Optimal Social Insurance with Individual Private Insurance and Moral Hazard

Brandon Lehr, MIT *

November 16 2009

Abstract

This paper characterizes optimal social insurance in an economy where competitive firms also provide insurance to workers facing uncertain outcomes. An ex-ante heterogeneous population of workers exerts effort to increase the likelihood of high outcome events. The effort is unobservable to competitive private insurers who offer insurance contracts as in a standard moral hazard problem. The private insurance provided by firms for each worker impacts the optimal role of a benevolent social planner in providing insurance and redistribution across workers. This paper is novel in its joint consideration of two sources of heterogeneity, two potential sources of insurance, and an endogenous ex-post distribution of outcomes. The introduction of ex-ante heterogeneity in the presence of optimal private insurance changes the optimal prescription for social insurance away from zero. In many cases, standard theory overstates the level of optimal social insurance by ignoring the presence of private insurance. However, it is possible in this setting for optimal social insurance to be higher in the presence of private insurance markets. Moreover, the relative source of the variation in outcomes due to ex-ante heterogeneity and ex-post shocks plays a significant role in the welfare loss associated with setting optimal social insurance without recognizing the presence of private insurance.

*Brandon Lehr, Department of Economics, Massachusetts Institute of Technology, 50 Memorial Drive E52-391, Cambridge, MA 02142, blehr@mit.edu. Thanks to helpful comments and suggestions from Peter Diamond, Amy Finkelstein, Emmanuel Saez, Ivan Werning, and participants at the MIT Public Finance Seminar.
1 Introduction

Social insurance provides the government with the ability to insure individuals against risk and also to redistribute across agents in the economy. If we accept that in a market economy the social planner cannot directly control or exclude the provision of private insurance, it is important for the government to internalize the endogenous response of private insurance markets and individual behavior to social policy. Failure to recognize the presence of privately provided insurance can lead to incorrect prescriptions for optimal policy that may result in substantial welfare losses. This paper therefore takes up the task of characterizing analytically and numerically optimal social insurance in the presence of endogenously supplied private insurance.

This paper builds a model of the economy in which a heterogeneous population of workers differ with respect to productivity and also face an ex-post shock to earnings. Workers exert effort to increase the likelihood of a high shock. This effort is unobservable to insurers, both private and public. The important assumption, however, is that private insurers have better information than the government and in particular can observe the productivity of each agent. Thus, private insurance insures risk for a single agent, while social insurance can redistribute across agents.

In the first part of the paper we obtain analytical results to characterize the impact of private insurance markets on the provision of optimal social insurance. In order to do so, we assume that the realization of the ex-post shock is observable to the social planner. A natural interpretation is one in which the negative shock is that of entering unemployment. In that case, private insurance can take the form of a severance package from firms or private unemployment insurance as provided by a union, for example. We could also apply this setting to natural disaster insurance, disability/health insurance, or even life insurance with an appropriate interpretation of utility in the low state (death) as corresponding to surviving family members. In each of these cases, agents exert effort to prevent a low outcome event and actuarially fair private insurance helps to smooth consumption in case of an accident. The private insurers offer type-contingent insurance contracts, whereas the government only observes the realization of a low outcome event and provides the same social insurance contract to everyone.

In the second part of the paper we allow for more general shocks that do not need to be observable by the government. In particular, ex-post shocks take the form of productivity shocks and private insurance comes in the form of a compressed wage structure. This insurance is equivalent to the solution of a firm solving a standard moral hazard problem. The additional social insurance is enacted via a linear income tax schedule. This extends the standard optimal income taxation literature stemming from the “hidden skill” model of Mirrlees (1971)[16] in which wages are exogenously determined by the productivity of each agent. Instead, firms offer contracts to workers which impact realized earnings and are conditional on the tax program in the economy.

This paper characterizes the optimal social insurance in the presence of individual private insurance and numerically simulates the importance of such
considerations. The main analytical result is that the desirability of social insurance in this model depends on the covariance between income and risk. In particular, if higher earners have a lower probability of bad shocks relative to the poor, the case for social insurance is strengthened. This result is analogous to the conclusions in Rochet (1991)\cite{18} and Cremer and Pestieau (1996)\cite{5}, despite the more complex setting in this paper of allowing for explicit moral hazard concerns.

In many cases the optimal level of social insurance is reduced by taking into account the presence of private insurance. Chetty and Saez (2009)\cite{4}, Golosov and Tsyvinski (2007)\cite{12}, and Kaplow (1991)\cite{15} make this point theoretically. Empirical work by Cutler and Gruber (1996)\cite{7}\cite{8} and Gruber and Simon (2008)\cite{13} suggests that crowdout is significant, with the implications that less social insurance can be welfare improving.\footnote{An exception, Finkelstein (2004)\cite{11} shows in the health insurance context that there is little impact of the partial insurance program of Medicare on the private insurance market.} There exist parameterizations of the model in this paper, however, in which social insurance should be higher with the presence of private insurance. This is due to the impact of insurance on effort, which in turn determines the ex-post distribution of earnings. Private insurance reduces effort, increasing the likelihood of low states of nature, and increasing the need for social insurance in some settings.

Moreover, the difference between optimal social insurance with and without private insurance depends on the ex-ante variance of productivity types. When ex-ante heterogeneity is large, firms provide little in the way of insurance across the economy, so the optimal social insurance is largely unaffected by the presence or absence of private insurance. However, when ex-ante heterogeneity is small, private insurance for a single individual, who is largely representative of the average worker in the economy, obviates the need social insurance across individuals.

The analysis here extends a previous literature on optimal insurance. There exist a class of models in which there is only a single source of heterogeneity. For example, agents are ex-ante identical, but experience exogenous “luck,” which introduces uncertainty and taxation therefore serves as social insurance (e.g. Diamond, Helms, and Mirrlees (1980)\cite{9}, Eaton and Rosen (1980)\cite{10}, and Varian (1980)\cite{19}). Kaplow (1991)\cite{15} allows for both private and social insurance, where unobservable actions impact the probability of a high outcome event, but all agents are ex-ante identical. Chetty and Saez (2009)\cite{4} build a similar model with a single source of earnings heterogeneity, but they provide formulas for the welfare gains from government intervention when private insurance is not optimal. Another literature considers optimal social insurance in the presence of two sources of heterogeneity due to differences in earnings ability and exogenous ex-post shocks to income. Mirrlees (1990)\cite{17} is able to characterize the optimal linear tax in such a setting, while Rochet (1991)\cite{18} and Cremer and Pestieau (1996)\cite{5} consider jointly optimal income taxation and social insurance without moral hazard.

The paper most closely related to the analysis of this paper is Boadway et
al. (2003) in which there also exist two sources of heterogeneity and both social and private insurance in the presence of moral hazard. Their application is the health insurance market and they allow for the government to offer both a linear tax system and an insurance program, as opposed to the analysis here in which taxation is the only government tool, which acts to potentially both insure and redistribute. Moreover, private insurance does not affect taxable earnings in their analysis. This may be justifiable in the health care context, but is certainly not a reasonable assumption when considering more general insurance in the form of wage compression for agents facing uncertain productivity outcomes. The analysis in this paper also has the advantage of characterizing the welfare impact of increasing (decreasing) social insurance with expressions that are in terms of observable elasticities (in the spirit of Chetty (2009)). In addition, this paper addresses the question of quantifying welfare losses when a social planner fails to recognize the scope of the private insurance market.

The paper is structured as follows. Section 2 presents the setup of the model. Section 3 presents a simple case in which agents are ex-ante identical as a benchmark for the general analysis in section 4. Section 5 provides numerical simulations of the optimal social insurance and Section 6 concludes. Supplementary proofs and figures are relegated to the Appendix.

2 The Model

The economy consists of workers, indexed by productivity $\theta$, with a distribution given by $f(\theta)$. Each worker inelastically supplies labor hours, but a variable level of effort $a \in [0, 1]$. A worker’s effort results in either a high or low state relative to the individual’s innate productivity. In particular, the possible earnings for an individual of type $\theta$ is assumed to either be $M(\theta)$ or $m(\theta)$ where $M(\theta) > m(\theta)$ for all $\theta$. The probability of the high outcome is given by $\alpha$. Workers are risk averse and have a standard separable utility function over consumption, $c$, and effort given by

$$U(c, a) = u(c) - h(a)$$

where $u(\cdot)$ is continuously differentiable, increasing, and concave and $h(\cdot)$ is differentiable, increasing, and convex, with $h'(0) = 0$ and $h'(1) = \infty$.

Workers choose effort to maximize expected utility. Denote consumption in the two possible states for a worker of type $\theta$ by $c_1(\theta)$ and $c_0(\theta)$ for the high and low realizations, respectively. We assume throughout that $c_1(\theta) > c_0(\theta)$ so that there is never full insurance. The first order condition for the agent’s choice of effort is given by

$$u(c_1) - u(c_0) = h'(a)$$

One interpretation of this economy is that in which the low state of nature is unemployment and the high state is employment. Since the states of nature are observable to the government, while the worker’s innate productivity is not,
the government provides insurance to all workers in the form of a tax, \( t \), when employed and a benefit, \( b \), when unemployed.\(^2\) We also assume the existence of actuarially fair private insurance contracts that can supplement the public insurance program.\(^3\) The agent’s innate productivity is public information to the firm, while the effort supplied by the agent is unobservable and noncontractable. Thus, private insurance contracts are type-dependent and consist of a tax, \( t_p(\theta) \), and benefit, \( b_p(\theta) \) for a worker of type \( \theta \).\(^4\) The benefit in the low state provided by the private insurance contract can be understood as a severance payment or privately provided unemployment insurance benefit.

The information structure of this problem is one in which firms have better information than the government. Although the assumption of observable type may not be wholly realistic, it allows us to focus on the moral hazard problem of the firm over the adverse selection issue. The market failure induced by adverse selection introduces a role for the government and this analysis is concerned with understanding whether or not social insurance is beneficial even without the informational problem of adverse selection in the private market.\(^5\) Moreover, it is reasonable that firms would have better information about the types of their own workers relative to the tax authority.

In the following sections, we characterize optimal social insurance in the presence of optimal private insurance contracts under two cases.\(^6\) First, we assume a degenerate ex-ante distribution of types in the economy so that all workers are equally productive. The only heterogeneity in the population is due to the ex-post productivity realization. This case provides a useful benchmark in recognizing that with only one dimension of heterogeneity, the government cannot improve upon the private insurance contracts. Next, we turn to the general case in which there are two sources of heterogeneity and characterize the impact of such ex-ante heterogeneity on optimal insurance as determined by a utilitarian social planner.

\(^2\)Note that the government does not use earnings data to infer an individual’s type, either because the earnings functions \( M(\theta) \) and \( m(\theta) \) are not invertible, or the administrative cost of truthfully collecting such data outweighs the gain of constructing a more sophisticated social insurance program.

\(^3\)The implicit assumption is that firms are large so that the law of large numbers implies firm profits are in fact equal to zero, and not simply zero in expectation. This allows for welfare to depend simply on the utility of the consumers in the economy.

\(^4\)As in standard problems of moral hazard, the private insurer perfectly predicts the effort level of each agent, but the incentive scheme is necessary to induce the optimal effort supply.

\(^5\)Boadway et. at. (2006)[2] note that the case for social insurance can be strengthened in the presence of both adverse selection and moral hazard.

\(^6\)By “optimal”, we assume throughout the analysis a utilitarian social planner. The allocations in the private economy are Pareto efficient and the role of the social planner is to redistribute to achieve a more equitable allocation on the Pareto frontier.
3 Ex-Ante Homogeneous Agents

3.1 Optimal Social Insurance without Private Insurance

Consider a population of identical agents with productivity $\theta$. Let the government be the only provider of insurance in this economy, with a tax, $q$, and a benefit, $s$. Thus, consumption for agents is given by $c_1 = M(\theta) - q$ and $c_0 = m(\theta) + s$. The government must break even, imposing the constraint that $aq = (1 - a)s$. The government thus chooses $s$ to maximize

$$W = au \left( M(\theta) - \frac{1 - a}{a} s \right) + (1 - a)u(m(\theta) + s) - h(a)$$

(3)

Since $a$ maximizes utility, by the envelope theorem we have that the first order condition with respect to $s$ is:

$$0 = \frac{dW}{ds} = (1 - a)u'(c_0) - au'(c_1) \left( \frac{1 - a}{a} + s \frac{d(1 - a)}{ds} \right)$$

$$= (1 - a)u'(c_1) \left( \frac{u'(c_0) - u'(c_1)}{u'(c_1)} - \frac{\varepsilon_{1-a,s}}{a} \right)$$

(4)

where we define $\varepsilon_{1-a,s}$ to be the elasticity of the probability of the low consumption event with respect to the benefit, $s$. For a nonzero level of effort, an increase in the benefit reduces effort, as seen by the worker’s first order condition in (2). In order to satisfy the rule for optimal social insurance provision in (4) we must have that $\varepsilon_{1-a,s} > 0$, so it follows that $s > 0$. Equation (4) yields the following proposition.

**Proposition 1** The optimal social insurance contract with ex-ante homogeneous agents and no private insurance is characterized by

$$\frac{u'(c_0) - u'(c_1)}{u'(c_1)} = \frac{\varepsilon_{1-a,s}}{a}$$

(5)

**Proof** Immediate from (4).  

The left hand side of (5) measures the marginal value of insurance as the difference between marginal utilities across the two states. The right hand side measures the marginal cost of insurance via the behavioral distortion. At the optimum, these must be equal. This analysis provides a benchmark for understanding how the introduction of private insurance impacts the role of social insurance.

3.2 Optimal Social Insurance in the Presence of Optimal Private Insurance

With the introduction of private insurance, denote the government’s social insurance contract by $(b, t)$. Private insurers and workers take this as given. Firms
respond by setting the optimal private insurance contract, \((b_p, t_p)\). Expected firm profits must be zero, thus imposing the constraint that \(at_p = (1-a)b_p\). It is also useful to define the crowdout parameter, \(r = -\frac{db_p}{dt}\), to measure the response of the private insurers to the government insurance contract. When \(r = 0\) there is no crowdout and with \(r = 1\), there is 100% crowdout of private insurance. In this setting with ex-ante homogeneous agents, we have the following result, which is analogous to the analysis in Chetty Saez (2009)[4].

**Proposition 2** When private insurance is set optimally for a population of ex-ante homogeneous agents,

1. The optimal social insurance contract is \(b = 0\)
2. The marginal effect on welfare of an increase in the benefit \(b\) is given by

\[
\frac{dW}{db} = -\frac{1-a}{a} u'(c_1) (1-r) \frac{b}{b + b_p} \varepsilon_{1-a,b+b_p} \\
= -\frac{1-a}{a} u'(c_1) \varepsilon_{1-a,b} \\
\]

(6)

where \(\varepsilon_{1-a,b}\) is the elasticity of the low probability event with respect to \(b\), taking into account the response of \(b_p\) to \(b\).

**Proof**

**Optimal Private Contract.** Firms, taking \((b, t)\) as given choose \(b_p\) to maximize

\[
W = au \left( M(\theta) - t - \frac{1-a}{a} b_p \right) + (1-a)u(m(\theta) + b + b_p) - h(a) \\
\]

(7)

Using the envelope theorem for the agents’ optimal choice of \(a\), we have that

\[
0 = \left. \frac{dW}{db_p} \right|_{b,t} = (1-a)u'(c_0) - au'(c_1) \left( \frac{1-a}{a} + b_p \frac{d(\frac{1-a}{a})}{db_p} \right) \\
= (1-a)u'(c_1) \left( \frac{u'(c_0) - u'(c_1)}{u'(c_1)} \frac{\varepsilon_{1-a,b+b_p}}{a} \right) \\
\]

(8)

We can define \(s = b + b_p\) to be the total insurance provided to individuals and note that \(\varepsilon_{1-a,b+b_p} = \frac{d(1-a)}{db_p + db} \frac{1-a}{1-a} = \frac{d(1-a)}{db_p + db} \frac{1-a}{b + b_p} = \varepsilon_{1-a,s} \frac{b_p}{b+b_p}\). The optimal private insurance contract, conditional on \((b, t)\), is therefore characterized by

\[
\frac{u'(c_0) - u'(c_1)}{u'(c_1)} = \varepsilon_{1-a,s} \frac{b_p}{a} \frac{b}{b + b_p} \\
\]

(9)

Note that in the case in which the government is not providing any insurance \((b = 0)\), the formula for optimal private insurance reduces to the same rule in (5) for optimal social insurance without private insurance.

**Social Insurance in the Presence of Optimized Private Insurance.** The government chooses benefit \(b\) and tax \(t = \frac{1-a}{a} b\), to maximize social welfare, taking into
account $b_p$ set as in (9). To simplify the problem, note that $c_1 = M(\theta) - \frac{1-a}{a} s$ and $c_0 = m(\theta) + s$ where again $s = b + b_p(b)$. With a change of variables, it is therefore equivalent for the government to choose $s$ instead of $b$ directly. In particular, from (7)

$$W = au \left( M(\theta) - \frac{1-a}{a} s \right) + (1-a)u(m(\theta) + s) - h(\alpha)$$

and

$$\frac{dW}{db} = \frac{dW}{ds} \frac{ds}{db} = (1-r) \frac{dW}{ds}$$

$$= (1-r)(1-a)u'(c_1) \left( \frac{u'(c_0) - u'(c_1)}{u'(c_1)} - \frac{\varepsilon_{1-a,s}}{a} \right)$$

where the expression for $\frac{dW}{ds}$ follows from (4). Plugging in the expression from (9), we have that

$$\frac{dW}{db} = (1-r)(1-a)u'(c_1) \left( \frac{\varepsilon_{1-a,s}}{a} \frac{b_p}{b + b_p} - \frac{\varepsilon_{1-a,s}}{a} \right)$$

$$= - (1-r) \frac{1-a}{a} u'(c_1) \varepsilon_{1-a,s} \frac{b}{b + b_p}$$

This expression establishes that $\frac{dW}{db}(b = 0) = 0$. To show that $b = 0$ is the global maximum, it is sufficient to establish that $r < 1$. If an increase in $b$ were exactly offset by a decrease in $b_p$, maintaining a constant level of total insurance, effort would be unchanged and we would violate (9). Thus, $b_p$ must fall by less than the increase in $b$ to satisfy (9). Moreover, we can simplify the above expression by noting that $\varepsilon_{1-a,s} = \frac{d(1-a)}{ds} \frac{s}{b + b_p} = \frac{\varepsilon_{1-a,s}}{b + b_p} \frac{1}{1-r}$. Therefore, we have

$$\frac{dW}{db} = - \frac{1-a}{a} u'(c_1) \varepsilon_{1-a,s}$$

as claimed in (6). ■

Proposition 2 establishes that when individuals are ex-ante homogeneous, there is no role for social insurance in the presence of optimally provided private insurance. To better understand the intuition for this result, first note the optimal private insurance as characterized in expression (9). We see that in the case in which there is no government intervention ($b = 0$), (9) reduces to (5) and firms provide the same level of optimal social insurance in the absence of private insurance contracts. The effect on welfare of increasing $b$ at zero equals the deadweight burden of greater taxation, implying that the government should do nothing. Hence, the optimal level of insurance can be equivalently provided exclusively by the private sector or the social planner.

The impact of social policy on welfare, $\frac{dW}{db}$, is expressed in terms of two different elasticities in (6). The first is an expression in terms of the behavioral
elasticity with respect to total insurance, \( s = b + b_p \). This elasticity, \( \varepsilon_{1-a,s} \), however, is less empirically useful since we do not observe the total change in insurance and thus need an estimate of the crowdout parameter, \( r \). The second expression employs the fact that \( \varepsilon_{1-a,b} = (1 - r) \frac{b}{b + b_p} \varepsilon_{1-a,s} \). This expression relates the observable elasticity of behavior with respect to policy to the more fundamental elasticity of behavior with respect to total insurance, which may not be observable.\(^7\) The intuition for this identity is that an increase in \( b \) leads to a smaller change in \( s \) via two channels. The first is that an increase in \( b \) leads to crowdout of \( b_p \), so we must scale the fundamental elasticity by \((1 - r)\). In addition, there is a mechanical effect of adjusting the impact of \( b \) on \( s \) by the relative magnitude of \( b \) to \( s \). Hence, when \( b \) is small relative to the private insurance, \( b_p \), the proportional effect of \( b \) on \( s \) is diminished. This second channel mechanically implies by the definition of an elasticity that \( \varepsilon_{1-a,b}(b = 0) = 0 \), verifying the intuition for \( \frac{dW}{db}(b = 0) = 0 \) as explained above.

When the government provides positive social insurance in the presence of optimal private insurance, there is a negative marginal effect on welfare. The intuition for this result is as follows. Given some \( b > 0 \), suppose firms topped up the insurance by setting \( b_p \) such that \( b + b_p \) were equal to the optimal level of social insurance in the absence of private contracts. Such a private insurance contract will satisfy (5) by definition, but will then not satisfy (9) as required for optimality. In fact, private insurers perceive a lower marginal cost of providing insurance given \( b \), and will increase \( b_p \) so as to satisfy (9). Thus, total insurance is greater than would be optimal with a single social insurance contract. This additional insurance reduces effort, which in turn requires greater taxes from the government to balance its budget. The failure of private insurers to internalize their impact on the government’s budget constraint leads to a strict welfare loss.

An analogous argument explains why \( \frac{dW}{db}(b < 0) > 0 \). Suppose the government were to exacerbate the difference in earnings levels in the two states of nature by setting \( b < 0 \). It would not be optimal for private insurers to compensate for the lack of social insurance by setting a total level of insurance equal to the optimal insurance when provided by a single source. In order to satisfy (9), private insurers would have to decrease \( b_p \). Thus, when \( b < 0 \), total insurance is less than when there is a single insurer. This decreased insurance increases effort, allowing the government to lower taxes for the same level of \( b \). Since private insurers do not internalize this impact of insurance and its induced change in effort on the government budget constraint, there is a welfare improvement by increasing \( b \). Note that an increase in \( b \) crowds out private insurance as before so that the total level of insurance increases, but by less than the change in \( b \). This increase in \( s \) reduces effort. However, somewhat counterintuitively, \( \varepsilon_{1-a,b}(b < 0) < 0 \). This is simply because when working with elasticities as opposed to derivatives, an increase in \( b \) when \( b \) is negative corresponds to a percentage decrease in \( b \).

Chetty and Saez (2009)[4] enrich a similar environment by analyzing the

---

\(^7\)I adopt the terminology of the elasticity with respect to \( s \) as the “fundamental elasticity” from the analysis in Chetty and Saez (2009)[4].
optimal social insurance contract in the presence of not necessarily optimized private insurance. Instead, this paper is focused on understanding how optimal social insurance is affected by the introduction of two sources of heterogeneity. In fact, with ex-ante heterogeneity, we can break the result of Proposition 2. We turn to this analysis in the following section.

4 Ex-Ante Heterogeneous Agents

4.1 Optimal Social Insurance without Private Insurance

We now allow for a general distribution of types in the population. With the government as the only provider of insurance, the insurance acts to insure individual workers and redistribute across types. The government must again break even, so a contract consisting of a benefit \( s \) and tax \( q \) must satisfy

\[
q = s(1 - \frac{s}{\bar{a}})
\]

where \( \bar{a} = \int a(\theta) f(\theta) d\theta \). We also define \( \varepsilon_{1-\pi,s} \) as the elasticity of the average probability of low events occurring across agents with respect to the benefit level, \( s \). We are now ready to state the analogous result to Proposition 1 as follows.

**Proposition 3** The marginal effect of an increase of a benefit \( s \) on welfare is given by:

\[
\frac{dW}{ds} = E((1 - a)u'(c_0)) - E(au'(c_1)) \frac{1 - \frac{s}{\bar{a}}}{1 + \frac{\varepsilon_{1-\pi,s}}{\bar{a}}}
\]

and the optimal social insurance contract with no private insurance is characterized by

\[
\frac{\bar{a}}{1 - \bar{a}} \frac{E((1 - a)u'(c_0))}{E(au'(c_1))} - 1 = \frac{\varepsilon_{1-\pi,s}}{\bar{a}}
\]

**Proof** The government chooses \( s \) to maximize

\[
W = \int \left( a(\theta)u \left( M(\theta) - \frac{1 - \frac{s}{\bar{a}}}{\bar{a}}s \right) + (1 - a(\theta))u(m(\theta) + s) - h(a(\theta)) \right) f(\theta) d\theta
\]

Since \( a \) maximizes utility, by the envelope theorem we have that the first order condition with respect to \( s \) is:

\[
\frac{dW}{ds} = E((1 - a)u'(c_0)) - E(au'(c_1)) \left( \frac{1 - \frac{s}{\bar{a}}}{\bar{a}} + s \frac{d \left( \frac{1 - \frac{s}{\bar{a}}}{\bar{a}} \right)}{ds} \right)
\]

\[
= E((1 - a)u'(c_0)) - E(au'(c_1)) \frac{1 - \frac{s}{\bar{a}}}{\bar{a}} \left( 1 + \frac{\varepsilon_{1-\pi,s}}{\bar{a}} \right)
\]

Setting \( \frac{dW}{ds} = 0 \), we can easily solve for (11). □
The addition of ex-ante heterogeneity does not change the fundamental notion that optimal social insurance should equate the expected marginal benefit, as measured by the relative difference between the marginal utilities across states, and the marginal cost, as measured by the average behavioral response. Without heterogeneity, \( \bar{\pi} = a \), the expectation operators drop out, and (11) reduces to (5). A notable difference is that the expectations of the marginal utilities are weighted by the relative probability of the high or low state occurring for a given type of individual. This is because each type of worker chooses a different effort level and thus realizes different states of nature with a different probability distribution. When those with the highest consumption experience the least risk (high effort), the left hand side of (11), which measures the marginal benefit of social insurance, is greater. Thus, the case for social insurance is greater when risk and income are negatively correlated. This confirms the same intuition from other models in which there exists heterogeneity with respect to ability and risk (Rochet (1991)[18] and Cremer and Pestieau (1996)[5]).

Such endogeneity of the distribution is in contrast to the analysis of optimal taxation as social insurance in Mirrlees (1990)[17]. Mirrlees allows for workers to be ex-ante different, but then experience a productivity shock that is independent of ability and labor choice. With a fixed distribution of productivity realizations, he is able to provide a simple approximation to the optimal tax rate as a linear combination of the ex-ante and ex-post variances. The simplicity of the characterization in terms of only the variances of the different sources of heterogeneity relies, however, on the assumption that both variances and tax rates are small. This paper does not impose such restrictions.

We now turn to the fully general case in which two sources of insurance work to smooth consumption within and across individuals.

4.2 Optimal Social Insurance in the Presence of Optimal Private Insurance

The agents in this economy can be individually insured by the firm at which they are employed. The government can augment this insurance with its own contract, which has the added benefit of being able to redistribute across agents. Thus, firms offer contracts \((b_p(\theta), t_p(\theta))\) to each worker in which \(a(\theta)t_p(\theta) = (1 - a(\theta))b_p(\theta)\). Firms and agents take the government contract, with a benefit of \(b\) and tax of \(t\) as given. We also must allow for differential crowdout effects for each type of worker, so define the crowdout of the private insurance contract for type \(\theta\) as \(r(\theta) = -\frac{db_p(\theta)}{db}\). The optimal utilitarian social insurance contract can be characterized as follows.

Proposition 4 When private insurance is set optimally, the marginal effect of an increase in the benefit \(b\) is given by

\[
\frac{dW}{db} = -\frac{1 - \bar{\pi}}{\pi} E(au'(c_1))_{\pi} E_{\pi,b} - \frac{\text{cov}(a, u'(c_1))}{\pi} (12)
\]
**Proof**

*Optimal Private Contract.* Firms, taking $b$ and $t = 1 - a(\theta)$ as given choose $b_p(\theta)$ to maximize

$$W(\theta) = a(\theta)u\left(M(\theta) - t - \frac{1 - a(\theta)}{a(\theta)}b_p(\theta)\right)$$

$$+ (1 - a(\theta))u(m(\theta) + b + b_p(\theta)) - h(a(\theta))$$  \hspace{1cm} (13)

Using the envelope theorem for the agents’ optimal choice of $a$, we have that

$$\frac{dW}{db_p(\theta)}|_{b,t} = (1 - a(\theta))u'(c_1(\theta)) \left(\frac{u'(c_0(\theta)) - u'(c_1(\theta))}{u'(c_1(\theta))} - \frac{\varepsilon_{1-a(\theta),b_p(\theta)|b,t}}{a(\theta)}\right)$$  \hspace{1cm} (14)

This expression is the analog of (8). Each firm separately provides an optimal insurance contract for their worker subject to earning zero profits. The optimal private insurance contract, conditional on $b$, is therefore characterized by

$$\frac{u'(c_0(\theta)) - u'(c_1(\theta))}{u'(c_1(\theta))} = \frac{\varepsilon_{1-a(\theta),b_p(\theta)|b,t}}{a(\theta)}$$  \hspace{1cm} (15)

*Social Insurance in the Presence of Optimized Private Insurance.* The government chooses $b$ to maximize social welfare, taking into account $b_p(\theta)$ set as in (15). In particular, welfare is given by

$$W = \int \left(a(\theta)u\left(M(\theta) - \frac{1 - \alpha}{\alpha}b - \frac{1 - a(\theta)}{a(\theta)}b_p(\theta)\right)\right.$$

$$+ (1 - a(\theta))u(m(\theta) + b + b_p(\theta)) - h(a(\theta))) \, f(\theta) \, d\theta$$  \hspace{1cm} (16)

Employing the envelope theorem for the agents’ utility maximizing choice of $a$, we have that

$$\frac{dW}{db} = \int ((1 - a(\theta))u'(c_0(\theta))(1 - r(\theta))$$

$$- a(\theta)u'(c_1(\theta)) \frac{d}{db} \left(\frac{1 - \alpha}{\alpha}b + \frac{1 - a(\theta)}{a(\theta)}b_p(\theta)\right)) \, f(\theta) \, d\theta$$  \hspace{1cm} (17)

Moreover,

$$\frac{d}{db} \left(\frac{1 - \alpha}{\alpha}b + \frac{1 - a(\theta)}{a(\theta)}b_p(\theta)\right)$$

$$= \frac{1 - \alpha}{\alpha} + b \frac{d}{db} \left(\frac{1 - \alpha}{\alpha}\right) - r(\theta) \frac{1 - a(\theta)}{a(\theta)} + b_p(\theta) \frac{d}{db} \left(\frac{1 - a(\theta)}{a(\theta)}\right)$$

$$= \frac{1 - \alpha}{\alpha} + \frac{1 - \alpha}{\alpha^2} \varepsilon_{1-\alpha,b} - r(\theta) \frac{1 - a(\theta)}{a(\theta)} + \frac{b_p(\theta) \frac{1 - a(\theta)}{a(\theta)}}{b} \frac{1 - a(\theta)}{a(\theta)^2} \varepsilon_{1-a(\theta),b}$$  \hspace{1cm} (18)
Plugging (18) into (17), we have that
\[ \frac{dW}{db} = E((1-a)u'(c_0)(1-r)) - E(au'(c_1)) \left( \frac{1-\frac{\pi}{a}}{a} \right) \left( 1 + \frac{\varepsilon_{1-\pi,b}}{a} \right) \\
+ E((1-a)u'(c_1)r) - E \left( \frac{1-a}{a}u'(c_1) \frac{b_p}{b} \varepsilon_{1-a,b} \right) \] (19)

Adding and subtracting the term \( E((1-a)u'(c_1)) \) to (19) we have that
\[ \frac{dW}{db} = E((1-a)u'(c_0)(1-r)) - E((1-a)u'(c_1)(1-r)) + E((1-a)u'(c_1)) \\
- E(au'(c_1)) \left( \frac{1-\frac{\pi}{a}}{a} \right) \left( 1 + \frac{\varepsilon_{1-\pi,b}}{a} \right) - E \left( \frac{1-a}{a}u'(c_1) \frac{b_p}{b} \varepsilon_{1-a,b} \right) \] (20)

To impose the constraint that firms are offering optimal insurance conditional on \( b \), we now utilize (15). Multiply both sides of (15) by \((1-a(\theta))(1-r(\theta))u'(c_1(\theta))\) and take the expectation over productivity types to obtain
\[ E((1-a)(1-r)(u'(c_0) - u'(c_1))) = E \left( \frac{1-a}{a}u'(c_1)(1-r)\varepsilon_{1-a,b}|b,t \right) \] (21)

Finally, we note that \( \varepsilon_{1-a(\theta),b} = \varepsilon_{1-a(\theta),b_0|b,t}(1-r(\theta)) \frac{b_p}{b_0r(\theta)} \). Plug this expression into (21) and note that (21) can be substituted in for the first two terms of (20), which then cancels with the last term of (20). This yields
\[ \frac{dW}{db} = E((1-a)u'(c_1)) - E(au'(c_1)) \left( \frac{1-\frac{\pi}{a}}{a} \right) \left( 1 + \frac{\varepsilon_{1-\pi,b}}{a} \right) \] (22)

We can further transform this expression by converting to one of covariances in the following way:
\[ \frac{dW}{db} = E(u'(c_1)) - E(au'(c_1)) \left( \frac{1-\frac{\pi}{a}}{a} \left( 1 + \frac{\varepsilon_{1-\pi,b}}{a} \right) \right) \\
= -\frac{1-\frac{\pi}{a}}{a} E(au'(c_1)) \varepsilon_{1-\pi,b} + E(u'(c_1)) - \frac{E(au'(c_1))}{a} \\
= -\frac{1-\frac{\pi}{a}}{a} E(au'(c_1)) \varepsilon_{1-\pi,b} - \frac{cov(a,u'(c_1))}{a} \]

This is the expression in (12) which we desired. 

It is clear that Proposition 4 provides a generalization of the result in Proposition 2. In particular, with ex-ante homogeneity of types, (12) reduces to (6) since \( a = \bar{a} \), the expectation operator drops out, and the covariance between effort and marginal utility in the high state is zero. It is no longer necessarily the case, however, that it is optimal for the government to set \( b = 0 \) with ex-ante heterogeneous agents.

The first term in (12) is of the same form as in (6). Providing a positive level of social insurance contract reduces effort and crowds out private insurance. With homogeneity, firms provide all of the necessary insurance, and this
negative effect of a marginal increase in \( b > 0 \) is the only effect that matters. By introducing heterogeneity, we add the second term in (12). We cannot sign the covariance without more structure on the nature of the shocks faced by workers. When the covariance between effort and marginal utility is positive, the marginal impact of \( b \) on welfare becomes more negative. A positive covariance is consistent with agents who consume more also exerting less effort. Thus, the ex-post distribution of production across the economy is compressed, mitigating the need for government redistribution. If we restrict \( b \geq 0 \), we are then at a corner solution in which the government should not provide any insurance.

If, however, the covariance between effort and marginal utility in the high state is negative, there is scope for government redistribution. It is reasonable that under some specifications, higher types will have larger consumption in the high state and also exert more effort. A negative covariance implies a positive term added to the first negative term in (12). Intuitively, if risks are negatively correlated with income so that the high earners have low probability of bad outcomes, there is a greater motive for redistribution through higher taxation. Although the setting is different than in previous analyses, we find that the introduction of moral hazard and an efficient private insurance market does not change the robust result that the covariance between risk and ability is a primary factor in determining the desirability of redistribution. Greater risk aversion also increases the magnitude of this covariance, whereas risk neutrality implies a zero covariance and there is hence no role for social insurance.

The expression in (12) can also be estimated by the econometrician to determine the marginal impact of more social insurance on welfare. In the unemployment context, although effort is not observable, the realized fraction of the population that is unemployed, \( 1-\bar{\pi} \), is observable. In addition to estimating the elasticity of unemployment with respect to social insurance, the econometrician must estimate the covariance between effort and marginal utility of consumption. This could be estimated by considering differential unemployment rates among different classes of earners. This makes the analysis of practical and empirical relevance.

It is also instructive to compare the optimal social insurance with and without private insurance. Define \( s^* \) to be the solution to \( \frac{dW}{ds} = 0 \) in (10) and \( b^* \) to be the solution to \( \frac{dW}{db} = 0 \) in (22). They are reprinted for clarity:

\[
\frac{dW}{ds} = E((1-a)u'(c_0)) - E(au'(c_1)) \frac{1-\bar{\pi}}{\bar{\pi}} \left( 1 + \frac{\bar{\pi}_{1-s}}{\bar{\pi}} \right) \tag{10}
\]

\[
\frac{dW}{db} = E((1-a)u'(c_1)) - E(au'(c_1)) \frac{1-\bar{\pi}}{\bar{\pi}} \left( 1 + \frac{\bar{\pi}_{1-b}}{\bar{\pi}} \right) \tag{22}
\]

Despite the similarity in form of the two expressions, we cannot directly compare them since the \( c_0, c_1, \) and \( a \) are different in the two cases. Although we expect that \( \bar{\pi}_{1-s} > \bar{\pi}_{1-b} \) due to crowding out, we also expect that average effort is lower with private insurance. This makes comparing the second terms in the above equations impossible without additional structure on the problem. If \( \frac{dW}{ds} > \frac{dW}{db} \) due to a potentially larger first term (if for example, \( c_1 = M(\theta) - t - t_p > m(\theta) + s = c_0 \)) and larger differences in average effort than in the elasticities,
we would obtain $s^* > b^*$. This result suggests that without firms providing insurance to individual workers, there exists a greater role for the government to provide insurance both within and across workers and the optimal benefit rate is greater than in the case in which firms are endogenously providing some insurance to individuals. If, however, $\bar{a}$ is much more responsive to $s$ relative to $b$, due to a large crowdout effect or large ex-ante variance, we may have that $s^* < b^*$. We turn to numerical simulations to better understand how the prescription for optimal policy depends on the presence of endogenously provided private insurance to individual agents.

4.3 Partial Private Insurance

It is reasonable to expect that not all individuals in an economy are privately insured. In this section we explore the impact of social insurance in a setting in which workers are exogenously determined to be either optimally insured by a private firm or receive no private insurance. To simplify the exposition, we assume in particular, that a share, $\alpha$ of each type of worker is privately insured. Given government policy, the problem for private insurers is unaffected in this new setting. The social planner’s problem is now essentially a convex combination of the objective functions in Propositions 3 and 4. The only difference is that the government’s budget constraint is affected by the fact that privately insured workers will exert less effort than their non-insured counterparts. We introduce the notation of superscripts of $I$ and $N$ on the variables $a$, $c_1$, and $c_0$ to indicate the differential efforts and consumptions of insured and non-insured individuals, respectively. It is also useful to define $\overline{a^I} = \int a^I(\theta) f(\theta) d\theta$ and $\overline{a^N} = \int a^N(\theta) f(\theta) d\theta$ as the average effort among the privately insured and non-insured, respectively. Finally, we denote the elasticities of the average probability of the low event with respect to the government benefit, $b$, for each class of workers in the standard fashion: $\varepsilon_{1-a^I,b}$ and $\varepsilon_{1-a^N,b}$.

**Proposition 5** With $\alpha$ share of the population privately insured, the marginal impact of a benefit $b$ on welfare is given by

$$
\frac{dW}{db} = \alpha E((1 - a^I)u'(c^I_1)) + (1 - \alpha)E((1 - a^N)u'(c^N_0)) \nonumber
$$

$$
- \frac{\alpha E(a^I u'(c^I_1)) + (1 - \alpha)E(a^N u'(c^N_1))}{\alpha \overline{a^I} + (1 - \alpha)\overline{a^N}} \nonumber
$$

$$
\times \left[ \alpha (1 - \overline{a^I}) + (1 - \alpha)(1 - \overline{a^N}) + \frac{\alpha (1 - \overline{a^I}) \varepsilon_{1-a^I,b} + (1 - \alpha)(1 - \overline{a^N}) \varepsilon_{1-a^N,b}}{\alpha \overline{a^I} + (1 - \alpha)\overline{a^N}} \right] \quad (23)
$$

**Proof** See Appendix
The expression in (23) reduces to (10) and (22) when $\alpha$ is zero and one, respectively. This generalization characterizes the effect of $\alpha$ on the marginal benefit to society of increasing social insurance benefits. As discussed in the previous subsection, it is not possible to order $\frac{dW}{db}$ for $\alpha \in \{0, 1\}$. Although we cannot sign the effect of $\alpha$ on $\frac{dW}{db}$, the root of the expression in (23) maps out the change in optimal policy as $\alpha$ changes continuously. We will explore this issue numerically in the following section.

Despite the somewhat cumbersome expression, it follows the standard intuition for optimal taxation. The marginal benefit is captured by the gap in marginal utilities in the low and high states. In particular, the first term in (23) is the weighted average of the the expected marginal utility in the high state for the privately insured and the marginal utility in the low state for the uninsured weighted respectively by the probability of the low state. The marginal utility is taken at the high state as opposed to the low state for the insured due to the presence of optimal private insurance. The first part of the second term is the average expected marginal utilities in the high states for both classes of agents weighted by the probability of the high state. The gap between these two terms is a measure of the benefit to society of greater insurance. We must also adjust this second term, however, by the last term in square brackets in (23) to capture the behavioral effect of greater social insurance and its subsequent impact on the government’s budget constraint. The form of the expression follows from the balanced budget constraint, which requires a tax of

$$t = \frac{\alpha(1 - a^I) + (1 - \alpha)(1 - a^N)}{\alpha a^I + (1 - \alpha)a^N}b$$

A richer extension would allow for the distribution of types who are privately insured to be endogenously determined. In the unemployment insurance context we could think of some occupations in which employment is at will, while others offer explicit severance packages or private unemployment insurance as negotiated by a union. Optimal policy may be different if individuals can search and self-select occupations and firms in turn can determine the optimal level of insurance to attract different types of workers.\footnote{Guerrieri, Shimer, and Wright (2009)\cite{footnote1} characterize equilibrium properties in a search model with ex-ante heterogeneity.}

## 5 Numerical Simulations

The preceding analysis provided a sharp characterization of the impact of government insurance in the presence of privately provided insurance. The tractability of the problem was due in part to the assumption that social insurance took an additive form. Thus, all agents received the same level of taxes and benefits, independent of their wages. In reality, however, we may imagine a social insurance system that relies on setting marginal tax rates. In particular, consider a social planner setting a single marginal income tax rate and rebating
the proceeds equally to all agents. This has the advantage of not requiring the
government to observe whether an agent is in the high or low state of nature.
Hence, although the unemployment context was an appropriate interpretation
of the additive social insurance setting, with wage taxes we can think of agents
experiencing high and low productivity shocks that do not lead to unemploy-
ment and firms offering insurance via a compressed wage structure. Such a
system, however, imposes the complication of workers responding to tax policy
via both income and substitution effects.

Consider the same environment as in Section 4 with no private insurance. A
single social insurance system consists of a wage tax, \( s \), and a rebate equal to
\( z \) where \( z \) is total output. Moreover,

\[
\frac{dW}{ds} = z \left( 1 - \varepsilon_{z,1-s} \right) \left( E(au'(c_1) + (1 - a)u'(c_0)) \right) - E(au'(c_1)M + (1 - a)u'(c_0)m) \quad (24)
\]

and the optimal social insurance contract with no private insurance is charac-
terized by

\[
\frac{s}{1 - s} = \frac{1}{\varepsilon_{z,1-s}} \left[ 1 - \frac{1}{z} E(au'(c_1)M + (1 - a)u'(c_0)m) \right] \quad (25)
\]

**Proof** See Appendix

Proposition 6 provides a characterization of the optimal marginal tax rate.
The moral hazard problem coupled with income and substitution effects makes
this expression which depends on the interaction of risk, heterogeneity, income,
and curvature of the utility function difficult to interpret relative to the expres-
sion for the analogous problem with lump sum redistribution in (11). In fact,
we will see in the following simulations how the optimal tax rate can depend in
complex ways on the various factors.

We can similarly consider an environment in which firms offer private insur-
ance in the same manner as before with type-contingent contracts \( (b_p(\theta), t_p(\theta)) \)
where \( t_p(\theta) = \frac{1 - a(\theta)}{a(\theta)} b_p(\theta) \). The government imposes a wage tax of \( b \) on wages
after the imposition of private insurance. The rebate is then equal to \( bw \) where
\( w = \int (a(\theta)(M(\theta) - t_p(\theta)) + (1 - a(\theta))(m(\theta) + b_p(\theta))) f(\theta) d\theta \) is aggre-
gate earnings. The Appendix provides a setup of the social planner problem
in this setting and an expression for \( \frac{dW}{db} \), but it is uninformative for providing
intuition as to how optimal social insurance is affected by the presence of
private insurance. Numerical simulations, however, are instructive in this more
complex environment.
5.1 The Impact of Private Insurance on the Level of Social Insurance

The previous theoretical and empirical literature has noted that in the presence of optimal private insurance, the optimal level of social insurance is lower (see Chetty and Saez (2009)[4], Cutler and Gruber (1996)[7], Golosov and Tsyvinski (2007)[12], and Kaplow (1991)[15]). The intuition for such a result is clear; social insurance provides insurance and redistribution across agents and with private insurance markets, part of the government’s objective is achieved privately. The previous theoretical analyses have not allowed, however, for endogenous distributions of ex-post outcomes, as in this paper. This allows for parameterizations of the model in which optimal social insurance is higher in the presence of optimal private insurance. The intuition for such a result is that private insurance compresses the wage distribution for each worker facing uncertain productivity shocks. This compression is welfare improving by smoothing consumption, but reduces the effort supplied by each agent. Lower effort choices increase the likelihood of low states of nature being realized. This exacerbates the need for redistribution to the lowest types in the economy, making the optimal social insurance level higher than if no private insurance existed.

To verify the above intuition, consider the following parameterization of the economy. Let utility, \( u(\cdot) \) be given by CRRA preferences with coefficient of relative risk aversion of 3. Let the marginal disutility of effort be given by \( h'(a) = -\eta \ln(1 - a) \) for \( \eta > 0 \). This somewhat nonstandard utility representation is not significant for the results, but simply ensures an interior solution to the consumer’s effort choice problem. Moreover, we will follow Mirrlees (1990)[17] in defining the shock process to be given by \( M(\theta) = M\theta \) and \( m(\theta) = m\theta \).

Let there be five uniformly distributed ex-ante types in the economy with \( E(\theta) = 1 \). Finally, we let \( \eta = 1, M = 1.25, \) and \( m = 0.75 \), but the result is robust to many alternative specifications.

Figure 1 plots the optimal tax rate in the presence and absence of optimal private insurance as a function of the variation in ex-ante heterogeneity (the distance between the highest and lowest skilled types, while fixing \( E(\theta) = 1 \).) For small variation in the ex-ante heterogeneity, we confirm the predictions of previous analyses with an optimal level of social insurance lowered by the presence of private insurance. The optimal tax rate is increasing in the heterogeneity in the population, as the redistributive motive is greater with greater variation. As the ex-ante variation increases relative to the ex-post shocks, redistribution becomes more important relative to insurance. Since private insurance cannot redistribute between workers, the difference between optimal taxes with and without private insurance diminishes. The novel feature of this setting is that in fact, the two plots cross. It is at this point that the presence of private insurance hampers the government’s redistributational goals. For large enough ex-ante heterogeneity, a social planner redistributing income with a linear wage tax simultaneously provides sufficient insurance for the uncertain

---

\[9\] Various other specifications of the shocks yield analogous results with only slight modifications to the levels.
outcomes. Additional private insurance further depresses effort and this greater probability of low outcomes reduces welfare more than the marginal increase in insurance. Thus, the social planner finds it optimal to redistribute more to the lowest earners with a higher marginal tax rate.

From a policy perspective, this result suggests that for economies with large variation in ability, generous insurance can exacerbate inequality, requiring higher taxation. Insurance programs for disability, for example, may induce low ability workers to choose low effort so as to receive disability benefits. Greater welfare could be potentially achieved by reducing the generosity of such benefits and lowering income tax rates.

When determining optimal policy it is therefore important to not only consider the availability of private insurance for individuals, but also the ex-ante heterogeneity of types in the economy. The welfare loss from failing to account for the existence of private insurance can be dramatic with more homogenous populations, but becomes small for more heterogeneous populations. In the next section we further investigate the role of heterogeneity on optimal policy.

5.2 Optimal Social Insurance without Private Insurance

To better understand the expression in (25) for the optimal marginal tax rate without private insurance, we continue with a slightly different numerical exer-
cise. In the previous subsection we saw that optimal taxation without private insurance is increasing in the variance of the ex-ante heterogeneity. Such monotonicity, however, does not obtain for an increase in the variance of ex-post shocks. To see this, continue with the same parameterization as in the previous subsection. Consider two different levels of ex-ante heterogeneity. In one case let the variance be large, with a distance between the highest and lowest productivity types, $\bar{\theta} - \underline{\theta} = 1$ and in another case $\bar{\theta} - \underline{\theta} = 0.2$. Again, we fix a uniform distribution of types with $E(\theta) = 1$. Figure 2 plots the optimal marginal tax rate in these two cases without private insurance as we vary the ex-post variation in shocks. In particular, we vary $m$ and $M$ such that $\frac{m+M}{2} = 1$ so that shocks are symmetric around an agent’s type.\(^{10}\) Figure 2 displays the simulation.

Figure 2: Optimal Taxation without Private Insurance

Note that we do not have monotonicity in the ex-post variance. This is in contrast to Mirrlees (1990)\(^{[17]}\) where it is shown that with two sources of heterogeneity and no private insurance, the optimal linear tax rate is increasing in both ex-ante and ex-post variation. That result does not hold in this model since the labor choice impacts the distribution of realized outcomes. To understand the intuition for such non-monotonicity, fix the ex-ante heterogeneity while introducing risk. When there is no ex-post risk and agents receive the

\(^{10}\)This assumption does not drive the results. For example, fixing $M = 1$ and varying $m$ from 0 to 1 yields similar results.
same consumption regardless of effort, effort is optimally zero. In such a case, there is no labor distortion to taxation and the government sets a confiscatory tax rate in which incomes are equated across all individuals in the economy. Greater variation in outcomes for each type of agent increases effort, and hence a distortionary cost of taxation to the social planner. The optimal tax rate must then initially decrease. As efforts increase, the rate at which different types of agents increase effort for a given increase in risk varies. When ex-ante heterogeneity is not very large, the lowest type productivity workers exert greater effort initially relative to their richer counterparts since they need to insure themselves more against the worst outcomes. This leads to a decrease in the average output, $\bar{z}$. The lower risk for the low types due to self-insurance reduces the motive for redistribution. As the shocks increase in magnitude, higher types start increasing effort relatively more than low types who are already exerting high effort. This leads to an increase in average output in the economy. With higher output and lower risk for the highest earners due to their greater effort, the redistributive motive increases and leads to higher optimal social insurance when this outweighs the cost of the behavioral response on the government budget constraint. In addition, when the ex-ante variance is large, it requires larger negative shocks for the highest types to increase effort sufficiently high so as to induce the social planner to increase redistribution.

The preceding analysis is depicted graphically in Figures 6 and 7 in the Appendix and follows the same intuition we have already seen in our analytical results. Namely, when risk and income are sufficiently negatively correlated, the motive for redistribution is strengthened. Although the importance of this correlation has been established in the literature, this numerical exercise permits a better understanding of how it translates into the level of optimal taxation for various parameterizations. Note also that as the magnitude of the shocks increases, the difference between optimal social insurance at different levels of ex-ante heterogeneity decreases since the ex-post variation dominates any initial ex-ante differences.

5.3 The Impact of Ex-Ante Heterogeneity on Optimal Policy in the Presence of Private Insurance

It is also instructive to simulate how the optimal tax policy is affected by the presence of optimal private insurance for different levels of ex-ante heterogeneity. We continue with the same parameterization as in the previous subsection. Figure 3 plots the optimal marginal tax rate in the two cases of large ($\theta - \bar{\theta} = 1$) and small ($\theta - \bar{\theta} = 0.2$) ex-ante heterogeneity with and without private insurance as we vary the ex-post variation in shocks. Note that the plots of optimal

\[\text{As evidenced by Figure 7 in the case of large ex-ante variance, high types may initially increase effort more than low types since the same shock is proportionally larger for the high types, inducing greater effort. This initially drives down the covariance between risk and income, but the covariance increases as low types exert more effort before decreasing dramatically as the high types once again exert more effort than low types for increasing ex-post shock variation.}\]
taxation without private insurance are the same as in Figure 2. The plot also reinforces the already noted observation from Figure 1 that an increase in ex-ante heterogeneity leads to an increase in the optimal tax rate at all levels of ex-post shocks.

![Figure 3: Optimal Taxation](image)

With optimal private insurance, the optimal tax rate is in general lower than without private insurance, but not always, as discussed previously. As the ex-post variance increases, the optimal tax rate monotonically declines in both cases. Although greater variation may increase the value of redistributive taxation, the source of that variation is important. For a fixed population of workers, as ex-post shocks increase in magnitude, insurance is the primary welfare improving tool relative to redistribution. Since firms are able to insure agents against such uncertainty, the government’s role as a redistributing agent is diminished and optimal taxation declines. To better understand the underlying mechanism for this result we can also examine how effort changes as the magnitude of the ex-post shocks increases. Although effort is increasing, the high type workers are better privately insured and therefore exert less effort than the low types. Since the highest types have more risk of low outcome events, the covariance between effort and marginal utility is positive (and an order of magnitude larger than the largest covariance without private insurance), thereby weakening the case for social insurance. The Appendix provides the plots of these observations in Figures 8 and 9.
What is most dramatic about the numerical simulation is the gap between the optimal tax rate with and without private insurance at the two different levels of ex-ante heterogeneity. When ex-ante heterogeneity is large ($\bar{\theta} - \bar{\theta} = 1$), there is very little difference between optimal taxation levels with and without private insurance for moderate variation in ex-post shocks. Most of the variation in the economy is due to ex-ante differences, which the private market cannot smooth across. In such a case, standard optimal taxation formulas that neglect private insurance do not lead to substantial welfare losses. For $M > 1.3$ the gap becomes significant, in which case, there is a potential large welfare gain to lowering the level of social insurance. This gap, however, is clearly smaller than the analogous gap for smaller ex-ante heterogeneity. When ex-ante heterogeneity is small, optimal taxation in the presence of private insurance converges to zero quickly. The intuition for such a result lies in the fact that with ex-ante similar individuals experiencing productivity shocks, almost all of the variation in the economy is due to ex-post shocks and each agent has a very similar set of possible outcomes as their cohorts. Thus, there is very little scope for redistribution. In such an economy, there may be substantial welfare losses by failing to recognize the role of private insurance.

Figure 4: Loss In Welfare Associated with Suboptimal Social Insurance

The results of Figure 3 can be translated into a welfare analysis by comparing welfare under two scenarios. Assume that the economy is one in which there is a well-functioning private insurance market. First, consider the optimal social insurance policy set by a utilitarian social planner who takes into account the
presence of private insurers. Second, consider a social planner setting optimal policy as if there were no private insurance market. As shown in Figure 3, this will in general lead to too much insurance relative to the optimum and reduce welfare. The percentage loss in welfare relative to the level of welfare in the first case is plotted in Figure 4. The parameterization of the economy is unchanged from that which generated Figure 3 and the two plots refer to the cases in which ex-ante heterogeneity is large ($\theta - \bar{\theta} = 1$) and small ($\theta - \bar{\theta} = 0.2$). As expected from the prior discussion, the welfare loss from failing to recognize endogenous private insurance markets is increasing in the variance of the ex-post shocks since private insurance is effective at smoothing consumption for a given agent across high and low states of nature. Moreover, a decrease in ex-ante heterogeneity implies that the primary source of heterogeneity is through ex-post shocks, which again can be optimally insured in the private market. A social planner ignoring such private insurance thus has a larger negative impact on welfare. In summary, welfare losses due to ignoring the private insurance market are decreasing in ex-ante heterogeneity, but increasing in ex-post heterogeneity.

5.4 Partial Private Insurance

In an economy in which only some workers are privately insured, it is clear that optimal tax rates should be decreasing in the fraction of the population which is insured (conditional on optimal taxes being lower in the presence of private insurance). When more individuals are able to smooth consumption with private insurance contracts, the government’s tax policy serves only a redistributive role as opposed to also acting as insurance. The following simulation considers the case we have already analyzed in which ex-ante heterogeneity is large ($\theta - \bar{\theta} = 1$). We have already computed the optimal tax rates in which everyone is privately insured and no one is privately insured for various levels of heterogeneity in the ex-post shocks. In Figure 5, six additional intermediate cases are considered. For three of the cases, 50% of the population is privately insured and in the other three, 90% of the population is privately insured. Given the aggregate number of insured, we then consider three sub cases in which insurance is evenly distributed among the population and when insurance is perfectly positively and negatively correlated with ability. The most reasonable case is that in which the correlation is positive. High ability workers are more likely to have access to private insurance. This may be due, for example, to employers who offer high ability workers severance packages in the case of unemployment. And although adverse selection has been assumed away in this model, in actual insurance markets, we expect to see higher type individuals to have greater access to private insurance.

The results in Figure 5 confirm the intuition that optimal taxation is indeed decreasing in the fraction of the population with private insurance. It is interesting to note the concave relationship between the fraction of the population covered and optimal marginal tax rate. When a small fraction of the population is privately insured, an increase in private insurance coverage has a smaller impact on reducing the optimal tax rate than when most of the population is
already insured. Consider the case in which $M = 1.6$. Figure 5 shows that randomly insuring the first half of the population leads to a decrease in the optimal tax rate of 4 percentage points, while insuring the next half leads to an additional 12 percentage point drop in the level of the optimal marginal tax rate. The intuition for this result follows from the endogeneity of the private insurance contracts. Consider private firms dropping insurance coverage for one percent of the population. When few are privately insured, the government can make up for the loss in private insurance by offering more generous insurance to all with a small increase in the marginal tax rate. This has a minimal impact on the small number of existing private insurance contracts. When most individuals in the economy are privately insured, however, providing the same additional insurance now requires a greater increase in taxes since the increase in taxation is less effective as it crowds out private insurance throughout the economy.

From a policy perspective, this result suggests two issues. First, our estimates of the welfare loss from setting taxes without accounting for endogenous private insurance markets may be overstated even with robust private insurance. Second, policies which encourage private insurance provision will lead to increasing declines in the optimal tax rate.

The numerical simulation additionally reveals that introducing a positive or negative correlation between ability and private insurance coverage while fixing
the total fraction of those privately insured yields only small adjustments to the level of the optimal tax rate. We see in fact that the plots for optimal policy in the cases of correlated private insurance coverage track and bound the optimal tax rate when the correlation is zero. In particular, relative to the uniform distribution case, a positive correlation implies that more high types are insured, reducing the average effort of the high types in the population. Hence, the covariance between risk and ability is increased, weakening the redistributive role of taxation. Similarly, a negative correlation implies that the highest types are least likely to be insured and therefore exert more effort on average. This decreases the covariance between risk and ability, making more redistribution optimal.

Although the preceding observations are based on numerical simulations, they allow for a richer understanding of how the presence of private insurance markets and the degree of heterogeneity in a population can dramatically affect optimal social insurance policy. The parameterizations have been stylized to highlight important features of the model. An important next step would be an attempt to calibrate this model to the U.S. economy. The challenge inherent in any such task, however, is disentangling observed earnings into their two components of ex-ante productivity and ex-post productivity shock. As explained, the relative magnitudes of these two sources of heterogeneity is integral to the computation of optimal social insurance.

6 Conclusion

This paper has characterized optimal social insurance in an economy with endogenous individual private insurance and a richer notion of heterogeneity than in previous theoretical analyses. In particular, agents differ in their ex-ante productivity and subsequently experience a shock, resulting in either a high or low state of nature, relative to their innate productivity. Analytical results are achieved in a setting in which the government can observe the nature of the ex-post shock, as in unemployment. The paper shows that with ex-ante homogeneous agents, the introduction of optimal private insurance contracts obviates the need for any social insurance. Private and public insurers are equally effective at providing insurance to smooth consumption for a single type of agent. However, when individuals differ ex-ante, the social planner has the extra degree of freedom of being able to redistribute across agents. Thus, there may be scope for government intervention. The optimal social insurance benefit depends on the covariance between effort and marginal utility of consumption in the high state of nature. Intuitively, if the agents who exert more effort also consume more upon the realization of the high state, the social planner can improve a utilitarian welfare objective function by redistributing to those with lower earnings.

To supplement the analytical results, the paper also provides numerical simulations of the economy in a more general and less tractable setting in which the governing authority needs only to observe realized earnings and not the
type of ex-post shock in order to implement its policy. The natural justification for this setting is one in which the government imposes a linear income tax on a wage distribution that has been compressed by firms offering insurance to implement optimal effort provision. The simulations highlight a number of interesting features. The first is that although in many cases the presence of private insurance reduces optimal social insurance, this may not always hold. Since private insurance reduces effort, making low state events more likely, the redistributive motive may increase in the presence of private insurance when ex-ante heterogeneity is sufficiently large. Also, the welfare loss from optimizing social insurance, but ignoring private insurance markets is greatest when agents are ex-ante similar, but face large shocks, since in that case, the primary welfare improving tool is consumption smoothing within workers and not redistribution across workers. Thus, the presence of private insurance is effective at providing all of the insurance for the economy.

There exist interesting directions to expand this line of research theoretically. A simplifying assumption of the analysis in this paper is that private insurers are able to observe innate productivity. Such an observation eliminates the adverse selection issue. The introduction of adverse selection interacting with the moral hazard problem is an important step in better understanding the endogenous response of private insurance contracts to social insurance provision. In addition, if productivity is imperfectly observable or there are administrative costs, private insurance may not be provided optimally as assumed throughout this paper. Non-optimized private insurance may also come in the form of informal insurance mechanisms available to individuals, such as the ability to borrow from friends or rely on spousal labor income, as documented by Cullen and Gruber (2000)[6]. Implicit throughout this analysis and much of the related literature on optimal insurance in the presence of private insurance is that private insurers earn zero profits. One could consider instead a situation in which firms are not risk neutral and there exists a risk sharing contract between worker and firm. In such a setting, the firm becomes an agent in the economy and one could investigate if social insurance is as effective at redistributing across workers or simply affects the extent of risk sharing between firm and worker.

This paper is an important step in better understanding the factors which affect the determination of optimal social insurance when private insurance contracts and individual behavior respond endogenously. An important, but challenging direction for future empirical work is to calibrate the model with taxation as social insurance to a real economy. Doing so would require a disentangling of the observed variation in realized earnings into their unobserved ex-ante variation in skills and ex-post shocks to productivity. The implication

---

12Boadway et. al. (2006)[2] consider a case with adverse selection and ex-post moral hazard (conditional on being in the low state, agents can spend money to improve their welfare in that state) as opposed to the ex-ante moral hazard problem in this paper (effort is chosen before uncertainty is realized to impact the probability of the high state).

13Chetty and Saez (2009)[4] investigate non-optimized private insurance in a simpler setting with only one source of heterogeneity.
for policy is that we may in fact be able to improve welfare by reducing taxes that serve as a redistributive mechanism.

7 Appendix

7.1 Supplemental Proofs

Proof of Proposition 5 Employing the notation introduced prior to the statement of Proposition 6, private insurers solve the same problem as in Proposition 4 given a social insurance contract, \((b, t)\). In particular, optimal private insurance is characterized by

\[
\frac{u'(c_1^0(\theta)) - u'(c_1^1(\theta))}{u'(c_1^1(\theta))} = \frac{\varepsilon_{1-a^l(\theta), b_p(\theta)} b_p(\theta)}{a^l(\theta)} \tag{15}
\]

The social planner chooses \(b\) to optimize social welfare given the budget constraint that \(t = \frac{\alpha(1-a^l) + (1-\alpha)(1-a^N)}{\alpha a^l + (1-\alpha) a^N}\).

\[
W = \alpha \int \left( a^l(\theta) u \left( M(\theta) - t - \frac{1-a^l(\theta)}{a^l(\theta)} b_p(\theta) \right) \right) f(\theta) d\theta
\]

\[
+ (1-a^l(\theta)) u(m(\theta) + b + b_p(\theta)) - h(a^l(\theta)) f(\theta) d\theta
\]

Employing the envelope theorem for the agents’ utility maximizing choice of \(a\), we have that

\[
\frac{dW}{db} = \alpha \int \left( (1-a^l(\theta)) u'(c_0^1(\theta))(1-r(\theta)) \right.
\]

\[
- a^l(\theta) u'(c_1^1(\theta)) \frac{d}{db} \left( t + \frac{1-a^l(\theta)}{a^l(\theta)} b_p(\theta) \right) f(\theta) d\theta
\]

\[
+ (1-\alpha) \int \left( (1-a^N(\theta)) u'(c_0^N(\theta)) - a^N(\theta) u'(c_1^N(\theta)) \frac{dt}{db} \right) f(\theta) d\theta
\]

Moreover,

\[
\frac{d}{db} \left( t + \frac{1-a^l(\theta)}{a^l(\theta)} b_p(\theta) \right) = \frac{\alpha(1-a^l) + (1-\alpha)(1-a^N)}{\alpha a^l + (1-\alpha) a^N}
\]

\[
+ \frac{\alpha(1-a^l) \varepsilon_{1-a^l, b_p} + (1-\alpha)(1-a^N) \varepsilon_{1-a^N, b_p}}{(\alpha a^l + (1-\alpha) a^N)^2}
\]

\[
- r(\theta) \frac{1-a^l(\theta)}{a^l(\theta)} + \frac{b_p(\theta)}{b} \frac{1-a^l(\theta)}{a^l(\theta)^2} \varepsilon_{1-a^l(\theta), b}
\]

28
Combining the previous two expressions and using (15) for optimal private insurance as in the proof of Proposition 4, it follows that
\[
\frac{dW}{db} = \alpha E((1 - a^I)u'(c^I_1)) + (1 - \alpha)E((1 - a^N)u'(c^N_0)) - \frac{\alpha E(a'u'(c^I_1)) + (1 - \alpha)E(a^N u'(c^N_1))}{\alpha a^I + (1 - \alpha) a^N} \times \left[ \alpha(1 - a^I) + (1 - \alpha)(1 - a^N) + \frac{\alpha(1 - a^I)\varepsilon_{1-a^I} + (1 - \alpha)(1 - a^N)\varepsilon_{1-a^N}}{\alpha a^I + (1 - \alpha) a^N} \right]
\]
as desired.

**Proof of Proposition 6**  The social planner chooses \(s\) to maximize
\[
W = \int [a(\theta)u(M(\theta)(1 - s) + s\varepsilon) + (1 - a(\theta))u(m(\theta)(1 - s) + s\varepsilon) - h(a(\theta))] f(\theta)d\theta
\]
Using the envelope theorem for consumer utility maximization we have that
\[
\frac{dW}{ds} = \int \left[ a(\theta)u'(c_1(\theta)) \left( \varepsilon - M(\theta) - s \frac{d\varepsilon}{d(1 - s)} \right) + (1 - a(\theta))u'(c_0(\theta)) \left( \varepsilon - m(\theta) - s \frac{d\varepsilon}{d(1 - s)} \right) \right] f(\theta)d\theta
\]
\[
= a u'(c_1) + (1 - a)u'(c_0) \left( \varepsilon - s \frac{d\varepsilon}{d(1 - s)} \right) - E(au'(c_1)M + (1 - a)u'(c_0)m)
\]
\[
= \varepsilon \left( 1 - \varepsilon \frac{s}{1 - s} \right) (E(au'(c_1) + (1 - a)u'(c_0))) - E(au'(c_1)M + (1 - a)u'(c_0)m)
\]
where the last equality follows from the observation that \(\frac{d\varepsilon}{d(1 - s)} s = \varepsilon \frac{s}{1 - s} \varepsilon\).
This is the expression in (24) as desired and (25) follows immediately from setting \(\frac{dW}{ds} = 0\).

**Impact of Marginal Tax Rate on Welfare in Ex-Ante Heterogeneous Economy with Optimal Private Insurance**

As discussed in Section 5, one can perform a similar analysis to that in Proposition 4 when taxes are proportional and not lump sum. The result is presented here, although numerical simulations provide more insight than the following analytical expression.

**Optimal Private Contract.** Firms choose \(b_p(\theta)\) to maximize welfare, taking a marginal tax rate of \(b\) and rebate of \(G = b\varepsilon\) as given, where
\[ \bar{w} = \int \left( a(\theta)(M(\theta) - \frac{1-a(\theta)}{a(\theta)}b_p(\theta)) + (1-a(\theta))(m(\theta) + b_p(\theta)) \right) f(\theta) d\theta. \]

\[ W(\theta) = a(\theta)u \left( \left( M(\theta) - \frac{1-a(\theta)}{a(\theta)}b_p(\theta) \right)(1-b) + G \right) \]

\[ + (1-a(\theta))u((m(\theta) + b_p(\theta))(1-b) + G) - h(a(\theta)) \]

Using the envelope theorem, we have that

\[ \frac{dW}{db_p(\theta)}|_{b,G} = (1-b)(1-a(\theta))u'(c_1(\theta)) \left( \frac{u'(c_0(\theta)) - u'(c_1(\theta))}{u'(c_1(\theta))} - \frac{\varepsilon_{1-a(\theta),b_p(\theta)}|_{b,G}}{a(\theta)} \right) \]

This expression is the analog of (14) with the only change being the multiplicative term of (1-b). The optimal private insurance contract, conditional on \( b \neq 1 \), is therefore characterized by

\[ \frac{u'(c_0(\theta)) - u'(c_1(\theta))}{u'(c_1(\theta))} = \frac{\varepsilon_{1-a(\theta),b_p(\theta)}|_{b,G}}{a(\theta)} \] (26)

Note that this expression is identical to that in (15) for the optimal private contract given tax policy.

**Social Insurance in the Presence of Optimized Private Insurance.** The government chooses \( b \) to maximize social welfare, taking into account \( b_p(\theta) \) set as in (26). In particular, welfare is given by

\[ W = \int \left[ a(\theta)u \left( \left( M(\theta) - \frac{1-a(\theta)}{a(\theta)}b_p(\theta) \right)(1-b) + b\bar{w} \right) \right. \]

\[ + (1-a(\theta))u((m(\theta) + b_p(\theta))(1-b) + b\bar{w}) - h(a(\theta))] f(\theta) d\theta \]

Employing the envelope theorem, we have that

\[ \frac{dW}{db} = \int \left[ a(\theta)u'(c_1(\theta)) \left( \bar{w} - b - \frac{\bar{w}}{d(1-b)} - M(\theta) + \frac{1-a(\theta)}{a(\theta)}b_p(\theta) \right. \right. \]

\[ - (1-b) \left( - \frac{1-a(\theta)}{a(\theta)}r(\theta) + b_p(\theta) \frac{d(1-a(\theta))}{a(\theta) db} \right) \]

\[ + (1-a(\theta))u'(c_0(\theta)) \left( \bar{w} - b - \frac{\bar{w}}{d(1-b)} - m(\theta) - b_p(\theta) \right. \]

\[ - (1-b)r(\theta)) f(\theta) d\theta \]

Note that \( da \bar{w}/db = \varepsilon_{1-b,1-b} \frac{b}{1-b} \bar{w} \) and \( \frac{da(\theta)}{db} = \varepsilon_{1-a(\theta),\theta} \frac{1-a(\theta)}{a(\theta)} \). This implies

\[ \frac{dW}{db} = \bar{w} \left( 1 - \varepsilon_{1-b,1-b} \frac{b}{1-b} \right) E \left( au'(c_1) + (1-a)u'(c_0) \right) \]

\[ - E \left( ((1-a)u'(c_0) (m + b_p + (1-b)r) \right) \]

\[ - E \left( au'(c_1) \left( M - \frac{1-a}{a} \left( b_p + (1-b)r - \frac{1-b}{b} \frac{b_p}{a} \varepsilon_{1-a,b} \right) \right) \right) \]
From (26) we know that

\[ E(a(u'(c_1) - u'(c_0))) = -E(u'(c_1)\varepsilon_{1-a,b_p|b,G} = -E \left( u'(c_1)\varepsilon_{1-a,b_p \frac{b_p}{b(1-r)}} \right) \]

Thus,

\[
\frac{dW}{db} = \bar{w}(1 - \varepsilon_{1-b,1-b} \frac{b}{1-b}) E \left( u'(c_0) - u'(c_1)\varepsilon_{1-a,b_p \frac{b_p}{b(1-r)}} \right) \\
- E((1-a)u'(c_0)(m + b_p + (1 - b)r)) \\
- E \left( a u'(c_1) \left( M - \frac{1-a}{a} \left( b_p + (1 - b)r - \frac{1-b}{b} b_p \varepsilon_{1-a,b} \right) \right) \right) \]

This expression does not collapse to some simpler expression and setting \( \frac{dW}{db} = 0 \) does not provide clear insight into the comparative statics and comparisons made in the numerical simulation.
7.2 Supplemental Figures for Simulation in Figure 3

Figure 6: Optimal Effort with No Private Insurance
Figure 7: Effort-Marginal Utility Covariance with No Private Insurance

Figure 8: Effort-Marginal Utility Covariance with Private Insurance
Figure 9: Optimal Effort with Private Insurance
References


