Working Papers / Documents de travail

Social Program Substitution and Optimal Policy

Nicholas Lawson

WP 2014 - Nr 17

Social Program Substitution and Optimal Policy

Nicholas Lawson*
Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS
nicholas.lawson@univ-amu.fr

May 16, 2014

Abstract

A growing literature on substitution between social programs provides consistent evidence that changes in the generosity of one program can lead to changes in enrollment on other programs. However, this evidence has been ignored in welfare analyses of social insurance programs. I demonstrate that substitutions between programs can dramatically alter conclusions about optimal policy, with a particular focus on optimal unemployment insurance (UI) when there is substitution between UI and disability insurance (DI). If more generous UI reduces enrollment on DI, the result is a reduction in government spending on DI, and I show that this effect can significant increase the optimal UI replacement rate from 3% to 85%.

Keywords: fiscal interactions, program substitution, optimal unemployment insurance, disability insurance

*I am very grateful to David Lee for many helpful comments and suggestions, and to Charles Beach, Cecilia Garcia-Peñalosa, and Marc Sangnier for their comments. I also wish to thank the participants of seminars at Princeton University, the Aix-Marseille School of Economics, the 2014 Royal Economic Society Conference, and the 2014 Society of Labor Economists Conference. Any errors or omissions are the responsibilty of the author.

1 Introduction

While a great deal of effort has been put into understanding the impacts of social programs on labour market outcomes such as labour force participation, unemployment, and wages,[1] until recently the literature studying interactions between social programs has been much smaller. However, a growing literature now documents significant evidence of substitution and other interactions between different government programs, particularly in the area of social insurance, where there is consistent evidence that changing the generosity of one program affects enrollment and spending on other programs. This evidence has been ignored by existing welfare analyses of social insurance programs. In this paper, I demonstrate that accounting for substitution between programs can dramatically alter conclusions about optimal policy. In particular, I study the optimal generosity of unemployment insurance (UI) benefits and show that the optimal benefit is greatly increased once substitution between UI and disability insurance (DI) is accounted for.

Several existing studies have found evidence of substitution between UI and DI, with increased generosity of one program reducing enrollment on the other: Petrongolo (2009), Lindner (2012), and Lammers, Bloemen, and Hochguertel (2013) document such substitution in the UK, USA, and Netherlands respectively. Other work confirms substitution or other interactions between DI and other social insurance programs, including Staubli (2011) and Borghans, Gielen, and Luttmer (2014), and Fortin and Lanoie (1992) is one of a number of papers finding substitution between UI and workers' compensation programs. In principle, these programs are generally designed to be mutually exclusive; in the United States, the focus of my analysis, UI is aimed at individuals who lose their job but are physically able to work, while DI is targetted at individuals who are physically unable to work. However, the disability evaluation is a subjective process (as might be the evaluation of eligibility for other programs), and for the US DI system Benitez-Silva, Buchinsky, and Rust (2004) provide evidence for significant quantities of both false negatives (rejections of disabled individuals) and false positives (non-disabled individuals being approved for DI).

In a second-best world in which we cannot perfectly target transfers at the intended

[1] See, for example, the survey of social insurance in Krueger and Meyer (2002).

states of the world, some amount of substitution between programs will be unavoidable, and will have important fiscal effects. For example, increases in UI benefits reduce enrollment on DI and thus reduce government spending on DI. However, this fiscal interaction between UI and DI has been ignored in welfare analyses. Indeed, in the area of social insurance more broadly, only Li (2014) has considered the welfare consequences of a fiscal interaction between programs, but focusing on a unique channel which differs from the one considered in this paper: Li presents a calibrated model which implies that the Affordable Care Act will improve health and thus reduce DI eligibility.[2] The standard assumption when studying optimal policy for a particular social insurance program is that the program in question is the only fiscal responsibility of government, completely ignoring all other programs. Thus, optimal UI studies such as Hansen and İmrohoroğlu (1992) and Chetty (2008) assume that the government uses a small payroll tax to pay for UI benefits, abstracting from all other activities of government, and a similar simplifying assumption is made for many studies of the welfare implications of DI (see Golosov and Tsyvinski (2006), for example) and Social Security (see Feldstein (1985), for example).[3]

I provide the first welfare analysis of unemployment insurance that accounts for an important source of substitution, specifically disability insurance. Previous analysis of optimal UI has used a variety of methods, all of which aim to balance the consumption-smoothing benefits of UI with the moral hazard costs of increased unemployment, with a median result featuring a replacement rate on the order of 50%.[4] Lawson (2013) demonstrates that these results are altered once we account for fiscal externalities, or effects of UI on the tax base, with the optimal replacement rate possibly dropping as far as zero. In this paper, using a plausibly small estimate of the substitution effect calculated by Lindner (2012), I show

[2]Lawson (2014) also incorporates reductions in spending on social insurance and corrections into a welfare analysis of post-secondary tuition subsidies.

[3]Lawson (2013) points out that this abstraction ignores the important role of fiscal externalities generated by the impact of a particular program on the tax base, but does not consider interactions with other programs. Some studies of social insurance programs do account for the existence of other government spending in their analyses, by estimating the fiscal effects of the program in question using an estimated real-world tax rate; examples include İmrohoroğlu, İmrohoroğlu, and Joines (2003) on Social Security and Bound, Cullen, Nichols, and Schmidt (2004) on Disability Insurance. However, 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.

[4]See Lawson (2013) for a more detailed description of the optimal UI literature.

that the optimal generosity of UI can be dramatically raised when the latter effect is taken into account: much more generous unemployment benefits may be optimal if they prevent individuals from applying for (and receiving) DI, with the optimal replacement rate rising to 85%. Lindner's is the only estimate of the substitution between UI and DI that I am aware of for the US, so the results should be interpreted with caution; further empirical work on the subject would allow us to reach a more definite conclusion.

I then present a general model which can be applied to any program of state-contingent transfers. The model is solved for a derivative of social welfare with respect to the generosity of any individual program, with a simple and intuitive result that depends directly on the magnitude of fiscal externalities and program interaction effects; I thereby show that the equation derived in the initial UI-DI example generalizes to this setting. I show that recognizing the full size of government causes the estimated 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 be applied, identifying important empirical and theoretical areas for future research.

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, derives analytical results, and briefly summarizes literatures of particular importance for future research, and section 4 concludes.

2 Optimal UI With Substitution to DI

For individuals who qualify as disabled but who are not in fact unable to work, 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. Of course, there are other programs which could be substitutes or complements with respect to UI, including social assistance programs such as TANF or SSI, but to demonstrate the importance of program substitution for policy analysis I simplify my focus to DI as the largest program that is plausibly substitutable with UI.

I model this substitution using a simple 2-period model of unemployment and disability; however, although the model is quite simple, the analysis in section 3 demonstrates that the same results can be obtained from a more general model, including a variety of extensions and a range of forms of (ex-post) heterogeneity. I begin with the presentation and solution of the model in the first two subsections, followed by a subsection explaining the empirical quantities used and providing the numerical results.

2.1 Modified Baily (1978) Model

I base my analysis on a version of the model from Baily (1978), modified to include DI. Time is continuous and can be divided into two periods of equal length. The model features an ex-ante identical population of individuals who work at a wage which I normalize to one in the first period, and who face a risk of losing their job at the end of the first period. Individuals lose their job with probability γ, which is divided into three separate reasons for job loss: with probability γ_1, the individual is simply laid off; with probability γ_2, the individual has become disabled and is now unable to work; and with probability γ_3, the individual is laid off but also develops a partial disability which does not prevent them from working, where $\gamma_1 + \gamma_2 + \gamma_3 = \gamma$.

Any individual who loses their job has 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 $F(\delta)$, where δ is restricted to be between zero and the value of DI benefits. However, I assume that non-disabled individuals will always be rejected (and thus will never apply), whereas completely disabled individuals will always be granted DI benefits; partially-disabled individuals are assumed to be approved with some positive probability α, and I denote the fraction of type-3 unemployed individuals who apply for DI using θ. Individuals who are approved received DI benefits b_D (including the value of Medicare coverage) during the entire second period. Meanwhile, unemployed individuals who don't apply for DI, or who are rejected, receive UI benefits b_U for some fraction s of the second period, and then resume employment at a wage equal to one for the remainder of the period.[5] s is chosen by the individual and subjects them to a utility cost of search $h(s)$ that

[5]By exogenously fixing the re-employment wage, I assume that there are no effects of UI on subsequent

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 equal and denoted by r. Because the rates are equal, and because there is no uncertainty within a period, consumption choices will be constant while in a particular state within 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 $\int_x^y e^{-rt}dt = r_x^y$. The representative individual seeks to maximize expected utility, and the decision problem can therefore be written as:

$$\begin{aligned}\max_{c_1,c_2,c_D,c_U,c_N,s,\theta} V = {} & r_0^1 U(c_1) + (1-\gamma) r_1^2 U(c_2) \\ & + \gamma_1 \left[r_1^{1+s} U(c_U) + r_{1+s}^2 U(c_N) - h(s) \right] + \gamma_2 \left[r_1^2 U(c_D) - \delta \right] \\ & + \gamma_3 \left[\theta\alpha r_1^2 U(c_D) + (1-\theta\alpha)(r_1^{1+s} U(c_U) + r_{1+s}^2 U(c_N) - h(s)) - \theta\delta \right] \\ & - \lambda_1 \left[r_0^1 c_1 + r_1^2 c_2 - r_0^2 (1-\tau) \right] - \lambda_2 \left[r_0^1 c_1 + r_1^2 c_D - r_0^1 (1-\tau) - r_1^2 b_D \right] \\ & - \lambda_3 \left[r_0^1 c_1 + r_1^{1+s} c_U + r_{1+s}^2 c_N - (r_0^1 + r_{1+s}^2)(1-\tau) - r_1^{1+s} b_U \right]\end{aligned}$$

where τ is the tax rate, the λ terms are Lagrange multipliers on the budget constraints, and $U(c)$ follows the usual properties of $U' > 0$, $U'' < 0$.

The government has to finance UI and DI, as well as potentially a lump-sum of spending on other programs which I denote as G, so the government budget constraint is:

$$\tau \left[r_0^2 - (\gamma_1 + \gamma_3(1-\theta\alpha)) r_1^{1+s} - (\gamma_2 + \gamma_3\theta\alpha) r_1^2 \right] = (\gamma_1 + \gamma_3(1-\theta\alpha)) r_1^{1+s} b_U + (\gamma_2 + \gamma_3\theta\alpha) r_1^2 b_D + G$$

which, if I denote $T_U = (\gamma_1 + \gamma_3(1-\theta\alpha)) r_1^{1+s}$ and $T_D = (\gamma_2 + \gamma_3\theta\alpha) r_1^2$ as the expected discounted amounts of time spent on UI and DI respectively, can be rewritten as:

$$\tau \left(r_0^2 - T_U - T_D \right) = T_U b_U + T_D b_D + G.$$

2.2 Welfare Analysis

I assume that the government wants to maximize ex-ante expected utility, and I am focusing on the optimal UI problem, so I hold b_D fixed and allow the government to choose the optimal

wages, to simplify the analysis. As noted in Lawson (2013), the recent empirical literature has tended to support this assumption.

value of b_U. Because of an envelope condition under which the partial derivative of V with respect to any individual choice is zero, 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:

$$\frac{\partial V}{\partial b_U} = \lambda_3 r_1^{1+s} = T_U U'(c_U)$$

$$\frac{\partial V}{\partial \tau} = -\lambda_1 r_0^2 - \lambda_2 r_0^1 - \lambda_3(r_0^1 + r_{1+s}^2) = -r_0^1 U'(c_1) - (1-\gamma) r_1^2 U'(c_2) - (\gamma_1 + \gamma_3(1-\theta\alpha)) r_{1+s}^2 U'(c_N)$$

where I use the individual's first-order conditions to substitute for the multipliers. If I define c_E such that $(r_0^2 - T_U - T_D)U'(c_E) = r_0^1 U'(c_1) + (1-\gamma) r_1^2 U'(c_2) + (\gamma_1 + \gamma_3(1-\theta\alpha)) r_{1+s}^2 U'(c_N)$, so that $U'(c_E)$ is the discounting-weighted average marginal utility among employed individuals, then I can rewrite $\frac{\partial V}{\partial \tau} = -(r_0^2 - T_U - T_D)U'(c_E)$.

Finally, the derivative of the government budget constraint is:

$$\begin{aligned}\frac{d\tau}{db_U} &= \frac{1}{r_0^2 - T_U - T_D}\left[T_U + (b_U + \tau)\frac{dT_U}{db_U} + (b_D + \tau)\frac{dT_D}{db_U}\right] \\ &= \frac{T_U}{r_0^2 - T_U - T_D}\left[1 + \left(1 + \frac{\tau}{b_U}\right)\varepsilon_{b_U}^{T_U} + \frac{T_D(b_D + \tau)}{T_U b_U}\varepsilon_{b_U}^{T_D}\right]\end{aligned}$$

where ε_x^y represents the elasticity of y with respect to x. Combining the marginal utility terms and $\frac{d\tau}{db_U}$, I get the following for $\frac{dV}{db_U}$:

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

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:

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

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

$$\frac{dW}{db_U} = T_U \left[R\frac{\Delta c}{c_E} - \left(1 + \frac{\tau}{b_U}\right) \varepsilon_{b_U}^{T_U} - \frac{T_D(b_D + \tau)}{T_U b_U} \varepsilon_{b_U}^{T_D} \right] \tag{2}$$

where $\tau = \frac{T_U b_U + T_D b_D + G}{r_0^2 - T_U - T_D}$.

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. Of course, reduced spending on DI also means individuals are receiving less in the form of DI benefits, but those lost benefits themselves have no first-order welfare impact, because individuals are assumed to make privately optimal decisions. Therefore, anyone induced by increased UI to not apply for DI must have been indifferent between the two options, and so the foregone DI benefits are replaced by UI benefits and savings of application costs, with no loss to the individuals in question. This is one of the strengths of the sufficient statistic approach: by assuming private optimization, I only need to consider the direct effect of changes in policy variables, and so we can indeed think about reduced DI spending as a cost savings to the government.

2.3 Sufficient Statistics and Numerical Results

To apply the formula in (2), I simply need estimates of each of the quantities appearing in the equation (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.

I begin by computing the baseline values of T_U and T_D, remembering that I must deflate both due to discounting, for which I assume that interest/discount rate is 3% per year. I will denote baseline values using hats, for example $\hat{T_U}$ and $\hat{T_D}$. 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 $\gamma\theta\alpha = \frac{2\times 12.53}{161.4} = 0.1553$.[6] I will make a period equal to 14 years in my model, which means that $r = 0.5126$, and therefore $\hat{T_D} = 0.1553 r_1^2 = 0.0728$. Then, I use an unemployment rate of 5.4% as in Lawson (2013), which means that the total amount of unemployment is $\gamma(1-\theta\alpha)s = 0.108$; if I use the job-losing rate from one of the intermediate cases in Lawson (2013), specifically $\gamma = 0.54$, then this implies that the duration of unemployment is $s = 0.2807$, and therefore $\hat{T_U} = \gamma(1-\theta\alpha)r_1^{1+s} = 0.0602$.

As in Lawson (2013), 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 $\hat{b_U} = 0.46\left(\frac{12.64}{24.3}\right)0.85 = 0.2034$.[7] 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, $b_D = \left(1 + \frac{e_{8/7}^2}{e_1^2}\frac{600}{963}\right)0.46 = 0.6961$.

The elasticities are calculated as follows: I use $\hat{\varepsilon}_{b_U}^{T_U} = 0.2544$ as in Lawson (2013), and Lindner (2012) finds that a \$100 increase in monthly UI benefits, or about a 10% increase, leads to 2700 fewer new DI beneficiaries per year, so $\hat{\varepsilon}_{b_U}^{T_D} = \frac{-27}{895} = -0.0302$; 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.

I assume $R = 2$, a standard value in the optimal UI literature, and $\frac{\Delta c}{c_E} = 0.222 - 0.265rr$,

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

[7] I assume a takeup rate of 80%, the approximate average rate found by Ebenstein and Stange (2010) over 1990-2005, and a ratio of compensated UI duration to total unemployment duration of $\frac{15.8}{24.3}$ as found by Chetty (2008), and $0.8 \times \frac{15.8}{24.3} = \frac{12.64}{24.3}$.

where rr is the UI replacement rate, as in Gruber (1997). Finally, for the baseline tax rate, I consider two cases: in the first, I assume that UI and DI are the only programs to be financed, i.e. $G = 0$, in which case the baseline $\hat{\tau}$ is the value that balances the budget, which is 0.0563. In the second case, I allow for a positive value of G by assuming 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 (the mean age of individuals in the SIPP sample of Chetty (2008)); I can then back out $G = 0.2289$.

The entire set of sufficient statistics is summarized in Table 1. To give a first indication of what they imply, consider what happens if the UI replacement rate is increased by 10% points, from 0.46 to 0.56; based on the statistics discussed above, total UI spending would increase by \$12.68 billion per year, due both to the increased benefit amount and the increased duration of unemployment. However, spending on DI would decline by \$1.54 billion per year. Lindner (2012) stated that the decline in DI spending from a UI expansion would not be enough to pay for the UI benefits, and clearly he was correct; but for purposes of social well-being, that isn't the relevant calculation, because the increased UI spending is, in and of itself, just a transfer and not a social cost, whereas reductions in DI spending can be treated as simple cost savings for the reasons described earlier. The relevant calculations are displayed in Table 2, where column 1 presents the results when G is assumed to be zero and column 2 uses $G = 0.2289$. Panel A shows the welfare derivative at the baseline replacement rate of 46%, and panel B the optimal replacement rate, both in the case in which I ignore interactions between UI and DI and the case in which I take them into account.

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 reproduce the findings of Lawson (2013) on the importance of fiscal externalities. The effect of fiscal externalities appears slightly less dramatic here than in Lawson (2013) mostly because the results in column 1 already include DI spending in the fiscal responsibilities of government.

Table 1: Sufficient Statistics

Statistic	Definition	Value(s)
$\hat{T_U}$	baseline fraction of time on UI	0.0602
$\hat{T_D}$	baseline fraction of time on DI	0.0728
$\hat{b_U}$	adjusted UI benefit	0.2034
b_D	DI benefit	0.6961
$\hat{\varepsilon}_{b_U}^{T_U}$	elasticity of UI duration	0.2544
$\hat{\varepsilon}_{b_U}^{T_D}$	cross-elasticity of DI	$\{0, -0.0302\}$
R	coefficient of relative risk-aversion	2
$\frac{\Delta c}{c_E}$	consumption drop on UI	$0.222 - 0.265rr$
r	discount and interest rate	0.5126
G	other government spending	$\{0, 0.2289\}$
$\hat{\tau}$	baseline tax rate	$\{0.0563, 0.261\}$

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

	(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
$\hat{\varepsilon}_{b_U}^{T_D} = -0.0302$	0.0084	0.0233
	B. Optimal Replacement Rate	
$\varepsilon_{b_U}^{T_D} = 0$	0.3315	0.0289
$\hat{\varepsilon}_{b_U}^{T_D} = -0.0302$	0.6036	0.8499

The main result, 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. 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 $\hat{\varepsilon}_{b_U}^{T_D}$, 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 that would be required to balance the budget if the tax base was unaffected by UI, i.e. the tax increase needed simply to pay for an expansion in UI benefits. 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.

Figure 1: Budget-Balancing Tax Rates

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

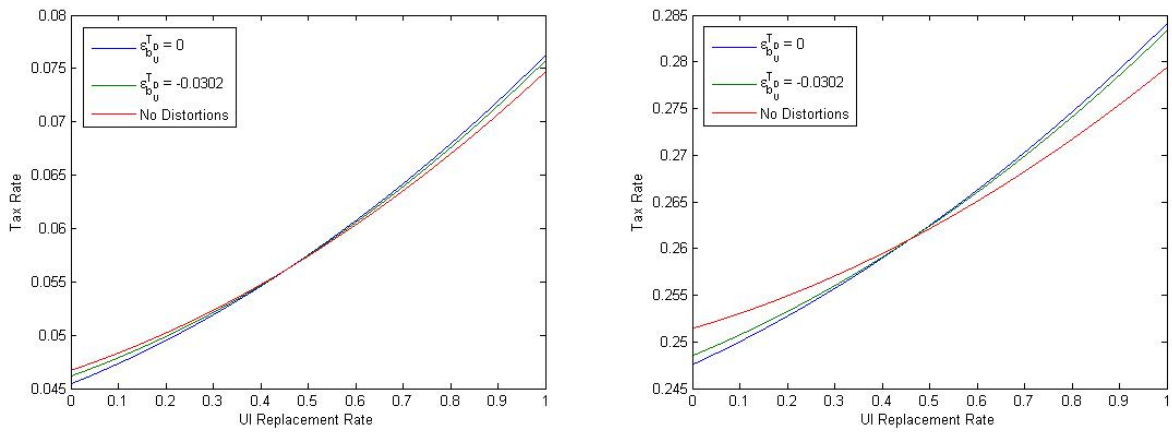
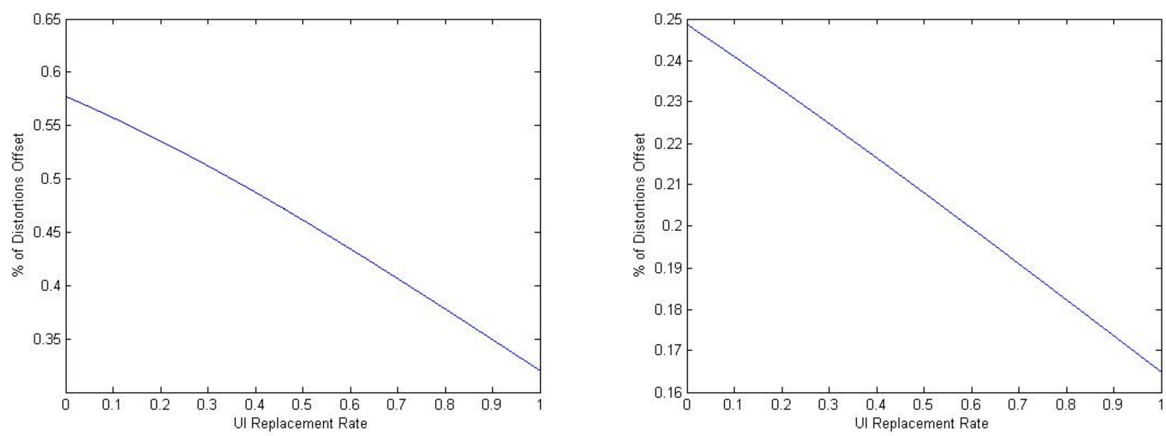


Figure 2: % of Tax Increase to Pay for Distortions Offset by Substitution Effect

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



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 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, because it includes Medicare and tends to be received for a longer period of time.[8] Therefore, preventing unemployed individuals from dropping out of the labour force and receiving DI is worth the increased distortions caused by UI on the job search margin.

3 Analysis of General Model

The results in the previous section clearly demonstrate the importance of taking into account program interaction effects when considering social insurance programs like UI. In this section, I generalize the analysis to a model with a potentially large number of programs. I begin the theoretical analysis with a description of the general model; I then solve for the welfare derivative, and conclude with a discussion of the results and areas where further research would be beneficial.

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 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, $t \in [0, 1]$, 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,

[8] If the Affordable Care Act reduces the value of Medicare coverage by making alternative coverage more easily and cheaply available to low-skill individuals, this could reduce the need for increased UI. Another way of thinking about this is that, by making insurance available to the unemployed, the Affordable Care Act essentially raises the value of UI relative to DI, accomplishing part of the policy change recommended by my analysis.

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 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 y provides a channel through which a program, through its effects on x, can affect labour market income. 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), com-

[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.

plicated 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 τ is the 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 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 ex-ante 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:

$$\frac{dV}{db_j}|_{db_{-j}=0} = \frac{\partial V}{\partial b_j} + \frac{\partial V}{\partial \tau} \frac{d\tau}{db_j}|_{db_{-j}=0} \tag{3}$$

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).

3.2 Calculation of Welfare Derivative

The next step is to evaluate (3), 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) 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, $\forall i, t, \omega_t$:*

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

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

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 (2006) argues that assumptions of this sort are typically satisfied in models in which different sources of income are fungible.

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 maximizes 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:

$$\frac{\partial V}{\partial b_j} = D_jE_j[U'(c)]$$
$$\frac{\partial V}{\partial \tau} = -E^r[yU'(c)]$$

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 program j,[10] 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:

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

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

The elasticity terms in brackets are simple to interpret: I 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. 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.

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

$$\begin{aligned} \frac{dV}{db_j}|_{db_{-j}=0} &= D_j E_j[U'(c)] - E^r[yU'(c)] \frac{D_j}{\bar{y}} \left[1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}} \right) \right] \\ &= D_j \left[E_j[U'(c)] - E_{\bar{y}}[U'(c)] \left(1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}} \right) \right) \right] \end{aligned} \tag{4}$$

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

[10] 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.

I normalize the welfare derivative by $E_{\bar{y}}[U'(c)]$, I can also define:

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

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 the leakiness of the bucket in the terminology of Okun (1975).

At the optimum, $\frac{dV}{db_j}|_{db_{-j}=0}$ must be equal to zero,[11] which means:

$$E_j[U'(c)] = E_{\bar{y}}[U'(c)] \left(1 + \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}} \right) \right). \tag{5}$$

3.3 Analysis of Welfare Derivative

Clearly, to use equations (4) and (5) 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 (4) translates directly into (1) 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 (4) and (5). I will now proceed to provide a series of results that parallel the analytical results in Lawson (2013). To begin with, let me denote $\frac{dV}{db_j}(b_j; b_{-j})$ 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 $\varepsilon_{b_j}^{\bar{y}} - \frac{\sum_{l \neq j} D_l b_l \varepsilon_{b_j}^{D_l}}{\sum_{l \neq j} D_l b_l}$.*

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

[11] 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.

$\frac{dV}{db_j}(b_j; b_{-j}) > \frac{dV}{db_j}(b_j; 0)$ 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 welfare gain from 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 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 (4) 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 (5) 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.[12]

A second issue is that I do not know what values the quantities in (5) will take if I change b_j.[13] 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 (2013) and Lawson (2014). Chetty (2009) suggests statistical extrapolation as an

[12] 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.

[13] 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.

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 (4) and (5) will respond to changes in b.[14] 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 (4) and (5) 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 I 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 (4) and/or (5), 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 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. 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.

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.

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

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 $\varepsilon_{b_j}^{\bar{D}_l2} > \varepsilon_{b_j}^{\bar{D}_l1}$, $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l2}) < \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l1})$.

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) $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l1}) = E_{\bar{y}}[u'(c)] \left(\sum_{l=1}^{M} \frac{D_l b_l}{b_j}\right) \left(\varepsilon_{b_j}^{\bar{D}_l1} - \varepsilon_{b_j}^{\bar{D}_l2}\right)$, 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 $\varepsilon_{b_j}^{\bar{D}_l}$, $b_j^*(\varepsilon_{b_j}^{\bar{D}_l2}) < b_j^*(\varepsilon_{b_j}^{\bar{D}_l1})$.

The proof to Corollary 2 is analogous to that for Corollary 1, and the results are straightforward: 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* $\varepsilon_{b_j}^{\bar{D}_l2} > \varepsilon_{b_j}^{\bar{D}_l1}$, $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l1}) < \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\bar{D}_l2}) - \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\bar{D}_l1})$.

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) $\left[\frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\bar{D}_l2}) - \frac{dV}{db_j}(b_j; 0, \varepsilon_{b_j}^{\bar{D}_l1})\right] = 0$ and $\frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l2}) - \frac{dV}{db_j}(b_j; b_{-j}, \varepsilon_{b_j}^{\bar{D}_l1}) < 0$. □

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 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 unrealistic assumptions. However, Proposition 1 and Corollary 1 indicate that $b_j^*(b_{-j}) - b_j^*(0)$ 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 $\varepsilon_{b_j}^{\bar{D}_l}$, a larger size of government increases the optimal b_j, and vice-versa for small values of $\varepsilon_{b_j}^{\bar{y}}$ and large values of $\varepsilon_{b_j}^{\bar{D}_l}$.

3.4 Areas for Future Research

From the generality of the model described above, it is clear that just about any program which can be described as a state-contingent transfer would fit into this framework. Two of the most obvious areas are social insurance and human capital development.

Social insurance programs have been the subject of extensive economic literatures, but as described earlier, interactions between social insurance programs have not been incorporated in any substantive welfare analysis except for the analysis of health insurance and DI in Li (2014). However, numerous empirical papers have considered such interactions, mostly between UI, DI and workers' compensation. Along with studies of optimal UI, welfare analyses of DI and other programs could benefit from greater attention to fiscal interactions, which appear to be substantial: Borghans, Gielen, and Luttmer (2014), for example, find that after a DI reform in the Netherlands in 1993, each dollar reduction in DI benefits was replaced by 31 cents of other social insurance support. Thus, the welfare implications of social insurance program interactions could be an important area for future research.

Additionally, program interaction effects could be substantial among programs designed to support the development of human capital, specifically education and job-training programs. Lochner (2011) finds important non-production benefits of education, stating 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, although 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.[15] Further empirical work documenting these

[15]Trostel (2010) estimates that, in the U.S., direct public expenditures on college education 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

interactions, as well as policy analysis that takes it into account, would fill an important gap in the literature. The same is true of training programs; 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, but empirical examination of these effects have been limited.

4 Conclusion

In this paper, I have demonstrated that fiscal interactions between social programs can have important impacts of optimal policy analysis. In particular, I show that the optimal generosity of unemployment insurance benefits may be far higher than previously recognized if more generous UI reduces enrollment on disability insurance. I then present a general model that allows for the consideration of interaction effects between a wide range of transfer programs; 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. I argue that this phenomenon deserves greater attention, both empirically and in policy analysis, in the areas of social insurance and human capital development.

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.

lifetime tax revenues of $197000.

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.

Borghans, L., A. C. Gielen, and E. F. Luttmer (2014): "Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform," *American Economic Journal: Economic Policy*, forthcoming.

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

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.

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

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

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

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

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

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.

İmrohoroğlu, A., S. İmrohoroğlu, and D. H. Joines (2003): "Time-Inconsistent Preferences and Social Security," *Quarterly Journal of Economics*, 118(2), 745–784.

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. (2013): "Fiscal Externalities and Optimal Unemployment Insurance," Working Paper No. 2013-57, Aix-Marseille School of Economics.

——— (2014): "Liquidity Constraints, Fiscal Externalities and Optimal Tuition Subsidies," Unpublished Paper, Aix-Marseille School of Economics.

Li, Y. (2014): "The Affordable Care Act in an Economy with Federal Disability Insurance," Unpublished Paper, University of Pittsburgh.

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.

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.