OPTIMAL TAXATION AND RISK SHARING

Hamish Low
Daniel Maldoom

THE INSTITUTE FOR FISCAL STUDIES
WP 01/00

Optimal Taxation and Risk Sharing*

Hamish Low

University of Cambridge, Institute for Fiscal Studies and CEPR

Daniel Maldoom

DotEcon Ltd.

January 2000

Abstract

This paper analyses the trade-off between the incentive effects of increased uncertainty and the welfare benefits of risk-sharing in the design of optimal tax schedules. We use numerical methods to characterise the tax schedule and to give comparative static results of changing risk aversion, uncertainty and the cost of effort. Increased uncertainty may increase effort for precautionary reasons, but leads to greater risk sharing in the optimal tax schedule. Similarly, a reduced cost of effort leads to greater risk sharing. Incentives to work are induced through punishment at low output realisations if risk aversion is high, and through reward of high output if risk aversion is low. We also consider introducing extra randomisation into the tax schedule to further incentivise individuals. This is only optimal if the form of the tax schedule is constrained, for example to be linear.

*We are grateful to Tim Besley, Jayasri Dutta and Jim Mirrlees for helpful discussions and to seminar participants at Cambridge and Essex for helpful comments. Correspondence to: Hamish Low, Faculty of Economics and Politics, University of Cambridge, Sidgwick Avenue, Cambridge, CB3 9DD, UK. email: hamish.low@econ.cam.ac.uk or Dan Maldoom, DotEcon Ltd, 16 Hanover Sq, London, W1R 9AJ, UK. email: dan.maldoom@dotecon.com

Executive Summary

- Uncertainty can make individuals work longer hours for precautionary reasons, and so social insurance which reduces this uncertainty can lead to lower labour supply and hence less output in the economy. However, social insurance provides risk sharing benefits and so there is a trade-off between increasing uncertainty to incentivise people and reducing uncertainty to provide risk-sharing. This paper analyses the extent to which this trade-off should be exploited by the government in providing social insurance.

- Marginal tax rates increase with gross income if risk aversion is high, but decrease with gross income if risk aversion is low. Further, incentives to work are induced through punishment at low output realisations if risk aversion is high, and through reward of high output if risk aversion is low.

- Increases in the degree of uncertainty about income may increase effort for precautionary reasons, but leads to greater risk sharing in the optimal tax schedule. Similarly, a reduced cost of effort leads to greater risk sharing.

- We also consider introducing extra uncertainty (through randomisation in the tax schedule) to further incentivise individuals. If the government is unconstrained in the choice of tax schedule, extra randomisation is not optimal. Introducing extra uncertainty can be optimal if the form of the tax schedule is constrained, for example to be linear.

1 Introduction

The aim of this paper is to analyse the trade-off between incentives and risk-sharing in the design of optimal tax schedules. Redistributive tax schemes reduce the variance of net income, providing risk sharing benefits. However, such schemes may reduce the expected marginal returns from effort and, since uncertainty may induce people to increase effort for precautionary reasons, risk sharing itself may have a further blunting effect on incentives. The first contribution of this paper is to provide a method of setting up principal-agent problems of this form and to use this method to characterise the trade-off between uncertainty and risk sharing, partly through numerical solutions. The second contribution is to show the value of introducing additional randomisation into payoffs under different assumptions about the complexity of the tax system.

This issue underlies a number of questions in the literature. For example, Varian (1980) addressed the question of how much risk sharing should be provided through the tax system: progressive taxation provides insurance by reducing the variance of net income, but this insurance reduces the incentive to self-insure through saving. Varian presents some comparative static results, but these are only approximations holding endogenous variables fixed. A second question, first raised by Stiglitz (1982b), is whether the tax schedule should be subject to randomisation. One interpretation of randomising the tax schedule is that it is a way of introducing additional uncertainty to incentivise individuals at the cost of some risk sharing benefits. The notion that risk increases effort for precautionary reasons is analogous to the precautionary saving behaviour analysis of Kimball (1990). Holmstrom (1979, 1982) also discusses introducing noise into the payoffs of a principal-agent model and concludes that it is only beneficial to condition an agent's compensation on a random variable if that random variable provides information on the agent's effort. This result arises because the principal does not face a budget constraint.

The formulation in this paper follows the framework of Varian (1980) in solving for the general solution to the tax schedule. Comparative static results on this tax schedule require the complete solution of the optimisation problem and we use numerical methods to derive comparative static results on the extent of risk sharing, varying risk aversion, the cost of effort and uncertainty. The trade-off at the heart of the model is between inducing particular individuals to exert greater effort by making their income uncertain and providing risk sharing benefits for all individuals to increase social welfare. In this model, individuals are ex-ante identical and so the individuals who exert greater effort are also benefiting from the risk sharing. The key results of the comparative statics are: first, increases in the amount of exogenous uncertainty have a direct effect leading to greater effort for precautionary reasons with a given tax schedule. However, the optimal tax schedule changes to provide greater risk sharing because less information on effort is learnt from a given output level, and this leads to reduced effort. Second, a greater cost of effort leads to reduced effort. This leads to reduced risk sharing as the tax schedule tries to provide greater incentives to work to compensate for this reduced effort. Third, as individuals dislike of uncertainty increases (increasing risk aversion and increasing convexity of marginal utility), they will exert greater effort, thus the incentive effects of uncertainty are sharper. Offsetting this, the social welfare benefit of risk sharing is greater and so the effect on risk sharing through the tax schedule is ambiguous. When individuals are very risk averse, individuals are incentivised through penalties at low output, with significant risk sharing elsewhere. When individuals are less risk averse, they are incentivised through rewards for high output, with significant risk sharing elsewhere.

The framework in the paper can also be used to address the question of whether the tax schedule should be subject to randomisation because imposed randomisation is analogous to a reduction in risk sharing in the tax schedule. We show that introducing such randomisation is optimal only if the tax schedule is constrained, for example to be linear. If the tax schedule is unconstrained and income is already uncertain, then it is never optimal to introduce additional randomisation. This is because the incentive effects of randomisation can already be exploited through limiting the amount of insurance against income shocks. This has implications for the size of "groups" where the tax schedule is chosen to distribute output within the group. The implication of the no-randomisation result is that we should not restrict group size: restricting group size introduces aggregate risk, which is analogous to randomisation, and so is only welfare increasing if we are constrained in the set of tax instruments we can use. This conclusion can, however, be viewed in the opposite way: if we are constrained to use linear taxation, restricting group size or introducing randomisation may be ways of exploiting the incentive effects of uncertainty.

Section 2 presents the model with fully diversifiable risk. This section also describes a general way of characterising the problem which rules out the possibility of Mirrlees penalties. Section 2.3 then provides a numerical characterisation of the tax schedule for a range of parameter values. Section 3 demonstrates the non-optimality of randomisation if the tax schedule is unconstrained and section 4 shows the situations in which randomisation can be welfare increasing. Section 5 concludes.

2 The model

We consider a simple principal-agent model in which there is a continuum of agents. Each agent produces an output that depends on his or her own effort and an idiosyncratic shock. All output risks are potentially fully diversifiable, although it will not be optimal to do so because of incentive problems.

Agents are indexed by $i \in [0,1]$. Agent i's output, x_i, is a function of effort, e_i, and a random shock. Let $F(x_i \mid e_i)$ be the distribution function of x_i if effort e_i is expended; we assume that this is continuously distributed with continuously differentiable p.d.f. $f(x_i \mid e_i)$. The outputs of agents are independently distributed. All agents have identical utility functions, which are additively separable in effort and income; let $u(y)$ be the utility of income y and $g(e)$ be the disutility of effort e. We assume that u and g are twice continuously differentiable.

We consider optimal mechanisms for redistributing output between agents. We focus on mechanisms that preserve the veil of ignorance, in that they do not identify any particular individual. Formally, this means that the allocation mechanism should yield the same income for each agent even if the agents are relabelled (i.e. the indices i are permuted). Thus a mechanism may condition an agent's income on that agent's output and the *distribution* of all agents' outputs, but may not condition on any *particular* agent's output. In the case of an infinite number of agents with uncorrelated output shocks, the law of large numbers implies that in equilibrium the distribution of all agents' outputs is non-stochastic. Therefore, we only need to consider mechanisms in which each agent's income may be written as a function $s(x)$ of that agent's output x alone.[1]

We require that the sharing rule, $s(x)$, must be feasible, so that the total payment to agents is no more than total output in any contingency. We also assume that the rule only implements Pareto optimal distributions (to avoid the problem of ex-post renegotiation), so that total payments exactly equal total output. Even without the possibility of ex-post renegotiation, it is optimal to redistribute all output. If a suggested tax schedule withheld part

[1] In order to consider the incentive compatibility of the optimal sharing rule, we will need to consider hypothetical changes in effort by a single agent. Even under such hypothetical deviations we need only consider the dependency of the sharing rule on an agents' own output and may supress any dependency on other agents' outputs.

of the output, an alternative schedule could allocate the surplus to the top earner and increase welfare without worsening incentives.

The optimal sharing rule is found by solving the optimisation problem:

$$\max_{s(\cdot),e^*} \int u(s(x))\, dF(x \mid e^*) - g(e^*) \tag{1}$$

$$\text{s.t.} \quad e^* \in \arg\max_e \int u(s(x))\, dF(x \mid e) - g(e) \tag{2}$$

$$\int s(x)\, dF(x \mid e^*) = \int x\, dF(x \mid e^*). \tag{3}$$

where equation (2) is the incentive compatibility constraint and equation (3) is the adding-up condition. This framework follows the standard principal-agent framework first used by Mirrlees (1974). As discussed further below, we restrict the class of distributions $F(x \mid e^*)$ such that the first order approach is valid and so the incentive compatibility constraint can be replaced by a first order condition:

$$\int u(s(x))\, dF_e(x \mid e^*) = g'(e). \tag{4}$$

Therefore, the Lagrangian of this problem (2) can be written

$$\begin{aligned} L &= \int [u(s(x)) + \mu x - \mu s(x)]\, dF(x \mid e^*) - g(e^*) \\ &\quad + \lambda \left\{ \int u(s(x))\, dF_e(x \mid e^*) - g'(e) \right\} \end{aligned} \tag{5}$$

where λ is the Lagrange multiplier on the incentive compatibility constraint and μ that on the adding-up condition; both μ and λ should be positive (Mirrlees, 1974; Varian, 1980). Then $s(\cdot)$ must be chosen to maximise the integral

$$\int \{[u(s(x)) - \mu s(x)]\, f(x \mid e) + \lambda\, u(s(x))\, f_e(x \mid e)\}\, dx,$$

pointwise maximisation of which gives the first order condition

$$u'(s(x))\,[f(x \mid e) + \lambda f_e(x \mid e)] - \mu f(x \mid e) = 0.$$

This gives a solution for $s(\cdot)$ in terms of three real variables e, λ and μ:

$$u'(s(x)) = \frac{\mu}{1 + \lambda h(x \mid e)} \quad \text{where} \quad h(x \mid e) = \frac{f_e(x \mid e)}{f(x \mid e)}. \tag{6}$$

The unknown real variables in equation (6) are determined by the first-order condition with respect to effort,

$$\int [u(s(x)) + \mu(x - s(x))] f_e(x \mid e)\, dx$$
$$+\lambda \int u(s(x))\, f_{ee}(x \mid e)\, dx - g'(e) - \lambda g''(e) = 0 \tag{7}$$

the adding-up constraint (3) and the incentive compatibility constraint (4). Although it is possible to give an analytic expression for $s(x)$ in terms of the unknown variables, it is not possible to solve analytically for the actual solution to these unknowns.

Solution (6) corresponds to the solution in Varian (1980). The first point to note is that if the incentive compatibility condition were not to bind (i.e. $\lambda = 0$) then the sharing rule is constant and there is complete risk sharing. The incentive problem means that some risk-sharing will be lost to induce greater effort. Whether an individual does better or worse than under complete risk sharing depends on the sign of the term $h(x \mid e)$. This term is the likelihood ratio (Holmstrom, 1979): the derivative of the log-likelihood function of output with respect to (unobservable) effort. Given an observed x, it provides information on whether the individual exerted effort e. If this derivative is close to zero, the choice of effort by the individual makes little difference to the probability of observing x and so there is little incentive benefit to be gained from varying $s(x)$ rapidly according to x. However, as the derivative deviates from zero, additional effort can be more reliably identified and it is optimal for the sharing rule to provide less risk-sharing. For any particular x, a large positive value of $h(x \mid e)$ means it is more likely that the individual put in higher effort; in equation (6), this means that $u'(s(x))$ will be small and so the share corresponding to this x will be large. Similarly, a large negative value for the derivative means that the probability of observing x with high effort is low, and so the individual will receive a lower share.

This raises the issue, discussed by Milgrom (1981), of how the likelihood ratio changes as x increases. If we want the sharing rule to be increasing in x, then $h(x \mid e)$ must be increasing in x (the monotone likelihood ratio property).[2] Additionally, we wish to rule out the possibility that it is possible to make perfect, or asymptotically perfect, inferences about effort from observing output. Without such a restriction, the optimal solution may involve large fines on agents (Mirrlees, 1974). This means that $h(x \mid e)$ must be bounded on the support of output. If the likelihood ratio is unbounded, then effort can be identified almost perfectly in certain cases. It is well known that this requirement rules out a number of obvious candidates for specifying distributions of output. For example, it is not possible to use additive shocks ($x = \theta + e$), multiplicative shocks ($x = \theta e$) or transforms of either (e.g. $x = \Psi(\theta + e)$ or $x = \Psi(\theta e)$) without running into solutions with Mirrlees-style penalties. Each of these formulations leads to the well known result that we can come arbitrarily close to the first best by putting a very large penalty on the worst outcome, while allowing full risk sharing for any other realisation of x. This solution invalidates the interior solution (6) because the sharing rule is discontinuous.

2.1 The deformation method

The approach we take in this paper is to specify the family of distributions of output as the solution to a partial differential equation representing a continuous deformation of the output distribution as effort changes. In particular, we assume that output is distributed over a finite domain and

[2] The MLRP is sufficient for the sharing rule to be increasing in x only if utility is additively separable between effort and output.

that the family of density functions $f(x \mid e)$ satisfies the PDE

$$f_e(x| e) = -\frac{\partial}{\partial x}\left(\tau(x| e) f(x| e)\right), \tag{8}$$

where the distribution of output at zero effort $f(x \mid 0)$ is specified as a boundary condition. We call τ the *transition* function and interpret it in the following way: $\tau(x \mid e)$ measures the marginal effectiveness of effort in shifting the distribution of output to the right. In particular, if an agent increases his or her effort by δe, then a proportion $\tau(x \mid e) f(x \mid e)\,\delta e$ of the probability mass at output x flows to the right. The change in the probability of a particular output level if effort is increased is given by the combination of probability mass flowing *in* from lower output levels and the probability mass flowing *out* to higher output levels. Therefore,

$$\left(f(x| e+\delta e) - f(x| e)\right)\delta x = -\left(\tau(x+\delta x| e) f(x+\delta x| e) - \tau(x| e) f(x| e)\right)\delta e$$

which in the limit gives (8).

This formulation has a number of highly convenient properties, as demonstrated in propositions 1 and 2 (which are proved in the appendix). First we impose some minor restrictions. We assume that:

1. output at zero effort is distributed over a compact support $[\underline{x}, \overline{x}]$;
2. τ takes the multiplicatively separable form $\tau(x \mid e) = a(x)\,b(e)$, where a and b are continuously differentiable; and
3. $a(x) \to 0$ as $x \to \underline{x}, \overline{x}$.

The first assumption is made to avoid problems with almost perfect inferences about effort that can arise with infinite or semi-infinite domains. The second assumption is made for convenience as it permits a simple solution of the PDE; it may well be that this assumption could be relaxed. The third assumption is to ensure that when the distribution of output is deformed as effort increases, no probability mass 'leaks' outside the domain $[\underline{x}, \overline{x}]$ so that the solution of (8) is a well-defined probability density for all e.

Proposition 1 *If* $a(x)\frac{f_x(x|0)}{f(x|0)}$ *is bounded on* $[\underline{x}, \overline{x}]$*, then the likelihood ratio* $h(x \mid e)$ *is bounded on* $[\underline{x}, \overline{x}]$ *for any given* e*.*

Proposition 2 *If* $a'(x)+a(x)\frac{f_x(x|0)}{f(x|0)}$ *is monotone decreasing in* x*, then the MLRP is satisfied.*

The value of this method of specifying a family of distributions of output at different effort levels is that it gives a tractable example that can be solved by the use of numerical techniques.

2.2 A simple parameterisation

The solution (6) raises a number of key questions, for example: how progressive is the tax schedule; how does this progressivity change as uncertainty increases or as we change the degree of risk aversion or the cost of effort? We cannot answer these questions without solving explicitly for the optimal level of effort and the value of the two Lagrange multipliers. Varian (1980) does provide some answers to these comparative static questions but his answers assume that the values of the endogenous variables remain fixed as the model parameters change. As he recognises, this is at best an approximation.

Since the model is analytically intractable in its general form, we restrict the analysis to a particular, yet flexible parameterisation. The numerical solution requires a utility function and the relationship between output and effort to be specified. We assume that the utility of consumption and disutility of effort have the forms

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma},\ \gamma > 1 \quad \text{and} \quad g(e) = e^{\eta},\ \eta > 1.$$

Comparative statics are then performed on γ and η. In our numerical results, effort is always less than 1 and so increases in η mean a reduced utility cost of effort. Similarly the effect of changing γ on the utility of consumption depends on the level of consumption. Consumption lies between 0 and 1 in our simulations and so an increase in γ means the marginal utility of consumption increases. Further, prudence and risk aversion both increase with γ. The different properties of utility functions are tied together in this formulation, and both are determined by the value of γ. It is not clear, however, what intuition could be given for increasing prudence without changing risk aversion. We interpret an increase in γ as indicating individuals have a greater dislike of uncertainty and so will be willing to give up effort to increase consumption. A higher γ also means that the social welfare cost of uncertainty over utility is greater.

We specify the family of distributions of output and the effect of effort on the distribution using the deformation method described in section 2.1 and construct a family of output distributions with support on $[0, 1]$. We suppose output at zero effort has a Beta-distribution:

$$F(x|0) = CDF\left[\beta_{q,r}(x)\right]$$

which is characterised by two parameters q and r.

The parameters q and r are chosen so that the distribution is symmetric at zero effort, but as effort increases, probability mass moves to the right, and the distribution becomes right skewed. Underlying this choice for the distribution is the recognition that we are dealing with agents of identical abilities, and so the relevant distribution is the distribution of income for people of the same ability. It is hard to disentangle this distribution from observed income distributions which reflect differences in ability as well as differences in effort. The characterisation used here means that effort can help move income up the income distribution given the individuals' ability. This movement in the distribution as effort increases means that effort leads to increases in the average value of output. Further, the increase in effort leads to a reduction in the variance of output.

Using this formulation, we solve numerically for e^*, λ and μ using the first-order condition (7), and the constraints (3) and (4). We solve for a solution using the first order condition for incentive compatibility, but then check numerically that the first order approach is in fact valid. Parameters used are given in table 1.

Table 1: Parameters for Numerical Solution

Parameter	*Baseline*	*Alternatives*
γ	1.5	1.2,3.5, 5.0
σ^2	0.028	0.017, 0.056
η	3.0	2.0, 4.0

Values of the variance are for the distribution at 0 effort. Changes in the value of the variance are introduced through changing the parameters of the Beta Distribution (q and r), keeping the mean constant.

2.3 Characterising the Tax Schedule

The key question we now address is the extent to which the tax schedule provides risk sharing benefits. This is reflected in the way the average tax rate changes with income: if the average tax rate rises with income, this means the tax schedule is progressive and reduces the variance of net income, thus providing insurance. There is no presumption that the tax schedule should be progressive and it is not even necessary that marginal taxes be positive: it may, for example, be optimal to introduce greater uncertainty over payoffs than the degree of uncertainty under autarky and this can be done through having negative marginal tax rates at the top of the income scale paid for by high marginal tax rates at low income.

In each of the scenarios below, we present results on the average and marginal tax rates across the income distribution and also the optimal sharing rule. In presenting the marginal tax rates, it is necessary to be clear that these tax rates determine how much of any additional output accrues to the individual but that this is different from the expected marginal benefit of extra effort. The expected marginal benefit of effort is reduced by higher marginal taxes (changing the average return to extra effort), but the expected marginal benefit is also affected by the degree of risk sharing, which changes the variance of the return to extra effort. In other words, the incentive costs of taxation come from two sources: first, the expected return from a marginal increase in effort may be reduced by taxation and second, the variance of the return to effort may be reduced.

Tax Progressivity: increasing risk aversion

Before describing the optimal tax schedule in detail, it is instructive to analyse the individual's optimal choice of effort as we increase the coefficient of relative risk aversion, γ. In the absence of taxation, individuals choose their effort levels knowing that they will have to consume their own output. This effort incurs the marginal utility cost, $\eta e^{\eta-1}$, and has some expected marginal utility benefit which is determined by the distribution of shocks and by the coefficient of relative risk aversion, γ. If there was no uncertainty, the parameters γ and η determine the "technology" for translating disutility of effort into the utility of consumption. If an increase in γ means that effort becomes more productive in this sense, then there are two initial effects: a given amount of effort will lead to greater consumption utility (an income effect) and the return to higher effort will increase (a substitution effect). The income effect leads to reduced effort, the substitution effect leads to increased effort. Further, when output is uncertain, effort will change not only the mean of the distribution, but it may also change the variance, skewness and other moments. If effort reduces the variance, then as we increase γ (making individuals dislike uncertainty more) this effect will lead to greater effort (a risk-reducing effect). These effects suggest that effort is unlikely to change monotonically as γ increases even in the absence of any taxation.

Allowing for the optimal tax schedule introduces the further complication that the optimal schedule will change as γ increases: an increase in γ means individuals are more prudent and so a given amount of uncertainty induces greater effort. In other words, a higher value of γ increases the incentive effects of uncertainty. However, in terms of social welfare, a higher value of γ means a given amount of uncertainty has a greater welfare cost because individuals are more risk averse. The first effect suggests the tax schedule should display less risk sharing, whereas the second effect suggests it should display greater risk sharing. These ambiguities are reflected in the numerical solutions given in figure 1 and tables 2 and 3.

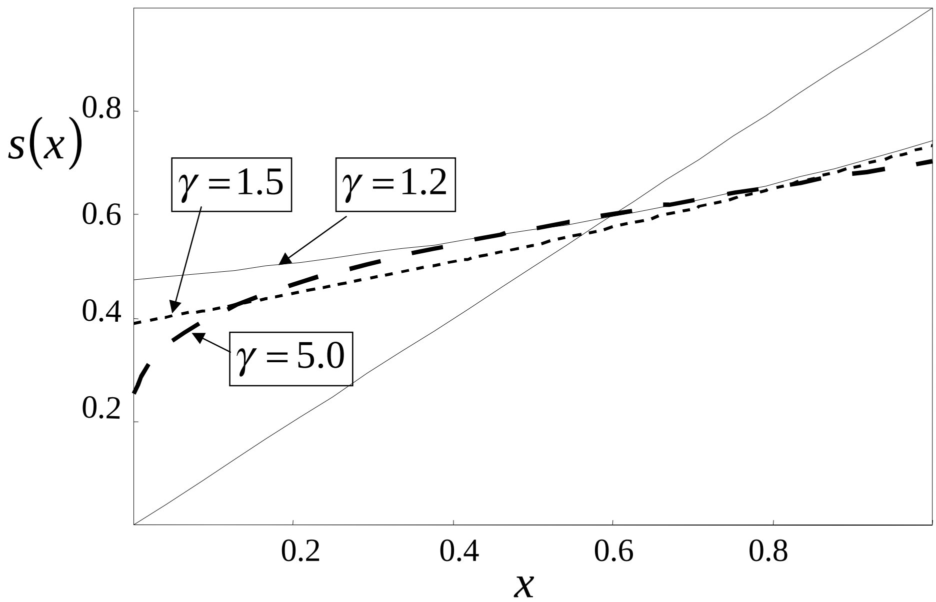
Table 2: Marginal Tax Rates $(1 - \frac{\partial s(x)}{\partial x})$ Varying γ

		$F(x\|e)$						
Model	*Effort*	*0.01*	*0.1*	*0.25*	*0.5*	*0.75*	*0.90*	*0.99*
$\gamma = 1.2$	0.61	0.81	0.77	0.74	0.71	0.67	0.63	0.58
$\gamma = 1.5$	0.40	0.71	0.69	0.67	0.65	0.63	0.61	0.58
$\gamma = 3.5$	0.45	0.60	0.66	0.69	0.70	0.71	0.71	0.71
$\gamma = 5.0$	0.60	0.56	0.68	0.72	0.74	0.74	0.75	0.75

Other parameters: $q = 4, r = 4$, and $\eta = 3.0$

Effort first decreases as γ increases and then starts to increase as γ

Figure 1: Sharing Rule Varying γ



Other parameters: $q = 4, r = 4$, and $\eta = 3.0$

Table 3: Average Tax Rates ($1 - \frac{s(x)}{x}$) Varying γ

		$F(x \| e)$						
Model	*Effort*	*0.01*	*0.1*	*0.25*	*0.5*	*0.75*	*0.90*	*0.99*
$\gamma = 1.2$	0.61	-1.19	-0.32	-0.09	0.06	0.15	0.19	0.23
$\gamma = 1.5$	0.40	-1.26	-0.36	-0.12	0.04	0.13	0.18	0.23
$\gamma = 3.5$	0.45	-1.18	-0.36	-0.12	0.04	0.14	0.20	0.26
$\gamma = 5.0$	0.60	-1.07	-0.33	-0.11	0.05	0.14	0.20	0.26

Other parameters: $q = 4, r = 4$, and $\eta = 3.0$.

becomes sufficiently large. This behaviour of effort as γ changes mirrors behaviour in the absence of a tax system, although the level of effort differs. If there were no uncertainty, we could interpret this as the income effect of "more productive" effort dominating the substitution effect as γ increases at low values, but the substitution effect dominating at higher values. With uncertainty over income, the increase in effort as γ increases will also be due to increased prudence.

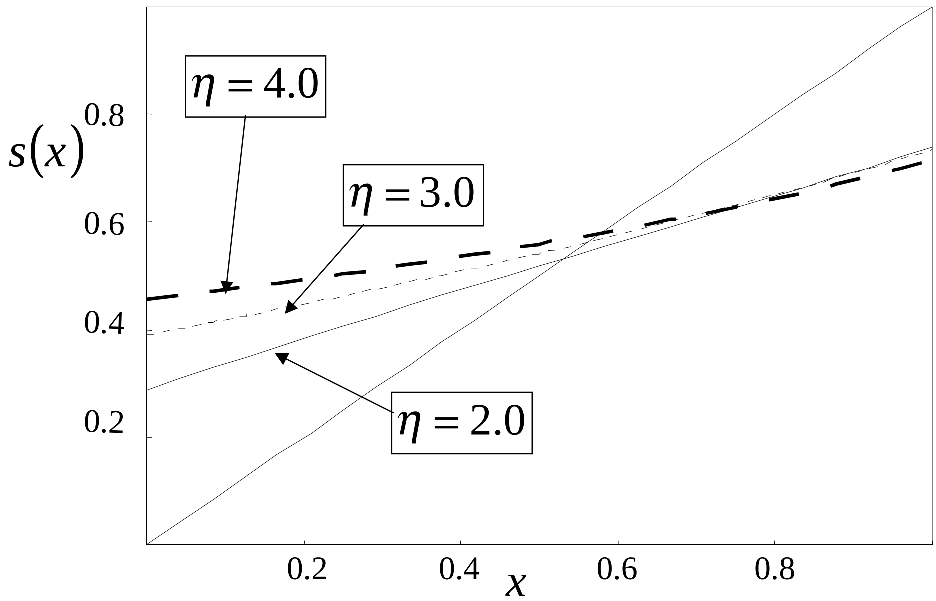
As γ becomes high (e.g. 3.5 or 5.0), the sharing rule becomes concave (marginal tax rates increase with x), in contrast to the convex rules for lower values of γ (marginal tax rates decrease with x). When γ is high individuals are incentivised through penalties at low output, with significant risk sharing elsewhere; however, when individuals are less risk averse, they are incentivised through rewards for high output, with significant risk sharing elsewhere. Further, the effect of increasing γ on marginal tax rates is dependent on the point in the income distribution: marginal tax rates are lower with higher values of γ at the bottom of the distribution, but higher with higher values of γ at the top of the distribution.

Tax Progressivity: increasing cost of effort

Increasing η is somewhat analogous to changing γ because it changes the rate of transformation of disutility of effort into the utility of consumption. A lower value of η means the utility cost of effort is greater because effort is less than 1. Thus, an increase in η will lead to increased effort for a given tax schedule. This generates extra revenue which results in changes to the optimal tax schedule: the extra revenue is partially spent on increased risk-sharing, offsetting the increase in effort. Further, risk sharing has less of an incentive cost with a higher η because individuals can easily be induced to supply effort. Thus, risk sharing and effort are both greater when $\eta = 4.0$.

Figure 2 shows the sharing rule, and tables 4 and 5 show the marginal and average tax rates for different values of η.

Figure 2: Sharing Rule Varying the Cost of Effort η



Other parameters: $q = 4, r = 4$, and $\gamma = 1.5$

Tax Progressivity: increasing the variance of shocks

Increasing the variance of the underlying shocks in the absence of taxation has the unambiguous effect of increasing effort if the marginal utility of consumption is convex. This arises because the expected marginal utility of consumption will increase with the variance, leading to an increase in effort in order to satisfy the Euler equation. In other words, precautionary motives for effort increase. However, if we allow the optimal tax schedule to vary, then the schedule will change according to the level of uncertainty. For

Table 4: Marginal Tax Rates $(1 - \frac{\partial s(x)}{\partial x})$ Varying η

		$F(x\|e)$						
Model	*Effort*	*0.01*	*0.1*	*0.25*	*0.5*	*0.75*	*0.90*	*0.99*
$\eta = 2.0$	0.18	0.54	0.55	0.56	0.56	0.56	0.56	0.56
$\eta = 3.0$	0.40	0.71	0.69	0.67	0.65	0.63	0.61	0.58
$\eta = 4.0$	0.50	0.80	0.77	0.75	0.73	0.70	0.68	0.65

Other parameters: $q = 4, r = 4$, and $\gamma = 1.5$

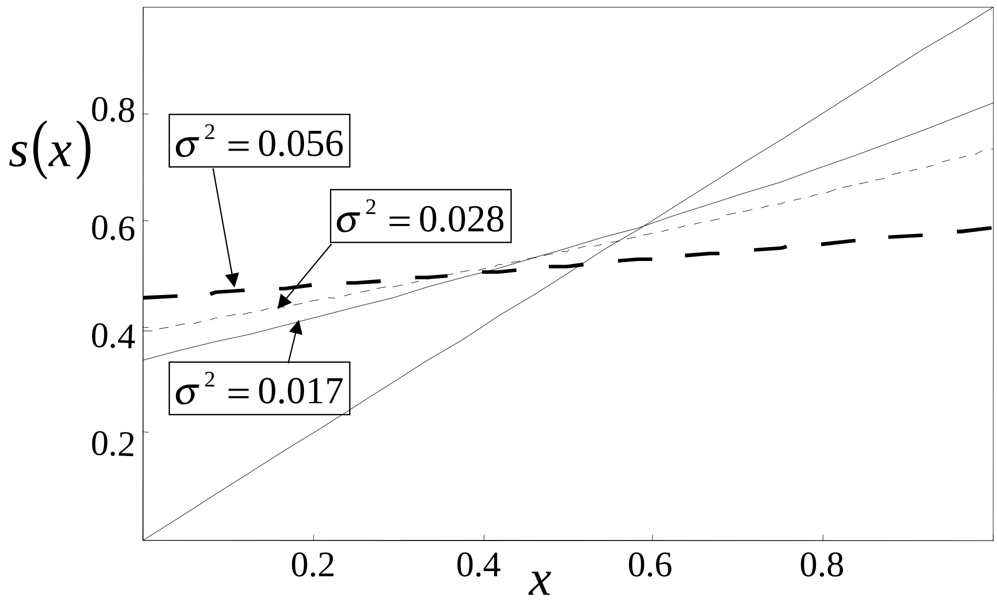
Table 5: Average Tax Rates $(1 - \frac{s(x)}{x})$ Varying η

		$F(x\|e)$						
Model	*Effort*	*0.01*	*0.1*	*0.25*	*0.5*	*0.75*	*0.90*	*0.99*
$\eta = 2.0$	0.18	-1.22	-0.38	-0.14	0.01	0.11	0.17	0.22
$\eta = 3.0$	0.40	-1.26	-0.36	-0.12	0.04	0.13	0.18	0.23
$\eta = 4.0$	0.50	-1.30	-0.37	-0.11	0.05	0.15	0.20	0.25

Other parameters: $q = 4, r = 4$, and $\gamma = 1.5$.

example, an increase in uncertainty is likely to lead to greater risk sharing because a given realisation of x gives less information about the amount of effort incurred. This greater risk sharing reduces effort, offsetting the direct effect of uncertainty increasing effort. In the range considered here, this indirect effect through the sharing rule dominates and effort falls and risk sharing increases as uncertainty increases. This means that the total output available to redistribute falls as uncertainty increases. The overall effects on the sharing rule and the tax rates are shown in figure 3 and tables 6 and 7.

Figure 3: Sharing Rule Increasing the Variance of Shocks



The baseline case with $q = 4$ and $r = 4$ is where $\sigma^2 = 0.028$. Values of the variance are for the distribution at 0 effort. Parameters q and r are altered to give the different values of the variance, holding the mean constant. Other parameters: $\eta = 3.0$ and $\gamma = 1.5$

Table 6: Marginal Tax Rates ($1 - \frac{\partial s(x)}{\partial x}$) Increasing the Variance

		$F(x \mid e)$						
Model	*Effort*	*0.01*	*0.1*	*0.25*	*0.5*	*0.75*	*0.90*	*0.99*
$\sigma^2 = 0.017$	0.47	0.57	0.55	0.53	0.50	0.48	0.46	0.42
$\sigma^2 = 0.028$	0.40	0.71	0.69	0.67	0.65	0.63	0.61	0.58
$\sigma^2 = 0.056$	0.20	0.88	0.88	0.87	0.87	0.86	0.86	0.85

The baseline case with $q = 4$ and $r = 4$ is where $\sigma^2 = 0.028$. Other parameters: $\eta = 3.0$ and $\gamma = 1.5$

Table 7: Average Tax Rates ($1 - \frac{s(x)}{x}$) Increasing the Variance

		$F(x \mid e)$						
Model	*Effort*	*0.01*	*0.1*	*0.25*	*0.5*	*0.75*	*0.90*	*0.99*
$\sigma^2 = 0.017$	0.47	-0.52	-0.18	-0.06	0.02	0.08	0.11	0.15
$\sigma^2 = 0.028$	0.40	-1.26	-0.36	-0.12	0.04	0.13	0.18	0.23
$\sigma^2 = 0.056$	0.20	-7.25	-1.28	-0.38	0.05	0.25	0.34	0.40

The baseline case with $q = 4$ and $r = 4$ is where $\sigma^2 = 0.028$. Other parameters: $\eta = 3.0$ and $\gamma = 1.5$

3 Non-Optimality of Random Taxes

In this section we show that it is never optimal to introduce additional randomisation to the tax schedule given that the tax schedule is unconstrained optimal. In the next section, we discuss why constraints on the choice of tax schedule mean randomisation will sometimes be Pareto improving.

This result on the non-optimality of randomisation is similar to the sufficient statistic result of Holmstrom (1979). The sufficient statistic result states that if a random variable does not reveal information about the agent's effort, then the sharing rule should be independent of that random variable. In other words, introducing randomisation to the sharing rule is never optimal. However, that result does not hold in models where the principal faces an adding up condition and so the proof given by Holmstrom (1979) does not apply in the current paper because the government faces an adding-up constraint in our model. In this case, the question arises of whether the incentive effects of risk could make randomisation desirable. If randomisation leads to increased effort, it increases the amount of output available for distribution between agents, thus relaxing the adding up constraint. Randomisation has no value when the principal does not face the adding up constraint because the principal is concerned only with monitoring how much effort a particular agent has made and not with aggregate output.

Stiglitz (1982a, 1982b) and further papers by Brito *et al.* (1991, 1995) present two alternative arguments in favour of randomisation: the first is that randomisation helps with self-selection when there is an adverse selection problem. The second result is more relevant to the current paper. Stiglitz (1982b) argues that if individuals have to choose their labour supply prior to knowing their (linear) income tax rate, they will choose a higher labour supply than if they knew they would face the average expected tax rate for certain. This means that a lower average tax rate can be used to meet a given revenue requirement. A higher labour supply arises if risk increases labour supply (which requires marginal utility to be convex in the wage): in other words, if individuals have precautionary motives. Individuals are worse off because they face the risk arising from the random tax, but they are better off because they face a lower average tax. This result is analogous to the result on optimal redistribution discussed in Varian (1980) and extended above: uncertainty increases labour supply because of prudence and so increases total output, but there is a risk sharing cost associated with this. This trade-off is exploited in the Varian (1980) model using the uncertainty due to shocks to income, whereas in Stiglitz (1982b) the trade-off is exploited by introducing randomisation into a model with no uncertainty and a linear income tax. Further, this trade-off can help explain the lack of complete risk-sharing in villages in the developing world documented by Townsend (1994).

This discussion raises the question of whether introducing randomisation is welfare increasing in the model presented in the current paper which already has exogenous uncertainty. Randomisation incentivises people through extra risk, but the uncertainty over gross income means individuals already face risk. If the government has the ability to set an entire income tax *schedule*, there is no benefit from randomising taxes. Any incentive benefits of risk may be fully exploited through the appropriate design of the tax schedule, since the shape of this schedule may be adjusted to alter the degree to which an agent is insured against output risk and the extent to which the agent is incentivised by uncertainty. This result is shown in proposition 3 below.

Proposition 3 *If the choice of tax schedule is unconstrained, introducing randomisation of the tax schedule is welfare reducing.*

Proof: Suppose that we have a "random" sharing rule. We represent this in a general manner as a stochastic process $S(x)$ giving the net income of an individual with gross income x. It is not necessary that $S(x_1)$ and $S(x_2)$

are independent for $x_1 \neq x_2$, but for the adding up rule to be satisfied, it must be the case that

$$E\left(S\left(x\right)\right)=E\left(x\right).$$

We show that such a rule cannot be optimal. Define a new sharing rule $\overline{s}\left(x\right)$ by

$$u\left(\overline{s}\left(x\right)\right)=E\left(u\left(S\left(x\right)\right)\right),$$

where the expectation is taken over possible values of $S\left(x\right)$ for a given gross income x. Given *any* particular effort level, an agent will receive the same expected utility under the deterministic sharing rule $\overline{s}$ as under the stochastic sharing rule S. Therefore it follows that if S was incentive compatible, so was $\overline{s}$. Moreover, expected utility is the same under $\overline{s}$ as under S.

Now since u is strictly concave, it follows that

$$\overline{s}\left(x\right)<E\left(S(x)\right)$$

for any x. Therefore, if $S\left(x\right)$ satisfied the adding-up condition

$$\int E\left(S(x)\right)f\left(x\mid e\right)=\int xf\left(x\mid e\right)$$

then

$$0<D=\int xf\left(x\mid e\right)-\int \overline{s}\left(x\right)f\left(x\mid e\right).$$

Therefore there exists an alternative sharing rule $\tilde{s}\left(x\right)$ that is incentive compatible, allocates the surplus output, D, and gives strictly higher expected utility than $\overline{s}$ and hence S. For example, define $\tilde{s}$ equal to $\overline{s}$ for all x except the highest possible gross income $\overline{x}$, where $\tilde{s}\left(\overline{x}\right)=\overline{s}\left(\overline{x}\right)+D$. Since the marginal return from effort at any effort level is at least as high under $\tilde{s}$, optimal effort is at least as great under $\tilde{s}$ as under $\overline{s}$. Therefore, $\tilde{s}$ dominates the stochastic sharing rule S.■

The property necessary for the conclusion that randomisation is welfare reducing is that the social planner must have sufficient flexibility in determining how post-tax income depends on output (pre-tax income). In the case that the social planner is able to choose any function mapping pre-tax into post-tax income, we find that randomisation is never optimal.

In the next section we consider the possibility of limits on the complexity of the sharing rule. In particular, if the sharing rule is linear, then it may in some cases be optimal to randomise taxes. We derive a condition for randomisation to be optimal and show this is related to the prudence of the individual. It is interesting to consider intermediate cases, where there may be some limits on the complexity of the tax system, but these limits are fairly weak. For example, suppose that post-tax income may be a piece-wise linear function of pre-tax income, with n pieces. Then as n becomes large, this piecewise linear schedule increasingly approximates the unrestricted optimal tax system, and by continuity it follows that randomisation is not desirable. Thus we may conclude that, in this model, the desirability of random taxes requires two conditions to be satisfied: that agents are sufficiently prudent; and that there is some limit on the complexity of the tax system (for example a limit, or a cost, to the number of tax bands that may be used).

It is also interesting to consider, in this context, optimal team size. By the size of the "team" we mean the size of the group of people who are risk sharing. Holmstrom (1982) considers the case of joint production where problems of free-riding occur and cases when relative performance evaluation is possible. If we abstract from these issues, however, restricting team size is a means of introducing extra randomisation into payoffs through making the total output available stochastic. Therefore, according to proposition 3, restricting team size cannot increase welfare. Returning to the issue of why risk sharing in the developing world (Townsend, 1994), the analysis above suggests the key point is not why we observe only partial risk sharing within villages, but why we observe limited attempts to insure against village-wide risk. Clearly monitoring of the output of other villages is more costly than monitoring within the village.

4 Linear Tax Schedules

The aim of this section is to present results on the choice of tax schedule when the social planner is constrained to use a linear tax. We first show that randomisation may be desirable: randomisation with a linear tax schedule can mimic some of the benefits of having a nonlinear tax schedule. We then present a numerical example showing how the tax rate and benefit of randomisation change with prudence.

Consider a linear sharing rule of the form $s_i(x) = \sum_j a_{ij} x_j + b_i$ where the a's and the b's are constants. The adding-up constraint requires that $\sum_i \sum_j a_{ij} x_j + \sum_i b_i = \sum_i x_i$ for all possible output vectors $x = (x_1, x_2, \ldots, x_n)$. It is easy to show that this condition is satisfied if and only if

$$\sum_i b_i = 0 \quad \text{and} \quad \forall j \quad \sum_i a_{ij} = 1.$$

If we additionally impose the condition that fines are not possible, i.e. $s_i(x) \geq 0$, then this is satisfied if and only if $a_{ij} \geq 0$ for all i, j and $b_i = 0$ for all i. If we also impose symmetry in that the permutations of the identities of agents should leave the contract unchanged, then we must have that $a_{ij} = a_{ik}$ for any $j, k \neq i$. Thus there is just one form of symmetric linear sharing rule, namely one of the form

$$s_i(x) = \beta x_i + \frac{1-\beta}{n-1} \sum_{j \neq i} x_j. \tag{9}$$

For the 'no-fines' condition to be satisfied, we must have that $\beta \in [0, 1]$.

In this section, we specify output to be additive in shocks, i.e. $x = e + \theta$. The restriction to linear sharing rules means that the additive nature of shocks will not lead to a violation of the first-order approach. We also assume, for simplicity, that $n = \infty$.

The optimisation problem faced by the social planner is therefore,

$$\max_{e,\beta} \int u(e + \beta\theta) f(\theta) \, d\theta - g(e)$$

subject to[3]

$$\beta \int u'(e + \beta\theta) f(\theta) \, d\theta - g'(e) \geq 0 \tag{10}$$

The adding-up constraint is included through the definition (9).

The Lagrangian for this problem is

$$L = \int u(e + \beta\theta) f(\theta) \, d\theta - g(e) + \lambda \left\{ \beta \int u'(e + \beta\theta) f(\theta) \, d\theta - g'(e) \right\} \tag{11}$$

giving first-order conditions,

$$\frac{\partial L}{\partial \beta} = \int \left[\theta \left(u'(e + \beta\theta) + \lambda\beta u''(e + \beta\theta) \right) + \lambda u'(e + \beta\theta) \right] f(\theta) \, d\theta = 0 \tag{12}$$

$$\frac{\partial L}{\partial e} = \int u'(e + \beta\theta) f(\theta) \, d\theta + \lambda\beta \int u''(e + \beta\theta) f(\theta) \, d\theta = g'(e) + \lambda g''(e) \tag{13}$$

Before specifying functional forms to yield an explicit solution for β and e, we define the condition necessary for randomisation to be optimal. Since β is endogenous, we need to solve the model explicitly to see if this noise criterion is satisfied and further, to establish comparative static properties such as whether the criterion is more likely to be satisfied with a higher degree of risk aversion.

Definition 1 *If the "noise criterion" is positive, then randomisation is welfare improving. If the tax schedule is constrained to be linear, the noise*

[3] The second order condition requires restrictions on the utility and distribution functions such that

$$\beta^2 \int u''(e + \beta\theta) f(\theta) \, d\theta - g''(e) < 0$$

which is satisfied if utility is concave in consumption and convex in effort.

criterion is defined as:

$$\frac{1}{1-\beta}-\frac{-\frac{\int u'''(e+\beta\theta)f(\theta)d\theta}{\int u''(e+\beta\theta)f(\theta)d\theta}}{\frac{g''(e)}{g'(e)}-\frac{\int u''(e+\beta\theta)f(\theta)d\theta}{\int u'(e+\beta\theta)f(\theta)d\theta}}>0 \tag{14}$$

This criterion is derived by introducing randomisation to the tax schedule by adding small amounts of individual specific noise, z distributed as $\Phi(z \mid \theta)$ with mean zero for each θ. We consider the introduction of noise at the margin by using the parameter α. We then wish to know how the optimised value of the following problem changes as α is increased from 0.

$$\max_{\beta,e} \iint u(e+\beta\theta+\alpha z)\, d\Phi(z \mid \theta)\, dF(\theta) - g(e)$$

$$\text{s.t.} \quad \beta \iint u'(e+\beta\theta+\alpha z)\, d\Phi(z \mid \theta)\; dF(\theta) = g'(e) \tag{15}$$

The Lagrangian of this problem is

$$\begin{aligned} L(\alpha) &= \iint u(e+\beta\theta+\alpha z)\,\phi(z \mid \theta)\, f(\theta)\, dz d\theta - g(e) \\ &\quad +\lambda\left\{\beta \iint u'(e+\beta\theta+\alpha z)\,\phi(z \mid \theta)\; f(\theta)\, dz d\theta - g'(e)\right\} \end{aligned}$$

We proceed by taking a second-order Taylor expansion of the Lagrangian around $\alpha = 0$. Since noise has mean zero, $\left.\frac{\partial L(\alpha)}{\partial \alpha}\right|_{\alpha=0} = 0$ and so the sign of $L(\alpha) - L(0)$ is determined by the second-order term,

$$\begin{aligned} \left.\frac{\partial^2 L(\alpha)}{\partial \alpha}\right|_{\alpha=0} &= \iint z^2 u''(e+\beta\theta)\,\phi(z \mid \theta)\, f(\theta)\, dz d\theta \\ &\quad +\lambda\beta \iint z^2 u'''(e+\beta\theta)\,\phi(z \mid \theta)\; f(\theta)\, dz d\theta \end{aligned}$$

Hence, the "noise criterion" is defined by this term being positive. Rearranging gives,

$$\iint z^2 \phi(z \mid \theta)\; dz \left\{u''(e+\beta\theta) + \lambda\beta u'''(e+\beta\theta)\right\} f(\theta)\, d\theta$$

If the distribution $\phi(z \mid \theta)$ is independent of θ, then the noise criterion becomes

$$\int u''(e+\beta\theta) f(\theta)\, d\theta + \lambda\beta \int u'''(e+\beta\theta) f(\theta)\, d\theta \tag{16}$$

The Lagrange multiplier, λ, can be substituted out of this criterion by dividing the first-order conditions with respect to e equation (13) by the incentive compatibility constraint (10) to give

$$\lambda = \frac{(1-\beta)}{\beta}\left(\frac{g''(e)}{g'(e)} - \frac{\int u''(e+\beta\theta) f(\theta)\, d\theta}{\int u'(e+\beta\theta) f(\theta)\, d\theta}\right)^{-1}$$

which, on substitution into equation (16) and rearranging, gives the noise criterion (14) in the definition.

4.1 A numerical example

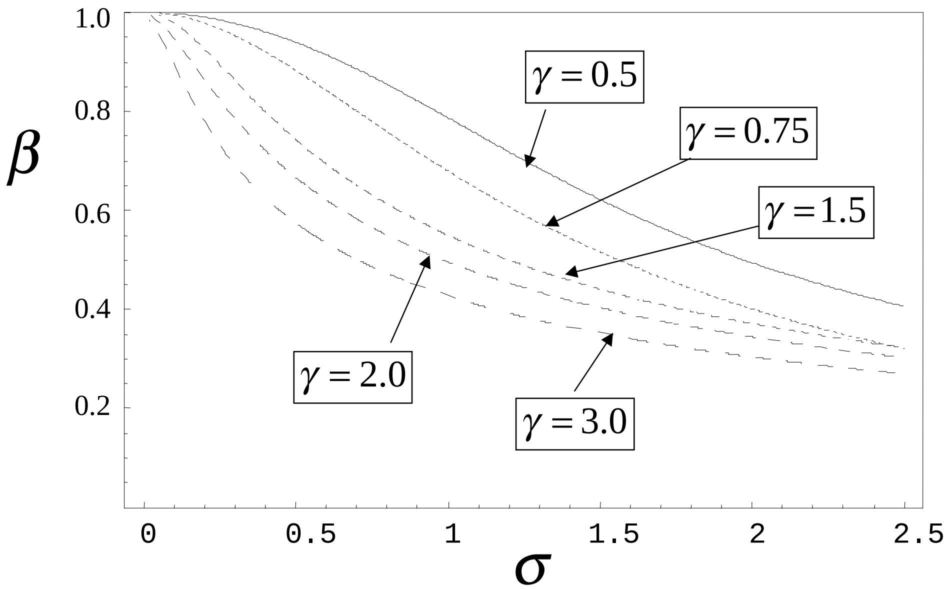
To give some indication of when the noise criterion is likely to be satisfied and to perform comparative statics, we specify a numerical example. Shock are assumed to be log-normally distributed, $\ln\theta \sim N\left(-\frac{\sigma^2}{2}, \sigma^2\right)$ and utility is as in section 2.3,

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma},\ \gamma > 0 \quad \text{and} \quad g(e) = e^{\eta},\ \eta > 1$$

We first show some comparative statics on the choices of β and e as we change σ,γ and η. A lower value of β means greater risk sharing because net income is less sensitive to the particular individual's output. Figures 4 and 5 show the optimal choice of β as σ increases for different values of γ and η. These comparative statics are fairly intuitive when $\eta = 1$: increases in γ or in σ mean there is greater risk sharing. When $\eta = 2$, the degree of risk sharing increases with γ if σ is sufficiently large, but decreases with γ if σ is very small.

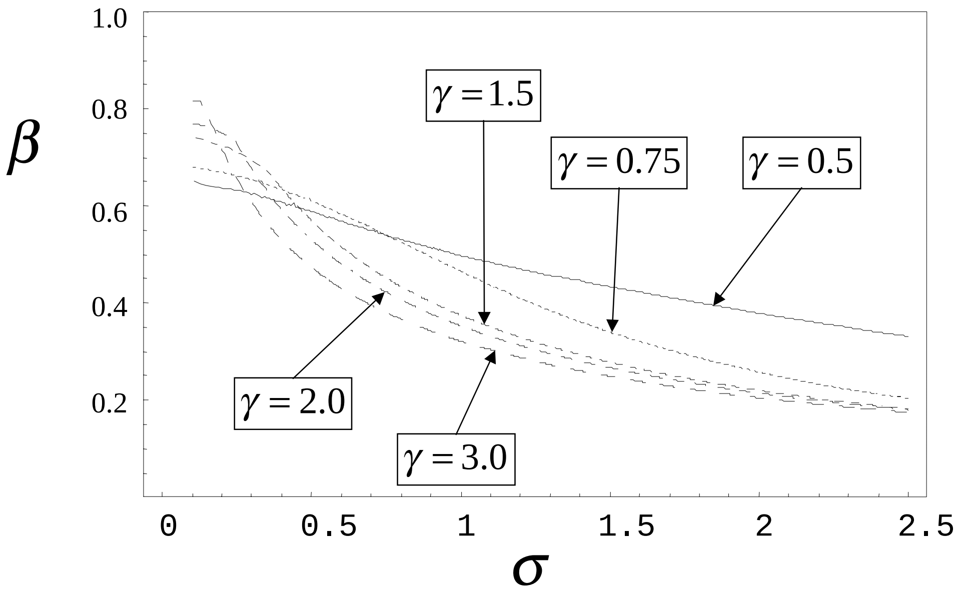
We now address the question of when the noise criterion will be satisfied and hence, extra randomisation will be welfare improving. If there is already

Figure 4: Values of β as σ Increases, Varying γ



Other parameters: $\eta = 1.0$

Figure 5: Values of β as σ Increases, Varying γ



Other parameters: $\eta = 2.0$

a large amount of uncertainty in the model, then introducing randomisation is less likely to be optimal. This leads to the following proposition:

Proposition 4 *The value of the noise criterion decreases as the variance of the shocks, σ, increases. This means that for any values of γ and η, there is a value $\bar{\sigma}$ such that randomisation is welfare increasing if $\sigma < \bar{\sigma}$, but randomisation is welfare reducing if $\sigma > \bar{\sigma}$.*

It is difficult to show this analytically because β is a function of σ. Instead, the value of the noise criterion was calculated for σ between 0.0 and 3.0, for the values of γ and η given above, and there was a unique value $\bar{\sigma}$ (dividing the σ domain into 1000 values). Table 8 shows the values of $\bar{\sigma}$ for a range of values of γ and η.

Table 8: Values of $\bar{\sigma}$ Varying η and γ

	$\eta = 1$	$\eta = 2$
$\gamma = 0.5$	0.91	0.72
$\gamma = 0.75$	0.92	0.73
$\gamma = 1.5$	1.33	1.25
$\gamma = 2.0$	1.64	1.79
$\gamma = 3.0$	2.35	3.00

When σ is less than the value of $\bar{\sigma}$ given in the table, local randomisation will be welfare increasing.

As γ increases, the ratio of prudence to risk aversion falls which means that adding noise to the tax schedule is beneficial even with a high variance of effort shocks. The most noteworthy point shown by this table is that the noise criterion is satisfied in a wide range of scenarios. Large output shocks are required before adding extra noise fails to increase welfare.

5 Conclusions

This paper provides a way of specifying principal-agent problems that permits tractable numerical solutions. This technique is used in characterising the optimal tax schedule under uncertainty: risk-sharing clearly increases social welfare, but risk-sharing reduces worker effort by firstly reducing the average marginal product of effort and secondly weakening precautionary motives for effort. Therefore, there is an optimal amount of risk-sharing. We characterise the nature of this trade-off numerically, and show comparative static results. The advantage of using numerical methods here is that we are able to characterise complete changes rather than only taking approximations holding endogenous variables fixed.

There are two, interrelated extensions that are of interest. First, we have only considered idiosyncratic shocks here. Allowing for aggregate shocks is somewhat analogous to introducing randomisation of the tax schedule and this will change the optimal sharing rule. This leads to the question of how the sharing rule should change across time and across the business cycle. This relates to the second extension which is to allow differences in ability: if we allow the tax schedule to vary across time, then it may be optimal to condition on past output and use past output to reveal information about ability. This clearly raises a number of additional problems, such as time inconsistency.

6 Appendix: The Deformation Method

This appendix describes a general way of constructing a family of distributions which ensures that the likelihood ration is bounded and increasing. We assume that output is distributed over a finite support $[\underline{x}, \bar{x}]$. We suppose that as effort increases we can think of the probability distribution of x shifting toward higher values. Consider the probability mass on a small range of outcomes $(x, x+\delta x)$, which is approximately $f(x|\, e)\, \delta x$, and then suppose that effort increases by δe. Probability mass moves to the right, with the *proportionate* rate of movement being given by $\tau(x|\, e)\, \delta e$. Thus, the range $(x, x+\delta x)$ receives an inflow of probability from the left of $\tau(x|\, e)\, f(x|\, e)\, \delta e$ and produces an outflow to the right of $\tau(x+\delta x|\, e)\, f(x+\delta x|\, e)\, \delta e$. Therefore,

$$(f(x|\, e+\delta e) - f(x|\, e))\, \delta x = -(\tau(x+\delta x|\, e)\, f(x+\delta x|\, e) - \tau(x|\, e)\, f(x|\, e))\, \delta e$$

which defines the change in probability mass in the region $(x, x+\delta x)$ following an increase in e. Taking the limits of δx and δe,

$$f_e(x|\, e) = -\frac{\partial}{\partial x}(\tau(x|\, e)\, f(x|\, e)) \tag{17}$$

In order for $f(x|\, e)$ to be a properly defined family of density functions, we need to ensure that the total density sums to 1. Since we are simply reallocating probability mass within the domain of f, this will be true providing that we ensure that no mass flows out of the edges of the support. We require $\tau(\underline{x}, e) = \tau(\bar{x}, e) = 0$. Then

$$\begin{aligned} \frac{\partial}{\partial e}\int_{\underline{x}}^{\bar{x}} f(x \mid e)\, dx &= \int_{\underline{x}}^{\bar{x}} f_e(x \mid e)\, dx \\ &= \tau(\underline{x} \mid e)\, f(\underline{x} \mid e) - \tau(\bar{x} \mid e)\, f(\bar{x} \mid e) \\ &= 0. \end{aligned}$$

Thus, suppose $e = 0$ is the lowest possible effort level, if

$$\int_{\underline{x}}^{\bar{x}} f(x \mid 0)\, dx = 1$$

then we have that

$$\int_{\underline{x}}^{\bar{x}} f(x \mid e)\, dx = 1 \quad \text{for all } e.$$

The partial differential equation (17) for the family of probability distributions f can be explicitly solved in certain simple cases. First, integrating

(17) with respect to x gives that

$$\int_{\underline{x}}^{x} f_e(s \mid e)\, ds = -f(x \mid e)\, \tau(x \mid e)$$

and so introducing the cumulative distribution F we obtain that

$$F_e = -F_x \tau \tag{18}$$

This PDE can be solved explicitly if we assume that τ takes the separable form

$$\tau(x \mid e) = a(x)\, b(e)$$

where $a(\underline{x}) = a(\bar{x}) = 0$. In this case we can separate (18) to obtain

$$-\frac{F_e(x \mid e)}{b(e)} = F_x(x \mid e)\, a(x)\,.$$

This has an 'almost additive' general solution of the form

$$F(x \mid e) = \phi\left(A(x) + B(e)\right)$$

where ϕ is any function and

$$A'(x) = \frac{1}{a(x)} \qquad\qquad B'(e) = -b(e)\,.$$

Thus the general solution of (18) is

$$F(x \mid e) = \phi\left[\int_{x^*}^{x} \frac{1}{a(s)} ds - \int_0^e b(s)\, ds\right]. \tag{19}$$

where x^* is some fixed point in the domain $(\underline{x}, \bar{x})$. The function ϕ is determined by providing a boundary condition, for example specifying the distribution of output at zero effort, i.e.

$$F(x \mid 0) = \phi\left[\int_{x^*}^{x} \frac{1}{a(s)} ds\right].$$

Notice that ϕ has the whole real line as its domain. Since a goes to zero at $\underline{x}$ and $\bar{x}$ we have that

$$A(\bar{x}) = \int_{x^*}^{\bar{x}} \frac{1}{a(s)} ds = +\infty \qquad \text{and} \qquad A(\underline{x}) = \int_{x^*}^{\underline{x}} \frac{1}{a(s)} ds = -\infty.$$

We can now derive sufficient conditions for $\frac{f_e(x|e)}{f(x|e)}$ to be bounded, thus proving proposition 1. Differentiating (19) gives that

$$f(x \mid e) = \phi' \left[\int_{x^*}^{x} \frac{1}{a(s)} ds - \int_0^e b(s)\, ds \right] \frac{1}{a(x)}$$

and so

$$f_e(x \mid e) = -\phi'' \left[\int_{x^*}^{x} \frac{1}{a(s)} ds - \int_0^e b(s)\, ds \right] \frac{b(e)}{a(x)}.$$

Thus

$$h(x \mid e) = -\frac{\phi'' \left[\int_{x^*}^{x} \frac{1}{a(s)} ds - \int_0^e b(s)\, ds \right]}{\phi' \left[\int_{x^*}^{x} \frac{1}{a(s)} ds - \int_0^e b(s)\, ds \right]} b(e) \tag{20}$$

Thus, $h(x \mid e)$ is bounded if $\phi''(y) / \phi'(y)$ is bounded.

We now need sufficient conditions for this. Since ϕ is defined in terms of a and the boundary condition $F(x \mid 0)$, we need to impose conditions on these functions. Now

$$f_x(x \mid 0) = \phi'' \left[\int_{x^*}^{x} \frac{1}{a(s)} ds \right] \frac{1}{a(x)^2} - \phi' \left[\int_{x^*}^{x} \frac{1}{a(s)} ds \right] \frac{a'(x)}{a(x)^2}$$

and so

$$\frac{f_x(x \mid 0)}{f(x \mid 0)} = \frac{\phi''}{\phi'} \frac{1}{a(x)} - \frac{a'(x)}{a(x)}.$$

Thus we obtain

$$\frac{\phi''}{\phi'} = a'(x) + a(x) \frac{f_x(x \mid 0)}{f(x \mid 0)}. \tag{21}$$

Now we have both a and a' tending to zero at the boundaries of the domain, so a' and a are bounded above and below on $(\underline{x}, \bar{x})$. Thus provided that

$$a(x) \frac{f_x(x \mid 0)}{f(x \mid 0)}$$

is bounded for all x in the domain, then $h(x \mid e)$ is bounded, proving proposition 1. This is a matter of ensuring that $a(x) \frac{f_x(x|0)}{f(x|0)}$ tends to a constant as $x \to \underline{x}, \bar{x}$.

In addition, it follows directly from equation (20) that $h(x \mid e)$ will be monotone increasing in x and the MLRP satisfied providing that $\phi''(x) / \phi'(x)$

is monotone decreasing in x. From equation (21) it follows that a necessary and sufficient condition for the MLRP is that $a'(x)+a(x)\frac{f_x(x|0)}{f(x|0)}$ is decreasing in x. This expresses the MLRP in terms of the transition function a and the boundary condition, proving proposition 2.

The easiest way to specify an example which satisfies these conditions is to define the mapping $A(x) : (\underline{x}, \bar{x}) \to \mathbf{R}$ such that $\underline{x}$ goes to $-\infty$ and $\bar{x}$ goes to $+\infty$. We do this in the numerical solution in section 2.3 by setting $\underline{x} = 0$, $\bar{x} = 1$,

$$\begin{aligned} A(x) &= \ln\left(\frac{x}{1-x}\right) \qquad a(x) = x(1-x) \\ B(e) &= -e \qquad\qquad\quad b(e) = 1 \end{aligned} \tag{22}$$

For the boundary condition, we specify the distribution of output at zero effort as having a Beta Distribution. In other words,

$$\begin{aligned} F(x \mid 0) &= \phi\left[\int_{x^*}^{x} \frac{1}{a(s)} ds\right] = CDF\left[\beta_{q,r}(x)\right] \\ f(x \mid 0) &= \phi'\left[\int_{x^*}^{x} \frac{1}{a(s)} ds\right] \frac{1}{a(x)} = \beta_{q,r}(x) \end{aligned}$$

Note that the domain of ϕ is the real line, whereas the domains of $F(x \mid 0)$ and $\beta_{q,r}(x)$ are the interval $[0,1]$. Therefore,

$$\begin{aligned} f(x \mid e) &= \phi'\left[\int_{x^*}^{x} \frac{1}{a(s)} ds - \int_0^e b(s)\, ds -\right] \frac{1}{a(x)} \\ &= \beta_{q,r}\left[\frac{\exp(A(x)+B(e))}{1+\exp(A(x)+B(e))}\right] \end{aligned}$$

where the transform of $A(x)+B(e)$ is to map from the real line back into $[0,1]$.

With this formulation, we have that

$$\frac{f_x(x \mid 0)}{f(x \mid 0)} = \frac{q}{x} - \frac{r}{1-x}$$

so

$$a(x)\frac{f_x(x \mid 0)}{f(x \mid 0)} = q(1-x) - rx$$

which is bounded on $[0,1]$, implying that the likelihood ratio h is bounded. Also

$$a'(x) + a(x)\frac{f_x(x \mid 0)}{f(x \mid 0)} = 1 + q - (2 + q + r)x$$

which is monotone decreasing in x, implying that the MLRP is satisfied.

References

[1] Brito, D. L., Hamilton, J. H., Slutsky, S. E., and Stiglitz, J. E. (1991), Dynamic optimal income taxation with government commitment. *Journal of Public Economics*, 44:15–35.

[2] Brito, D. L., Hamilton, J. H., Slutsky, S. E., and Stiglitz, J. E. (1995), Randomisation in optimal income tax schedules. *Journal of Public Economics*, 56:189–223.

[3] Coate, S. and Ravallion, M. (1993), Reciprocity without commitment: characterisation and performance of informal insurance arrangements. *Journal of Development Economics*, 40:1–24.

[4] Cremer, H. and Pestieau, P. (1996), Redistributive taxation and social insurance. *International Tax and Public Finance*, 3(3):281–295.

[5] Grossman, S. J. and Hart, O. D. (1983), An analysis of the principal-agent problem. *Econometrica*, 51:7–45.

[6] Hart, O. and Holmstrom, B. (1987), The theory of contracts. In Bewley, T., editor, *Advances in Economic Theory: Fifth World Congress of the Econometric Society*, pages 71–155. Cambridge University Press.

[7] Holmstrom, B. (1979), Moral hazard and observability. *Bell Journal of Economics*, 10:74–92.

[8] Holmstrom, B. (1982), Moral hazard in teams. *Bell Journal of Economics*, 13:324–340.

[9] Jewitt, I. (1990), Justifying the first order approach to principal-agent problems. *Econometrica*, 56:1177–1190.

[10] Kimball, M. (1988), Farmers cooperatives as behaviour toward risk. *American Economic Review*, 78:224–232.

[11] Kimball, M. (1990), Precautionary saving in the small and in the large. *Econometrica*, 58:53–73.

[12] Lazear, E. and Rosen, S. (1981), Rank order tournaments as optimum labour contracts. *Journal of Political Economy*, 89:841–864.

[13] Milgrom, P. (1981), Good news and bad news: Representation theorems and applications. *Bell Journal of Economics*, 12:380–391.

[14] Mirrlees, J. A. (1971), An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38:175–208.

[15] Mirrlees, J. A. (1974), Notes on welfare economics, information and uncertainty. In Balch, M., McFadden, D., and Wu, S., editors, *Essays in Equilibrium Behaviour and Uncertainty*. North-Holland.

[16] Mirrlees, J. A. (1986), The theory of optimum taxation. In Arrow, K. J. and Intriligator, M. D., editors, *Handbook of Mathematical Economics*, volume 3. North-Holland.

[17] Rogerson, W. (1985), The first order approach to principal-agent problems. *Econometrica*, 53:1357–1368.

[18] Shavell, S. (1979), Risk sharing and incentives in the principle and agent relationship. 10:55–73.

[19] Stiglitz, J. (1982), Utilitarianism and horizontal equity. *Journal of Public Economics*, 18:1–33.

[20] Townsend, R. M. (1994), Risk and insurance in village india. *Econometrica*, 62(3):539–591.

[21] Tressler, J. H. and Menezes, C. F. (1980), Labour supply and wage rate uncertainty. *Journal of Economic Theory*, 23:425–436.

[22] Varian, H. (1980), Redistributive taxation as social insurance. *Journal of Public Economics*, 14:49–68.