The Welfare Implications of Fiscal Interactions Between Social Programs

Nicholas Lawson*
Aix-Marseille School of Economics

nicholas.lawson@univ-amu.fr

October 28, 2013

Abstract

The labour market impacts and welfare implications of social programs have been the focus of vast empirical and theoretical literatures. However, welfare analyses have largely failed to recognize the importance of fiscal linkages between government programs. In this paper, I demonstrate that fiscal interactions generated when programs act as substitutes or complements to each other can have dramatic effects on welfare calculations: changing the generosity of one program can lead to changes in enrollment on other programs, with important fiscal effects. In particular, I show that substitution between unemployment insurance (UI) and disability insurance can significantly increase the optimal UI replacement rate, from as low as 3% to 85%. I then present a general model which can be applied to any program of state-contingent transfers, and solve 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 interactions. I conclude by suggesting areas of research to which this idea and approach could be applied.

*I am very grateful to David Lee for many helpful comments and suggestions, and to Marc Sangnier for his comments. I also wish to thank the participants of the Industrial Relations Section Graduate Lunch seminar at Princeton University and the Ecolunch Seminar at the Aix-Marseille School of Economics. Any errors or omissions are the responsibilty of the author.

1 Introduction

The labour market impacts and welfare implications of social programs have been the focus of vast empirical and theoretical literatures.[1] Programs in areas from social insurance to education have been found to have important effects on a variety of labour market outcomes. However, the literature that performs welfare analyses of these programs has largely failed to recognize the importance of fiscal linkages between government programs. In this paper, I demonstrate that fiscal interactions generated when programs act as substitutes or complements for each other from the perspective of individual participants can have dramatic effects on welfare calculations.

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. A small empirical literature documents the fact that various social insurance programs exhibit patterns of substitution across programs, but the welfare implications have never been studied before.[2] I begin by considering the specific case of optimal unemployment insurance (UI) when some unemployed individuals may also be eligible for disability insurance (DI). If those individuals substitute between programs based on their relative generosity, then an increase in the generosity of UI benefits will reduce the number of DI claims, with beneficial effects on the government's fiscal situation. Using a plausibly small estimate of the substitution effect calculated by 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, with the optimal replacement rate rising from as low as 3% to around 85%.

I then present a general model which can be applied to any program that involves state-contingent transfers, and I show that results from the initial example generalize to this setting. The model can be solved for a derivative of social welfare with respect to any indi-

[1] For example, Krueger and Meyer (2002) survey the empirical literature on the labour market impacts of social insurance programs, while Lawson (2013a) and Lawson (2013b) discuss the empirical and welfare analysis literatures on social insurance and tuition subsidy programs respectively.

[2] Lawson (2013b) is a rare example of a paper that even accounts for the possibility of cross-program substitution, incorporating reductions in spending on social insurance and corrections into a welfare analysis of post-secondary tuition subsidies.

vidual program, with a simple and intuitive result that depends directly on the magnitude of fiscal externalities, or effects of a program on income tax revenues, and program interaction effects. I show that recognizing the full size of government and taking into account substitution effects with other programs causes our estimate of the optimal generosity of a transfer program to increase if and only if the effect of that program on income is greater than its effect on spending on other programs. Finally, I examine some of the areas of research to which this approach could profitably be applied, identifying important empirical and theoretical areas for future research.

This paper is a complement to earlier work on fiscal linkages in social programs in Lawson (2013a) and Lawson (2013b), where I demonstrate the importance of fiscal externalities in areas of UI and post-secondary tuition subsidies. Those papers combine with the current paper to provide us with a better understanding of the fiscal linkages between government programs, and together they demonstrate that modelling the full extent of government spending and taxation, and taking into account the way in which specific programs interact, has a first-order impact on welfare analysis of social programs.

The rest of the paper proceeds as follows. Section 2 presents an examination of optimal UI when individuals may substitute to or from DI. Section 3 then presents the general model and derives analytical results. Section 4 briefly summarizes literatures of particular importance where the method can be applied, and section 5 concludes and suggests areas for future research.

2 Optimal UI With Substitution to DI

Numerous papers have studied the question of optimal unemployment insurance, from Baily (1978) to Hansen and İmrohoroğlu (1992) to Chetty (2008); this literature seeks to balance the consumption-smoothing benefits of UI with the moral hazard costs of increased unemployment, and while the results have varied, the most typical result features an optimal replacement rate (or percentage of previous earnings received as benefits) on the order of 50%, which is close to the current generosity in most U.S. states.

In Lawson (2013a), 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. In particular, I show that the optimal replacement rate could drop significantly, perhaps to zero, once fiscal externalities are taken into account, unless wages are positively affected by UI generosity.

Here, I introduce another new element to optimal UI analysis: the interaction of UI and DI, among individuals for whom these programs 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 remaining in the labour force and receiving some combination of UI and employment income is another, and changes in the generosity of one program may affect enrollment on the other.

UI and DI are, in principle, mutually exclusive in that the former is aimed at individuals who are physically able to work but currently unemployed, while the latter is targetted at individuals who are physically unable to work. However, the disability evaluation is a subjective process, and Benitez-Silva, Buchinsky, and Rust (2004) provide evidence not only for false positives (non-disabled individuals being approved for DI) but also a significant amount of false negatives (rejections of disabled individuals). Meanwhile, the fact that UI and DI appear to be substitutes has been mentioned by Bound and Burkhauser (1999), and documented empirically by Petrongolo (2009) and Lammers, Bloemen, and Hochguertel (2013); Lindner (2012), in particular, estimates that a \$100 per month increase in UI benefits in the U.S. 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. In this section of the paper, I will rectify this omission. I begin in the first subsection by presenting the model that I will use, while the second subsection explains the empirical quantities used and provides numerical results.

2.1 Modified Baily (1978) Model

I base my analysis on a version of Baily (1978), as in section 3 of Lawson (2013a), but now the model will be modified to include DI. The model focusses on an ex-ante identical population of individuals whose labour market experience consists of two periods: in the first period, the representative 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. Individuals lose their job with probability γ, and if they lose their job, they have the option of applying for disability insurance and realizing a utility loss δ from stigma or effort costs of applying; in the population, δ will follow some distribution $F(\delta)$. The choice of applying for DI is denoted by $\theta = 1$, and if the individual is approved, which happens with probability α, 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.[3] s is chosen by the individual and subjects the individual to a utility cost of search $h(s)$ that is decreasing and convex in s (a less intense, less costly search will take longer); 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 discounting notation, I also denote $\int_x^y e^{-rt}dt = e_x^y$. The representative individual seeks to maximize expected utility, and the decision problem can therefore be written as:

$$\begin{aligned}
\max_{c_1,c_2,c_D,c_U,c_N,s,\theta} V = {} & e_0^1 U(c_1) + (1-\gamma)e_1^2 U(c_2) \\
& + \gamma \left[\theta\alpha e_1^2 U(c_D) + (1-\theta\alpha)(e_1^{1+s}U(c_U) + e_{1+s}^2 U(c_N) - h(s)) - \theta\delta\right] \\
& - \lambda_1 \left[e_0^1 c_1 + e_1^2 c_2 - e_0^2(1-\tau)\right] - \lambda_2 \left[e_0^1 c_1 + e_1^2 c_D - e_0^1(1-\tau) - e_1^2 b_D\right] \\
& - \lambda_3 \left[e_0^1 c_1 + e_1^{1+s} c_U + e_{1+s}^2 c_N - (e_0^1 + e_{1+s}^2)(1-\tau) - e_1^{1+s} b_U\right]
\end{aligned}$$

where τ is the tax rate and $U(c)$ follows the usual properties of $U' > 0$, $U'' < 0$.

The government is assumed to want to maximize ex-ante expected utility, which is equivalent to equally-weighted utility for all individuals across the range of potential outcomes.[4]

[3] By exogenously fixing the re-employment wage, I assume that there are no effects of UI on subsequent wages, to simplify the analysis; additionally, I explain in Lawson (2013a) that the recent empirical literature has tended to support that conclusion.

[4] Thus, the model could be rearranged into a one-period steady-state in which individuals are randomly

I am focussing on the optimal UI problem, so I hold b_D fixed and allow the government to choose the optimal value of b_U. As in Lawson (2013a), 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:

$$\frac{\partial V}{\partial b_U} = \lambda_3 e_1^{1+s} = \gamma(1-\theta\alpha)e_1^{1+s}U'(c_U)$$

$$\frac{\partial V}{\partial \tau} = -\lambda_1 e_0^2 - \lambda_2 e_0^1 - \lambda_3(e_0^1 + e_{1+s}^2) = -e_0^1 U'(c_1) - (1-\gamma)e_1^2 U'(c_2) - \gamma(1-\theta\alpha)e_{1+s}^2 U'(c_N)$$

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

$$\tau\left[e_0^2 - \gamma(1-\theta\alpha)e_1^{1+s} - \gamma\theta\alpha e_1^2\right] = \gamma(1-\theta\alpha)e_1^{1+s}b_U + \gamma\theta\alpha e_1^2 b_D$$

which, if I denote $T_U = \gamma(1-\theta\alpha)e_1^{1+s}$ and $T_D = \gamma\theta\alpha e_1^2$ as the expected discounted amounts of time spent on UI and DI respectively, can be rewritten as:

$$\tau\left(e_0^2 - T_U - T_D\right) = T_U b_U + T_D b_D \tag{1}$$

and therefore the derivative of the government budget constraint is:

$$\begin{aligned}\frac{d\tau}{db_U} &= \frac{1}{e_0^2 - T_U - T_D}\left[T_U + (b_U+\tau)\frac{dT_U}{db_U} + (b_D+\tau)\frac{dT_D}{db_U}\right]\\ &= \frac{T_U}{e_0^2 - T_U - T_D}\left[1 + \left(1+\frac{\tau}{b_U}\right)\varepsilon_{b_U}^{T_U} + \frac{T_D(b_D+\tau)}{T_U b_U}\varepsilon_{b_U}^{T_D}\right]\end{aligned}$$

where ε_x^y represents the elasticity of y with respect to x.

If I define c_e such that $(e_0^2 - T_U - T_D)U'(c_e) = e_0^1 U'(c_1) + (1-\gamma)e_1^2 U'(c_2) + \gamma(1-\theta\alpha)e_{1+s}^2 U'(c_N)$, so that $U'(c_e)$ is 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 $\frac{d\tau}{db_U}$, I get the following for $\frac{dV}{db_U}$:

$$\frac{dV}{db_U} = T_U U'(c_U) - T_U U'(c_e)\left[1 + \left(1+\frac{\tau}{b_U}\right)\varepsilon_{b_U}^{T_U} + \frac{T_D(b_D+\tau)}{T_U b_U}\varepsilon_{b_U}^{T_D}\right] \tag{2}$$

assigned into three categories: some are employed, some are unemployed and receiving UI, and some are on DI. The important margins that shift with UI are the probability of ending up on DI and the probability of being on UI at a particular point in time.

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

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

Finally, 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:

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

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

$$\frac{dW}{db_U} = T_U \left[R(c_e) \frac{\Delta c}{c_e} - \left(1 + \frac{\tau}{b_U}\right) \varepsilon_{b_U}^{T_U} - \frac{T_D(b_D + \tau)}{T_U b_U} \varepsilon_{b_U}^{T_D} \right] \qquad (3)$$

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

Inside the square brackets, the tradeoff is between the gain from consumption smoothing, which is increasing in the level of risk-aversion and the magnitude of the drop in consumption, and the fiscal effects of UI: more generous UI increases T_U, leading to longer durations of benefit payments and less time working and paying taxes, whereas if $\varepsilon_{b_U}^{T_D}$ is negative, more generous UI also reduces DI enrollment, with offsetting fiscal benefits.

To apply this formula, I simply need estimates of each of the quantities in (3) - the sufficient statistics - and then 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 unemployment benefits.

2.2 Sufficient Statistics and Numerical Results

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 so the average duration on DI of 14 years in Autor and Duggan (2006) implies a steady-state of 12.53 million, which makes $\gamma\theta\alpha = \frac{2\times 12.53}{161.4} = 0.1553$.[5] I will make a period equal to 14 years in my model, which means that $r = 0.5126$, and therefore $T_D = 0.1553e_1^2 = 0.0728$. Then, I use an unemployment rate of 5.4% as in Lawson (2013a), which means $\gamma(1-\theta\alpha)s = 0.108$; if I use the job-losing rate from one of the intermediate cases in Lawson (2013a), 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 Lawson (2013a), 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 $b_U = 0.46\left(\frac{12.64}{24.3}\right)0.85 = 0.2034$.[6] 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, applying discounting to the future Medicare benefits, $b_D = \left(1 + \frac{e_{8/7}^2}{e_1^2}\frac{600}{963}\right)0.46 = 0.6961$. This implies that the budget-balancing tax rate is $\tau = 0.0563$.

The elasticities are calculated as follows: I use $\varepsilon_{b_U}^{T_U} = 0.2544$ as in Lawson (2013a),[7] 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 $\varepsilon_{b_U}^{T_D} = \frac{-27}{895} = -0.0302$; I will also consider a value of zero for comparison. 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 of Lawson (2013a), and $\frac{\Delta c}{\bar{c}_e} = 0.222 -$

[5] I multiply by 2 because all of the DI spells in the model occur in the second period.

[6] As in Lawson (2013a), I assume a takeup rate of 80% and a ratio of compensated UI duration to total unemployment duration of $\frac{15.8}{24.3}$ as found by Chetty (2008), and $0.8 \times \frac{15.8}{24.3} = \frac{12.64}{24.3}$.

[7] Chetty (2008) finds an elasticity of unemployment durations with respect to UI of 0.53, but this number is multiplied by 0.48 as in Gruber (1997) to account for the fact that not all unemployed individuals receive benefits; 0.48 is the derivative of UI benefit receipt to benefit eligibility in Gruber's sample.

$0.265rr$, where rr is the UI replacement rate, as in Gruber (1997). The entire set of sufficient statistics is summarized in Table 1. If I put all of these estimates together, I get the results displayed in column 1 of Table 2. 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.

Table 1: Sufficient Statistics

Statistic	Definition	Value(s)
$\hat{T_U}$	baseline UI duration	0.0602
$\hat{T_D}$	baseline DI duration	0.0728
$\hat{b_U}$	adjusted UI benefit	0.2034
b_D	DI benefit	0.6961
$\varepsilon_{b_U}^{T_U}$	elasticity of UI duration	0.2544
$\varepsilon_{b_U}^{T_D}$	cross-elasticity of DI	$\{0, -0.0302\}$
R	coefficient of relative risk-aversion	2
$\frac{\Delta c}{c_e}$	consumption drop on UI	$0.222 - 0.265rr$
G	other government spending	$\{0, 0.2289\}$
$\hat{\tau}$	baseline tax rate	$\{0.0563, 0.261\}$

To this point I have ignored the fiscal externalities that were found to be so important in Lawson (2013a), but it is easy to incorporate them; if I allow for an additional amount of exogenous government spending equal to G, the government budget constraint (1) becomes $\tau\left(e_0^2 - T_U - T_D\right) = T_U b_U + T_D b_D + G$, and (3) is unchanged except that this new value of τ 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 $G = 0.2289$. Results in this case can be found in column 2 of Table 2.

In considering the results, it is first interesting to compare the optimal replacement rates in the two columns when $\varepsilon_{b_U}^{T_D} = 0$; 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 Lawson (2013a) on the importance of fiscal externalities.[8]

[8]The effect of fiscal externalities appears less dramatic here partly because of changes to various assump-

Table 2: Results from Sufficient Statistics and Extrapolation using (3)

	(1) $G=0$	(2) $G=0.2289$
	A. Estimate of $\frac{dW}{db_U}$ at $rr=0.46$	
$\varepsilon_{b_U}^{T_D}=0$	-0.0075	-0.0229
$\varepsilon_{b_U}^{T_D}=-0.0302$	0.0084	0.0233
	B. Optimal Replacement Rate	
$\varepsilon_{b_U}^{T_D}=0$	0.3315	0.0289
$\varepsilon_{b_U}^{T_D}=-0.0302$	0.6036	0.8499

The main result in this section, however, is the fact that the welfare derivative and the optimal level of UI both go up significantly when substitution onto DI is considered, because the fiscal benefits of lowering UI are greatly reduced. In column 1, the welfare derivative switches sign, and the optimal replacement rate increases from 33% to 60%, as more generous UI leads to a lower tax rate than would have been expected when $\varepsilon_{b_U}^{T_D}$ was assumed to be zero. However, the results are most dramatic when fiscal externalities are also accounted for; in column 2, since the tax rate is large, the tax savings from moving people off of DI are considerable, and the optimal replacement rate jumps from 3% to 85%. Thus, the optimal UI benefit is nearly double the baseline value, and becomes much more comparable to the size of the DI benefit.

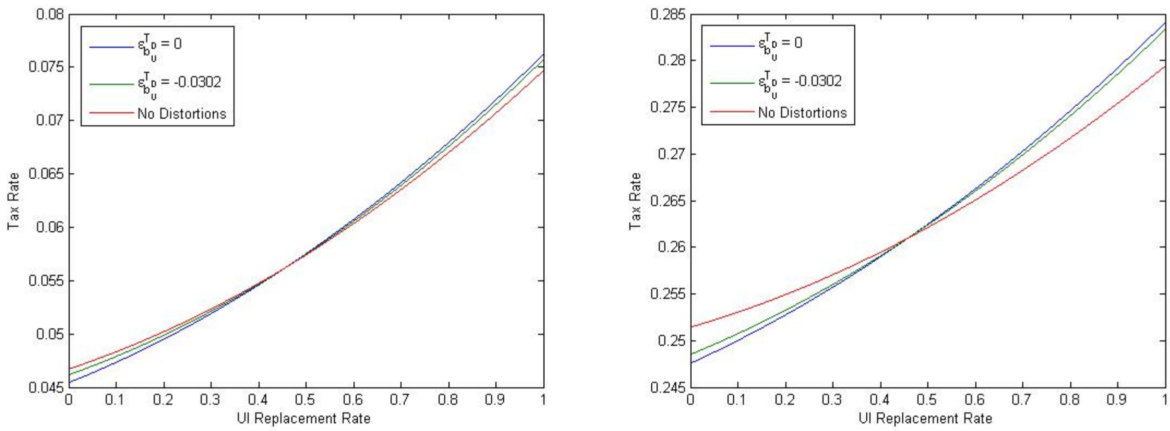
An illustration of the fiscal benefits from the interaction effect can be found in Figures 1 and 2. In Figure 1, the budget-balancing tax rates are compared across the two values of $\varepsilon_{b_U}^{T_D}$, and it can be seen that when there are substitution effects, the tax rate is less steep in rr. However, the difference might seem small until one considers the third line in Figure 1, which I label "No Distortions"; this illustrates the tax rate if the tax base was unaffected by UI. As seen in Figure 2, the substitution effects from DI are enough to offset a significant fraction of the distortions from UI: between 30% and 60% when G is assumed to be zero, or 15% to 25% with a positive G.

To summarize, if partially disabled individuals' decisions about whether or not to apply for DI are affected by the generosity of UI, because they are unemployed and/or may expect to be unemployed again in the future, generous UI may keep those individuals in the labour

tions, but mostly because the results in column 1 already include DI spending in the fiscal responsibilities of government.

Figure 1: Budget-Balancing Tax Rates

$G = 0$ $\qquad\qquad$ $G = 0.2289$



force, preventing them from receiving more generous DI benefits and providing an incentive to remain employed for at least part of the time. Both UI and DI are costly to the government to provide, but DI is considerably more costly, largely because it includes Medicare, and tends to be received for a longer period of time. Therefore, seeing unemployed individuals drop out of the labour force and receive DI is a bad situation becoming worse from the government's perspective, and, in the context of Aesop's fable, they will have an incentive to increase the generosity of the "frying pan" (UI) to prevent people from jumping to the "fire" (DI).

3 Theoretical Analysis of General Model

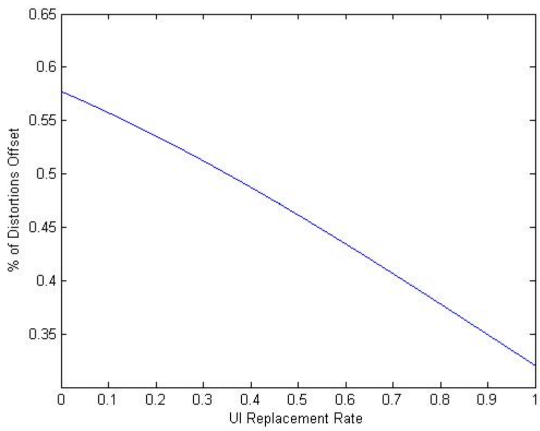
The results in the previous section show clearly the importance of taking into account program interaction effects when considering social insurance programs like UI. In this section, I move to a general case with a potentially large number of programs. I begin the theoretical analysis with a description of the general model; I then proceed to solve for the welfare derivative, and conclude with a discussion of the results.

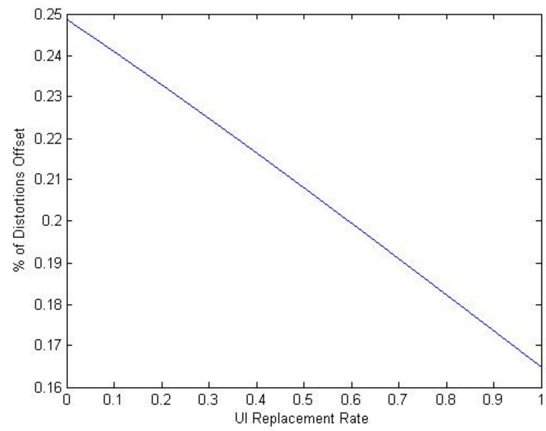
3.1 General Model

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

Figure 2: % of Tax Increase to Pay for Distortions Offset by Substitution Effect

$G = 0$ $G = 0.2289$





how the insights obtained may apply outside of the context of the basic social insurance problem.

As in section 2, the model features an ex-ante identical population of individuals, who may experience stochastic events across time, which is continuous with a unit duration, ie. $t \in [0,1]$, and represents the individual's working life (or some portion thereof). ω_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 ω_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 ω_t. As in the earlier example, I assume that the interest and discount rates are both equal to r.

Instead of focussing 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 ω_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.[9] I define labour market income as $y(t, \omega_t, x)$, which can vary across different states of the world and individual decisions; allowing the individual's actions to influence this level of income 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;[10] 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 τ is the percentage tax rate on labour market income,[11] 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

[9] I 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.

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

[11] I 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 necessarily follows.

time:

$$A(1,\omega_1) \geq A_{term}, \qquad \forall w_1$$

$$g_{i\omega t}(c,x;b,\tau) \geq \bar{k}_{i\omega t}, \qquad i=1,...,N.$$

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:

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

$$\int_t \int_{\omega_t} \lambda_{\omega t}[\log(1+r)A(t,\omega_t) + f(t,\omega_t) + (1-\tau)y(t,\omega_t) + \sum_{j=1}^{M} P_j(t,\omega_t) b_j - c(t,\omega_t) - \dot{A}(t,\omega_t)] dF_t(\omega_t) dt$$

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

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:

$$\frac{dV}{db_j}|_{db_{-j}=0} = \frac{\partial V}{\partial b_j} + \frac{\partial V}{\partial \tau}\frac{d\tau}{db_j}|_{db_{-j}=0} \tag{4}$$

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

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).[12]

3.2 Calculation of Welfare Derivative

The next step is to evaluate (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 (4) in terms of marginal utilities, and doing so requires some assumptions about how b and τ 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,* $\forall i, t, \omega_t$*:*

$$\frac{\partial g_{i\omega t}}{\partial b_j} = -P_j(t, \omega_t)\frac{\partial g_{i\omega t}}{\partial c(t, \omega_t)}$$
$$\frac{\partial g_{i\omega t}}{\partial \tau} = y(t, \omega_t)\frac{\partial g_{i\omega t}}{\partial c(t, \omega_t)}$$

[12] In the current paper, I seek context-specific welfare improvements 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. 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.

$$\frac{\partial g_{i\omega t}}{\partial c(s,\omega_s)} = 0 \quad \forall t \neq s.$$

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 τ 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:

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

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

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:

$$\begin{aligned}\frac{\partial V}{\partial b_j} &= D_j E_j[U'(c)] \\ \frac{\partial V}{\partial \tau} &= -E^r[yU'(c)]\end{aligned}$$

where $E_j[U'(c)] = \frac{\int_t \int_{\omega_t} P_j(t,\omega_t) e^{-rt} U'(c(t,\omega_t) dF_t(\omega_t) dt}{D_j}$ is the expected discounted value of $U'(c(t,\omega_t))$

over the times and states in which the agent is enrolled in the program,[13] and $E^r[yU'(c)] = \int_t \int_{\omega_t} y(t,\omega_t)e^{-rt}U'(c(t,\omega_t))dF_t(\omega_t)dt$ is the expected discounted 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 the 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 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:

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

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

$$= \frac{D_j}{\bar{y}}\left[1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j}\left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}}\right)\right] \quad (5)$$

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.[14] 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 increase in the tax rate.[15]

[13] Although I have suppressed x from this notation, if the marginal utility of consumption varies 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.

[14] If government programs do not interact with each other, equation (5) tells us that it is appropriate to evaluate each one individually.

[15] These two effects are therefore analogous, respectively, to the "duration" effect and the "revenue" effects in Lawson (2013a).

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

$$\begin{aligned}\frac{dV}{db_j}|_{db_{-j}=0} &= D_j E_j[U'(c)] - E^r[yU'(c)]\frac{D_j}{\bar{y}}\left[1 + \sum_{l=1}^{M}\frac{D_l b_l}{D_j b_j}\left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}}\right)\right] \\ &= D_j\left[E_j[U'(c)] - E_{\bar{y}}[U'(c)]\left(1 + \sum_{l=1}^{M}\frac{D_l b_l}{D_j b_j}\left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}}\right)\right)\right] \qquad (6)\end{aligned}$$

where $E_{\bar{y}}[U'(c)] = \frac{E^r[yU'(c)]}{\bar{y}}$ is the expected discounted income-weighted marginal utility. If I normalize the welfare derivative by $E_{\bar{y}}[U'(c)]$, I can also define:

$$\frac{dW}{db_j} \equiv \frac{\frac{dV}{db_j}|_{db_{-j}=0}}{E_{\bar{y}}[U'(c)]} = D_j\left[\frac{E_j[U'(c)] - E_{\bar{y}}[U'(c)]}{E_{\bar{y}}[U'(c)]} - \sum_{l=1}^{M}\frac{D_l b_l}{D_j b_j}\left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}}\right)\right]$$

As in Lawson (2013b), 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 effect, or what Okun (1975) would describe as the leakiness of the bucket.

At the optimum, $\frac{dV}{db_j}|_{db_{-j}=0}$ must be equal to zero,[16] which means:

$$E_j[U'(c)] = E_{\bar{y}}[U'(c)]\left(1 + \sum_{l=1}^{M}\frac{D_l b_l}{D_j b_j}\left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}}\right)\right). \qquad (7)$$

3.3 Analysis of Welfare Derivative

Clearly, to use equations (6) and (7) for practical policy-evaluation purposes, further assumptions are needed, and in the particular case of substitution between UI and DI studied earlier it is straightforward to show that (6) translates directly into (2) once the necessary assumptions are made: $D_j = T_U$, $E_j[U'(c)] = U'(C_u)$, $E_{\bar{y}}[U'(c)] = U'(C_e)$, and $-\sum_{l=1}^{M}\frac{D_l b_l}{D_j b_j}\varepsilon_{b_j}^{\bar{y}} = \frac{\tau}{b_U}\varepsilon_{b_U}^{T_U} + \frac{\tau T_D}{T_U b_U}\varepsilon_{b_U}^{T_D}$. 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 (6) and (7). I will now proceed to provide a series of results that parallel the analytical results in Lawson (2013a). To begin with, let me denote

[16] This is a necessary condition for a maximum; for $\frac{dV}{db_j}|_{db_{-j}=0} = 0$ 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.

$\frac{dV}{db_j}(b_j;b_{-j})$ 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 1. *For $b_{-j}>0$, $\frac{dV}{db_j}(b_j;b_{-j})-\frac{dV}{db_j}(b_j;0)$ has the same sign as $\varepsilon^{\bar{y}}_{b_j}-\frac{\sum_{l\neq j}D_l b_l \varepsilon^{D_l}_{b_j}}{\sum_{l\neq j}D_l b_l}$.*

Proof. Simple algebra gives $\frac{dV}{db_j}(b_j;b_{-j})-\frac{dV}{db_j}(b_j;0)=-D_j E_{\bar{y}}[U'(c)]\sum_{l\neq j}\frac{D_l b_l}{D_j b_j}\left(\varepsilon^{D_l}_{b_j}-\varepsilon^{\bar{y}}_{b_j}\right)=\frac{E_{\bar{y}}[U'(c)]\sum_{l\neq j}D_l b_l}{b_j}\left(\varepsilon^{\bar{y}}_{b_j}-\frac{\sum_{l\neq j}D_l b_l \varepsilon^{D_l}_{b_j}}{\sum_{l\neq j}D_l b_l}\right)$, and every term but the latter is positive. □

$\frac{dV}{db_j}(b_j;b_{-j})>\frac{dV}{db_j}(b_j;0)$ if and only if $\varepsilon^{\bar{y}}_{b_j}>\frac{\sum_{l\neq j}D_l b_l \varepsilon^{D_l}_{b_j}}{\sum_{l\neq j}D_l b_l}$; 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 also reducing spending in other areas if it reduces substitution from UI onto DI or social assistance. A straightforward corollary of proposition 1 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), $\frac{dV}{db_j}$ will increase if and only if $\frac{\sum_{l\neq j}D_l b_l \varepsilon^{D_l}_{b_j}}{\sum_{l\neq j}D_l b_l}<0$.

Equation (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 1, thus, tells us something about the local effect of program j on welfare around the baseline b. However, to evaluate (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. Although one may wish to find the optimal design and generosity for each program, 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 (7) will take if I change

b_j.[17] Therefore, I propose approximating those values using the method of statistical extrapolation that I used in section 2, and which has also been used by Baily (1978), Gruber (1997), Lawson (2013a) and Lawson (2013b). Chetty (2009) suggests statistical extrapolation as an alternative to calibrating and simulating a structural model: the available data and intuition are used to form the best estimate of how each of the sufficient statistics in (6) and (7) will respond to changes in b.[18] That is, if $\chi = \{E_j[U'(c)], E_{\bar{y}}[U'(c)], D_1, ..., D_M, \varepsilon_{b_j}^{D_1}, ..., \varepsilon_{b_j}^{D_M}, \varepsilon_{b_j}^{\bar{y}}\}$ represents all sufficient statistics in (6) and (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 1 leads directly to the subsequent corollary about the optimal generosity of program j, $b_j^*(b_{-j})$.

Corollary 1. *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* $\varepsilon_{b_j}^{\bar{y}} > \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_j}^{D_l}}{\sum_{l \neq j} D_l b_l}$ *in between* $b_j^*(0)$ *and* $b_j^*(b_{-j})$*.*

Proof. If we use a statistical extrapolation to find $b_j^*(0)$, then the estimate of $\frac{dV}{db_j}(b_j^*(0), b_{-j})$ using the same statistical extrapolation takes the same sign as $\varepsilon_{b_j}^{\bar{y}} - \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_j}^{D_l}}{\sum_{l \neq j} D_l b_l}$. If this is positive, then strict quasi-concavity implies that $b_j^*(b_{-j}) > b_j^*(0)$, and that $\varepsilon_{b_j}^{\bar{y}} - \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_j}^{D_l}}{\sum_{l \neq j} D_l b_l}$ will continue to be positive at least until b_j reaches $b_j^*(b_{-j})$; vice-versa if $\varepsilon_{b_j}^{\bar{y}} < \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_j}^{D_l}}{\sum_{l \neq j} D_l b_l}$. $\square$

Therefore, if two researchers are attempting to implement (6) and/or (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 $\varepsilon_{b_j}^{\bar{y}} > \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_j}^{D_l}}{\sum_{l \neq j} D_l b_l}$; 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 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

[17] 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.

[18] My assumption that V is strictly quasi-concave will place implicit restrictions on the permissable statistical extrapolations.

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.[19]

Next, let me define $\varepsilon_{b_j}^{\bar{D}_l} = \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_j}^{D_l}}{\sum_{l \neq j} D_l b_l}$, and I can prove a few simple results about $\varepsilon_{b_j}^{\bar{y}}$ and $\varepsilon_{b_j}^{\bar{D}_l}$, which follow below.

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

Proof. (i) Some simple algebra immediately gives us $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{y}2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{y}1}) = E_{\bar{y}}[u'(c)] \left(\sum_{l=1}^{M} \frac{D_l b_l}{b_j}\right) \left(\varepsilon_{b_j}^{\bar{y}2} - \varepsilon_{b_j}^{\bar{y}1}\right)$, and all of the terms on the right-hand side are positive.
(ii) $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l 2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l 1}) = E_{\bar{y}}[u'(c)] \left(\sum_{l=1}^{M} \frac{D_l b_l}{b_j}\right) \left(\varepsilon_{b_j}^{\bar{D}_l 1} - \varepsilon_{b_j}^{\bar{D}_l 2}\right)$, and only the final term is negative. □

Corollary 2. *(i) For statistical extrapolations that do not vary with the assumed value of* $\varepsilon_{b_j}^{\bar{y}}$, *i.e.* $\chi(b_j; \varepsilon_{b_j}^{\bar{y}}) = \chi(b_j)$, $b_j^*(\varepsilon_{b_j}^{\bar{y}2}) > b_j^*(\varepsilon_{b_j}^{\bar{y}1})$.
(ii) For statistical extrapolations that do not vary with the assumed value of $\varepsilon_{b_j}^{\bar{D}_l}$, $b_j^*(\varepsilon_{b_j}^{\bar{D}_l 2}) < b_j^*(\varepsilon_{b_j}^{\bar{D}_l 1})$.

The proof to Corollary 2 is analogous to that for Corollary 1, and the results are 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, and vice-versa if b_j requires a greater spending increase in other areas. Finally, these results can be combined to show the following.

Proposition 3. *(i) For* $b_{-j} > 0$ *and* $\varepsilon_{b_j}^{\bar{y}2} > \varepsilon_{b_j}^{\bar{y}1}$, $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{y}2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{y}1}) > \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\bar{y}2}) - \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\bar{y}1})$.
(ii) For $b_{-j} > 0$ *and* $\varepsilon_{b_j}^{\bar{D}_l 2} > \varepsilon_{b_j}^{\bar{D}_l 1}$, $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l 2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l 1}) < \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\bar{D}_l 2}) - \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\bar{D}_l 1})$.

[19] 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.

Proof. (i) From the proof to Proposition 2, it is immediate that $\left[\frac{dV}{db_j}(b_j;0,\varepsilon_{b_j}^{\bar{y}2}) - \frac{dV}{db_j}(b_j;0,\varepsilon_{b_j}^{\bar{y}1})\right]$ $= 0$, and therefore $\left[\frac{dV}{db_j}(b_j;b_{-j},\varepsilon_{b_j}^{\bar{y}2}) - \frac{dV}{db_j}(b_j;b_{-j},\varepsilon_{b_j}^{\bar{y}1})\right] = E_{\bar{y}}[u'(c)]\left(\sum_{l=1}^{M}\frac{D_l b_l}{b_j}\right)\left(\varepsilon_{b_j}^{\bar{y}2} - \varepsilon_{b_j}^{\bar{y}1}\right) > 0$.

(ii) $\left[\frac{dV}{db_j}(b_j;0,\varepsilon_{b_j}^{\bar{D}_l 2}) - \frac{dV}{db_j}(b_j;0,\varepsilon_{b_j}^{\bar{D}_l 1})\right] = 0$ and $\frac{dV}{db_j}(b_j;b_{-j},\varepsilon_{b_j}^{\bar{D}_l 2}) - \frac{dV}{db_j}(b_j;b_{-j},\varepsilon_{b_j}^{\bar{D}_l 1}) < 0$. □

In words, this means that the absolute values of the effects of $\varepsilon_{b_j}^{\bar{y}}$ and $\varepsilon_{b_j}^{\bar{D}_l}$ on the welfare derivative are increasing in b_{-j}; thus, when I increase the amount of other spending that I account for in my analysis of a government program, the question of whether or not the program increases total lifetime income or spending on other programs becomes far more important. One might also like to know whether the effect of $\varepsilon_{b_j}^{\bar{y}}$ or $\varepsilon_{b_j}^{\bar{D}_l}$ on the optimal value of b_j is increasing in b_{-j}, but this cannot be shown without additional unintuitive assumptions. However, Proposition 1 and Corollary 1 indicate that $b_j^*(b_{-j}) - b_j^*(b_{-j})$ follows a single-crossing property in both $\varepsilon_{b_j}^{\bar{y}}$ and $\varepsilon_{b_j}^{\bar{D}_l}$: for large values of $\varepsilon_{b_j}^{\bar{y}}$ and small values of $\varepsilon_{b_j}^{\bar{D}_l}$, fiscal externalities increase the optimal b_j, and vice-versa for small values of $\varepsilon_{b_j}^{\bar{y}}$ and large values of $\varepsilon_{b_j}^{\bar{D}_l}$.

4 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.

4.1 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.[20] 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).

The literature performing welfare analysis of social insurance programs has made little recognition of fiscal externalities, as documented in Lawson (2013a). But it is also the case that interactions between social insurance programs have 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 the duration of accident compensation; Campolieti and Krashinsky (2003) find evidence of substitution between WC and DI in Canada, but McInerney 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 2, highlights that the welfare implications of social insurance program interactions could be an important area for future research.

[20] As described in Lawson (2013a), some empirical research also finds evidence that more generous UI leads to increased wages upon re-employment, but this remains a controversial question.

4.2 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.[21]

Lawson (2013b) 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 (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, based on estimates in Trostel (2010).

The analysis in Lawson (2013b) demonstrates that, when considering post-secondary education policy, it is 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 em-

[21] 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.

pirical 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 considers the full range of fiscal benefits of training programs, with the goal of evaluating impacts on social welfare.

5 Discussion and Conclusion

In this paper, I have considered the possibility that social programs can interact with each other, so that changes in one program can lead to changes in enrollment on other programs. I examine the importance of this program interaction effect on social welfare and optimal policy analysis, focussing 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 by Lawson (2013a), 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 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.

One area of study that should be considered in greater detail in the future is mentioned in the final section: empirical and welfare analysis of the program interaction effects of social insurance and human capital development programs. There is also a larger question that needs further study: what do my results tell us about the optimal shape of social insurance policy? That is, are we limited to considering programs as they currently exist today? Are program interaction effects unavoidable, or is it possible to target programs more effectively at the states they are designed to subsidize or insure? My analysis indicates that the generosity of UI should be increased to approach that of DI, and thus given the administrative costs and the waiting time involved in applying for and being evaluated for DI, understanding how DI, UI and other social insurance programs can best be designed or perhaps combined is a promising subject for future work.

References

Autor, D. H., and M. G. Duggan (2006): "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding," *Journal of Economic Perspectives*, 20(3), 71–96.

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

Benitez-Silva, H., M. Buchinsky, and J. Rust (2004): "How Large are the Classification Errors in the Social Security Disability Award Process?," Working Paper no. 10219, NBER, Cambridge, MA.

Blau, D. M., and R. M. Goodstein (2010): "Can Social Security Explain Trends in Labor Force Participation of Older Men in the United States?," *Journal of Human Resources*, 45(2), 328–363.

Borghans, L., A. C. Gielen, and E. F. Luttmer (2012): "Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform," Working Paper no. 18261, NBER, Cambridge, MA.

Bound, J., and R. V. Burkhauser (1999): "Economic Analysis of Transfer Programs Targeted on People with Disabilities," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter, and D. Card, vol. 3C, Elsevier, pp. 3417–3528.

Campolieti, M., and H. Krashinsky (2003): "Substitution between Disability Support Programs in Canada," *Canadian Public Policy*, 29(4), 417–430.

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

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

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

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

——— (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–487.

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

Fortin, B., and P. Lanoie (1992): "Substitution Between Unemployment Insurance and Workers' Compensation," *Journal of Public Economics*, 49(3), 287–312.

Greenberg, D. H., C. Michalopoulos, and P. K. Robins (2003): "A Meta-Analysis of Government-Sponsored Training Programs," *Industrial and Labor Relations Review*, 57(1), 31–53.

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

Gruber, J., and B. C. Madrian (2004): "Health Insurance, Labor Supply, and Job Mobility: A Critical Review of the Literature," in *Health Policy and the Uninsured*, ed. by C. G. McLaughlin, pp. 97–177. Urban Institute Press.

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

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999): "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter, and D. Card, vol. 3A, Elsevier, pp. 1865–2097.

Inderbitzin, L., S. Staubli, and J. Zweimüller (2012): "Extended Unemployment Benefits and Early Retirement: Program Complementarity and Program Substitution," Unpublished Paper.

Karlström, A., M. Palme, and I. Svensson (2008): "The Employment Effect of Stricter Rules for Eligibility for DI: Evidence from a Natural Experiment in Sweden," *Journal of Public Economics*, 92(10-11), 2071–2082.

Krueger, A. B., and B. D. Meyer (2002): "Labor Supply Effects of Social Insurance," in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. Feldstein, vol. 4, Elsevier, pp. 2327–2392.

LaLonde, R. J. (1995): "The Promise of Public Sector-Sponsored Training Programs," *Journal of Economic Perspectives*, 9(2), 149–168.

Lammers, M., H. Bloemen, and S. Hochguertel (2013): "Job Search Requirements for Older Unemployed: Transitions to Employment, Early Retirement and Disability Benefits," *European Economic Review*, 58, 31–57.

Lawson, N. (2013a): "Fiscal Externalities and Optimal Unemployment Insurance," Unpublished Paper, Aix-Marseille School of Economics.

——— (2013b): "Fiscal Externalities, Liquidity Constraints and Grants to Post-Secondary Students," Unpublished Paper, Aix-Marseille School of Economics.

Lindner, S. (2012): "How Do Unemployment Insurance Benefits Affect the Decision to Apply for Social Security Disability Insurance?," Unpublished Paper, The Urban Institute.

Lipsey, R. G. (2007): "Reflections on the General Theory of Second Best at its Golden Jubilee," *International Tax and Public Finance*, 14(4), 349–364.

Lochner, L. (2011): "Nonproduction Benefits of Education: Crime, Health, and Good Citizenship," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Woessmanm, vol. 4, Elsevier, pp. 183–282.

Madrian, B. C. (2005): "The U.S. Health Care System and Labor Markets," *Federal Reserve Bank of Boston Research Conference Series*, 50, 137–163.

McInerney, M., and K. Simon (2012): "The Effect of State Workers' Compensation Program Changes on the Use of Federal Social Security Disability Insurance," *Industrial Relations*, 51(1), 57–88.

Okun, A. M. (1975): *Equality and Efficiency: The Big Tradeoff*. The Brookings Institution, Washington, D.C.

Petrongolo, B. (2009): "The Long-Term Effects of Job Search Requirements: Evidence from the UK JSA Reform," *Journal of Public Economics*, 93(11-12), 1234–1253.

Rutledge, M. S. (2011): "The Impact of Unemployment Insurance Extensions on Disability Insurance Application and Allowance Rates," Working Paper no. 2011-17, Center for Retirement Research at Boston College.

Staubli, S. (2011): "The Impact of Stricter Criteria for Disability Insurance on Labor Force Participation," *Journal of Public Economics*, 95(9-10), 1223–1235.

Trostel, P. A. (2010): "The Fiscal Impacts of College Attainment," *Research in Higher Education*, 51(3), 220–247.