>
<td>$9343$</td>
</tr>
<tr>
<td>C. Welfare Gains from Moving to Optimum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>$1563 (5.6%)$</td>
<td>$3648 (13.1%)$</td>
</tr>
<tr>
<td>1.8</td>
<td>$5220 (18.7%)$</td>
<td>$6252 (22.4%)$</td>
</tr>
</tbody>
</table>

than $1000, and the baseline result no longer involves the complete abolition of tuition, but still calls for significantly reduced out-of-pocket costs.

Table 2.15: Results from Calibration and Simulation for $\hat{b} = 18$

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{Sb}$</th>
<th>$\frac{1}{3}$</th>
<th>$\frac{1}{3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{L}$</td>
<td>$\hat{L}$</td>
</tr>
<tr>
<td>A. Numerical Estimate of $\frac{dW}{db}$ at $b = 18$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>0.0700</td>
<td>0.1990</td>
</tr>
<tr>
<td>1.8</td>
<td>0.3098</td>
<td>0.4391</td>
</tr>
<tr>
<td>B. Optimal Student Grants $b = b - 16$</td>
<td>$\hat{b}$</td>
<td>$\hat{b}$</td>
</tr>
<tr>
<td>0.9</td>
<td>$3061$</td>
<td>$5219$</td>
</tr>
<tr>
<td>1.8</td>
<td>$4994$</td>
<td>$6624$</td>
</tr>
<tr>
<td>C. Welfare Gains from Moving to Optimum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td>$149 (0.5%)$</td>
<td>$1263 (4.5%)$</td>
</tr>
<tr>
<td>1.8</td>
<td>$1861 (6.7%)$</td>
<td>$4082 (14.6%)$</td>
</tr>
</tbody>
</table>

**Income Uncertainty**

The final extension is a case with uncertainty about future incomes. To keep the problem simple, I assume that all uncertainty is resolved after the first period. Thereafter, educated individuals receive either $Y_{1tH} = (1 + g)^{t-1}Y_{11H}$ in each period or $Y_{1tL} = (1 + g)^{t-1}Y_{11L}$, each with probability 0.5, where $Y_{11H} > Y_{11L}$ and $\frac{Y_{11H} + Y_{11L}}{2} = Y_{11}$. Meanwhile, uneducated workers begin with $Y_{01}$ in the first period, and thereafter receive $Y_{0tH} = (1 + g)^{t-1}Y_{01H}$ or
$Y_{0L} = (1 + g)^{t-1}Y_{01L}$, each with probability 0.5, where $\frac{Y_{01H} + Y_{01L}}{2} = Y_{01}$. The corresponding consumption values will be denoted as $c_{v1}^1$ and $c_{v1}^1$ for educated workers and $c_{v1}^0$ and $c_{v1}^0$ for uneducated workers, with $c_{v1}^0$ representing the consumption of first-period workers.

In deriving $\frac{dW}{dv}$, the only meaningful change will come from the fact that $\frac{\partial V}{\partial \tau}$ takes a different form, specifically:

$$\frac{\partial V}{\partial \tau} = -\frac{\gamma_2^2}{2} S \left( v'(c_{v1}^1)Y_{11L} + v'(c_{v1}^1)Y_{11H} \right)$$

$$- \frac{1}{2} \left( v'(c_{v1}^0)(Y_{01} + \gamma_2 Y_{01L}) + v'(c_{v1}^0)(Y_{01} + \gamma_2 Y_{01H}) \right).$$

However, this equation cannot be used in its current form, and the most reasonable simplification is still $\bar{Y} v'(c^*)$, where $\bar{Y}$ remains equal to $S \gamma_2 Y_{11} + (1 - S) \gamma_1 Y_{01}$, so that (2.4) holds in this case as well, and the results are unchanged.

I will therefore focus on the structural analysis. The calibration proceeds largely as before, except that $A$ and $\theta$ must be chosen simultaneously to generate consumption choices which match $E(c_{v1}^1) = 1.275E(c_{v1}^0)$ and $u'(c_{v1}) = (\hat{L} + 1)v'(c_{v1}^0)$. For the variability of of income, I collect data on the median and interquartile range of income for high school and university graduates from the CPS in the 4th quarter of 2012. Then I consider three cases: one case in which I choose the values of $\{Y_{0L}, Y_{0H}, Y_{1H}, Y_{1L}\}$ that produce the same interquartile range, specifically 74.3% for high school graduates and 81.5% for PSE graduates, one case in which I cut the high school variance in half, and one in which I cut the PSE variance in half. The results are displayed in Table 2.16, and the optimal grants and welfare gains are larger in every case, though there does not appear to be one unambiguous pattern of results across the three cases. The baseline result features the abolition of tuition accompanied by a stipend of about $600 to $2000 per year.
Table 2.16: Results from Calibration and Simulation with Uncertain Income

<table>
<thead>
<tr>
<th>Value of $\varepsilon_{sb}$</th>
<th>CPS Variance</th>
<th>Low HS Variance</th>
<th>Low PSE Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>$\frac{1}{3}$</td>
<td>0</td>
</tr>
<tr>
<td>A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.1567</td>
<td>0.2858</td>
<td>0.1319</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4168</td>
<td>0.5464</td>
<td>0.4200</td>
</tr>
<tr>
<td>B. Optimal Student Grants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$4556$</td>
<td>$6904$</td>
<td>$5183$</td>
</tr>
<tr>
<td>0.2</td>
<td>$7569$</td>
<td>$8200$</td>
<td>$7839$</td>
</tr>
<tr>
<td>C. Welfare Gains from Moving to Optimum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>$767$</td>
<td>$2724$</td>
<td>$697$</td>
</tr>
<tr>
<td>0.2</td>
<td>$4016$</td>
<td>$6765$</td>
<td>$4682$</td>
</tr>
</tbody>
</table>

125
Chapter 3

Interactions Between Government Social Programs: An Application and General Formula

3.1 Introduction

There has been a great deal of attention paid by economists to the labour market effects and welfare implications of government social programs. For example, Krueger and Meyer (2002) survey the empirical literature on the labour market impacts of social insurance programs, while the first two chapters of this dissertation have discussed the empirical and welfare analysis literatures on social insurance and tuition subsidy programs respectively.

However, this literature has been incomplete in several important respects. Fiscal externalities, or the idea that government programs that impact labour market outcomes can have important effects on income tax revenues, have received limited attention in many areas of research, as documented in chapter 1. In that paper, I perform an analysis of optimal unemployment insurance and show that fiscal externalities can significantly alter estimates of the optimal policy. I also examine the effect of fiscal externalities on post-secondary tuition subsidy policy in chapter 2, and show that they can justify extensive subsidies which
effectively eliminate tuition at most public universities.

In this paper, I focus on another aspect of government transfer programs which has received limited attention: interactions between two or more programs, where those programs act as substitutes or complements from the perspective of individual participants. In a second-best world in which it is not possible to perfectly target transfers at states for which they are intended, it is important to consider that changing the generosity of one program can lead to changes in enrollment on other programs, with important fiscal effects.

I demonstrate that this can be an important issue, starting with a specific example in which I consider unemployment insurance (UI) when some unemployed individuals may also be eligible for disability insurance (DI). Using an estimate of the change in DI enrollment after a change in UI from Lindner (2012), I show that the optimal generosity of UI can be dramatically altered when DI is taken into account: much more generous UI may be indicated if doing so prevents individuals from applying for (and receiving) DI.

I then present a general model which can be applied to any program that involves state-contingent transfers, showing that the results from the initial example generalize to this setting. The model can be solved for a derivative of social welfare with respect to any individual program, with a simple and intuitive result that depends directly on the magnitude of fiscal externalities and program interaction effects. I conclude by examining some of the areas of research to which this approach could profitably be applied, and describe some unanswered questions.

The rest of the paper proceeds as follows. Section 3.2 presents an examination of optimal UI when individuals may substitute to or from DI. Section 3.3 provides a more general discussion of program interaction effects as well as fiscal externalities, and section 3.4 then presents
the general model and derives analytical results. Section 3.5 briefly summarizes literatures of particular importance where the method can be applied, and section 3.6 concludes and suggests areas for future research.

### 3.2 Optimal UI With DI as Substitute Pathway

Numerous papers have studied the question of optimal unemployment insurance, from Baily (1978) to Hansen and İmrohoroğlu (1992) to Chetty (2008), and the first chapter of this dissertation, where I demonstrate that this literature has ignored an important aspect of the optimal UI problem: the effect of UI on tax revenues, which I refer to as a “fiscal externality.”

Here, I introduce another new element to optimal UI analysis: the interaction of UI and DI, among individuals for whom UI and DI are substitutes in the area of social insurance. That is, for individuals who may qualify as disabled, going on DI is one possible income pathway, while some combination of UI and employment is another, and changes in the generosity of one program may affect enrollment on the other. The fact that UI and DI appear to be substitutes has been mentioned by Bound and Burkhauser (1999), and documented by Petrongolo (2009) and Lammers, Bloemen, and Hochguertel (2013); Lindner (2012), in particular, estimates that a $100 per month increase in UI benefits leads to about 2700 fewer new DI spells per year. However, the welfare implications of this substitution, and the consequences for optimal policy, have not previously been considered.

I will base my analysis on a version of Baily (1978), modified to include DI. The model consists of two periods: in the first period, the individual is employed at a wage which I normalize to one, and at the end of the first period they face a risk of losing their job. If they become unemployed, which happens with probability $\gamma$, they have the choice of applying
for disability insurance and realizing a utility loss $\delta$ from stigma or effort costs of applying; in the population, $\delta$ will follow some (unspecified) distribution. The choice of applying for DI is denoted by $\theta = 1$, and if the individual is approved, which happens with probability $\alpha$, they receive DI benefits $b_D$ (including the value of Medicare coverage) during the entire second period. Meanwhile, individuals who don’t apply for DI, or who are rejected, remain unemployed and receive UI benefits $b_U$ for some fraction of the second period, denoted by $s$, and then resume employment at a wage equal to one for the remainder of the period. By exogenously fixing the re-employment wage, I assume that there are no effects of UI on subsequent wages, since I argued in chapter 1 that the recent empirical literature has tended to support that conclusion. $s$ is chosen by the individual and subjects the individual to a utility cost $h(s)$ that is decreasing and convex in $s$; I abstract from uncertainty in unemployment duration.

I assume that the interest and discount rates facing the individual are both equal to $r$, which will be the equivalent of 3% per year. Because the rates are equal, and because there is no uncertainty, consumption choices will be constant while in a particular state in a particular period: I use $c_1$ and $c_2$ to represent consumption on the original job in periods 1 and 2, respectively, $c_U$ and $c_D$ for consumption on UI and DI, and $c_n$ as consumption on the new job if an individual was unemployed. To simplify the notation, I also denote
\[ \int_x^y e^{-rt} dt = e^y_x. \] The individual’s decision problem can therefore be written as:

\[
\max_{c_1,c_2,c_D,c_U,c_n,s,\theta} V = e^{1}_{0}U(c_1) + (1 - \gamma)e^{2}_{1}U(c_2) + \gamma [\theta \alpha e^{2}_{1}U(c_D) + (1 - \theta \alpha)(e^{1+}_{1}sU(c_U) + e^{2}_{1+s}U(c_n) - h(s)) - \theta \delta] - \lambda_1 [e^{1}_{0}c_1 + e^{2}_{1}s_2 - e^{2}_{0}(1 - \tau)] - \lambda_2 [e^{1}_{0}c_1 + e^{2}_{1}c_D - e^{2}_{0}(1 - \tau) - e^{2}_{1}b_D] - \lambda_3 [e^{1}_{0}c_1 + e^{1+}_{1}s_U + e^{2}_{1+s}c_n - (e^{1}_{0} + e^{2}_{1+s})(1 - \tau) - e^{1+}_{1}s_U]
\]

where \( \tau \) is the tax rate.

As in chapter 1, I can write the derivative of social welfare as:

\[
\frac{dV}{db_U} = \frac{\partial V}{\partial b_U} + \frac{\partial V}{\partial \tau} \frac{d\tau}{db_U}
\]

and the marginal utility terms can be expressed as:

\[
\frac{\partial V}{\partial b_U} = \lambda_3 e^{1+}_{1}s = \gamma(1 - \theta \alpha)e^{1+}_{1}sU'(c_U)
\]

\[
\frac{\partial V}{\partial \tau} = -\lambda_1 e^{2}_{0} - \lambda_2 e^{1}_{0} - \lambda_3 (e^{1}_{0} + e^{2}_{1+s}) = -e^{1}_{0}U'(c_1) - (1 - \gamma)e^{2}_{1}U'(c_2) - \gamma(1 - \theta \alpha)e^{2}_{1+s}U'(c_n)
\]

If UI and DI are the only two government programs to be financed (I will incorporate fiscal externalities later), the government budget constraint is:

\[
\tau \left[ e^{2}_{0} - \gamma(1 - \theta \alpha)e^{1+}_{1}s - \gamma \theta \alpha e^{2}_{1} \right] = \gamma(1 - \theta \alpha)e^{1+}_{1}s_U + \gamma \theta \alpha e^{2}_{1}b_D
\]

which, if I denote \( T_U = \gamma(1 - \theta \alpha)e^{1+}_{1}s_U \) and \( T_D = \gamma \theta \alpha e^{2}_{1} \) as the expected discounted amounts of time spent on UI and DI respectively, can be simplified to:

\[
\tau \left( e^{2}_{0} - T_U - T_D \right) = T_U b_U + T_D b_D
\]

and therefore the derivative of the government budget constraint is:

\[
\frac{d\tau}{db_U} = \frac{1}{e^{2}_{0} - T_U - T_D} \left[ T_U + (b_U + \tau) \frac{dT_U}{db_U} + (b_D + \tau) \frac{dT_D}{db_U} \right] = \frac{T_U}{e^{2}_{0} - T_U - T_D} \left[ 1 + \left( 1 + \frac{\tau}{b_U} \right) \frac{T_U}{b_U} + \frac{T_D(b_D + \tau)}{T_U b_U} \right]
\]
where $\varepsilon^y_x$ represents the elasticity of $y$ with respect to $x$.

If I define $c_e$ such that $(e_0^2 - T_U - T_D)U'(c_e) = e_0^1U'(c_1) + (1 - \gamma)e_1^2U'(c_2) + \gamma(1 - \theta \alpha)e_{1+s}^2U'(c_n)$, so that $U'(c_e)$ is equal to the discounting-weighted average marginal utility among employed individuals, then I can rewrite $\frac{\partial V}{\partial \tau} = -(e_0^2 - T_U - T_D)U'(c_e)$. Then, combining the marginal utility terms and $\frac{df}{db}$, I get the following for $\frac{dV}{db}$:

$$\frac{dV}{db_U} = T_UU''(c_U) - T_UU'(c_e) \left[ 1 + \left( 1 + \frac{\tau}{b_U} \right) \varepsilon^T_U + \frac{T_D(b_D + \tau)}{T_Ub_U} \varepsilon^T_D \right]$$

Finally, I normalize the welfare derivative by $U'(c_e)$ to get:

$$\frac{dW}{db_U} \equiv \frac{dV}{U'(c_e)} = T_U \left[ \frac{U'(c_U) - U'(c_e)}{U'(c_e)} - \left( 1 + \frac{\tau}{b_U} \right) \varepsilon^T_U - \frac{T_D(b_D + \tau)}{T_Ub_U} \varepsilon^T_D \right].$$

To put the marginal utility term into an empirically measurable form, I use a Taylor series expansion:

$$U'(c_U) \simeq U'(c_e) + U''(c_e)(c_U - c_e)$$

so therefore:

$$\frac{U'(c_U) - U'(c_e)}{U'(c_e)} \simeq -c_eU''(c_e) \frac{c_U - c_e}{c_e} = R(c_e) \frac{\Delta c}{c_e}$$

where $R$ is the coefficient of relative risk-aversion, and $\Delta c = c_e - c_U$.

Therefore the welfare derivative is:

$$\frac{dW}{db_U} = T_U \left[ R(c_e) \frac{\Delta c}{c_e} - \left( 1 + \frac{\tau}{b_U} \right) \varepsilon^T_U - \frac{T_D(b_D + \tau)}{T_Ub_U} \varepsilon^T_D \right]$$

where $\tau = \frac{T_Ub_U + T_Db_D}{e_0^2 - T_U - T_D}$.

Given estimates of each of the quantities in (3.3) - the sufficient statistics - I can calculate an estimated welfare gain from increasing $b_U$ in terms of dollars of consumption; then, using statistical extrapolations, I can approximate the values of the sufficient statistics out of sample and find the optimal level of UI benefits.
I begin by computing the baseline values of $T_U$ and $T_D$, remembering that I must deflate both due to discounting. I start with the fact that the size of the US labour force was about 154 million in 2008, and that 7.4 million people were receiving DI by the end of 2008, as reported by the Social Security Administration. Therefore, the size of the relevant population is 161.4 million. However, based on current flows onto DI, the fraction of people on DI appears to be below the steady-state value, and since I am looking at the effect of UI on flows into DI, the steady-state number is the relevant one; in 2008, about 895000 new awards were made, and the average duration on DI of 14 years in Autor and Duggan (2006) indicates a steady-state of 12.53 million, which makes $\gamma \theta \alpha = \frac{2 \times 12.53}{161.4} = 0.1553$.\footnote{I multiply by 2 because all of the DI spells occur in the second period.} Given that the average duration on DI is 14 years, I will make a period equal to 14 years in my model, which means that $r = 0.5126$, and therefore $T_D = 0.1553 e_1^2 = 0.0728$. Then, I use an unemployment rate of 5.4\% as in chapter 1, which means $\gamma(1 - \theta \alpha)s = 0.108$; if I use the job-losing rate from one of the cases in chapter 1, specifically $\gamma = 0.54$, then this implies that $s = 0.2807$, and therefore $T_U = \gamma(1 - \theta \alpha)e_1^{1+s} = 0.0602$.

As in chapter 1, I use a baseline UI replacement rate of 46\% and adjust benefits for takeup and finite duration, along with a tax rate applied to UI income, which I assume here is just a federal income tax rate of 15\%, to get $b_U = 0.46 \left( \frac{12.64}{24.3} \right) 0.85 = 0.2034$. Rutledge (2011) finds that before-tax average UI and DI benefits are of comparable magnitude, $\$233$ per week for UI and $\$963$ per month for DI in his sample, so I assume that they are equal, but DI recipients also receive Medicare after two years, with average benefits of about $\$7200$ per year according to Rutledge (2011); DI benefits are not subject to tax unless recipients have significant outside income, and therefore $b_D = (1 + \frac{600}{963}) 0.46 = 0.7466$. This implies
that the budget-balancing tax rate is $\tau = 0.0596$.

The elasticities are calculated as follows: I use $\varepsilon_{b_U}^{T_U} = 0.48 \times 0.53 = 0.2544$ as in chapter 1, and Lindner (2012) finds that a $100$ increase in monthly UI benefits, about a 10% increase, should lead to 2700 fewer new DI beneficiaries per year, so $\varepsilon_{b_D}^{T_D} = \frac{-27}{895} = -0.0302$. When extrapolating out of sample, I assume that the derivatives $\frac{dT_U}{db_U}$ and $\frac{dT_D}{db_U}$ stay constant at their baseline values, rather than assuming that the elasticities themselves stay constant, as the latter implies unrealistic behaviour of $T_U$ and $T_D$ as $b_U$ approaches zero.

Finally, I assume $R = 2$ as in the baseline case studied in chapter 1, and $\frac{\Delta c}{c_e} = 0.222 - 0.265rr$, where $rr$ is the UI replacement rate, as in Gruber (1997). If I put all of these estimates together, I get the results displayed in column 1 of Table 3.1. Panel A shows the welfare derivative at the baseline replacement rate of 46%, and panel B the optimal replacement rate, in the case in which I ignore interactions between UI and DI and the case in which I take them into account.

To this point I have ignored the fiscal externalities that were found to be so important in the first chapter of this dissertation, but it is easy to incorporate them; if I allow for an additional amount of exogenous government spending equal to $G$ beyond UI and DI, the government budget constraint (3.1) becomes $\tau (c_0^2 - T_U - T_D) = T_U b_U + T_D b_D + G$, and (3.3) is unchanged except that this new value of $\tau$ must be used in the calculations. To pin down $G$, I assume a 26.1% tax rate on earned income, incorporating a 15% federal rate, a typical 5% state tax, 2.9% for the Medicare tax, and 3.2% as the marginal OASDI tax rate calculated by Cushing (2005) for 37-year-olds, which implies $G = 0.2252$. Results in this case can be found in column 2 of Table 3.1.

In considering the results, it is first interesting to compare the optimal replacement
Table 3.1: Results from Sufficient Statistics and Extrapolation using (3.3)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$G = 0$</td>
<td>$G = 0.2252$</td>
</tr>
<tr>
<td>A. Estimate of $\frac{\partial W}{\partial b_1}$ at $rr = 0.46$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_{TD}^T = 0$</td>
<td>-0.0077</td>
<td>-0.0229</td>
</tr>
<tr>
<td>$\varepsilon_{b_D}^T = -0.0302$</td>
<td>0.0091</td>
<td>0.0239</td>
</tr>
<tr>
<td>B. Optimal Replacement Rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_{TD}^T = 0$</td>
<td>0.3271</td>
<td>0.0289</td>
</tr>
<tr>
<td>$\varepsilon_{b_D}^T = -0.0302$</td>
<td>0.6166</td>
<td>0.8580</td>
</tr>
</tbody>
</table>

rates in the two columns when $\varepsilon_{b_D}^T = 0$; in both cases, it is assumed that DI spending is accounted for, but that changes in $b_U$ have no effect on DI spending. When $G = 0$, I find an optimal replacement rate of about 33%, which drops to 3% when the large amount of other government spending is accounted for; these results confirm the findings of chapter 1 on the importance of fiscal externalities.

The main result in this section, however, is the fact that the optimal level of UI goes up significantly when substitution onto DI is considered; the fiscal benefits of lowering UI are greatly reduced, and the increase is most dramatic when fiscal externalities are also accounted for. In the latter case, the tax rate is large, and the tax savings from moving people off of DI are considerable. If partially disabled individuals’ decision about whether or not to apply for DI are affected by the generosity of UI, because they may expect to be unemployed on more than one occasion in the future, generous UI may keep those individuals in the labour force, providing an incentive to remain employed for at least part of the time.
3.3 Understanding Program Interaction Effects & Fiscal Externalities

The term “fiscal externalities” has been used in several contexts, as explained in chapter 1, but the context in which it is used here is the same as that described in the first two chapters: an effect of one government program on labour market outcomes and therefore on income tax revenues, affecting the government’s ability to fund other programs. My first chapter points out that the importance of fiscal externalities remains largely unrecognized in a number of areas of research, including most research on social insurance.

However, while I incorporate fiscal externalities in my analysis, both in the specific example described above and in the general model to come, the important new contribution of this paper is to consider the effects of one program on spending on other programs, to complete the picture of the overall fiscal impact of transfer programs. Previous research has occasionally considered the possibility that particular government programs may not be targetted perfectly at the state that is intended to be insured or subsidized, and that therefore multiple programs may be substitutes or complements for each other in the eyes of participants; for example, the possibility that UI and DI could be substitute pathways for some individuals is mentioned by Bound and Burkhauser (1999). This means that raising the generosity of transfers on one program may raise or reduce the expenditures on another program, something which should be taken into account when considering the impact of the program in question on society. However, analyses of the welfare implications of such interaction have been few and far between - chapter 2 being a rare exception in considering reductions in spending on social insurance and corrections from higher levels of education -
even though the analysis in section 3.2 above demonstrates that it is a highly relevant issue.

This omission in previous research motivates not only the specific example in the previous section, but also the attempt in the next section to develop a framework that can accommodate a general range of programs and interactions, and to derive results that can inform us about the implications of both fiscal externalities and program interaction effects. In the current paper, I seek context-specific welfare improvements, or what Lipsey (2007) would call “piecemeal second-best policy,” in that I only consider changing only one policy at a time and offer an equation for determining whether welfare would increase or decrease with the value of this policy instrument. I aim for generality by allowing for a set of unspecified constraints on agents, but I do not attempt to solve for a global general equilibrium Second Best optimum, which would require modelling all the irreducible distortions in the economy, and which Lipsey (2007) persuasively argues to be impractical.\(^2\)

3.4 Theoretical Analysis of General Model

I now move to a general case with a potentially large number of programs. I begin the theoretical analysis with a description of the general model; I then proceed to solve for the welfare derivative, and conclude with a discussion of various results.

3.4.1 General Model

I begin with the model from the general case of Chetty (2006), but I will make several modifications. In particular, I apply a more general interpretation of the model, to demonstrate

\(^2\)Lipsey (2007) argues that ‘the set of realistic policy goals does not include achieving either an economy-wide, Pareto efficient allocation of resources or an economy-wide second best setting” for any particular distortion, as the information required is prohibitive. In specific contexts, however, such as optimal social insurance, in which a set of programs can reasonably be assumed to interact primarily with each other, it may be reasonable to undertake structural analysis aimed at jointly optimizing multiple policies.
how the insights obtained may apply outside of the context of the basic social insurance problem.

Time is continuous with a unit duration, i.e. \( t \in [0, 1] \), and represents the individual’s working life (or some portion thereof). \( \omega_t \) is a state variable containing the agent’s history up to time \( t \), which follows an arbitrary stochastic process for which the unconditional (at time 0) distribution function is \( F_t(\omega_t) \). This state variable, which may be a vector, can contain such information as the agent’s record of employment and earnings, time spent in education and training, health status, or any number of other quantities. The representative individual chooses consumption \( c(t, \omega_t) \) and a vector of other actions \( x(t, \omega_t) \) for each time \( t \) and state \( \omega_t \) to maximize expected utility, which is time-separable and described as the discounted double integral of \( U(c(t, \omega_t), x(t, \omega_t)) \) across \( t \) and \( \omega_t \). As in the earlier example, I assume that the interest and discount rates are both equal to \( r \).

Instead of focusing on the distinction between states of employment and unemployment and a single program depending on those states (i.e. unemployment insurance as in Chetty (2006)), I will consider participation in a generalized range of programs of state-contingent transfers. To be precise, let there be \( M \) programs, where participation in program \( j = \{1, ..., M\} \) is denoted by \( P_j(t, \omega_t, x) = 1 \), and where \( x \) represents the complete set of state- and time-contingent choices of \( x(t, \omega_t) \) over the individual’s lifetime. There may be idiosyncratic uncertainty in program participation status, since it is a function of \( \omega_t \), but it is also possible that it may be completely determined by the individual’s choice of \( x \); thus, the program can represent something as unpredictable as a sudden and unexpected diagnosis of a rare illness or as deterministic as enrollment in a training program open freely to all members of the public.
While enrolled in program $j$, the government provides the individual with a non-taxable transfer of $b_j$. I define labour market income as $y(t, \omega_t, x)$, which can vary across different states of the world and individual decisions; it is important that I allow for the individual’s actions to influence this level of income, as this provides a channel through which a program, through its effects on $x$, can affect labour market income, thereby allowing for fiscal externalities. Additionally, an agent may be required to pay costs of program participation to some third party (for instance, tuition in the case of post-secondary education, or private health care expenditures), or may receive some income from untaxed sources; these will be denoted generally as $f(t, \omega_t, x)$, where a cost corresponds to a negative value of $f$.

The agent’s and planner’s problems have the same basic form as in Chetty (2006), complicated slightly by discounting; suppressing $x$ where it appears as an argument, the agent’s dynamic budget constraint is:

$$\dot{A}(t, \omega_t) = \log(1 + r)A(t, \omega_t) + f(t, \omega_t) + (1 - \tau)y(t, \omega_t) + \sum_{j=1}^{M} P_j(t, \omega_t) b_j - c(t, \omega_t)$$

where $\tau$ is the percentage tax rate on labour market income, and $A$ is the level of assets, with $\dot{A}$ representing the derivative of $A$ with respect to time. The individual also faces a terminal condition on assets, and a set of $N$ additional general constraints in each state and

---

3I limit my focus to state-contingent transfers because I want to consider policies which influence the individual’s decisions but which ensure that I can still use their first-order conditions to solve the model. A coercive policy of, for example, enforcing consumption of a quantity of education, presents difficulties in this analysis in that the quantities chosen can be corner solutions.

4Discounting is included for generality, but the final equation for the welfare derivative is identical if $r$ is assumed to be zero.

5I assume a proportional income tax for simplicity. I make the standard implicit assumption that there are some constraints, perhaps political in nature, which make it undesirable for the government to use a lump-sum tax, and once some sort of proportionality is assumed, the general intuition of my result is unavoidable. A progressive tax, or whatever form of tax may be found to be optimal, would change the results to some extent, but I use a proportional tax as a first-order approximation to both the existing and the optimal tax system.
time:

\[ A(t, \omega_t) \geq A_{\text{term}}, \quad \forall \omega_t \]

\[ g_{i\omega t}(c, x; b, \tau) \geq \bar{k}_{i\omega t}, \quad i = 1, \ldots, N. \]

where \( c \) is the set of state- and time-contingent choices of \( c(t, \omega_t) \). The \( N \) additional constraints are meant to represent any number of possible non-policy-generated distortions, such as borrowing constraints while unemployed or hours constraints while employed, as discussed by Chetty (2006); I will later place some restrictions upon these constraints.

The agent’s problem is to choose \( \{c, x\} \) to:

\[
\max V = \int_t \int_{\omega_t} e^{-rt} U(c(t, \omega_t), x(t, \omega_t)) dF_t(\omega_t) dt + \int_{\omega_1} \lambda_{\omega_1} [A(1, \omega_1) - A_{\text{term}}] dF_1(\omega_1)
\]

\[
+ \int_t \int_{\omega_t} \lambda_{\omega_t} \left[ \log(1 + r) A(t, \omega_t) + f(t, \omega_t) + (1 - \tau) y(t, \omega_t)
\right.
\]

\[
+ \sum_{j=1}^M P_j(t, \omega_t) b_j - c(t, \omega_t) - \dot{A}(t, \omega_t) \left] dF_t(\omega_t) dt
\]

\[
+ \sum_{i=1}^N \int_t \int_{\omega_t} \lambda_{gi\omega t} [g_{i\omega t}(c, x; b, \tau) - \bar{k}_{i\omega t}] dF_t(\omega_t) dt.
\]

Chetty’s Assumptions 1 and 2 ensure that the agent’s problem has a unique global maximum in his case, and they are also sufficient as well as plausible in my case, so I make them as well: they are that total lifetime utility is smooth, increasing and strictly quasiconcave in \( (c, x) \), and that the set of \( \{(c, x)\} \) which satisfy all the constraints is convex. Assumption 3 in Chetty (2006), which states that the set of binding constraints at the agent’s optimum does not change for a perturbation of \( b \) in \( (b - \varepsilon, b + \varepsilon) \), allows use of the envelope theorem to obtain \( \frac{dV}{db} \), and I also make the same assumption.
The optimal value for the agent’s problem is then denoted as $V(b, \tau)$, and the social planner will maximize this subject to the government budget constraint, which takes the following form:

$$
\tau \int_t \int_{\omega_t} e^{-rt} y(t, \omega_t) dF_t(\omega_t) dt = \int_t \int_{\omega_t} e^{-rt} \sum_{j=1}^{M} P_j(t, \omega_t) b_j dF_t(\omega_t) dt
$$

If I define $\bar{y} = \int_t \int_{\omega_t} e^{-rt} y(t, \omega_t) dF_t(\omega_t) dt$ as average discounted lifetime labour market income and $D_j = \int_t \int_{\omega_t} e^{-rt} P_j(t, \omega_t) dF_t(\omega_t) dt$ as the expected discounted fraction of the agent’s life spent enrolled in program $j$, I can rewrite the budget constraint as:

$$
\tau \bar{y} = \sum_{j=1}^{M} D_j b_j.
$$

Through the envelope theorem, I know that while a change in any element of $b$ will change individual choices in $x$, this has no direct first-order welfare effect, because the individual maximizes utility with respect to those choices; therefore, the government’s marginal value of increasing $b_j$ is:

$$
\frac{dV}{db_j} \bigg|_{db_j=0} = \frac{\partial V}{\partial b_j} + \frac{\partial V}{\partial \tau} \frac{d\tau}{db_j} \bigg|_{db_j=0}
$$

(3.4)

where $b_{-j} = \{b_1, \ldots, b_{j-1}, b_{j+1}, \ldots, b_M\}$.

If the government is free to vary all $M$ programs, one equation for each $j$ should be satisfied at the optimum. Alternatively, if political or other constraints prevent changing other programs, this equation provides information on welfare-increasing changes to one program, in the spirit of “piecemeal second-best policy” a la Lipsey (2007).

### 3.4.2 Calculation of Welfare Derivative

The next step is to evaluate (3.4), to derive a form that can be used for policy analysis; however, I first need to be able to express the partial derivatives in (3.4) in terms of marginal
utilities, and doing so requires some assumptions about how $b$ and $\tau$ affect the $N$ extra constraints. The assumption below, which is analogous to Assumption 5 from Chetty (2006), summarizes the conditions I require.

**Assumption 1.** *The feasible set of choices can be defined using a set of constraints such that, $\forall i, t, \omega_t$:*

\[
\frac{\partial g_{i\omega t}}{\partial b_j} = -P_j(t, \omega_t) \frac{\partial g_{i\omega t}}{\partial c(t, \omega_t)} \\
\frac{\partial g_{i\omega t}}{\partial \tau} = y(t, \omega_t) \frac{\partial g_{i\omega t}}{\partial c(t, \omega_t)} \\
\frac{\partial g_{i\omega t}}{\partial c(s, \omega_s)} = 0 \quad \forall t \neq s.
\]

The third part of the assumption simply states that consumption at two different times do not enter the same constraint; the first two parts, however, are slightly more complicated. The key is to remember that these are partial derivatives of the constraints, so I do not need to be concerned here about behavioural responses to $b$ and $t$. I assume that, if the agent is on program $j$, then raising $b_j$ by one unit has the same effect on the constraints as reducing consumption by one unit; in this way, program payments enter each constraint in the same way as consumption while on the program. Similarly, raising $\tau$ by one unit reduces disposable income by $y$, which has the same effect on the constraints as increasing consumption by $y$ units. Chetty (2009) argues that an assumption of this sort is typically satisfied in models “in which the private-sector choices are second-best efficient subject to the resource constraints,” because of fungibility of resources.
I can now proceed to evaluate the partial welfare derivatives:

\[
\frac{\partial V}{\partial b_j} = \int_t \int_{\omega_t} \left[ \lambda_{\omega_t} P_j(t, \omega_t) + \sum_{i=1}^{N} \lambda_{g_{i\omega_t}} \frac{\partial g_{i\omega_t}}{\partial b_j} \right] dF_t(\omega_t) dt \\
= \int_t \int_{\omega_t} P_j(t, \omega_t) \left[ \lambda_{\omega_t} - \sum_{i=1}^{N} \lambda_{g_{i\omega_t}} \frac{\partial g_{i\omega_t}}{\partial c(t, \omega_t)} \right] dF_t(\omega_t) dt
\]

\[
\frac{\partial V}{\partial \tau} = \int_t \int_{\omega_t} \left[ -\lambda_{\omega_t} y(t, \omega_t) + \sum_{i=1}^{N} \lambda_{g_{i\omega_t}} \frac{\partial g_{i\omega_t}}{\partial \tau} \right] dF_t(\omega_t) dt \\
= -\int_t \int_{\omega_t} y(t, \omega_t) \left[ \lambda_{\omega_t} - \sum_{i=1}^{N} \lambda_{g_{i\omega_t}} \frac{\partial g_{i\omega_t}}{\partial c(t, \omega_t)} \right] dF_t(\omega_t) dt.
\]

But since the agent has already maximized with respect to \(c\), I know that (suppressing the \(x\) from my notation)

\[e^{-rt} U'(c(t, \omega_t)) = \lambda_{\omega_t} - \sum_{i=1}^{N} \lambda_{g_{i\omega_t}} \frac{\partial g_{i\omega_t}}{\partial c(t, \omega_t)},\]

and therefore these partial derivatives can be written as:

\[
\frac{\partial V}{\partial b_j} = D_j E_j[U'(c)] \\
\frac{\partial V}{\partial \tau} = -E'[yU'(c)]
\]

where \(E_j[U'(c)] = \int_t \int_{\omega_t} P_j(t, \omega_t) e^{-rt} U'(c(t, \omega_t)) dF_t(\omega_t) dt / D_j\) is the expected value of \(U'(c(t, \omega_t))\) over the times and states in which the agent is enrolled in the program,\(^6\) and \(E'[yU'(c)] = \int_t \int_{\omega_t} y(t, \omega_t) e^{-rt} U'(c(t, \omega_t)) dF_t(\omega_t) dt\) is the discounted expected value of \(yU'(c)\).

These expressions are actually quite intuitive, as both are written in terms of marginal utilities of consumption, weighted by the amount of income gained or lost. \(\frac{\partial V}{\partial b_j}\) is a marginal benefit of increasing \(b_j\) by one unit, and this is equivalent in welfare terms to a one dollar increase in consumption at those times when the individual is on the program. Meanwhile, the marginal cost of increasing \(b_j\) comes from the resulting change in taxes, and when taxes

\(^6\)Although I have suppressed \(x\) from this notation, if the marginal utility of consumption can vary with \(x\), it will be important to keep that in mind when implementing my final formula. On the other hand, if utility is separable in \(c\) and \(x\), it is okay to ignore \(x\).
increase by one unit, this is equivalent in welfare terms to the marginal welfare cost of losing
\(y(t, \omega_t)\) of consumption at all times.

Next, I differentiate the government budget constraint with respect to \(b_j\):

\[
\tau \frac{d\bar{y}}{db_j} + \frac{d\tau}{db_j} = D_j + \sum_{l=1}^{M} b_l \frac{dD_l}{db_j}
\]

\[
\frac{d\tau}{db_j} \bigg|_{db_j=0} = \frac{D_j}{\bar{y}} \left[ 1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_l} \varepsilon_{b_l} - \frac{\tau \bar{y}}{D_j b_j} \varepsilon_{b_j} \right]
\]

\[
= \frac{D_j}{\bar{y}} \left[ 1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left( \varepsilon_{b_l} - \varepsilon_{b_j} \right) \right]
\]

(3.5)

The elasticity terms in square brackets are simple to interpret: we need to add up the effect of program \(j\) on spending on other programs and the effect on total income to determine the overall budgetary impact of program \(j\).\(^7\) If a higher \(b_j\) encourages people to spend longer on program \(j\) or on complementary programs, this means more time spent receiving payments and a larger required tax increase, whereas if higher \(b_j\) increases total income, this means more tax revenues paid to government and a smaller tax increase.\(^8\)

Therefore, the marginal value of increasing \(b_j\) can be expressed as:

\[
\frac{dV}{db_j} \bigg|_{db_j=0} = D_j E_j[U'(c)] - E[yU'(c)] \frac{D_j}{\bar{y}} \left[ 1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_l} \left( \varepsilon_{b_l} - \varepsilon_{b_j} \right) \right]
\]

\[
= D_j \left[ E_j[U'(c)] - \bar{y} E_y[U'(c)] \right] \left[ 1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left( \varepsilon_{b_l} - \varepsilon_{b_j} \right) \right]
\]

(3.6)

where \(E_y[U'(c)] = \frac{E[yU'(c)]}{\bar{y}}\) is the expected discounted income-weighted marginal utility. If

\(^7\)Moving from an approach in which other government programs are ignored, or lumped together in an exogenous \(G\) quantity as in chapter 1, to the current approach, means that we have to take into account any cross-program interactions. If, however, government programs do not interact with each other, equation (3.5) tells us that it is perfectly appropriate to evaluate each one individually.

\(^8\)These two effects are therefore analogous, respectively, to the “duration” effect and the two “revenue” effects in chapter 1.
I normalize the welfare derivative by \( \bar{E} \), I can also define:

\[
\frac{dW}{db_j} \equiv \frac{dV}{db_j} \bigg|_{db_j = 0} = D_j \left[ \frac{E_j[U'(c)] - E_y[U'(c)]}{E_y[U'(c)]} - \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left( \bar{\epsilon} b_j - \bar{\epsilon} b_j \right) \right]
\]

As in the analysis of post-secondary education in chapter 2, this expression can easily be understood as a tradeoff between the redistribution and fiscal effects of the program in question; the marginal utility ratio measures the welfare gain from taking a dollar from one person and giving it to another, while the sum of elasticities represents the overall fiscal impact of the transfer. The latter also represents the efficiency effects, or what Okun (1975) would describe as the leakiness of the bucket.

At the optimum, \( \frac{dV}{db_j} \bigg|_{db_j = 0} = 0 \) must be equal to zero,\(^9\) which means:

\[
E_j[U'(c)] = E_y[U'(c)] \left( 1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left( \bar{\epsilon} b_j - \bar{\epsilon} b_j \right) \right).
\] (3.7)

### 3.4.3 Analysis of Welfare Derivative

Clearly, to use equations (3.6) and (3.7) for practical policy-evaluation purposes, further assumptions are needed, and in the particular case of substitution between UI and DI that was studied earlier it is straightforward to show that (3.6) translates directly into (3.2) once the necessary assumptions are made: \( D_j = T_U, E_j[U'(c)] = U'(C_u), E_y[U'(c)] = U'(C_e), \) and

\[
- \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \varepsilon^{\bar{\epsilon} b_j} = \frac{\tau_U}{b_U} \varepsilon^{\frac{T_U}{b_U}} + \frac{\tau_D}{b_U} \varepsilon^{\frac{T_D}{b_U}}.
\]

However, in the current analysis, instead of imposing specific assumptions, I will simply assume that I have some way of evaluating the expected utility terms, so that I can use equations (3.6) and (3.7). I will now proceed to provide a series of results that parallel the analytical results in chapter 1. To begin with, let me denote

\(^9\)This is a necessary condition for a maximum; for \( \frac{dV}{db_j} \bigg|_{db_j = 0} = 0 \) to be unique and thus a sufficient condition for the optimum, \( V \) must be strictly quasi-concave in \( b_j \), which I assume to be the case.
\( \frac{dV}{db} (b_j; b_{-j}) \) as the welfare derivative at \( b_j \) with a vector \( b_{-j} \) of payments on other programs; then my first result is as follows.

**Proposition 6.** For \( b_{-j} > 0 \), \( \frac{dV}{db} (b_j; b_{-j}) - \frac{dV}{db} (b_j; 0) \) has the same sign as \( \bar{\varepsilon}_b \) if \( \sum_{l \neq j} D_l b_l \varepsilon_{D_l b_l} \).

**Proof.** Simple algebra leads to

\[
\frac{dV}{db} (b_j; b_{-j}) - \frac{dV}{db} (b_j; 0) = -D_j \bar{\varepsilon}_b U'(c) \sum_{l \neq j} D_l b_l \varepsilon_{D_l b_l} (\varepsilon_{D_l b_l} - \varepsilon_{b_l})
\]

and every term but the latter is positive. \( \square \)

\( \frac{dV}{db} (b_j; b_{-j}) > \frac{dV}{db} (b_j; 0) \) if and only if \( \varepsilon_{b_j} \sum_{l \neq j} D_l b_l \varepsilon_{D_l b_l} < 0 \); in words, taking into account the existence of other government programs will increase the optimal generosity of program \( j \) if and only if the effect of program \( j \) on tax revenues is greater than the weighted average impact of \( j \) on other program spending, weighted by the size of each other program. These could both be negative; for example, a program like unemployment insurance might reduce tax revenues, while potentially also reducing spending in other areas if it reduces substitution from UI onto DI or social assistance. A straightforward corollary of proposition 6 is that if fiscal externalities have been taken into account, but the substitution effects with other programs are then added to the analysis (as earlier in my study of UI with substitution from DI), \( \frac{dV}{db} \) will increase if and only if \( \sum_{l \neq j} D_l b_l \varepsilon_{D_l b_l} \sum_{l \neq j} D_l b_l < 0 \).

Equation (3.6) can be evaluated using real-world estimates of the various relevant quantities, thereby providing an estimate of the welfare derivative at the current real-world value of \( b = \{b_1, ..., b_M\} \); proposition 6, thus, tells us something about the local effect of program \( j \) on welfare around the baseline \( b \). However, to evaluate (3.7) for the optimal level of benefits, and to derive any analytical results about optimal policy, requires further assumptions. To begin with, when analyzing one program, it is important to consider whether the parameters of the other \( M - 1 \) programs are to be held fixed or allowed to vary. Our ultimate goal would
presumably be to find the optimal design and generosity for each program, but to solve for such an optimum using the sufficient statistics method would require strong statistical assumptions about the interactions of various programs. My goal, therefore, will be to provide results about “piecemeal second-best policy,” as advocated by Lipsey (2007), and so I will focus on the optimal policy for program \( j \) holding the generosity of other programs fixed.

A second issue is that I do not know what values the quantities in (3.7) will take if I change \( b_j \). Therefore, I approximate those values using the method of statistical extrapolation that I used in section 3.2, and which has also been used by Baily (1978), Gruber (1997), and the first two chapters of this dissertation. Chetty (2009) suggests statistical extrapolation as an alternative to calibrating and simulating a structural model: the available data and intuition is used to form the best estimate of how each of the sufficient statistics in (3.6) and (3.7) will respond to changes in \( b \). That is, if \( \chi = \{ E_j[U'(c)], E_{\bar{y}}[U'(c)], D_1, \ldots, D_M, \varepsilon_{b_j}^{D_1}, \ldots, \varepsilon_{b_j}^{D_M}, \varepsilon_{\bar{y}} \} \) represents all sufficient statistics in (3.6) and (3.7) other than \( b \), then \( \chi(b_j; b_{-j}) \) represents the assumed values of those statistics for a given value of \( b_j \). This definition of statistical extrapolations and Proposition 6 leads directly to the subsequent corollary about the optimal generosity of program \( j \), \( b_j^*(b_{-j}) \).

**Corollary 6.** For statistical extrapolations that do not vary with the assumed value of \( b_{-j} \), i.e. \( \chi(b_j; b_{-j}) = \chi(b_j) \), \( b_j^*(b_{-j}) > b_j^*(0) \) if and only if \( \varepsilon_{\bar{y}} > \frac{\sum_{i \neq j} D_i b_i \varepsilon_{b_i}^{D_i}}{\sum_{i \neq j} D_i b_i} \) in between \( b_j^*(0) \) and \( b_j^*(b_{-j}) \).

**Proof.** If we use a statistical extrapolation to find \( b_j^*(0) \), then using the same statistical

\[10\] This is why, in the context of UI, Chetty (2008) limits himself to using his equation to make a local analysis of the welfare derivative; he only calculates whether \( b \) should be smaller or larger.

\[11\] My assumption that \( V \) is strictly quasi-concave will, of course, place implicit restrictions on the permissible statistical extrapolations.
extrapolation, the estimate of $\frac{dV}{db_j}(b_j^*(0), b_{-j})$ takes the same sign as $\varepsilon_{b_j}^\bar{g} = \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_l}^D}{\sum_{l \neq j} D_l b_l}$. If this is positive, then strict quasi-concavity implies that $b_j^*(b_{-j}) > b_j^*(0)$, and that $\varepsilon_{b_j}^\bar{y} = \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_l}^D}{\sum_{l \neq j} D_l b_l}$ will continue to be positive at least until $b_j$ reaches $b_j^*(b_{-j})$; vice-versa if $\varepsilon_{b_j}^\bar{y} < \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_l}^D}{\sum_{l \neq j} D_l b_l}$. \[\Box\]

Therefore, if two researchers are attempting to implement (3.6) and/or (3.7), and agree on the statistical extrapolations to be used but disagree about the existence or size of other programs, with one assuming $b_{-j} = 0$ and the other assuming positive values, the latter researcher will estimate a larger welfare gain from raising $b_j$ and a higher optimal value of $b_j$ if and only if $\varepsilon_{b_j}^\bar{y} = \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_l}^D}{\sum_{l \neq j} D_l b_l}$. In words, if and only if raising $b_j$ has a positive external fiscal effect, by raising earnings more than it raises spending on other programs. This result is quite intuitive and easy to understand: when higher $b_j$ leads to higher total taxable income, this helps offset the fiscal externality and is a beneficial effect of the program, and when other spending is accounted for, the fiscal externality is also large and the beneficial aspect of the program is amplified. Meanwhile, if higher $b_j$ draws some of program $j$’s new participants from other programs, the increased spending on $j$ is offset by reduced spending elsewhere and the cost of the program is less severe.\[12\]

Next, I can prove a few simple results about the role of $\varepsilon_{b_j}^\bar{y}$, which follow below.

**Proposition 7.** For $\varepsilon_{b_j}^{\bar{y}2} > \varepsilon_{b_j}^{\bar{y}1}$, $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{y}2}) > \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{y}1})$.

**Proof.** Some simple algebra immediately gives us $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{y}2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{y}1}) = E\bar{g}[u'(c)] \times \left(\sum_{t=1}^M \frac{D_t b_t}{b_j}\right) \left(\varepsilon_{b_j}^{\bar{y}2} - \varepsilon_{b_j}^{\bar{y}1}\right)$, and all of the terms on the right-hand side are positive. \[\Box\]

\[12\]A key aspect of the sufficient statistics approach is that we do not need to account for the costs to individuals of receiving less transfers from other programs; the assumption is that individuals make their choices to maximize utility, and so if fewer people apply for a particular program, we can conclude that they are not directly worse off from making such a choice.
Corollary 7. For statistical extrapolations that do not vary with the assumed value of $\bar{\varepsilon}_{b_j}$, i.e. $\chi(b_j; \varepsilon_{b_j}^\theta) = \chi(b_j), b_j^*(\varepsilon_{b_j}^{\theta 2}) > b_j^*(\varepsilon_{b_j}^{\theta 1})$.

The proof to Corollary 7 is analogous to that for Corollary 6, and the results are quite easy to understand; if the effect of an increase in $b_j$ on total income is more positive, this increases the welfare gain from increasing $b_j$ and raises the optimal value of $b_j$. Finally, these results can be combined to show the following.

Proposition 8. For $b_{-j} > 0$ and $\varepsilon_{b_j}^{\theta 2} > \varepsilon_{b_j}^{\theta 1}$, $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\theta 2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\theta 1}) > \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\theta 2}) - \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\theta 1})$.

Proof. From the proof to Proposition 7, clearly $\left[ \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\theta 2}) - \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\theta 1}) \right] = 0$, and therefore $\left[ \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\theta 2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\theta 1}) \right] = E \left[ u'(c) \right] \left( \sum_{l=1}^{M} \frac{D_l b_l}{b_j} \right) \left( \varepsilon_{b_j}^{\theta 2} - \varepsilon_{b_j}^{\theta 1} \right) > 0$. 

In words, this means that the effect of $\varepsilon_{b_j}^{\theta}$ on the welfare derivative is increasing in $b_{-j}$; thus, when I allow for the possibility of fiscal externalities in my analysis of a government program, the question of whether or not the program increases total lifetime income becomes far more important. One might also like to know whether the effect of $\varepsilon_{b_j}^{\theta}$ on the optimal value of $b_j$ is increasing in $b_{-j}$, but this cannot be shown without additional unintuitive assumptions. However, Proposition 6 and Corollary 6 indicates that, if there is a cutoff value of $\varepsilon_{b_j}^{\theta}$ where fiscal externalities have no effect on the optimal $b_j$, a higher value of $\varepsilon_{b_j}^{\theta}$ will lead to a positive welfare derivative at the $b_j$ in question, meaning a higher optimal $b_j$ due to quasi-concavity, and vice-versa for a lower value of $\varepsilon_{b_j}^{\theta}$. Therefore, $b_j^*(b_{-j}) - b_j^*(b_{-j})$ follows a single-crossing property in $\varepsilon_{b_j}^{\theta}$: for large values of $\varepsilon_{b_j}^{\theta}$, fiscal externalities increase the optimal $b_j$, and vice-versa for small values of $\varepsilon_{b_j}^{\theta}$.

148
3.5 Applications of the General Approach

From the generality of the model described in the previous section, it is clear that just about any program which can be described as a state-contingent transfer would fit into this framework. In this section, I will present a brief survey of programs in the area of social insurance and human capital development which are obvious candidates for this analysis, including an overview of the existing literatures in each area.

3.5.1 Social Insurance

Social insurance programs have been the subject of extensive economic literatures, both in the area of empirical research about their labour market impacts and in the area of welfare analysis. Krueger and Meyer (2002) provide a survey of the labour market effects of social insurance programs, including UI, DI, Social Security, Workers’ Compensation (WC), and public health insurance, and most of their general conclusions remain representative of the literature today. UI and WC are typically found to have significant negative effects on employment, with elasticities of time out of work with respect to benefits of 0.5 to 1.\textsuperscript{13} Social Security and DI, meanwhile, do appear to reduce labour force participation among affected populations, but the effects are generally regarded as smaller and insufficient to explain the entire pattern of decreased labour force participation among the disabled and elderly in recent decades; see, for example, Chen and van der Klaauw (2008) and Blau and Goodstein (2010). Finally, the consequences of health insurance policy for the labour market are considered in a number of surveys, including Gruber and Madrian (2004) and Madrian (2005).

\textsuperscript{13}As described in chapter I, some empirical research also finds evidence that more generous UI leads to increased wages upon re-employment, but this remains a controversial question.
The literature performing welfare analysis of social insurance programs has made little recognition of fiscal externalities, as documented in the first chapter of this dissertation. But it is also the case that interactions between social insurance programs has not been incorporated in any substantive welfare analysis, although numerous empirical papers have considered them, mostly between UI, DI and WC. For example, substitution between UI and DI has been documented by Petrongolo (2009) and Lammers, Bloemen, and Hochguertel (2013), while I cited Lindner (2012) for his estimate of this substitution effect; Inderbitzin, Staubli, and Zweimüller (2012), meanwhile, show that both substitution and complementarity between UI and DI can occur at different ages in Austria due to details of the programs. Karlström, Palme, and Svensson (2008) and Staubli (2011) find that tightened DI eligibility leads to significant increases in receipt of other social insurance programs, while Borghans, Gielen, and Luttmer (2012) find that after a DI reform in the Netherlands in 1993, each dollar reduction in DI benefits was replaced 31 cents of other social insurance support. WC in particular is studied by Fortin and Lanoie (1992), who find evidence of substitution between UI and WC in Canada, as more generous UI reduces accident durations; Campolieti and Krashinsky (2003) find evidence of substitution between WC and DI in Canada, but McNerney and Simon (2012) do not find such evidence in the US.

This discussion, combined with the results for the case of UI and DI presented in section 3.2, highlights that research into the welfare implications of social insurance program interactions could be a fruitful area for future research.
### 3.5.2 Human Capital Development

Another important set of government policies where both program interaction effects and fiscal externalities are likely to be important are those designed to support the development of human capital, specifically education and job-training programs. Such policies are generally explicitly aimed at improving labour market outcomes, but education is also commonly thought to provide important non-production benefits: the survey in Lochner (2011) finds that “Education has been shown to reduce crime, improve health, lower mortality, and increase political participation.” Lochner acknowledges that most of the literature has focussed on the high school level, but Trostel (2010) finds that post-secondary education appears to reduce participation in social assistance and insurance programs, along with less corrections spending, with important fiscal benefits.\(^{14}\)

The second chapter of this dissertation presents an analysis of optimal tuition subsidy policy at the post-secondary level; there, I focus on the fiscal externalities generated by education’s positive effects on income, along with liquidity constraints, finding that substantially increased subsidies would improve welfare, with an optimal policy roughly corresponding to abolishing tuition at public universities. Although I use a specific and simplified model of post-secondary education in that paper, the resulting welfare derivative is in fact exactly equivalent to (3.6). My primary focus in that paper is not on interactions between tuition subsidies and other transfer programs, but I do take them into account by assuming that state appropriations per student are perfectly offset by reductions in spending elsewhere.

\(^{14}\)Trostel (2010) estimates that, in the U.S., direct public expenditures on PSE are about $71000 per degree in present value 2005 dollars, which is more than offset by expenditure savings of $56000 per degree (largely from reduced spending on corrections, Medicaid and social assistance) and increased tax revenues of $197000.
Based on estimates in Trostel (2010). In appendix 2.9.3 from the second chapter, I explicitly model the program interaction effects, and it is straightforward to show that (3.6) is exactly equivalent to equation (2.5).

The analysis in chapter 2 demonstrates that, when considering post-secondary education policy, it is especially important to measure and take into account effects of policy on labour market outcomes and participation on other programs. The same will generally be true of job-training programs, as such programs are aimed at improving labour market outcomes, often of lower-skilled individuals, and may have beneficial effects of substituting individuals away from other social programs. Numerous surveys and meta-analyses summarize the empirical literature that estimates the effects of training programs on labour market outcomes, including LaLonde (1995), Heckman, LaLonde, and Smith (1999), Greenberg, Michalopoulos, and Robins (2003), and Card, Kluve, and Weber (2010). The effects are usually found to be positive but small; LaLonde (1995) states that, given the modest amount of public investment in such programs, it looks like “we got what we paid for.”

LaLonde (1995) and Heckman, LaLonde, and Smith (1999) point out the possibility that training programs may lead to a reduction in welfare benefits, as well as reduced criminal activity, and output may be produced while in training, all of which could have beneficial fiscal effects. However, empirical examination of these effects have been limited. Therefore, there is considerable scope for future work that consider the full range of fiscal benefits of training programs, with the goal of evaluating impacts on social welfare.
3.6 Discussion and Conclusion

In this paper, I have considered the possibility that government social programs can interact with each other, so that changes in one program can lead to changes in enrollment on other programs. I then examined the importance of this program interaction effect on social welfare and optimal policy analysis, focusing on the specific case of unemployment insurance when unemployed individuals may also qualify for disability insurance. I show that accounting for this substitution can dramatically affect the conclusions from welfare analysis; if reduced generosity of UI increases applications for and enrollment on DI, this weakens the fiscal benefits of reducing UI found in chapter 1, and may well indicate that UI should be made considerably more generous.

I then move on to present a general model that allows for the consideration of interaction effects between a wide range of government transfer programs, which I describe in the final section of the paper, and to combine interaction effects with fiscal externalities. I provide an equation for the derivative of social welfare with respect to transfer generosity, and provide general results about the effect of program interactions on welfare calculations.

For future research, an important question must be answered: when we consider transfer programs, are we limited to considering programs as they currently exist today? Put differently, are program interaction effects unavoidable, or is it possible to target programs more effectively at the states they are designed to subsidize or insure? For example, my analysis of UI and DI, which finds that the generosity of UI should be increased to approach that of DI, raises the question of whether these and possibly other social insurance programs should be combined into one comprehensive social insurance policy, saving on administrative costs. This and other related areas of research are promising subjects for work in the future.


