Optimal Taxation and Social Insurance with Endogenous Private Insurance*

Raj Chetty and Emmanuel Saez

UC Berkeley and NBER

October 6, 2008

Abstract

This paper characterizes the welfare gains from redistributive taxation and social insurance in an environment where the private sector provides partial insurance. We analyze stylized models in which adverse selection, pre-existing information, or imperfect optimization in private insurance markets create a role for government intervention. We derive simple formulas that map reduced-form empirical estimates into quantitative predictions for optimal tax and social insurance policy. Applications to unemployment and health insurance show that taking private market insurance into account matters significantly for optimal benefit levels given existing empirical estimates of the key parameters.

*E-mails: chetty@econ.berkeley.edu, saez@econ.berkeley.edu. We thanks Andrew Clark, Amy Finkelstein, Roger Gordon, Ken Judd, Narayana Kocherlakota, Pierre Pestieau, and the TAPES conference participants for very helpful comments and discussions. Financial support from NSF Grant SES-0134946 and the Sloan Foundation is gratefully acknowledged.

1 Introduction

A recent literature in public economics has begun to integrate the theory of optimal taxation and social insurance with empirical evidence on behavioral responses to these policies. Several studies have proposed "sufficient statistic" formulas to map elasticities estimated in the modern program evaluation literature into predictions about the welfare consequences of policies (e.g., Diamond 1998, Saez 2001, Shimer and Werning 2008, Chetty 2008a-b). One important limitation of existing sufficient statistic formulas is that they do not allow for private market insurance, implicitly assuming that the government is the sole provider of insurance. Previous theoretical and empirical studies have emphasized that the existence of private insurance can lower the optimal level of social insurance.[1] However, there is no method of mapping empirical evidence such as that of Cutler and Gruber (1996a) into quantitative statements about the optimal level of government intervention in models such as that analyzed by Golosov and Tsyvinski (2007).

The paper takes a step toward bridging this gap. We develop formulas for optimal taxation and social insurance in stylized models that allow for partial private insurance. The formulas are functions of reduced-form parameters that are frequently estimated in empirical studies, and can therefore be easily adapted to analyze policies ranging from optimal tax and transfer policy to unemployment and health insurance.

The starting point for our analysis is the specification of the limits of private market insurance and the potential role for government intervention. There are at least five reasons that government intervention could improve upon private insurance markets. First, private markets can only insure against shocks that occur *after* agents purchase private insurance. Only the government can provide redistribution across types revealed before private insurance contracts are signed. Second, informational asymmetries can lead to market unravelling through adverse selection (Akerlof 1970). Third, even when private markets function perfectly, individuals may suffer from behavioral biases such as myopia or overconfidence that lead them to underinsure relative to the optimum (Kaplow 1991, DellaVigna 2008, Spinnewijn 2008). Fourth, private firms generally cannot sign exclusive contracts, leading to inefficient outcomes because of multiple dealing (Pauly 1974). Finally, some studies have argued that the administrative and

[1] For example, Golosov and Tsyvinski (2007) make this point theoretically. Cutler and Gruber (1996a-b) present evidence that crowdout of private insurance is substantial in health care.

marketing costs of private insurance exceed those of public insurance (Woolhandler, Campbell, and Himmelstein 2003; Reinhardt, Hussey, and Anderson 2004) because of increasing returns and zero-sum strategic competition.

In this paper, we characterize the welfare gains from government intervention under the first three private market limitations.[2] We analyze models in which the agent's earnings vary across states. This variation can be interpreted as uncertainty due to shocks, as in a social insurance problem, or as variation in earnings ability behind the veil of ignorance, as in an optimal taxation problem. The model permits suboptimal choice of private insurance as well as market limitations due to pre-existing information or adverse selection. We derive a formula for the welfare gain from increasing the government tax rate (or social insurance benefit) that depends on five parameters: (1) the variation in consumption across states and risk types, (2) the curvature of the utility function, (3) the elasticity of effort with respect to the tax or benefit rate, (4) the size of the private insurance market, and (5) the crowdout of private insurance by public insurance. The first three parameters are standard elements of sufficient statistic formulas for optimal taxation and social insurance without private insurance; the last two are the new elements. In addition to offering a method of making quantitative predictions about welfare gains, our analysis yields three qualitative lessons.

First, and most importantly, standard optimal tax and social insurance formulas overstate the optimal degree of redistribution in the presence of private insurance that generates moral hazard. A planner may observe substantial income inequality and conclude based on classic optimal tax results that redistributive taxation would improve welfare. However, if the observed earnings distribution already reflects implicit insurance provided by the private sector – e.g. through wage compression by firms – then making the redistribution through taxation could reduce welfare via its effects on the private insurer's budget. In the extreme case where private insurance markets function optimally, the planner would end up strictly reducing welfare by implementing such redistributive taxes (Kaplow 1991). Intuitively, the government exacerbates the moral hazard distortion created by private-sector insurance, and must therefore take into account the amount of private insurance and degree of crowdout to calculate the optimal policy. Taking the observed earnings distribution as a reflection of marginal products

[2] We do not consider multiple dealing as it is treated in detail by Pauly (1974). We also do not consider administrative costs, but conjecture that it would be straightforward to extend the formulas we develop to incorporate such costs using estimates of loading factors for public and private insurance.

– as is standard practice both in the theory of optimal taxation and in policy debates – may therefore lead to misleading conclusions about optimal tax policy.

The second lesson is that it is important to distinguish private insurance mechanisms that generate moral hazard from those that do not. While "formal" arms-length insurance contracts are likely to generate as much moral hazard as public insurance, "informal" risk sharing arrangements – such as borrowing from close relatives or relying on spousal labor supply to buffer shocks – may involve much less moral hazard. When private insurance does not generate moral hazard, the formula for optimal government benefits coincides exactly with existing formulas that ignore private insurance completely. This is because the effect of informal private insurance is already captured in the smaller consumption-smoothing effect of public insurance. This point is of practical importance because existing empirical studies estimate the extent of formal and informal insurance without distinguishing their policy implications (e.g. Townsend 1994, Cullen and Gruber 2000, Attanasio and Rios-Rull 2000, Schoeni 2002).

The third lesson is that even when the level of private insurance enters the formula for optimal benefits, it matters for a very different reason than what is traditionally emphasized in the literature (e.g. Cutler and Gruber 1996a). The conventional wisdom is based on Okun's proverbial "leaky bucket:" more government expenditure is required to achieve a given increase in insurance coverage when there is crowdout of private insurance. Since the marginal cost of public funds exceeds one, this leakage lowers the welfare gain from public insurance. In our model, this channel does not operate because a formal private insurance contract generates exactly the same moral hazard distortion in effort as public insurance. The added deadweight burden generated by higher taxes to fund public insurance is exactly offset by the reduction in deadweight burden because of lower formal private insurance premiums.[3] The reason that the level of formal private insurance actually matters is that it imposes a fiscal externality on the public sector, as in Pauly (1974), where non-exclusive insurance contracts lead to over-provision of insurance.

To illustrate how our formula can be applied to obtain quantitative predictions about optimal policy, we present applications to unemployment insurance (UI) and health insurance. In the unemployment application, we focus on severance pay provided by private employers

[3] In practice, government insurance programs may be funded using more distortionary tools than private insurance programs. In this case, the deadweight burden offset is less than 100% and the bucket becomes leaky.

as a form of private insurance. Severance pay generates moral hazard because it can induce workers to shirk on the job, since they do not fully internalize the costs of being laid off. Using variation in UI benefit laws across states in the U.S., we estimate that a 10% increase in UI benefit levels reduces private insurance against job loss (severance pay) by approximately 7%. Plugging this estimate into our formula along with other parameter estimates from the existing literature, we find that there is a fairly wide range of parameters for which standard formulas that ignore private insurance and crowdout imply that raising the benefit level would raise welfare when in fact it would lower welfare.

Our second application explores the welfare gains from expanding public health insurance (e.g. Medicare and Medicaid) using existing estimates of behavioral responses to health insurance. Calibrations of our formula suggest that the aggregate level of public health insurance is near the optimum given the amount of private insurance and its response to public insurance. Accounting for private insurance is very important: the standard formula that ignores the private insurance provision overstates the welfare gain from an aggregate expansion of public health insurance by a factor of more than 100 with existing elasticity estimates. Note that these calibration results are based on a representative-agent model with elasticity estimates for the aggregate population. There are likely to be subgroups of the population that are underinsured, such as low income individuals, and others that are overinsured. Our analysis should therefore be interpreted not as a policy recommendation but rather as a call for further work estimating the key elasticities by subgroup to identify how public health insurance should be reformed.

This paper builds on and relates to several strands of the literature on optimal insurance. One theoretical literature has considered optimal insurance and government redistribution problems jointly (e.g., Blomqvist and Horn, 1984, Rochet, 1991, Cremer and Pestieau, 1996). These papers analyze models with heterogeneity in ability to earn (as in Mirrlees 1971) coupled with ex-post shocks to income (such as a health shock). In these models, the government is the sole provider of insurance, and chooses both an optimal income tax schedule and a social insurance program.[4] In contrast, our paper considers a simpler model with a single source of earnings heterogeneity, which does not distinguish between risk and ability but allows public

[4] An exception is Boadway et al. (2006), who allow for private insurance. However, they assume that private insurers observe ability while the government does not. In our model, private and public insurers have the same informational constraints.

and private insurance to coexist.

Models with private and public insurance have been considered in the literature on optimal health insurance (e.g., Besley 1989, Selden 1993, Blomqvist and Johansson 1997, Petretto 1999, Encinosa 2003, Barrigozzi 2006). We develop empirically implementable formulas for the welfare gains from public insurance in such models. Our formulas help to connect the theoretical work to the corresponding empirical literature on the interaction between private and public health insurance (e.g., Ginsburg 1988, Taylor et al. 1988, Wolfe and Godderis 1991, Cutler and Gruber 1996a-b, Finkelstein 2004). The formulas we derive are in the same spirit as recent sufficient statistic formulas in that they shed light on the essential features of the models that matter for welfare analysis. However, unlike typical sufficient statistic results, the formulas we derive here are based on stylized models and therefore may not be fully robust to modifications of the primitive structure, an issue that we discuss in greater detail in the conclusion.

The most closely related paper to our study is that of Einav, Finkelstein, and Cullen (2008), who develop a different method of characterizing welfare in a model with adverse selection in the private insurance market. Einav et al. show that the slopes of the demand and cost curves for private insurance are together sufficient statistics for welfare. Our formula depends instead on ex-post behavioral responses to change in government benefit levels. The two formulas are complements. Einav et al.'s method is easier to implement when exogenous price variation in insurance markets and demand and cost data are available; our formulas may be easier to implement when there is variation in government benefit levels that permits estimation of ex-post behavioral responses. Likewise, Einav et al.'s method will yield reliable results in markets where ex-ante insurance purchase decisions reveal true preferences, while our formulas rely on optimization of choices such as consumption and labor supply.

The remainder of the paper is organized as follows. We begin in section 2 by analyzing a model without private insurance as a benchmark. In Section 3, we introduce endogenous private insurance by studying a model in which the government and private sector have the same tools, but the level of private insurance is not necessarily set optimally. Section 4 considers a model with market limitations due to the inability of the private sector to insure pre-existing risks and adverse selection. The applications are presented in section 5. Section 6 concludes.

2 Benchmark: Second-Best Contract

We analyze models where the government uses tax or social insurance policies to redistribute income across individuals with different levels of earnings. Risk averse individuals would like to insure themselves against the risk of having low output realizations. If effort were observable, the "first-best" contract would provide full insurance and the agents would be required to exert the first best optimal level of effort. If effort is unobservable, the first-best contract is not incentive compatible. The optimal contract when effort is unobservable specifies consumption levels contingent on output realization that reduce the variance of consumption. In this section, we characterize this "second-best" contract, which reflects the optimal policy for a single insurer (government or private). This is the problem considered in previous studies that developed sufficient statistic formulas for taxation and social insurance (e.g. Saez 2001, Chetty 2006a).

We analyze this problem using two canonical models in the optimal tax and social insurance literatures. The first is a "moral hazard" model, in which the level of output is a stochastic function of the individual's effort choice. This model, which is equivalent to Varian's (1980) model of taxation, can be interpreted as a model of optimal social insurance against shocks such as unemployment or illness.[5] The second model is a "hidden skill" model, in which the uncertainty about output (skill) levels is resolved *before* individuals choose effort. This model is the same as that studied by Mirrlees (1971), and is more naturally suited to studying optimal redistribution across individuals with different skills. To simplify the analysis, we restrict attention to optimal linear contracts throughout, and show that the two models yield the same results for such contracts.

2.1 Model 1: Moral Hazard

Consider a model with a continuum of identical individuals of measure one who exert effort e to produce output z. There are two states of output, high and low, with $z_H > z_L$. The probability of producing z_H is e. Let $\bar{z} = e \cdot z_H + (1 - e) \cdot z_L$ denote the average level of output. The utility of consumption c is $u(c)$ and the disutility of effort is $\psi(e)$. Utility $u(c)$

[5] In the empirical applications in section 5, we translate the results from this model into the notation used in the social insurance literature. In the theoretical analysis, we use the tax notation to highlight similarities between the formulas in the two models.

is increasing and strictly concave and $\psi(e)$ is increasing and convex.

The individual chooses e to maximize his expected utility,

$$e \cdot u(c_H) + (1-e) \cdot u(c_L) - \psi(e)$$

taking the consumption levels as given. This leads to the first order condition $\psi'(e) = u(c_H) - u(c_L)$, which implicitly defines an effort supply function $e^*(c_H, c_L)$.

Any feasible contract (c_H, c_L) must leave the insurer with zero profits: $\bar{c} \equiv ec_H + (1-e)c_L = \bar{z}$. Because there are only two states, a contract (c_H, c_L) that satisfies the insurer's break-even constraint can be characterized by a single tax rate m along with a lump sum repayment of the tax revenue $m\bar{z}$:

$$c_H = (1-m)z_H + t\bar{z},\ c_L = (1-m)z_L + m\bar{z}.$$

The agent's effort e^* depends on the net of tax rate $1-m$. Thus, average earnings is a function of the net-of-tax rate $1-m$, which we denote by $\bar{z}(1-m)$.

The optimal second-best contract sets m to maximize the agent's expected utility:

$$W = e^* \cdot u((1-m)z_H + m \cdot \bar{z}(1-m)) + (1-e^*) \cdot u((1-m)z_L + m \cdot \bar{z}(1-m)) - \psi(e^*). \quad (1)$$

Our objective is to characterize the marginal welfare gain from increasing the tax rate $(\frac{dW}{dm})$ in terms of empirically estimable parameters. Since e^* is chosen to maximize expected utility, we can use the envelope theorem when differentiating (1) to obtain

$$\frac{dW}{dm} = e \cdot u'(c_H) \cdot \left(-z_H + \bar{z} - m\frac{d\bar{z}}{d(1-m)}\right) + (1-e) \cdot u'(c_L) \cdot \left(-z_L + \bar{z} - m\frac{d\bar{z}}{d(1-m)}\right).$$

Defining $\bar{u}' = eu'(c_H) + (1-e)u'(c_L)$ as the average marginal utility,

$$\begin{aligned}\frac{dW}{dm} &= -\bar{u}' \cdot m \cdot \frac{d\bar{z}}{d(1-m)} + \bar{u}' \cdot \bar{z} - [ez_H u'(c_H) + (1-e)z_L u'(c_L)] \\ &= -\bar{u}' \cdot m \cdot \frac{d\bar{z}}{d(1-m)} - cov(z, u')\end{aligned}$$

where $cov(z, u')$ denotes the covariance between earnings z and marginal utility u'. Let the elasticity of earnings with respect to the net-of-tax rate be denoted by $\varepsilon_{\bar{z},1-m} = [(1-m)/\bar{z}] \cdot d\bar{z}/d(1-m)$. We obtain the following equation for the welfare gain from raising the tax rate:

$$\frac{dW}{dm} = -\bar{z} \cdot \bar{u}' \left[\frac{m}{1-m}\varepsilon_{\bar{z},1-m} + cov(z/\bar{z}, u'/\bar{u}')\right] \quad (2)$$

It is important to recognize that all of the parameters in (2) are functions of m – that is, they are not policy invariant. Therefore, equation (2) can only be used for *local* welfare analysis: the welfare consequences of changes in tax rates around the point at which the elasticities are estimated. These standard properties of sufficient statistic formulas are discussed in more detail in Chetty (2008b).

The optimal tax rate satisfies the first order condition $\frac{dW}{dm}(m) = 0$, yielding the formula

$$\frac{m}{1-m} = \frac{1}{\varepsilon_{\bar{z},1-m}} \cdot \frac{-cov(z,u')}{\bar{z} \cdot \bar{u}'} \tag{3}$$

The elasticity $\varepsilon_{\bar{z},1-m}$ must be positive at the optimum; if $\varepsilon_{\bar{z},1-m} \leq 0$, increasing the tax rate would both increase redistribution and (weakly) increase tax revenue, strictly improving welfare. Again, because all the terms on the right hand side are functions of the the tax rate m itself, (3) should not be viewed as an explicit formula for the optimal tax rate. Rather, it provides a simple condition to test whether a given tax rate is optimal and highlights the parameters that matter for optimal taxation.

Three aspects of this formula deserve mention First, the elasticity $\varepsilon_{\bar{z},1-m}$ measures the *total* effect of a change in the tax rate on earnings, including the response to the change in the lump sum grant. Hence, this elasticity reflects a mix of substitution and income effects. To see this, let $\bar{z}^u(1-m, E)$ denote the agent's standard Marshallian average earnings supply function, which maps the tax rate m and lump sum grant E to a level of average earnings. Since $\bar{z}(1-m) = \bar{z}^u(1-m, m\bar{z}(1-m))$, it follows that $\frac{\partial \bar{z}}{\partial(1-m)} = \frac{\partial \bar{z}^u}{\partial(1-m)} + \frac{\partial \bar{z}^u}{\partial E} \cdot (-\bar{z} + m \cdot \frac{\partial \bar{z}}{\partial(1-m)})$ and hence

$$\varepsilon_{\bar{z},1-m} = \frac{\varepsilon^u_{\bar{z},1-m} - \bar{\eta}}{1 - \bar{\eta} \cdot m/(1-m)} \tag{4}$$

where $\varepsilon^u_{\bar{z},1-m} = \frac{1-m}{\bar{z}^u} \cdot \frac{\partial \bar{z}^u}{\partial(1-m)}$ is the uncompensated (Marshallian) elasticity of earnings with respect to $1-m$ and $\bar{\eta} = (1-m)\frac{\partial \bar{z}^u}{\partial E}$ measures the unearned income effect on labor supply. Note that in the conventional case where leisure is a normal good, $\bar{\eta} < 0$.[6] The elasticity $\varepsilon_{\bar{z},1-m}$ thus differs from the standard parameters estimated in the empirical literature, which typically estimate $\varepsilon^u_{\bar{z},1-m}$ or the pure substitution effect.

Second, when the agent is risk averse, $u'(c_H) < u'(c_L)$, and hence $cov(z,u') < 0$. Combined with the fact that $\varepsilon_{\bar{z},1-m} > 0$, this implies that the optimal tax rate m is strictly positive. If utility is linear and the agent is risk neutral, the optimal tax rate is zero. At the other extreme,

[6] As $\bar{c} = (1-m)\bar{z} + E$, $\frac{\partial \bar{c}}{\partial E} = 1 + \bar{\eta} < 1$ measures the average propensity to consume out of unearned income.

if risk aversion is infinite and $z_L = 0$, then the optimal tax rate is the revenue maximizing tax rate: $m/(1-m) = 1/\varepsilon_{\bar{z},1-m} = 1/\varepsilon^u_{\bar{z},1-m}$. More generally, between these extremes, the optimal tax rate m increases with risk aversion, holding $\varepsilon_{\bar{z},1-m}$ constant.

Third, (3) remains a valid formula for the optimal *linear* tax rate in a model with more than two states, including the case with a continuous state space typically analyzed in income tax models.

2.2 Model 2: Hidden Skills

Now consider a "hidden skill" model of optimal taxation where individuals differ in privately observed ability n and choose effort levels *after* ability is revealed. Let the distribution of skills be given by a density $f(n)$ and normalize the size of the population to 1. An individual who works l hours and has ability n earns $z = nl$.

To simplify the analysis, we eliminate income effects by considering the following form for utility (as in Diamond (1998)):

$$U(c, z, n) = u\left(c - h\left(\frac{z}{n}\right)\right), \tag{5}$$

where $u(.)$ is a concave and increasing function reflecting risk aversion, and $h(.)$ is increasing and convex and represents disutility of work. With a linear tax, $c = (1-m)z + R$ and the first order condition for z is $n(1-m) = h'(z/n)$. We denote by $\varepsilon = \frac{1-m}{z}\frac{dz}{d(1-m)} = \frac{h'(z/n)}{h''(z/n) \cdot z/n}$ the elasticity of earnings with respect to the net-of-tax rate $1-m$. Let $\bar{z} = \int z f(n) dn$ denote average earnings in the population and $\varepsilon_{\bar{z},1-m}$ the elasticity of average earnings $\bar{z}$ with respect to the net-of-tax rate.[7]

Suppose there is a single insurer who offers a linear insurance contract with tax rate m behind the veil of ignorance (i.e., before individuals learn about their ability n). As above, the tax revenue is rebated as a lump sum grant $m\bar{z}$. The insurer chooses m to maximize

$$W = \int u\left(z \cdot (1-m) + m\bar{z} - h\left(\frac{z}{n}\right)\right) f(n) dn$$

Because z maximizes individual utility, using the envelope theorem, the first order condition with respect to m is:

$$0 = \frac{dW}{dm} = \int u'(c) \cdot \left[(\bar{z} - z) - m\frac{d\bar{z}}{d(1-m)}\right] f(n) dn = -cov(z, u') - \frac{m}{1-m} \cdot \varepsilon_{\bar{z},1-m} \cdot \bar{u}' \cdot \bar{z},$$

[7] Formally, $\varepsilon_{\bar{z},1-m}$ is the average of the individual elasticities weighted by earnings.

which yields the optimal tax formula:

$$\frac{m}{1-m} = \frac{1}{\varepsilon_{\bar{z},1-m}} \cdot \frac{-cov(z,u')}{\bar{z}\cdot\bar{u}'} \tag{6}$$

Formulas (6) in the hidden skill model and formula (3) in the moral hazard model are identical. Hence, the optimal *linear* tax rate is driven by the same economic forces in both models.[8]

Both optimal tax formulas (3) and (6) can be derived heuristically as follows. Suppose the insurer increases m by dm. The direct utility cost for individual with earnings z is $-u'(c)\cdot zdm$. The behavioral response to dm does not generate a first order effect on utility because of the envelope theorem. Therefore, in aggregate, the direct welfare cost is $dW = -dm \int u' \cdot zf(n)dn$. The mechanical increase in tax revenue (ignoring behavioral responses) due to dm is $dM = \bar{z}dm$. The behavioral response in earnings reduces tax revenue by $dB = -m\cdot d\bar{z} = -\frac{m}{1-m}\varepsilon_{\bar{z},1-m}\cdot\bar{z}dm$. Hence, the lump sum grant increases by $dM + dB$, increasing welfare by $\bar{u}'\cdot(dM+dB) = \bar{u}'\cdot\bar{z}\cdot[1-\frac{m}{1-m}\varepsilon_{\bar{z},1-m}]dm$. At the optimum, these effects must all cancel so that $dW + \bar{u}'\cdot(dM+dB) = 0$, which yields (3) and (6).

3 Public Insurance with Endogenous Private Insurance

We now turn to the problem of optimal government insurance with endogenous private insurance. We first analyze the case in which the private insurance level is set arbitrarily, and then turn to the special case where it is set optimally. The level of private insurance level might not be optimized because of individual failures – left to their own devices, individuals may purchase too little insurance.

3.1 General Case: Private Insurance not Necessarily Optimized

Moral Hazard Model. We begin by introducing notation to distinguish the private and government insurance contracts. Let τ denote the tax rate chosen by the government and t the tax rate in the private insurance contract. Private insurance applies to raw output (z_H, z_L). We denote by $w_i = (1-t)z_i + t\bar{z}$ the net-of-private insurance incomes in each state. Government taxation applies to the net wage incomes w_i, and we denote by $c_i = (1-\tau)w_i + \tau\bar{w}$ final disposable income. Concretely, the private insurer can be thought of as a firm that compresses its wage structure $(w_H - w_L)$ relative to true marginal products $(z_H - z_L)$ to

[8] This equivalence would no longer hold in the case of nonlinear taxation (Varian 1980).

provide insurance. The government can only observe earnings, not true underlying marginal products, and hence sets taxes as a function of w.[9]

Let m denote the total tax rate on output, defined such that $1 - m = (1 - t)(1 - \tau)$ and $c = (1 - m)z + m\bar{z}$. If the private insurer and government cooperated to set m to maximize social welfare, the resulting contract would be identical to that described in the single insurer setting above. However, in practice private insurers take the government contract τ as given when they choose t. As a result, they fail to internalize the effects of their choice of t on the government's budget. Because of this externality effect, the government needs to follow a different rule in setting the optimal tax rate τ that adjusts for the private insurer's response.

Let $t(\tau)$ denote the private insurer's choice of t as a function of the government tax. In this subsection, we take the function $\tau \to t(\tau)$ as given, and do not assume that $t(\tau)$ is chosen optimally to maximize the agent's expected utility. Let $r = -d\log(1-t)/d\log(1-\tau)$ denote the rate at which public insurance crowds out private insurance. If $r = 0$, there is no crowdout. If $r = 1$, there is perfect crowdout.

The government chooses τ to maximize the agent's expected utility, taking into account the private insurer's response:

$$\max W = e^* \cdot u((1-m)z_H + m \cdot \bar{z}(1-m)) + (1-e^*) \cdot u((1-m)z_L + m \cdot \bar{z}(1-m)) - \psi(e^*), \quad (7)$$

where $m = t(\tau) + \tau - \tau \cdot t(\tau)$ is a function of τ.

Let $\varepsilon_{\bar{w},1-\tau} = \frac{d\log\bar{w}}{d\log(1-t)}$ denote the elasticity of average (post insurance) earnings $\bar{w}$ with respect to $1-\tau$, taking into account the endogenous response of private insurance t to a change in τ. The following proposition characterizes the marginal welfare gain from increasing the tax rate ($\frac{dW}{d\tau}$) in terms of empirically estimable elasticities and provides a condition for the optimal tax rate.

Proposition 1. *In the moral hazard model, the welfare gain from raising the government tax is*

$$\frac{dW}{d\tau} = -(1-r) \cdot \bar{z} \left[\left(\frac{\tau}{1-\tau} + t \right) \frac{\varepsilon_{\bar{w},1-\tau}}{1-r} + \frac{cov(w,u')}{\bar{w} \cdot \bar{u}'} \right] \quad (8)$$

[9] Because the individuals' effort decision depends solely on the return net of all taxes $(1-t)(1-\tau)$, the analysis below goes through with no changes if government taxes are levied on the true marginal products z and private insurance is based on net-of government tax incomes.

and the optimal tax rate satisfies

$$\frac{\tau}{1-\tau} = -t + \frac{1-r}{\varepsilon_{\bar{w},1-\tau}} \cdot \frac{-cov(w,u')}{\bar{w} \cdot \bar{u}'}. \tag{9}$$

Proof. The problem in (7) is identical to (1). Hence, (2) and (3) remain valid. The government sets τ so that the total tax rate m satisfies the standard formula. Because the government does not observe z directly, it is useful to rewrite (2) and (3) as a function of w instead of z. To do this, first note that $\bar{w} = \bar{z}$ and hence $\varepsilon_{\bar{w},1-m} = \varepsilon_{\bar{z},1-m}$. Second, $cov(w,u') = cov(z(1-t)+t\bar{z},u') = (1-t)cov(z,u')$. Third, $\varepsilon_{\bar{w},1-\tau} = \varepsilon_{\bar{w},1-m} \cdot (1-r)$ – a one percent increase in $1-\tau$ translates into a $1-r$ percent increase in $1-m$ because of crowdout effects. Similarly, $\frac{dm}{d\tau} = (1-r)(1-t)$ and hence, $\frac{dW}{d\tau} = (1-r)(1-t)\frac{dW}{dm}$. Finally $m/(1-m) = [\tau/(1-\tau)+t]/(1-t)$. Using these expressions, we can rewrite (2) as (8) and (3) as (9). QED.

Proposition 1 shows that private insurance affects the formula for the optimal tax rate in two ways. First, the added $-t$ term on the right side of (9) reflects the mechanical reduction in the optimal level of government taxation given the presence of private insurance. Formula (9) shows that the *sum* of private and public insurance should be set according to the standard formula, and hence the optimal τ is reduced in proportion to t. The second effect is that the inverse elasticity term is multiplied by $1-r$. Since $r > 0$, this effect also makes the optimal government tax rate smaller. Intuitively, the elasticity relevant for optimal taxation is the fundamental elasticity of output with respect to total taxes $\varepsilon_{\bar{z},1-m} = \varepsilon_{\bar{w},1-m}$, which measures the total moral hazard cost (to both the private and public insurers) of redistributing one dollar of tax revenue. To recover the fundamental output elasticity in the presence of crowdout, one must rescale the observed elasticity $\varepsilon_{\bar{w},1-\tau}$ by $\frac{1}{1-r}$. Therefore, the crowdout rate r only matters for rescaling the observed elasticity, and plays no fundamental economic role in the formula. If one could measure the fundamental elasticity $\varepsilon_{\bar{w},1-m}$ directly, an estimate of crowdout would be unnecessary.

This intuition about the relevance of crowdout differs from the conventional wisdom (see e.g., Cutler and Gruber 1996a-b). The conventional wisdom is that crowdout makes government intervention less desirable because it requires more government expenditure to achieve a given insurance level, and raising revenue to finance government expenditure is costly because of the deadweight burden of taxation. This intuition does not apply in our model because the tax levied by the private insurer generates exactly the same deadweight burden as government taxation. A natural counterargument is that private insurance may not generate moral hazard because private insurers can monitor effort more closely. However, we show in section 3.3 that in this case, crowdout is actually irrelevant because the optimal private contract does not generate externalities.

An important implication of (9) is that if the wage structure already reflects implicit insurance provided by the private sector, then standard optimal tax formulas that are functions of the observed wage distributions and wage earnings elasticities are invalid. Intuitively, the private sector already bears a moral hazard cost for the insurance it provides. The government exacerbates this pre-existing distortion by introducing additional insurance. Therefore, one must take into account the amount of private insurance to back out the optimal policy.[10]

In the simple model considered here, the planner can replicate the second-best optimal allocation by choosing the level of τ that generates the optimal m given the private insurer's response. The second best can be achieved because private and public insurance are identical tools in this model: the role for government emerges if private insurance is not set optimally. When the government and private sector have different tools – as in the adverse selection model analyzed in section 4 – the second best level of welfare can no longer be achieved.

Hidden Skill Model. In the hidden skill model, $\bar{z}(1-m) = \bar{z}((1-t)(1-\tau)) = \bar{w}((1-t)(1-\tau))$ and $\varepsilon_{\bar{w},1-\tau} \equiv \frac{1-\tau}{\bar{w}}\frac{d\bar{w}}{d(1-\tau)} = (1-r)\varepsilon_{\bar{z},1-m}$. With these definitions, equations (8) and (9) of Proposition 1 can be obtained by following a proof that has exactly the same steps.

3.2 Special Case: Optimized Private Insurance

To gain further insight into the effects of government intervention and connect our results to the existing literature, we now consider the special case where private insurance is chosen optimally to maximize expected utility. The following proposition characterizes the effects of government intervention on social welfare when private insurance is optimized in both the moral hazard and hidden skill models.

Proposition 2. *If private insurance is set optimally,*

[10] In redistributive taxation applications, it is challenging to measure the private insurance rate t because the difference between w and z is not directly observed. Conceptually, one must estimate the difference between observed wages and marginal products. In social insurance applications, private insurance typically involves explicit transfers, and can therefore be measured directly.

1) The welfare cost of introducing a small tax τ is second-order: $\frac{dW}{d\tau}(\tau = 0) = 0$.

2) Government intervention strictly reduces welfare when τ is positive: $\frac{dW}{d\tau}(\tau > 0) < 0$.

3) The effect of a tax increase on welfare is given by

$$\frac{dW}{d\tau} = -\varepsilon_{\bar{z},1-\tau} \cdot \frac{\tau}{1-\tau} \cdot \frac{\bar{u}' \cdot \bar{z}}{1 - \bar{\eta} \cdot \frac{t}{1-t}} \tag{10}$$

where $\bar{\eta} = (1-m)\frac{\partial \bar{z}^u}{\partial E}$ is the income effect and $\varepsilon_{\bar{z},1-\tau}$ is the elasticity of average earnings with respect to $1-\tau$, taking into account the endogenous response of t to τ. The same formula applies in the hidden skill model with $\bar{\eta} = 0$.

Proof. In the interest of space, we provide a proof only for the moral hazard model. An analogous derivation can be used to establish the same results for the hidden skill model.

Optimal Private Contract. In a competitive market where insurers are price takers, the private insurer takes the government tax rate τ and the lump sum grant $R = \tau\bar{w}$ as given when setting t. Therefore, t is chosen to maximize:

$$W = e^* \cdot u((1-t)(1-\tau)z_H + t(1-\tau)\bar{z} + R) + (1-e^*) \cdot u((1-t)(1-\tau)z_L + t(1-\tau)\bar{z} + R) - \psi(e^*). \tag{11}$$

Using the envelope condition for e^*, the first order condition with respect to t is:

$$0 = \frac{dW}{dt}|_{R,\tau} = -(1-\tau)(z_H u'(c_H) + z_L u'(c_L)) + (1-\tau)\bar{z}\bar{u}' + t(1-\tau)\bar{u}'\frac{d\bar{z}}{dt}|_{R,\tau}, \tag{12}$$

which implies

$$\frac{t}{1-t} = \frac{1}{\varepsilon_{\bar{z},1-t}|_{R,\tau}} \cdot \frac{-cov(z,u')}{\bar{z} \cdot \bar{u}'}. \tag{13}$$

This expression shows that the private insurer follows a rule analogous to the second-best single insurer choice in (3) in setting t.

Effect of Government Intervention with Optimized Private Insurance. Noting that the government grant is $R = \tau\bar{z}(1-m)$, it follows from (11) that

$$W = e^* \cdot u((1-m)z_H + m \cdot \bar{z}(1-m)) + (1-e^*) \cdot u((1-m)z_L + m \cdot \bar{z}(1-m)) - \psi(e^*).$$

Using the envelope condition for e^* and the equation $dm/d\tau = (1-t)(1-r)$, we obtain

$$\frac{dW}{d\tau} = (1-t)(1-r) \cdot \left[\bar{u}' \cdot \bar{z} - (e \cdot u'(c_H) \cdot z_H + (1-e) \cdot u'(c_L) \cdot z_L) - \bar{u}' \cdot m \cdot \frac{d\bar{z}}{d(1-m)}\right].$$

The first two terms in the square bracket can be rewritten using the first-order-condition in (12) to obtain

$$\frac{dW}{d\tau} = (1-t)(1-r) \cdot \bar{u}' \cdot \bar{z}\left[\frac{t}{1-t} \cdot \varepsilon_{\bar{z},1-t}|_{R,\tau} - \frac{m}{1-m} \cdot \varepsilon_{\bar{z},1-m}\right].$$

Introducing the Marshallian demand $\bar{z}^u(1-m,E)$, (4) implies $\varepsilon_{\bar{z},1-m} = \frac{\varepsilon^u_{\bar{z},1-m}-\bar{\eta}}{1-\bar{\eta}\cdot m/(1-m)}$ where $\bar{\eta} = (1-m)\frac{\partial \bar{z}^u}{\partial E}$. Since $c = z(1-t)(1-\tau)+t(1-\tau)\bar{z}+R$ and $E = t(1-\tau)\bar{z}+R$, we can write

$$\frac{d\bar{z}}{d(1-t)}|_{R,\tau} = \frac{\partial \bar{z}^u}{\partial(1-m)}\cdot(1-\tau)+\frac{\partial \bar{z}^u}{\partial E}\cdot\left(-(1-\tau)\cdot\bar{z}+t(1-\tau)\frac{d\bar{z}}{d(1-t)}|_{R,\tau}\right).$$

It follows that

$$\varepsilon_{\bar{z},1-t}|_{R,\tau} = \frac{\varepsilon^u_{\bar{z},1-m}-\bar{\eta}}{1-\bar{\eta}\cdot t/(1-t)}. \tag{14}$$

Consolidating terms yields

$$\frac{dW}{d\tau} = (1-t)(1-r)\cdot\frac{\bar{u}'\cdot\bar{z}}{1-\bar{\eta}\cdot t/(1-t)}\;\frac{\varepsilon^u_{\bar{z},1-m}-\bar{\eta}}{1-\bar{\eta}\cdot m/(1-m)}\cdot\left[\frac{t}{1-t}-\frac{m}{1-m}\right].$$

Finally, observing that $t/(1-t)-m/(1-m) = -\tau/((1-\tau)(1-t))$ and that $\varepsilon_{\bar{z},1-\tau} = (1-r)\varepsilon_{\bar{z},1-m} = (1-r)\frac{\varepsilon^u_{\bar{z},1-m}-\bar{\eta}}{1-\bar{\eta}\cdot m/(1-m)}$, we obtain (10). QED.

This proposition shows that the standard lessons of the theory of excess burden apply in our model: the welfare cost of taxation is proportional to the size of the behavioral response to taxation and the marginal cost of taxation increases linearly with the tax rate. However, there is one important difference relative to the traditional analysis. In standard models of taxation and social insurance without endogenous private insurance, the deadweight burden of taxation is an efficiency cost that the government would be willing to trade-off against the benefits of more insurance. In the present model, the level of redistribution through market insurance is already optimal given incentive constraints, and thus the net welfare gain *equals* the deadweight burden of taxation. Hence, a benevolent government should do precisely nothing.

When private insurance markets function optimally, government insurance via taxation or social insurance *strictly* reduces welfare because crowdout of private insurance is imperfect. To understand the intuition, it is helpful to consider an example. Suppose the government naively thinks it can achieve the second-best optimum by setting τ and R such that $(1-\tau)z_H+R = c^*_H$ and $(1-\tau)z_L+R = c^*_L$ where (c^*_L, c^*_H) denotes the optimal second-best insurance contract for a single insurer. This tax system would achieve the optimum described in Section 2 if the private insurance market chose not to provide any insurance. However, zero insurance is not the contract that will emerge in an economy with a private insurance market. Since $z_L(1-\tau)+R < z_H(1-\tau)+R$, individuals will find it worthwhile to purchase further insurance from private providers. This additional insurance would increase the expected utility of the agent and the insurance companies would break even. But the individual and private insurer do not internalize the effect of their choices on the government's budget. The added private insurance reduces effort below the second-best optimum and hence leads to a lower average output than the government was expecting. As a result, the government goes into deficit and its lumpsum grant R needs to be reduced.[11]

We are not the first to make the point that government intervention is undesirable when private insurance markets function efficiently. When private insurance is set optimally, the market equilibrium with no government intervention is information-constrained Pareto efficient. Abstractly, it is well known that when the private market equilibrium is constrained efficient, government price distortions lead to inefficiency (Prescott and Townsend 1984a-b). Kaplow (1991) makes a similar point about the welfare losses from government intervention in a simple social insurance model, and Blomqvist and Johannson (1997) and Barrigozzi (2006) obtain similar results in a model with public and private health insurance. The main contribution of the present paper is to develop formulas for the welfare gains from government intervention when the private market does not reach the second-best optimum.

Two additional remarks on Proposition 2 deserve mention. First, we expect that the relationship between the degree of crowdout r and risk aversion will generally have an inverted U-shape. When individuals are risk neutral ($u(c) = c$), there is no private insurance ($t = 0$) so there is by definition no crowdout ($r = 0$). At the other extreme, with infinite risk aversion and $z_L = 0$, $cov(z, u') = -\bar{z} \cdot \bar{u}'$ and (13) combined with (14) implies that $t/(1-t) = 1/\varepsilon^{u}_{\bar{z},1-m}$. If the uncompensated elasticity does not vary with the net-of-tax rate $1 - m$, the private insurance rate t is independent of τ. Therefore, in the case of infinite risk aversion, there is also no crowdout. For any interior case with non-zero, finite risk aversion, crowdout will be positive and incomplete. Hence, government intervention is likely to induce the greatest welfare losses when risk aversion is very low or very high, because these are the cases in which crowdout will be small and individuals will end up being most over-insured.

Second, a corollary of Proposition 2 is that in the absence of private market failures, public

[11] These results contrast with the analysis of Golosov and Tsyvinski (2007), who argue that government insurance has no effect on social welfare in static models because of 100% crowdout. The reason for the difference is that Golosov and Tsyvinski study government intervention where the government controls directly the final consumption outcomes whereas we consider the effects of distortionary taxation, which is typically the nature of policies used in practice.

goods should be financed via a uniform lump-sum tax that generates the desired amount of revenue, even if agents have different marginal utilities of income in each state. The private insurance market will then set redistribution to the optimal level. A distortionary tax to finance the public good would lead to lower expected utility than a lump-sum tax that generated an equivalent amount of revenue. This point underscores the substantial effect that endogenous private insurance arrangements can have on standard intuitions in the literatures on optimal taxation and social insurance.

3.3 Private Insurance Without Moral Hazard

Thus far, we have considered private insurance contracts that generate the same amount of moral hazard as public insurance. However, "informal" private risk-sharing mechanisms may involve less moral hazard. Examples include self insurance through spousal labor supply or insurance through relatives who can monitor effort.[12] In this section, we explore how the degree of moral hazard in private insurance affects our formulas.

Moral Hazard Model. Consider a model where government insurance generates moral hazard but private insurance does not. When private insurance does not generate moral hazard, reaching the first-best of full insurance would be feasible in principle, completely eliminating the role for government intervention. In practice, there are costs of informal insurance – such as limits to liquidity, costs of borrowing from relatives, or relying on spousal labor supply – that prevent full insurance. We model such costs by a loading factor on informal insurance transfers: transferring T to the low state requires a payment of $T + s(T)$ in the high state where $s(.)$ is an increasing and convex loading factor cost with $s(0) = 0$. The informal insurance contract involves no moral hazard, in the sense that the agent internalizes the effect of his effort choice on the insurer's budget. Hence, e and T are chosen simultaneously (taking τ and R as fixed) to maximize

$$eu[(z_H - (T + s(T))(1 - e)/e))(1 - \tau) + R] + (1 - e)u[(z_L + T)(1 - \tau) + R] - \psi(e).$$

The first-order-condition for T is:

$$\frac{u'(c_L) - u'(c_H)}{u'(c_H)} = s'(T). \tag{15}$$

[12] Another important example of informal insurance in developing economies is risksharing in villages (Townsend 1994).

The government chooses τ to maximize social welfare, taking into account that $R = \tau\bar{w}$ where $\bar{w} = ew_H + (1-e)w_L$ and w_i is post-private insurance earnings.

$$\max_\tau W(\tau) = e^* u[(z_H - (T^* + s(T^*))(1-e^*)/e^*))(1-\tau) + \tau\bar{w}] + (1-e^*)u[(z_L + T^*)(1-\tau) + \tau\bar{w}] - \psi(e^*).$$

Using the envelope conditions for e^* and T^*, we immediately obtain the following formula for the welfare gain for raising τ:

$$\frac{dW}{d\tau} = -\bar{w} \cdot \bar{u}' \left[\frac{\tau}{1-\tau}\varepsilon_{\bar{w},1-\tau} + cov(w/\bar{w}, u'/\bar{u}')\right] \tag{16}$$

This equation is the same as the standard formula (2) in the benchmark single insurer model. It shows that private insurance that does not generate moral hazard can be ignored in the calculation of optimal government redistribution. This result is consistent with Chetty (2006a), who shows that Baily's (1978) formula is robust to allowing for arbitrary choices in the private sector as long as they are constrained efficient. To be clear, this result does *not* imply that the optimal government benefit level is unaffected by the presence of private insurance. A higher informal private insurance benefit T will reduce the optimal τ by reducing the correlation between u' and w in (16).

Hidden Skill Model. In the hidden skill model, informal private insurers can make lump sum transfers T_n to individuals with ability n by paying a loading cost $s(T_n)$.[13] Note that $T_n < 0$ corresponds to a lump sum tax on type n. The private insurer and individuals take τ and R as given, and simultaneously choose z_n and T_n to maximize

$$\int u\left[(z_n - (T_n + s(T_n))(1-\tau) + R - h\left(\frac{z_n}{n}\right)\right] f(n)dn.$$

subject to the budget constraint $\int T_n f(n)dn \geq 0$ (multiplier λ). The first order condition for z_n implies that $h'(z_n/n) = n(1-\tau)$, i.e., the government tax reduces incentives to work but not the amount of informal insurance. The first order condition for T_n implies that $u'(c_n) \cdot (1 + s'(T_n)) \cdot (1-\tau) = \lambda$: marginal utility is equalized up to the loading factor.

Let us denote by $w = z - (T + s(T))$ post-insurance earnings. The government chooses τ to maximize:

$$W(\tau) = \int u\left[(z_n - (T_n + s(T_n))(1-\tau) + \tau\bar{w} - h\left(\frac{z_n}{n}\right)\right] f(n)dn.$$

[13] For consistency with the moral hazard model, we can assume that $s(T) = 0$ for $T \leq 0$ and $s(T) > 0$ for $T > 0$).

Using the envelope conditions for z_n, we have:

$$\begin{aligned}\frac{dW}{d\tau} &= \int u' \cdot \left[-w + \bar{w} - \tau \frac{d\bar{w}}{d(1-\tau)} - (1+s'(T_n))(1-\tau)\frac{dT_n}{d\tau}\right] f(n)dn \\ &= -\bar{w} \cdot \bar{u}' \left[\frac{\tau}{1-\tau}\varepsilon_{\bar{w},1-\tau} + cov(w/\bar{w}, u'/\bar{u}')\right] - \lambda \int \frac{dT_n}{d\tau} f(n)dn,\end{aligned}$$

where the last term is obtained using $u'(c_n) \cdot (1 + s'(T_n)) \cdot (1-\tau) = \lambda$. The last term cancels out as $\int T_n f(n)dn = 0$ is constant with τ. Hence, we obtain again the standard formula.

Why does only private insurance that generates moral hazard change the formulas for optimal taxation? When private insurance generates moral hazard, the changes in e induced by government intervention have a first-order externality on the private-insurer's budget and must therefore be taken into account directly in the formula. If private insurance does not generate moral hazard, the fiscal externality term disappears because e is chosen jointly with T to optimize W.

The irrelevance of informal insurance again challenges the existing literature, which views all crowdout of private insurance as reducing the welfare gains from government intervention. In the context of informal insurance, the conventional wisdom would posit that if distortionary government insurance partly displaces non-distortionary private insurance, then government intervention increases moral hazard efficiency costs and is less desirable. This reasoning misses the point that informal insurance itself must have costs if individuals are not fully insured in equilibrium. If informal private insurance is optimized, the marginal welfare loss from having less informal insurance must be exactly offset by the marginal gain from paying less for that insurance ($s'(T)$), as shown by (15). Since the costs and benefits cancel out at the margin, the displacement of informal insurance by public insurance does not lead to a "leaky bucket" despite the fact that the government achieves smaller increases in insurance coverage per dollar spent.

4 Market Failures

In this section, we extend our formulas to allow for two market failures: pre-existing information and adverse selection. Pre-existing information refers to *public* information about individuals' ability before insurance contracts are signed. Adverse selection refers to *private* information about individuals' ability before insurance contracts are signed. Because income effects in the effort decision greatly complicate the analysis, we focus exclusively on the hidden skills model without income effects in the labor supply decision.[14]

4.1 Pre-Existing Information

We model pre-existing public information as follows. The population is partitioned into K exogenous groups $k = 1, .., K$. For instance, the groups could represent health status or geographic region of birth. Group k contains a fraction p_k of the population and has (conditional) density of abilities $f_k(n)$ so that $f(n) = \sum_k p_k f_k(n)$. Group identity is known by both private insurers and individuals before signing insurance contracts, but the realization of n within a group is unknown to both insurers and individuals when signing the contract. As a result, private insurance is offered *within* each group k. The private insurer offers an insurance rate t_k conditional on group k. When there is a single group, this corresponds to the model in Section 3. At the other extreme, when the number of groups is as large as the number of individuals, all information is revealed before contracts can be signed and there is no scope for private insurance contracts. In contrast, the government can impose redistributive taxation on the entire population.

We assume that the government is restricted to using a uniform tax rate τ on earnings. This is a strong assumption in the present model because the government could potentially improve welfare by conditioning tax rates on groups, but it is useful to gain insight into the key features of the problem.[15]

We now introduce notation for the multiple-group case. Let m_k denote the total tax rate in group k such that $1 - m_k = (1-\tau)(1-t_k)$. Individual n in group k optimally chooses earnings such that $h'(z/n) = (1 - m_k)n$. We denote by $\bar{z}_k = \int z f_k(n)dn$ average earnings in group k and by $\varepsilon_k = \frac{1-m_k}{\bar{z}_k}\frac{d\bar{z}_k}{d(1-m_k)}$ the elasticity of $\bar{z}_k$ with respect to $1 - m_k$. As in Section 3, let $w = (1-t_k)z + t_k\bar{z}_k$ denote earnings post private insurance (but before government taxation). Note that $\bar{w}_k = \bar{z}_k$ and $\bar{w} = \bar{z}$. Let $c = (1-\tau)w + \tau\bar{w} = (1-\tau)(1-t_k)z + t_k(1-\tau)\bar{z}_k + \tau\bar{z}$ denote disposable income. The crowd-out rate r_k of private insurance in group k is $1 - r_k = \frac{1-\tau}{1-m_k}\frac{d(1-m_k)}{d(1-\tau)}$. Note that this implies that $\frac{dm_k}{d\tau} = (1-t_k)(1-r_k)$ and $\frac{dt_k}{d\tau} = -r_k\frac{1-t_k}{1-\tau}$. As above,

[14] Laffont and Martimort (Section 7.1.3, 1999) analyze a moral hazard model with adverse selection. Even with only two levels of effort, the analysis and formulas (written in terms of primitives) are complex.

[15] In practice, horizonal equity considerations appear to prevent the government from using pre-existing public information such as height, age, family background, or education for tax or redistribution purposes.

we denote by $\varepsilon_{\bar{w},1-\tau} = \frac{1-\tau}{\bar{z}} \frac{d\bar{z}}{d(1-\tau)} = \frac{1-\tau}{\bar{w}} \frac{d\bar{w}}{d(1-\tau)}$ denote the elasticity of aggregate earnings with respect to the net-of-government tax rate $1-\tau$, *taking into account crowdout.* We denote by $cov(\bar{u}'_k, \bar{w}_k) = \sum_k p_k \bar{u}'_k [\bar{w}_k - \bar{w}]$ the covariance between average marginal utilities ($\bar{u}'_k$) and wages ($\bar{w}_k$) across the K groups. Let $cov_k(u', w) = \int u' \cdot (w - \bar{w}_k) f_k(n) dn$ denote the covariance between u' and w *within* group k. Finally, in the case where private insurance does not generate moral hazard, it has the same lump-sum and loading factor design as in Section 3.3 within each group k.

The government chooses τ to maximize social welfare under a utilitarian criterion:

$$W = \sum_k p_k \int u\left(z \cdot (1 - m_k) + t_k(1-\tau)\bar{z}_k + \tau \bar{z} - h\left(\frac{z}{n}\right)\right) f_k(n) dn.$$

The following proposition characterizes the solution to this problem.

Proposition 3

1) In the general case where private insurance is not necessarily optimized, the optimal government tax rate τ satisfies

$$\frac{\tau}{1-\tau} = -\frac{1}{\varepsilon_{\bar{w},1-\tau}} \cdot \left[\sum_k p_k (1 - r_k) \frac{cov_k(u', w) + \varepsilon_k t_k \bar{u}'_k \cdot \bar{w}_k}{\bar{u}' \cdot \bar{w}} + \frac{cov(\bar{u}'_k, \bar{w}_k)}{\bar{u}' \cdot \bar{w}}\right]. \qquad (17)$$

2) When private insurance is optimized within each group, $cov_k(u', w) + \varepsilon_k t_k \bar{u}'_k \cdot \bar{w}_k = 0$ for each k and hence the optimal government tax rate τ satisfies

$$\frac{\tau}{1-\tau} = -\frac{1}{\varepsilon_{\bar{w},1-\tau}} \cdot \frac{cov(\bar{u}'_k, \bar{w}_k)}{\bar{u}' \cdot \bar{w}}. \qquad (18)$$

3) When private insurance does not generate moral hazard, the optimal government tax rate τ follows the standard formula:

$$\frac{\tau}{1-\tau} = -\frac{1}{\varepsilon_{\bar{w},1-\tau}} \cdot \frac{cov(u', w)}{\bar{u}' \cdot \bar{w}}. \qquad (19)$$

Proof: We start with the general case. Using the envelope condition for individual z, the first order condition with respect to τ is:

$$0 = \frac{dW}{d\tau} = \sum_k p_k \int u'(c) \cdot \left[-z \frac{dm_k}{d\tau} - t_k \bar{z}_k + (1-\tau) \frac{dt_k}{d\tau} \bar{z}_k + t_k (1-\tau) \frac{d\bar{z}_k}{d\tau} + \bar{z} + \tau \frac{d\bar{z}}{d\tau}\right] f_k(n) dn.$$

which we can rewrite as:

$$\frac{\tau}{1-\tau} \cdot \varepsilon_{\bar{w},1-\tau} \cdot \bar{u}' \cdot \bar{z} = \bar{u}' \cdot \bar{z} - \sum_k p_k \bar{u}'_k \cdot \bar{z}_k - \sum_k p_k (1 - t_k)(1 - r_k) cov_k(z, u') - \sum_k p_k \bar{u}'_k \cdot \bar{z}_k \cdot t_k (1 - r_k) \varepsilon_k.$$

Note that $\bar{z} = \bar{w}$, $\bar{z}_k = \bar{w}_k$, and $cov_k(w, u') = (1 - t_k)cov_k(z, u')$. Hence

$$\frac{\tau}{1-\tau} \cdot \varepsilon_{\bar{w},1-\tau} \cdot \bar{u}' \cdot \bar{w} = -c\bar{o}v(\bar{u}'_k, \bar{w}_k) - \sum_k p_k(1 - r_k)[cov_k(w, u') + \varepsilon_k t_k \bar{u}'_k \cdot \bar{w}_k],$$

which demonstrates (17).

When the private insurer sets t_k optimally, it takes τ and $R = \tau\bar{z}$ as given and chooses t_k to maximize

$$\int u\left(z \cdot (1 - t_k)(1 - \tau) + t_k(1 - \tau)\bar{z}_k + R - h\left(\frac{z}{n}\right)\right) f_k(n)dn.$$

This is the same problem as in Section 3.2 and the first order condition leads to the same formula:

$$\frac{t_k}{1 - t_k} = \frac{-cov_k(z/\bar{z}_k, u'/\bar{u}'_k)}{\varepsilon_k},$$

which can be rewritten as $t_k\varepsilon_k = -cov_k(w/\bar{w}_k, u'/\bar{u}'_k)$. Therefore, the first term in the square bracket expression in (17) vanishes, yielding the simpler formula (18).

If there is no moral hazard in the private insurer choice, then, as in Section 3.3, lumpsum transfers T_n^k and z_n are chosen simultaneously to maximize

$$\int u\left[(z_n - (T_n^k + s_k(T_n^k))(1 - \tau) + R - h\left(\frac{z_n}{n}\right)\right] f_k(n)dn.$$

subject to the budget constraint $\int T_n^k f_k(n)dn \geq 0$ (with multiplier λ_k). The first order condition for T_n^k implies that $u_n'^{k}) \cdot (1 + s_n'^{k})) \cdot (1 - \tau) = \lambda_k$. The government chooses τ to maximize:

$$W(\tau) = \sum_k p_k \int u\left[(z_n - (T_n^k + s_k(T_n^k))(1 - \tau) + \tau\bar{w} - h\left(\frac{z_n}{n}\right)\right] f_k(n)dn.$$

Using the envelope conditions for z_n, we have:

$$\begin{aligned} \frac{dW}{d\tau} &= \sum_k p_k \int u' \cdot \left[-w + \bar{w} - \tau\frac{d\bar{w}}{d(1-\tau)} - (1 + s_n'^{k}))(1 - \tau)\frac{dT_n^k}{d\tau}\right] f_k(n)dn \\ &= -\bar{w} \cdot \bar{u}'\left[\frac{\tau}{1-\tau}\varepsilon_{\bar{w},1-\tau} + cov(w/\bar{w}, u'/\bar{u}')\right] - \sum_k p_k\lambda_k \int \frac{dT_n^k}{d\tau} f_k(n)dn \end{aligned}$$

The last term cancels out as $\int T_n^k f_k(n)dn = 0$ for each k and any value of τ, yielding (19). QED.

Equation (17) shows that the welfare gain from government intervention in this more general model consists of two elements: (1) increased insurance within groups, which is captured by parameters analogous to those analyzed in the baseline case, and (2) increased insurance across groups, which reflects the gains from pooling risk across types. In the present model, (1) only is operative if private insurance is not optimized; if it was, this term would be zero and only the second term is operative.

When private insurance is optimized, the optimal tax rate follows a modified formula where individuals in group k are treated as a *single* individual with marginal utility and earnings equal to the average in group k. Intuitively, private insurance takes care of within group insurance as well as the government could. Therefore, the government should focus solely on redistribution *across* groups, which cannot be insured by the private sector because of pre-existing information. Finally, (19) shows that the government can again ignore "informal" private insurance that does not general moral hazard and apply the standard optimal tax formula.

An Alternative Representation. It is useful to simplify formula (17) by defining measures of the aggregate crowdout parameter $\hat{r}$ and aggregate private insurance rate $\hat{t}$. Define

$$\hat{r} = \frac{\sum_k p_k r_k cov_k(u', w)}{cov(u', w)} = \frac{\sum_k p_k r_k cov_k(u', w)}{\sum_k p_k cov_k(u', w)} \cdot (1 - \rho),$$

as the (weighted) average of the crowdout rates r_k, multiplied by the fraction of risk that is within-group (and hence can potentially be insured by the private sector). Conceptually, $\hat{r}$ is a measure of the *total* crowdout of private insurance by government insurance. Similarly, define

$$\hat{t} = \sum_k p_k t_k \frac{\varepsilon_{\bar{w}_k, 1-\tau} \cdot \bar{w}_k \cdot \bar{u}'_k}{\varepsilon_{\bar{w}, 1-\tau} \cdot \bar{w} \cdot \bar{u}'},$$

as the (weighted) average of the private insurance rates t_k. Using the standard covariance decomposition, we have

$$\begin{aligned} cov(u', w) &= \sum_k p_k \int u'(w - \bar{w}) f_k \\ &= \sum_k p_k \int u'(w - \bar{w}_k + \bar{w}_k - \bar{w}) f_k = \sum_k p_k cov_k(u', w) + cov(\bar{u}'_k, \bar{w}_k). \end{aligned}$$

Denote by $\rho = cov(\bar{u}'_k, \bar{w}_k)/cov(u', w)$ the fraction of the covariance between earnings w and marginal utility u' that is due to cross-group differences. We can then rewrite equation (17) as follows:

$$\frac{\tau}{1 - \tau} = -\hat{t} + \frac{1 - \hat{r}}{\varepsilon_{\bar{w}, 1-\tau}} \cdot \frac{-cov(u', w)}{\bar{u}' \cdot \bar{w}}. \tag{20}$$

This alternative representation of (17) has exactly the same form as equation (9) obtained in Section 2, which allowed for non-optimized private insurance but did not explicitly model market failures. This result shows that the baseline formula holds in the model with pre-existing information, provided that the tax rate and crowdout parameter are measured appropriately.[16]

4.2 Adverse Selection

Now suppose that each individual knows which group he belongs to before signing insurance contracts, but insurers do not have this information. This model is an extension of the classic adverse selection models proposed by Rothschild and Stiglitz (1976) and Wilson (1977). If private insurers were to offer a menu of competitive insurance contracts with insurance rates $(t_1, .., t_M)$, individuals in group k would self-select the insurance rate that yields the highest expected utility and private insurers would make negative profits. Hence, private insurers are forced to adjust contracts to respect incentive compatibility constraints. A strong equilibrium is defined as a set of contracts such that no firm can make positive profits by offering a new insurance contract. When such strong equilibria exist, they are always separating in the sense that each group k selects a single and specific insurance rate t_k (Rothschild and Stiglitz 1976). When such strong equilibria do not exist, it is necessary to weaken the concept of equilibrium to obtain existence of an equilibrium (Wilson 1977).

Here, we restrict attention to the region of the parameter space where an equilibrium exists. In this equilibrium, each group k is offered private insurance at tax rate t_k. The t_k can be common across some groups if there is partial pooling. Furthermore the t_k depend on the government tax rate τ. Assume that the t_k are smooth functions of τ so that crowdout rates r_k are well defined and welfare gains are smooth.[17] The revelation principle implies that in equilibrium each individual will truthfully reveal his type k. Hence, the equilibrium in this model can be viewed as a special case of the pre-existing information model analyzed above, where private insurance contracts were set arbitrarily. It follows that the general formula (17) in Proposition 3.1 and the alternative representation in (20) hold in this environment.

[16] Implementing (20) exactly would require knowledge about the structure of the groups and private insurance contracts to measure $\hat{t}$ and $\hat{r}$. The measurement of the amount of private insurance is complicated in tax models because it requires knowledge of marginal products. In social insurance applications, the amount of private insurance can be measured directly because there are explicit transfers.

[17] In cases with multiple equilibria, this assumption requires that the allocation never jumps across equilibria following a small change in τ.

Adverse selection does, however, affect the formula with optimized private insurance (18) because it distorts the provision of private insurance within groups. This leads to a violation of the within-group private insurance optimality condition $t_k^* = -cov_k(w/\bar{w}_k, u'/\bar{u}_k')/\varepsilon_k$. To obtain further intuition, rewrite (17) as:

$$\frac{\tau}{1-\tau} = -\frac{1}{\varepsilon_{\bar{w},1-\tau}} \cdot \frac{c\bar{o}v(\bar{u}_k', \bar{w}_k)}{\bar{u}' \cdot \bar{w}} + \frac{1}{\varepsilon_{\bar{w},1-\tau}} \cdot \sum_k p_k(1-r_k)\varepsilon_k \cdot (t_k^* - t_k) \cdot \frac{\bar{u}_k' \cdot \bar{w}_k}{\bar{u}' \cdot \bar{w}}, \tag{21}$$

where t_k^* is the optimal private insurance rate in the absence of adverse selection constraints. There are two components in (21). The first term reflects the value of redistribution across the groups, as in (18). The second term reflects the value of fixing the distortions created by adverse selection *within* each group. Typically, adverse selection leads to under-provision of private insurance: $t_k < t_k^*$ (Rothschild and Stiglitz 1976). Hence, there is value to increasing within-group redistribution even if private contracts are optimized in the presence of adverse selection.

More generally, (21) implies that the government tax rate τ should be *higher* than in the case with no adverse selection (18) because it fixes two market failures rather than one. This effect is captured in the general formula in (20) with non-optimized private insurance contracts, because one would observe a higher absolute value of $cov(w, u')$ if adverse selection were to limit the degree of private insurance. Intuitively, the reason that private insurance is under-provided in equilibrium – be it adverse selection, pre-existing information, or imperfect optimization – does not matter when measuring the welfare gains of private insurance in the class of models we have analyzed. Since (20) has the same form as (9) up to the weighting of t and r, we conclude that (9) and (8) can be used to obtain simple approximate measures of the welfare gains from government intervention with endogenous private insurance.

5 Empirical Applications

In this section, we apply the formulas derived above to analyze the welfare gains from social insurance. To do so, we first present an alternative representation of the formula for the marginal welfare gain from government insurance written in terms of the parameters estimated in empirical studies of social insurance.

5.1 Welfare Gain from Social Insurance

The moral hazard model with two states (z_L, z_H) is isomorphic to standard social insurance models (e.g. Baily 1978), where the low state is interpreted as a negative shock such as job loss or illness. In the social insurance literature, contracts are characterized by a benefit B paid to the individual in the low state and a premium T that the individual pays to the insurer in the high state. Since the insurer's break-even constraint is $T = (1-e)\cdot B/e$, the consumption levels are a function of the single parameter B:

$$c_H = z_H - \frac{1-e}{e}B,\ c_L = z_L + B.$$

Individual effort e^* is a function of the benefit level B. The second-best optimal contract chooses B to maximize

$$W = e^* u(z_H - \frac{1-e}{e}B) + (1-e^*)u(z_L + B) - \psi(e^*). \tag{22}$$

This problem is equivalent to the optimal taxation problem in (1) because any tax contract m can be mapped to an equivalent social insurance contract B by setting $B = (\bar{z} - z_L)m$. The only difference between the models is that they use different notation to describe the contracts.

Let $\varepsilon_{1-e,B}$ denote the elasticity of the probability of the low state $(1-e^*)$ with respect to B. The government sets an insurance benefit level b financed by a premium $\tau = b(1-e)/e$ and the private insurer sets a benefit level b_p financed by a premium $\tau_p = b_p(1-e)/e$. Effort e depends on total benefit $B = b + b_p$. The private insurance benefit level b_p depends on the government benefit b according to a function $b_p(b)$. Because the benefits are additive rather than multiplicative, it is convenient to define the crowdout parameter as $r = -\frac{db_p}{db}$ in the social insurance scenario. With endogenous private insurance, the government chooses b to maximize

$$W = e^* u(z_H - \frac{1-e}{e}(b_p(b) + b)) + (1-e^*)u(z_L + b_p(b) + b) - \psi(e^*). \tag{23}$$

Proposition 4. The welfare gain from raising the government social insurance benefit is:

$$\frac{dW}{db} = (1-e)(1-r)\cdot u'(c_H)\cdot\left[\frac{u'(c_L) - u'(c_H)}{u'(c_H)} - 1 - \frac{\varepsilon_{1-e,b}}{e}\cdot\frac{1 + b_p/b}{1-r}\right]. \tag{24}$$

Proof. The proof follows Proposition 1: (23) is identical to the problem in (22), with the total benefit $B = b_p(p) + b$ replacing b. Choosing b is equivalent to choosing the total benefit B.

Differentiating (22) yields $\frac{dW}{dB} = (1-e)u'(c_H)\left[\frac{u'(c_L)-u'(c_H)}{u'(c_H)} - \frac{\varepsilon_{1-e,B}}{e}\right]$. Observe that $dW/db = dW/dB\cdot(1-r)$. Likewise, $de/db = de/dB\cdot(1-r)$ and thus $\varepsilon_{1-e,B} = -(b+b_p)(de/dB)/(1-e) = (1+b_p/b)\cdot\varepsilon_{1-e,b}/(1-r)$. Plugging these expressions into the expression for $\frac{dW}{dB}$ yields (24). QED.

The first term in 24) measures the gap in marginal utilities across the two states, which captures the marginal value of insurance. The second term captures the cost of insurance through the behavioral response. Analogous to the tax scenario, private insurance amplifies the second term and makes $\frac{dW}{db}$ smaller through two channels: the crowdout effect $1-r$ in the denominator and the mechanical effect b_p/b in the numerator. The crowdout term reflects a rescaling to recover the fundamental elasticity $\varepsilon_{1-e,B}$ and the b_p/b term reflects the reduction in b required to achieve the optimal level of B. The same formula holds (with appropriate measures of aggregate crowdout and private insurance) a model that permits pre-existing information and adverse selection, as in section 4.

The expressions for $\frac{dW}{db}$ and $\frac{dW}{d\tau}$ derived above all measure the marginal welfare gain of changing taxes and social insurance benefits in utils. To convert these expressions into an interpretable money metric, we normalize the welfare gain from a \$1 (balanced budget) increase in the size of the insurance program by the welfare gain from raising the wage bill in the high state by \$1. In particular, define

$$G(b) = \frac{dW}{db}\frac{1}{1-e}/\frac{dW}{dz_H}\frac{1}{e} \tag{25}$$

$$= \frac{u'(c_L)-u'(c_H)}{u'(c_H)} - \frac{\varepsilon_{1-e,b}}{e}\frac{1+b_p/b}{1-r} \tag{26}$$

We now apply (26) to characterize the welfare gains from increasing unemployment and health insurance benefits. These calibrations are intended primarily to illustrate the potential impacts of endogenous private insurance on calculations of welfare gains from government intervention rather than for policy analysis. These simple calculations do not account for all margins of behavioral responses and for heterogeneity across individuals.

5.2 Application 1: Unemployment Insurance

The existing literature on optimal unemployment insurance essentially ignores the existence of private insurance. Much of private insurance against unemployment is provided through informal risk sharing that is unlikely to generate much moral hazard, and hence can be ignored in the calculation of optimal benefits according to the results in section 3.3. However, many private firms provide unemployment insurance in the form of severance payments – lump sum cash grants made by firms to workers who are laid off. Unlike government-provided unemployment benefits, severance pay does not distort job search behavior after job loss because it does not affect marginal incentives to search. Severance pay can, however, distort effort choices while working by changing the relative price of being unemployed relative to having a job.

In this section, we calibrate the welfare gain from raising the UI benefit level when the response of severance pay to UI benefits is taken into account. To adapt the optimal UI problem to our static framework above, we ignore the job search decision, treating search effort after job loss as invariant to the UI benefit level. Instead, we focus on the distortion in the probability of job loss (e.g. due to shirking) caused by UI benefits and severance pay. In our static model, both UI benefits and severance pay act as transfers to the unemployed state, and are financed by taxes in the employed state.[18]

Estimation of Crowdout Elasticity. As an illustration of the data and empirical strategy needed to implement our formula with endogenous private insurance, we begin by estimating the two key parameters – the size of the private insurance market (b_p/b) and the crowdout effect (r). To do so, we use data on severance pay from a survey conducted by Mathematica on behalf of the Department of Labor. The dataset (publicly available from the Upjohn Institute) is a sample of unemployment durations in 25 states in 1998 that oversamples UI exhaustees. We reweight the data using the sampling weights to obtain estimates for a representative sample of job losers. The dataset contains information on unemployment durations, demographic characteristics, and an indicator for receipt of severance pay. There are 3,395 individuals in the sample, of whom 508 report receiving a severance payment. See Chetty (2008a) for further details on the dataset and sample construction. We obtain data on mean unemployment benefits by state in 1998 from the Department of Labor.

To calculate b_p/b, first note that 15% of job losers report receiving severance pay in our data. According to calculations reported in Chetty (2008a), the mean severance payment

[18] A more precise calibration would take account of the fact that UI benefits are conditioned on duration, and thus are larger when a worse "state" is realized. This calibration would require separate estimates of the effect of UI benefits and severance pay on the probability of job loss.

conditional on receipt of severance pay is equal to 10.7 weeks of wages, the mean UI benefit level is 50% of the wage, and the mean unemployment duration is 15.8 weeks. Hence, in the aggregate population, the ratio of total private insurance to total public insurance is

$$\frac{b_p}{b} = \frac{0.15 \times 10.7}{0.5 \times 15.8} = 0.20.$$

To estimate r – the effect of an increase in the UI benefit level on severance pay – we exploit variation in UI benefit levels across states. Most states pay a fixed wage replacement rate up to a maximum, which varies considerably across states and thereby creates variation in UI benefit levels. The maximum benefit can be viewed as an instrument for individual benefit levels (Meyer 1990). We do not exploit the variation in benefit levels across individuals within a state because of endogeneity concerns. Although state benefit maximums are exogenous to individual characteristics, they are not orthogonal to all aspects of the economic environment. In particular, richer states (or those with a higher cost of living) provide both more public and private insurance. As a result, both state unemployment benefit maximums and the fraction of individuals receiving severance pay are positively correlated with mean wage rates in each state. To account for this confounding factor, we control for wages throughout our analysis using a flexible 10 piece spline for the individual log wage.

We begin with a simple graphical analysis to illustrate our estimation of the crowdout effect. Figure 1 plots the relationship between average severance pay receipt and the maximum UI benefit level, conditioning on wages. To construct this figure, we first regress the severance pay dummy on the wage spline and the maximum UI benefit level on the wage spline and compute residuals. We then compute mean residuals of both variables by state. The figure is a scatter plot of the mean residuals. We exclude states that have fewer than 50 individuals from this figure to reduce the influence of outliers on the graph; all observations are included in the regression analysis below. The figure shows that states with higher UI benefit levels have fewer severance payments, indicating that private insurance is crowded out to some extent by public insurance.

To quantify the amount of crowdout, we estimate a set of regression models of the following form:

$$sev_i = \alpha + \beta \log b_i + f(w_i) + \gamma X_i + \varepsilon_i \tag{27}$$

where sev_i is an indicator for whether individual i received a severance payment, b_i is a measure of the UI benefit level for individual i, $f(w_i)$ denotes the wage spline, and X_i denotes a vector of additional controls.

Specification 1 of Table 1 reports estimates of (27) without any additional (no X), with b_i equal to the maximum benefit level in the state where individual i lives. Standard errors in this and all subsequent specifications are clustered by state to adjust for arbitrary within-state correlation in errors. The estimated coefficient of $\beta = -0.075$ implies that a doubling the UI benefit maximum would reduce the fraction of individuals receiving severance pay by 7 percent. Specification 2 replicates 1 with the following individual-level covariates: job tenure, age, gender, household size, education, dropout, industry, occupation, and race dummies. The point estimate on the UI benefit level is not affected significantly by the inclusion of these controls.

Specifications 1 and 2 can be interpreted as "reduced form" regressions which show the effect of the instrument (maximum benefit levels) on severance pay. To obtain an estimate of the effect of a $1 increase in the benefit level on the probability of severance pay receipt, we estimate a two-stage least squares regression, instrumenting the log individual benefit level with the log state maximum. The estimate on the log individual benefit, reported in column 3 of Table 2, is $\beta = -0.105$. Doubling the UI benefit level would reduce severance receipt by 10.5 percentage points, relative to a mean value of 15%, implying $\varepsilon_{b_p,b} = -0.7$. We conclude that $r = -\frac{db_p}{db} = -\varepsilon_{b_p,b}\frac{b_p}{b} = 0.7 \times 0.2 = 0.14$.

The identification assumption underlying these regressions is that the cross-state variation in UI benefit maximums is orthogonal to other determinants of severance pay receipt conditional on wage levels. Most plausible endogeneity stories would work toward attenuating our estimate of the crowdout effect. For example, suppose states with higher UI benefit maximums are populated by individuals who are more risk averse and therefore place higher value on insurance. Such states would also have higher private insurance, biasing the correlation between the UI benefit level and severance pay receipt upward. Given these concerns about policy endogeneity, our simple empirical analysis should be viewed as illustrative. Future work should exploit within-state variation in UI benefits (as in Meyer 1990) to obtain a more credible and precise estimate of the crowdout effect.

Calibration. We now use the estimates above in conjunction with other estimates from the existing literature to calibrate the welfare gain from raising the UI benefit level. Using the normalization in (25) and the formula for $\frac{dW}{db}$ in (24), the welfare gain from increasing total government expenditure on unemployment insurance by \$1 ($\frac{dW}{db}/(1-e)$) relative to the welfare gain of a \$1 increasing the wage of the employed agent ($\frac{dW}{dz_H}/e = u'(c_H)$) is

$$G(b) = (\frac{dW}{db}\frac{1}{1-e})/(\frac{dW}{dz_H}\frac{1}{e}) = (1-r)(\frac{u'(c_L)}{u'(c_H)} - 1 - \frac{\varepsilon_{1-e,b}}{e}\frac{1+b_p/b}{1-r})$$

We calibrate $G(b)$ using the following inputs:

$$\begin{aligned}
e &= 0.95 \text{ from CPS statistics (5\% unemployment rate)} \\
r &= -\frac{db_p}{db} = 0.14 \text{ from calculations above} \\
\frac{b_p}{b} &= 0.2 \text{ from calculations above} \\
\frac{c_e}{c_u} &= \frac{1}{0.9} \text{ from Gruber (1997)} \\
\gamma &= 2 \text{ from Chetty (2006b)}
\end{aligned}$$

Under the approximation that utility exhibits constant relative risk aversion between c_u and c_e, $\frac{u'(c_u)}{u'(c_e)} = (\frac{1}{0.9})^\gamma$ where γ denotes the coefficient of relative risk aversion. The remaining parameter, for which we have no existing estimate, is $\varepsilon_{1-e,b}$ – the elasticity of the probability of job loss with respect to the UI benefit level b. Leaving this parameter unspecified and plugging in the remaining values into the formula for $\frac{dW}{db}$, we obtain

$$G(b) = (1-0.14)(0.23 - 1.47\varepsilon_{1-e,b}).$$

It follows that if the job loss elasticity $\varepsilon_{1-e,b} > 0.15$, $\frac{dW}{db} < 0$ at present UI benefit levels when crowdout is taken into account. In contrast, if we were to apply a formula that does not take crowdout of private insurance into account, we would obtain

$$G(b) = (0.23 - 1.05\varepsilon_{1-e,b}).$$

Hence, an analyst who ignores crowdout would conclude that the welfare gain from raising the UI benefit level is negative only if $\varepsilon_{1-e,b} > 0.25$ (ignoring distortions in unemployment durations). We conclude that in this application, there is a significant but modest range of parameters for which adjusting the formula for endogenous private insurance leads to different policy implications.

5.3 Application 2: Health Insurance

A Model of Health Shocks. We adapt our two-state analysis of social insurance to the case of health insurance using an extensive-margin model of health consumption. Suppose that purchasing health care costs $\$C$. There is a continuum of agents in the economy who differ only in their valuation of health care. Agent i gets a benefit from health care equivalent to v_i utils. Hence an agent buys health care iff

$$v_i > u(w - \tau - \tau_p) - u(w - C + b + b_p) = z$$

Valuation of healthcare v_i is distributed according to a cdf F. The fraction of agents who buy health care is

$$s = 1 - F(z) = \int_z^\infty dF$$

Let F^{-1} denote the inverse of F. Then $z = F^{-1}(1-s)$ and the aggregate utility gain from consumption of health care is

$$v(s) = \int_{F^{-1}(1-s)}^\infty v_i dF$$

Note that $v'^{-1}(1-s) = z = u(w - \tau - \tau_p) - u(w - C + b + b_p)$. Therefore $v(.)$ is increasing in s and $v'(.)$ is decreasing in s so that $v(.)$ is concave. Aggregating over the agents yields the following social welfare function:

$$W = (1-s)u(w - \tau - \tau_p) + su(w - C + b + b_p) + v(s)$$

The fraction of agents who consume health care s is effectively chosen to maximize W, taking the government and private insurance contracts as given. Since v is increasing and concave, this problem has the same structure as that analyzed in section 2, with e replaced by $1-s$ and $\psi(e)$ replaced by $-v(s)$. Applying (24) immediately yields the following formula for the aggregate utility gain from raising the public health insurance benefit level b by \$1:

$$\frac{dW}{db} = s(1-r)u'(c_H)(\frac{u'(c_L)}{u'(c_H)} - 1 - \frac{\varepsilon_{s,b}}{1-s}\frac{1+b_p/b}{1-r}) \tag{28}$$

To convert this expression into a money metric, we follow (25) and compare the welfare gain from increasing total government expenditure on health insurance by \$1 ($\frac{dW}{db}/s$)) to the welfare gain of spending \$1 on a wage increase for healthy individuals ($\frac{dW}{dz_H}/(1-s) = u'(c_H)$):

$$G(b) = (\frac{1}{s}\frac{dW}{db})/(\frac{1}{1-s}\frac{dW}{dz_H}) = (1-r)(\frac{u'(c_L)}{u'(c_H)} - 1 - \frac{\varepsilon_{s,b}}{1-s}\frac{1+b_p/b}{1-r})$$

Calibration. We calibrate the formula using the following inputs drawn from the empirical literature:

$$\begin{aligned}
\varepsilon_{s,C} &= -0.2 \text{ from Manning et al. (1987)} \\
s &= 0.1 \text{ from Manning et al. (1987) for inpatient usage rate} \\
r &= -\frac{db_p}{db} = 0.5 \text{ from Cutler and Gruber (1996a)} \\
\frac{b_p}{b} &= 0.89,\ \frac{b}{C} = 0.45 \text{ from National Health Care Statistics Table 6 (2006)} \\
\frac{c_e}{c_u} &= \frac{1}{0.85} \text{ from Cochrane (1991)} \\
\gamma &= 2 \text{ from Chetty (2006b)}
\end{aligned}$$

Under CRRA utility, these parameters imply that $\frac{u'(c_u)}{u'(c_e)} = (\frac{c_e}{c_u})^\gamma = (\frac{1}{0.85})^2 = 1.384$. Also note that $\varepsilon_{s,b} = -\varepsilon_{s,C}\frac{b}{C} = 0.2 \times .45 = 0.09$. Hence

$$G(b) = (1 - 0.5) \times (0.384 - \frac{0.2 \times .453}{0.9}\frac{1 + 0.89}{0.5}) = 0.0017$$

If we had ignored crowdout, we would have obtained

$$G(b) = (0.384 - \frac{0.2 \times .453}{0.9}) = 0.28$$

Taking crowdout into account lowers the estimate of $G(b)$ by a factor of more than 100. An analyst who ignored crowdout and applied existing formulas (e.g. Chetty 2006a) would infer that a \$100 million expansion in public health insurance programs would, loosely speaking, generate \$28 million in surplus net of the requisite tax increases. This analyst would mistakenly conclude that substantial expansions in the overall level of public health insurance are desirable. Taking crowdout into account implies that we are near the optimum in terms of *aggregate* public health insurance levels, as a \$100 million across-the-board expansion would generate only \$0.17 million in net social surplus.

There are several important caveats to this calibration that should be kept in mind when evaluating the policy implications of this simple calibration. First, this calculation does not fully account for adverse selection, as it neglects the benefits of redistributing across pre-existing risk types via government insurance reflected in the first term of (21). Second, the calibration ignores the correlation between health shocks and income inequality, which could potentially increase the welfare gains from health insurance (Cremer and Pestieau 1996). Third, private health insurance benefits are already tax subsidized in the United States, an issue that we have neglected in the calculation above. Finally, and perhaps most importantly, our aggregate welfare gain calculation ignores substantial heterogeneity across types of people, conditions, Medicare vs. Medicaid, etc. For some subgroups, such as the uninsured, there could clearly be substantial welfare gains from increasing public insurance benefits whereas for others there could be substantial welfare gains from cutting benefits. Equation (28) should be applied with group-specific estimates of the inputs to identify how government health insurance expenditures should be restructured to maximize welfare.

6 Conclusion

This paper has characterized the welfare gain from public insurance in the presence of endogenous private insurance. The formulas for optimal tax and social insurance policies derived here highlight two general parameters as the determinants of how private insurance impacts the welfare gains from social insurance: (1) the size of the formal private insurance market and (2) the crowdout of formal private insurance by public insurance. Like recent "sufficient statistic" formulas for welfare analysis, the formulas we have derived can be implemented using reduced-form empirical evidence without full identification of the model's primitives. However, unlike existing sufficient statistic formulas, we cannot make strong claims about the robustness of our formulas because they have been derived in stylized models. While experience with earlier sufficient statistic formulas suggests that the results may generalize (Chetty 2008b), an important difference in models with endogenous private insurance is that the envelope conditions that underlie the robustness of previous formulas may not hold when the private sector outcome is constrained inefficient. Thus, the source of the deviation from constrained efficiency – e.g. asymmetric information, imperfect optimization, or administrative costs – could potentially affect the formula. We believe that the parameters we have identified are likely to be relevant in more general models, but other factors may also prove be relevant.

We see three major directions for generalization of our analysis. First, one should characterize the effects of government intervention on the equilibria in the adverse selection model, as in Rothschild and Stiglitz (1976) and Wilson (1977). We have implicitly assumed that contracts and behavior will respond smoothly to changes in government policies, but in a setting with multiple equilibria, there could be jumps between equilibria that would invalidate our formula. Second, it would be useful to extend the analysis to allow for different loading factors for private and public insurance, microfounded via increasing returns or wasteful marketing costs. Conversely, one could allow for different levels of efficiency, reflecting the possibility that private insurers could be more efficient because of competitive pressure. Finally, in the simple model we analyzed here, the best policy is simply to rule out formal private insurance and have the government provide all insurance. However, there are some areas in which private insurers have an informational advantage relative to the government. For instance, employers have more information on effort on the job, making the moral hazard problem smaller for the employer. Characterizing the optimal mix of government and private insurance is an important next step.

If government and private insurers optimize along the lines described by our analysis, our model makes testable predictions about the pattern of insurance contracts we should observe. For instance, private insurance should be more prevalent in economies with low job mobility (such as Japan), where firms have the ability to insure shocks through a compressed wage structure without facing as much adverse selection. Another prediction is that government insurance should be more prevalent for shocks that occur prior to the point at which insurance contracts can be purchased, such as disability at birth, or for shocks where optimization of insurance purchases is unlikely. It would be valuable to test empirically whether observed contracts deviate systematically from these theoretical prescriptions.

References

Akerlof, George A. (1970) "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 84(3): 488-500.

Attanasio, Orazio and Jose-Victor Rios-Rull (2000) "Consumption smoothing in island economies: Can public insurance reduce welfare?" *European Economic Review*, 44(7): 1225-1258.

Baily, Martin (1978) "Some Aspects of Optimal Unemployment Insurance," *Journal of Public Economics,* 10: 379-402.

Barigozzi, Francesca (2006) "Supplementary Insurance with Ex-Post Moral Hazard: Efficiency and Redistribution," *Annales d'Economie et de Statistique*, 83-84 (July-December): 295-325.

Besley, Timothy (1989) "Publicly provided disaster insurance for health and the control of moral hazard," *Journal of Public Economics,* 39: 141–156.

Blomqvist, Ake and Henrik Horn (1984) "Public health insurance and optimal income taxation," *Journal of Public Economics*, 24: 352-371.

Blomqvist, Ake and P.O. Johannson (1997) "Economic efficiency and mixed public/private insurance," *Journal of Public Economics* 66: 505–516.

Boadway, Robin, Manuel Leite-Monteiro, Maurice Marchand, and Pierre Pestieau (2006) "Social Insurance and Redistribution with Moral Hazard and Adverse Selection," *Scandinavian Journal of Economics,* 108(2): 279-298.

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

Chetty, Raj (2006b) "A New Method of Estimating Risk Aversion," *American Economic Review,* 96 (December): 1821-1834.

Chetty, Raj (2008a) "Moral Hazard vs. Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy,* 116(2): 173-234.

Chetty, Raj (2008b) "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," UC-Berkeley mimeo.

Cochrane, John H. (1991) "A Simple Test of Consumption Insurance," *Journal of Political Economy*, 99: 957-76.

Cremer, Helmuth and Pierre Pestieau (1996) Redistributive taxation and social insurance (with H. Cremer), *International Taxation and Public Finance*, 3: 259-280.

Cullen, Julie Berry and Jonathan Gruber (2000). "Does Unemployment Insurance Crowd Out Spousal Labor Supply?" *Journal of Labor Economics*, 18(3): 546-72, July.

Cutler, David and Jonathan Gruber (1996a) "Does public insurance crowd out private insurance?" *Quarterly Journal of Economics,* 111(2): 391–430.

Cutler, David and Jonathan Gruber (1996b) "The effect of medicaid expansions on public insurance, private insurance and redistribution," *American Economic Review*, 86(2): 378–383.

DellaVigna, Stefano (2008) "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature,* forthcoming.

Diamond, Peter (1998) "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88: 83-95.

Encinosa, William (2003) "The economics of crowd out under mixed public/private health insurance," *Portuguese Economic Journal*, 2: 71–86.

Finkelstein, Amy (2004) "The interaction of partial public insurance programs and residual private insurance markets: evidence from the US Medicare program," *Journal of Health Economics,* 23(1): 1-24.

Ginsburg, P.B. (1988) "Public insurance programs: medicare and medicaid" In Frech, H.E. (Ed.), *Health Care in America: The Political Economy of Hospitals and Health Insurance.* Pacific Institute for Public Policy: San Francisco.

Golosov, Michael and Aleh Tsyvinski (2007) "Optimal Taxation with Endogenous Insurance Markets," *Quarterly Journal of Economics,* 121: 487-534.

Gruber, Jonathan (1997) "The Consumption Smoothing Benefits of Unemployment Insurance," *American Economic Review,* 87 (March): 192-205.

Kaplow, Louis (1991) "Incentives and Government Relief for Risk," *Journal of Risk and Uncertainty,* 3: 167-175.

Laffont, Jean-Jacques and David Martimort (1999) *The Theory of Incentives*, Princeton University Press: Princeton and Oxford.

Manning et al. (1987) "Health Insurance and the Demand for Medical Care," *American Economic Review,* 77: 251-277.

Meyer, Bruce (1990) "Unemployment Insurance and Unemployment Spells," *Econometrica,* 58 (July): 757-782.

Mirrlees, James A. (1971) "An exploration in the theory of optimum income taxation," *Review of Economic Studies,* 38: 175-208.

National Health Care Statistics (2006). Available at http://www.cms.hhs.gov/NationalHealthExpendData/downloads/tables.pdf

Pauly, Mark (1974) "Overinsurance and public provision of insurance: the roles of moral hazard and adverse selection," *Quarterly Journal of Economics,* 88(1): 44–62.

Petretto, Alessandro (1999) "Optimal Social Health Insurance with Supplementary Private Insurance," *Journal of Health Economics,* 18(6): 727-45.

Prescott, Edward C. and Robert M. Townsend (1984a) "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard," *Econometrica* 52(1): 21-45.

Prescott, Edward C. and Robert M. Townsend (1984b) "General Competitive Analysis in an Economy with Private Information," *International Economic Review* 25(1): 1-20.

Reinhardt, Uwe E., Peter S. Hussey and Gerard F. Anderson (2004) "U.S. Health Care Spending In An International Context," *Health Affairs,* 23(3): 10-25

Rochet, Jean-Charles (1991) "Incentives, redistribution and social insurance," *The Geneva Papers on Risk and Insurance,* 16: 143-165.

Rothschild, Michael and Joseph Stiglitz (1976) "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information " *Quarterly Journal of Economics* 90(4): 629-649.

Saez, Emmanuel (2001) "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68: 205-229.

Schoeni, Robert (2002) "Does unemployment insurance displace familial assistance?" *Public Choice,* 110: 99–119.

Selden, Thomas (1993) "Should the government provide catastrophic insurance?" *Journal of Public Economics,* 51: 241-247.

Shimer, Robert and Ivan Werning (2007) "Reservation Wages and Unemployment Insurance," *Quarterly Journal of Economics*, 122(3): 1145-1185.

Spinnewijn, Johannes (2008) "Unemployed, but Optimistic: Optimal Insurance Design with Biased Beliefs," Unpublished MIT mimeo.

Varian, Hal R. (1980) "Redistributive Taxation as Social Insurance," *Journal of Public Economics,* 14(1): 49-68.

Taylor, A.K., Farley Short, P., Horgan, C.M. (1988) "Medigap insurance: friend or foe in reducing medicare deficits?" In Frech, H.E. (Ed.), *Health Care in America: The Political Economy of Hospitals and Health Insurance*. Pacific Institute for Public Policy, San Francisco.

Townsend, Robert (1994) "Risk and Insurance in Village India," *Econometrica,* 62(3): 539-591.

Wilson, Charles (1977) "A model of insurance markets with incomplete information," *Journal of Economic Theory,* 16: 167-207.

Wolfe, J.R., Goddeeris, J.H. (1991) "Adverse selection, moral hazard, and wealth effects in the medigap insurance market," *Journal of Health Economics*, 10: 433-459.

Woolhandler, Steffie, Terry Campbell, and David Himmelstein (2003) "Costs of Health Care Administration in the United States and Canada," *New England Journal of Medicine,* 349: 768-775.

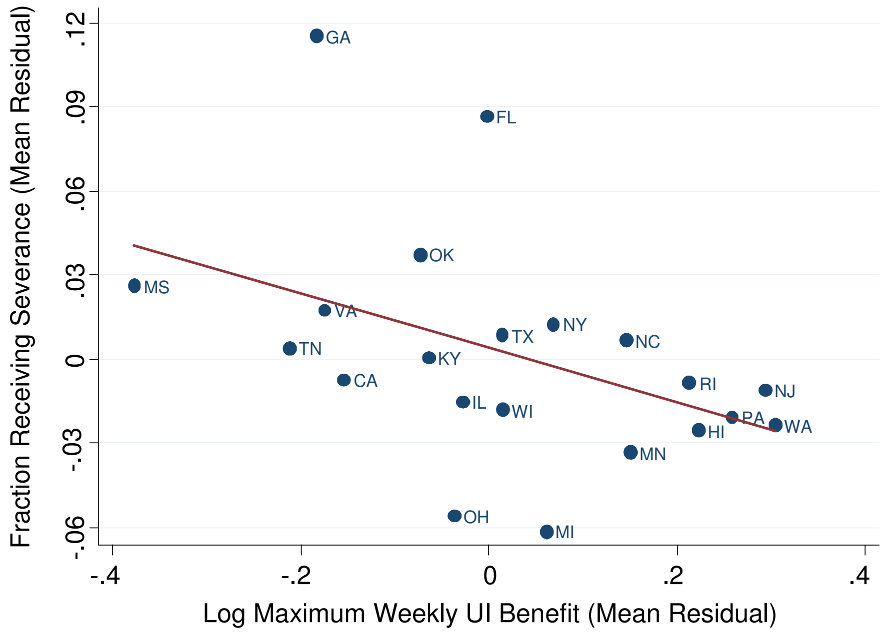
TABLE 1

Effect of UI Benefits on Severance pay: Regression Estimates

Dependent Variable: Severance pay			
	Reduced-Form OLS		TSLS
	No Controls	With Cntrls	With Cntrls
	(1)	(2)	(3)
log max UI benefit	-.074	-.065	
	(0.030)	(0.030)	
log individual UI benefit			-.105
			(0.054)
Sample Size	2,996	2,733	2,733

NOTE-Specifications 1 and 2 report estimates from an OLS regression; specification 3 reports estimates from a two-stage least squares regression using log state max benefit as an instrument for actual individual benefit reported in data. Specifications 2 and 3 include the following controls: job tenure, age, gender, household size, education, dropout, industry, occupation, and race dummies.

Figure 1
Effect of UI Benefits on Severance Pay



NOTE–Figure plots relationship between fraction of individuals receiving severance pay in each state vs. maximum state UI benefit level, conditioning on wages. Figure shows a scatter plot of the mean residuals by state from a regression of severance pay receipt and log maximum weekly benefit level on a log wage spline (see text for details). Data source: Mathmetica survey of UI Exhaustees in 25 States in 1998. States with fewer than 50 individual observations are excluded from this figure.