NBER WORKING PAPER SERIES

MORAL HAZARD AND OPTIMAL COMMODITY TAXATION

Richard Arnott

Joseph E. Stiglitz

Working Paper No. 1154

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

June 1983

The research reported here is part of the NBER's research program in Taxation. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #1154
June 1983

MORAL HAZARD AND OPTIMAL COMMODITY TAXATION

Richard Arnott
Joseph Stiglitz

ABSTRACT

The central result of this paper is that when moral hazard[+] is present, competitive equilibrium is almost always (constrained) inefficient. Moral hazard causes shadow prices to deviate from market prices. To remedy this market failure, the government could introduce differential commodity taxation. Moral hazard causes people to take too little care to prevent accidents. The corresponding deadweight loss can be reduced by subsidizing (taxing) those goods the consumption of which encourages (discourages) accident avoidance. At the (constrained) optimum, the sum of the deadweight losses associated with moral hazard, on the one hand, and differential commodity taxation, on the other, is minimized.

[+] Moral hazard is a generic phenomenon which occurs whenever the provision of insurance, be it explicit or implicit, affects the probability(ies) of the insured-against event(s). The classic papers in the literature include Arrow[1965], Spence and Zeckhauser [1971], Pauly [1974], and Mirrlees [1975].

Richard Arnott
Department of Economics
Queen's University
Kingston, Ontario
Canada K7L 3N6

Joseph Stiglitz
Department of Economics
Princeton University
Princeton, New Jersey 08544
U.S.A.

Moral Hazard and Optimal Commodity Taxation*

There has been increasing awareness over the past fifteen years of the importance and pervasiveness of problems of moral hazard. Risk-averse individuals purchase insurance which affects their incentives to undertake accident avoidance activities. With perfect (costless) information, insurance contracts would specify the actions to be undertaken and provide complete insurance. With costly information, however, insurance contracts provide only partial insurance, balancing at the margin the loss in reduced incentives from providing more insurance and the gain from risk-sharing. Moral hazard problems arise not only in insurance markets, but also in labor, product, and capital markets, in all of which elements of implicit or explicit insurance are prevalent in contractual relations.

In this paper we establish that, with more than one consumer good and costless government intervention, competitive equilibrium is almost always (constrained) inefficient when moral hazard is present; and to correct the market failure, differential commodity taxation is necessary.

The rationale for this result is as follows: Consider the extreme moral hazard situation, in which the insurer can observe only the outcome of an accident and has no information on either the underlying state of nature or the precaution taken by the individual to prevent the accident. In this case, the insurer can do no better than to provide insurance against the accident *per se* (i.e., he cannot make the payout contingent on either the insured's actions or the state of nature). As a result, the individual will typically

---

* We would like to thank the National Science Foundation and the Social Sciences and Humanities Research Council of Canada for financial support, Jim Gaisford for expositional criticism, and seminar participants at Stanford, Yale, MIT, and Tel-Aviv Universities for useful comments.

take less care than he would with either no insurance or perfect monitoring. Take as the hypothetical base case, competitive equilibrium with costless monitoring. Relative to this base case, competitive equilibrium with costly monitoring and therefore moral hazard entails a deadweight loss; because individuals' precautions are not monitored, they take "too little" care. (This in itself does not imply constrained inefficiency; all it says is that the unobservability of accident-prevention activities entails a cost.). Now consider the effect in the economy with moral hazard of subsidizing those goods that are complementary to accident-avoidance activities, and taxing those goods that are substitutable for them. The individual will undertake more accident-prevention activity, which will reduce the deadweight loss associated with being insufficiently careful. However, such commodity taxation, by causing prices to diverge from marginal costs, introduces a second source of deadweight loss. Constrained efficiency obtains when the sum of the two sources of deadweight loss is minimized, and we establish that this does involve differential commodity taxation.[1]

This point is intuitively appealing. It will be desirable to subsidize fire extinguishers if the social cost of the additional fire extinguishers bought from doing so is small relative to the expected reduction in fire damages. Similarly, it will be desirable to tax alcohol if the direct efficiency loss from doing so is outweighed by the fall in automobile accident damages.

---

1. It appears to be the conventional wisdom that moral hazard does not cause constrained inefficiency. This belief is based on the results of Pauly [1974], Stiglitz [1974], Helpman and Laffont [1975], Marshall [1976], and Shavell [1979 a,b] among others, all of which assume that there is a single consumer good, a linear production technology, and observability of an individual's total insurance purchases. In Arnott and Stiglitz [1982c] we showed that constrained efficiency obtains under these assumptions, but not when any one of them is relaxed.

The rationale for differential commodity taxation here is different from that in the conventional optimal tax literature (e.g., Diamond and Mirrlees [1971], Mirrlees [1971] and Atkinson and Stiglitz [1976]). There, the tax authorities would like to tax individuals according to say ability, but this is presumed to be unobservable and as a result the tax authorities must tax observable items, which include commodities, on the basis of their correlation with ability. If individuals were identical, it would be efficient to impose lump sum taxation; differential commodity taxation would be harmful. Here, however, even with identical individuals so that lump-sum taxation is possible, one will want to impose differential commodity taxation.[2]

Thus far and in the analysis that follows, we ignore the costs of government intervention. These may, however, be substantial and exceed any possible efficiency gains from different commodity taxation. In this case, even though shadow prices deviate from market prices, the market allocation, since it cannot be improved upon, must be said to be constrained efficient. Hence, the market failures we identify should be interpreted as <u>potential</u> market failures, becoming <u>actual</u> market failures only when the benefits of government intervention exceed the costs. Furthermore, we have assumed that the market price of each good equals its production cost. But if firms can tie the sale of various goods, they can replicate any allocation achievable via a tax system. For the case where this possibility is admitted (which we ignore

2. Viewed from another perspective, however, the conventional optimal tax problem and that considered here are structurally similar. Even though individuals in our model are identical <u>ex ante</u>, they differ <u>ex post</u> on the basis of the state of nature experienced. It is the impossibility of imposing differential lump-sum taxation on these <u>ex post</u> groups (because of the unobservability of the state of nature) that gives rise to the desirability of differential commodity taxation.

for most of the paper), our analysis should be interpreted as indicating how the market responds to moral hazard.

We organize our discussion as follows: Section 1 presents the general model, while section 2 examines a variety of special cases. In section 3, we discuss a few of the policy implications of our analysis.

1. The General Model

We start by considering a general formulation of the optimal tax problem with moral hazard. This will give insight into the structure of the problem. Unfortunately, the first-order conditions (as in the conventional optimal commodity tax problem), though interpretable, are sufficiently complex that they provide little guidance concerning the optimal tax rates on the various goods. As a consequence, in the next section we treat several special cases, each of which focuses on a different determinant of the optimal tax structure.

We assume that there is a single representative risk-averse individual in the economy for whom there are I possible outcomes indexed by i.[3,4] Outcomes are differentiated on the basis of which of a variety of kinds of accidents befall the individual, and the damage associated with each of the accidents. The outcome may affect the individual either directly (giving him pain or pleasure and affecting his tastes) or via his gross (of insurance premium and payout) income.

---

3. Thus, we rule out problems associated with adverse selection, which are treated in Rothschild and Stiglitz [1976] and Spence [1978], *inter alia*.

4. Alternatively, one may interpret the model as applying to an economy with a large number of *ex ante* identical individuals whose accident outcomes are statistically independent.

In principle, we can put any "commodity" into one of 30 categories on the basis of

i) whether or not it is taxable[5] (smoking in bed is not, fire extinguishers are);

ii) if it is taxable, whether an individual's consumption of it can be taxed non-linearly[6] (cigarettes cannot be, water consumption can);

iii) it is exchangeable or not (a fire extinguisher is, sleeping a sufficient amount of time that one is alert is not);

iv) it affects utility directly or not (smoking does, a fire extinguisher does not);

v) it affects the accident probability directly or not; and

vi) it is used before or after the accident outcome.

The following combinations of attributes are possible, where B denotes used before the accident, and A after.

	Affects Accident Probability						Does Not Affect Accident Probability			
	Affects utility directly				Does not affect utility directly		Affects utility directly			
	Exchangeable		Non Exchangeable		Exchangeable		Exchangeable		Non-Exchangeable	
	B	A	B	A	B	A	B	A	B	A
Taxable Non-linear										
Taxable Linear					(b)			(a)		
Non-taxable			(c)							

For simplicity, we limit ourselves to three categories:

(a) Linearly taxable exchangeable goods which affect utility but not accident probabilities and are used after the accident outcome; we refer to these

---

5. Taxability is closely related to observability at reasonable cost.

6. Non-linear taxation of a good is typically possible if its consumption can be observed. If only the exchange of the good can be observed (so that an individual's total purchases are unobservable) taxation has to be linear.

as consumer goods, index them by k, and let c_k^i be the quantity of consumer good k purchased by the individual with outcome i.

(b) Linearly taxable exchangeable goods which affect accident probabilities but not utility and are used before the accident outcome; we refer to these as accident-prevention goods, index them by ℓ, and let f^ℓ be the quantity of accident-prevention good ℓ purchased by an individual.

(c) Non-exchangeable, non-taxable goods which affect both utility and accident probabilities directly and are used before the accident outcome; we refer to these as types of accident-prevention effort, and let e^j be the amount of accident-prevention effort of type j expended by an individual.

In keeping with our assumption of linear taxation, we assume that the insurer, since he cannot observe an individual's total purchases of a commodity, cannot write his insurance contingent on an individual's purchases. Thus, the insurer can do no better than choose the net payout (payout minus premium) for each outcome, some of which may be negative. To simplify the analysis, we treat a linear production technology, and measure produced goods in such a way that all producer prices are unity.

A doubling of all consumer prices and net (of net payout) incomes in this economy has no effect. We therefore require another normalization. There is no obviously preferable one. We shall employ different normalizations in different parts of the paper. In this section, we normalize on the basis of the individual's net income if no accident befalls him.

<u>Notation</u>:

i index of the outcome; i = 0 corresponds to no accident

j index of the type of accident-prevention effort

k index of the type of consumer good

ℓ index of type of accident-prevention good

x_i the individual's net insurance payout with outcome i

y_i the individual's gross income with outcome i; $y_0 = 1$

$z_i \equiv y_i + x_i$ an individual's net-of-insurance income with outcome i

e^j quantity of accident-prevention effort of type j expended by the individual

c_i^k quantity of good k consumed with outcome i by the individual

f^ℓ quantity of accident-prevention good ℓ purchased by the individual

q^k consumer price of consumer good k

p^ℓ consumer price of accident-prevention good

π_i probability of outcome i

u_i utility of the individual with outcome i.

We let e, c_i, f, p, q, x, y, π and z denote vectors (since there will be no ambiguity, we shall not distinguish between row and column vectors).

We now characterize the social welfare optimum. The analysis proceeds in the same way as in the conventional optimal commodity tax problem. In the first stage, the individual chooses e, f, and $\{c_i\}$, treating consumer prices and the parameters of the insurance contract $\{x_i\}$ as given. This gives $e = e[p,q,z]$, $f = f[p,q,z]$ and $c_i = c_i[p,q,z]$.[7] (Throughout, we use square

7. It is possible that the dependent variables will change discontinuously with changes in p, q, and z. We ignore this problem. It can be handled, however, by the procedure developed in Grossman and Hart [1980].

brackets to enclose the argument of a function.) In the second stage, the planner chooses p, q, and z (or, more directly, x) to maximize expected utility, subject to an economy-wide resource constraint, and taking into account how the individual adjusts e, f, and $\{c_i\}$ in response to changes in p, q, and z.[8]

Note that in this formulation, the consumer does not decide how much insurance to buy at a price quoted by the planner; the net payout provided in each of the outcomes is instead decided by the planner. In Arnott-Stiglitz [1982b] we showed that if an insured's total purchases of insurance are observable, insurance contracts specify net incomes for each outcome and do <u>not</u> allow the individual to purchase as much insurance as he wants at the quoted price; if, however, an individual's total insurance purchases are unobservable, then contracts which (essentially) allow the individual to purchase as much insurance as he wishes at a parametric price will prevail. Thus, the way we have formulated the problem assumes that the government or an insurer can monitor individuals' total insurance purchases. We could have treated the other case; the argument for taxes and subsidies would be even stronger.

The first stage of the analysis, the consumer maximization problem, is now presented.

The probability of outcome i is

---

8. Since this maximization problem is not necessarily convex, random insurance policies and random taxation may be desirable. The former is discussed in Arnott and Stiglitz [1982b]; the latter in Weiss [1976] and Stiglitz [1982a]. We ignore these possibilities in this paper.

$$\pi_i = \pi_i[e,f], \tag{1}$$

while utility with outcome i is

$$u_i = u_i[c_i, e], \tag{2}$$

where $\frac{\partial u_i}{\partial c_i^k} > 0$, $\frac{\partial^2 u_i}{\partial (c_i^k)^2} < 0$, and $\frac{\partial u_i}{\partial e^j} < 0$. Note that we allow tastes to depend on the outcome. Expected utility is

$$EU = \sum_i \pi_i[e,f] u_i[c_i, e]. \tag{3}$$

With outcome i, the individual's budget constraint is[9]

$$z_i = qc_i + pf. \tag{4}$$

Where $\alpha_i \pi_i$ is the Lagrange multiplier on the individual's budget constraint with outcome i, the first-order conditions of his maximization problem are:

$$e^j: \quad e^j \left(\sum_i \left(\frac{\partial \pi_i}{\partial e^j} u_i + \pi_i \frac{\partial u_i}{\partial e^j} \right) \right) = 0 \text{ and } e^j \geq 0. \tag{5a}$$

$\sum_i \frac{\partial \pi_i}{\partial e^j} u_i$ is the expected marginal benefit in utiles associated with a unit increase in effort of type j, and $\sum_i \pi_i \frac{\partial u_i}{\partial e^j}$ is the corresponding marginal cost.

$$f^\ell: \quad f^\ell \left(\sum_i \left(\frac{\partial \pi_i}{\partial f^\ell} u_i - \alpha_i \pi_i p^\ell \right) \right) = 0 \text{ and } f^\ell \geq 0. \tag{5b}$$

9. We ignore the possibility of bankruptcy.

This equation has an interpretation analogous to (5a).[10]

$$c_i^k: \qquad c_i^k\left(\pi_i\left(\frac{\partial u_i}{\partial c_i^k} - \alpha_i q^k\right)\right) = 0 \text{ and } c_i^k \geq 0. \tag{5c}$$

(5c) states that with each outcome, the individual will choose the bundle of goods that maximizes his (<u>ex post</u>) utility.

(4), along with (5a) - (5c), yield

$$e = e[p,q,z], \; f = f[p,q,z], \text{ and } c_i = c_i[p,q,z]. \qquad (6a,6b,6c)$$

And substituting these equations into (3) gives

$$EU = \sum_i \pi_i[p,q,z]V_i[p,q,z],$$

where V_i is the indirect utility function with outcome i.

We now turn to the second-stage of the planning problem in which the planner chooses p, q, and x so as to maximize expected utility, taking into account how individuals respond to changes in these variables, and subject to the economy's resource constraint.

The economy-wide resource constraint facing the planner (recall footnote 4) is

$$\sum_i \pi_i(y_i - c_i - f) = 0. \tag{8}$$

Using the individual's budget constraint, we may write (8) alternatively as

10. Note that we have assumed here that if an individual purchases an accident-prevention good he always uses it. Complications which arise when this assumption is relaxed are treated in section 2.2.

$$\sum_i \pi_i x_i = \sum_i \pi_i((q-1)c_i+(p-1)f), \tag{8'}$$

which states that (expected) net insurance payouts must equal (expected) tax revenues.

The planner's problem may therefore be written as

$$\max_{q,p,x} \mathcal{L} = \sum_i \pi_i V_i - \lambda(\sum_i \pi_i (x_i-(q-1)c_i-(p-1)f)), \tag{9}$$

where λ is the Lagrange multiplier on (8'). We let X denote any element of $\{q^k\}$ or $\{p^\ell\}$, and $s_i \equiv c_i+f-y_i = x_i-(q-1)c_i-(p-1)f$ be the net social subsidy to an individual who experiences outcome i. We distinguish between subscripts or superscripts by using ~'s, e.g. $\frac{\partial f^{\tilde{\ell}}}{\partial p^\ell}$ denotes the change in accident-prevention good $\tilde{\ell}$ with respect to a change in the price of accident-prevention good ℓ. Finally, it turns out that the relevant "compensated" derivatives entail compensation such that, after the change, the individual is able to purchase his pre-change (outcome-contingent) bundles of goods and no more, which preserves expected utility;[10a] thus, where subscript θ indicates such compensation,

$$\left(\frac{\partial c_i^{\tilde{k}}}{\partial q^k}\right)_\theta \equiv \frac{\partial c_i^{\tilde{k}}}{\partial q^k} + c_i^k \frac{\partial c_i^{\tilde{k}}}{\partial x_i},$$

$$\left(\frac{\partial c_i^k}{\partial p^\ell}\right)_\theta \equiv \frac{\partial c_i^k}{\partial p^\ell} + f^\ell \frac{\partial c_i^k}{\partial x_i},$$

10a. For example, $\left(\frac{\partial EU}{\partial q^k}\right)_\theta = \frac{\partial EU}{\partial q^k} + \sum_i c_i^k \frac{\partial EU}{\partial x_i} = -\sum_i \alpha_i \pi_i c_i^k + \sum_i c_i^k \alpha_i \pi_i = 0.$

$$\left(\frac{\partial f^{\ell}}{\partial q^{k}}\right)_{\theta} \equiv \frac{\partial f^{\ell}}{\partial q^{k}} + \sum_{i} \pi_i c_i^k \frac{\partial f^{\ell}}{\partial x_i}$$

$$\left(\frac{\partial f^{\tilde{\ell}}}{\partial p^{\ell}}\right)_{\theta} \equiv \frac{\partial f^{\tilde{\ell}}}{\partial p^{\ell}} + \sum_{i} \pi_i f^{\ell} \frac{\partial f^{\tilde{\ell}}}{\partial x_i},$$

$$\left(\frac{\partial \pi_{\tilde{i}}}{\partial q^{k}}\right)_{\theta} \equiv \frac{\partial \pi_{\tilde{i}}}{\partial q^{k}} + \sum_{i} \frac{\partial \pi_{\tilde{i}}}{\partial x_i} c_i^k, \text{ and}$$

$$\left(\frac{\partial \pi_{\tilde{i}}}{\partial p^{\ell}}\right)_{\theta} \equiv \frac{\partial \pi_{\tilde{i}}}{\partial p^{\ell}} + \sum_{i} \frac{\partial \pi_{\tilde{i}}}{\partial x_i} f^{\ell}.$$

After some algebraic manipulation which is given in the Appendix, we obtain[11]

$$\{\sum_i \pi_i (\sum_k (q^k-1)(\frac{\partial c_i^k}{\partial \chi})_\theta + \sum_\ell (p^\ell-1)(\frac{\partial f^\ell}{\partial \chi})_\theta)\}$$

$$-\{\sum_i (\frac{\partial \pi_i}{\partial \chi})_\theta \; (x_i - \sum_k (q^k-1)c_i^k - \sum_\ell (p^\ell-1) f^\ell)\} = 0 \qquad (10)$$

for all χ. Now let $\left(\frac{\partial f}{\partial \chi}\right)$, $\left(\frac{\partial \pi}{\partial \chi}\right)$, and s denote vectors, and

$$(\frac{\partial c}{\partial \chi})_\theta = \begin{pmatrix} (\frac{\partial c_1^1}{\partial \chi})_\theta & (\frac{\partial c_1^2}{\partial \chi})_\theta & \cdots \\ (\frac{\partial c_2^1}{\partial \chi})_\theta & & \\ \vdots & & \end{pmatrix}, \text{ etc.}$$

Then we can write (10) alternatively as

$$\{\pi(\frac{\partial c}{\partial \chi})_\theta (q-1) + \pi(\frac{\partial f}{\partial \chi})_\theta (p-1)\} - \{(\frac{\partial \pi}{\partial \chi})_\theta \; s\} = 0 . \qquad (10')$$

---

11. For situations in which these first-order conditions do not characterize the optimum, see section 2.2.

Equation (10') has two intuitively appealing interpretations. First, it states that at the optimum, the change in net government revenue from a "compensated" unit increase in χ is zero. To obtain an alternative interpretation, take $(\frac{\partial\pi}{\partial\chi})_\theta$ s over to the right-hand side of (10'). The term remaining on the left-hand side gives the marginal excess burden associated with consumer prices diverging from marginal costs, π fixed. Meanwhile, $(\frac{\partial\pi}{\partial\chi})_\theta$ s is the marginal excess burden associated with moral hazard - the increase in expected net social subsidy due to changes in π. Thus, (10') confirms our claim in the introduction that, with the optimal set of taxes, the marginal excess burden associated with consumer prices diverging from marginal cost equals the marginal excess burden associated with moral hazard. The optimal tax system therefore entails some average of Ramsey pricing and pricing to reduce accident probabilities.

When there is no moral hazard, it can be shown that p=q=1 and full insurance are optimal.

We bring these results together in

<u>Proposition 1</u>: With an interior solution, the optimal tax cum insurance structure may be characterized in two ways. First, the change in net government revenue from a "compensated" unit increase in any tax rate must be zero. Second, for each tax rate, the marginal excess burden associated with moral hazard must equal that associated with the wedges between consumer and producer prices.

Since (9) is not necessarily a convex problem, equation (10) may characterize saddlepoints and minima, as well as maxima, and there may be multiple local maxima.

Let us briefly consider competitive attainability and decentralizability of the above optimum. We have assumed that an individual's total insurance purchases are observable, but that his total purchases of goods are not. Competitive decentralizability of the (constrained) optimum, in the strong sense that all agents respond only to parametric prices, with no government intervention, is impossible. The most decentralized institutional mechanism consistent with attainability of the optimum is as follows: The government taxes and subsidizes goods, using this revenue to provide lump-sum subsidies to individuals or insurance companies. Each insurance company, meanwhile, insists on being its clients' exclusive agent for all types of insurance; competition, subject to this government intervention, occurs in both goods and insurance markets. With p and q fixed at their optimal levels, insurance companies face the same (actually the dual) maximization problem as the planner. Depending on the way in which competitive insurance companies adjust their policies out of equilibrium and on the characteristics of the nonconvexities in the maximization problem, the resulting equilibrium might be only locally optimal. In this event, to achieve the global optimum, the government would have to take over the provision of insurance, as well as tax and subsidize goods. We have shown elsewhere (Arnott and Stiglitz [1982c]) that it is generally necessary for efficiency that an individual purchase all types of insurance from a single insurer. If this were not the case, there would be uninternalized externalities. For example, the firm providing fire insurance would neglect in its calculations that the terms of the contract it provided would affect its clients' efforts to avoid automobile accidents and hence the profits of firms providing automobile accident insurance. For similar reasons, it would be inefficient to have an individual obtain insurance against a particular accident from more than one insurer.[12]

The next section examines a few simple situations, so as to provide some insight into the optimal tax structure.

2. Some Simple Cases

2.1 Accident-prevention effort and equipment, and separable, event-independent utility

In this subsection, we consider an economy in which there is a single consumer good, a single accident-prevention good, and a single kind of effort, and in which there are only two possible outcomes: either a fixed-damage accident occurs or it does not.

As in the previous section, we assume that the individual's accident-prevention effort and his purchases of the consumer and accident-prevention goods are unobservable. As a result, insurance is provided against the accident and taxation is linear. We retain the assumption that both goods are produced at constant cost. And we normalize and choose units of measurement such that the producer price of each is unity. Since expected utility is homogeneous

12. We have implicitly assumed that an individual is perfectly informed concerning all insurance contracts offered. Such an assumption is typically justified on the basis of the consumer's shopping around. Here, however, shopping around is inconsistent with exclusivity. It might therefore be more realistic to assume that consumers have imperfect information about the available menu of contracts, in which case market equilibrium would entail a form of monopolistic competition.

of degree zero in consumer prices and income, we are allowed an additional normalization, and assume the consumer price of the consumer good to be unity.[13] Note that, with this normalization, either the accident-prevention good will be taxed and insurance subsidized or vice versa. We employ the same notation as in the previous section, except that, since we now have only one kind of each commodity, we drop superscripts j, k, and ℓ. We adopt the convention that $y_0 > y_1$. Thus, the size of the fixed-damage accident is $y_0 - y_1$. Finally, we assume that utility is separable and outcome-independent, by which we mean that $EU = (1-\pi)u[c_0]+\pi u[c_1]-e$; thus, $u_0 = u[c_0]-e$ and $u_1 = u[c_1]-e$.

Where π is the probability of accident, expected utility is

$$EU = (1-\pi[e,f])u[y_0+x_0-pf] + \pi[e,f]u[y_1+x_1-pf]-e, \tag{11}$$

and the economy's resource constraint is

$$(1-\pi[e,f])x_0+\pi[e,f]x_1-(p-1)f = 0. \tag{12}$$

Applying (10') to this problem gives[14]

$$(p-1)\ \left(\frac{\partial f}{\partial p}\right)_\theta = \left(\frac{\partial \pi}{\partial p}\right)_\theta\ (x_1-x_0). \tag{13a}$$

Since utility is outcome-independent, the individual will receive a larger net payout if the accident occurs than if it does not, i.e. $x_1 > x_0$. Ordinarily, one would expect $\left(\frac{\partial f}{\partial p}\right)_\theta < 0$ and $\left(\frac{\partial \pi}{\partial p}\right)_\theta > 0$; raising the price of the

13. Note that this is different from the normalization employed in the previous section.

14. If utility were not outcome-independent, (13) would not necessarily characterize the optimum. See section 2.2.

accident-prevention good will typically decrease the "compensated" demand for it, and increase the probability of accident. In this case, the accident-prevention good should be subsidized. It seems plausible, however, that there are situations where the accident-prevention good should be taxed; for example $(\frac{\partial f}{\partial p})_\theta < 0$ and $(\frac{\partial \pi}{\partial p})_\theta < 0$ may be a possibility - driving a safer car may make an individual so much more complacent at the wheel that the probability of accident increases. Let us investigate this possibility further.

It can be shown that $(\frac{\partial f}{\partial p})_\theta < 0$. Combining the result with $x_1 > x_0$ gives

$$\text{sign } (p-1) = \text{sign } (-(\frac{\partial \pi}{\partial p})_\theta). \tag{13b}$$

The individual's choices of effort and the quantity of accident-prevention goods to purchase are characterized by

$$\pi_e(u_1-u_0) - 1 = 0 \text{ (where } \pi_e \equiv \frac{\partial \pi}{\partial e}\text{, etc.) and} \tag{14a}$$

$$\pi_f(u_1-u_0) - p((1-\pi)u_0'+\pi u_1') = 0. \tag{14b}$$

From total differentiation of the (14a) and (14b), one obtains that $(\frac{\partial \pi}{\partial p})_\theta$ has the same sign as

$$\frac{\pi_f \pi_{ee}}{\pi_e} - \pi_{ef} + p\pi_e \frac{u_1'-u_0'}{u_1-u_0} \tag{13c}$$

We know that $\pi_e < 0$, $\pi_f < 0$, $u_0 > u_1$, $u_1' > u_0'$ (since utility is separable and outcome-independent and only partial insurance is provided) and $\pi_{ee} > 0$ (from the second-order conditions of the individual's choice problem). Thus, the expression in (13c) and therefore $(\frac{\partial \pi}{\partial p})_\theta$ are negative only if π_{ef} is much greater than the zero, i.e. if the accident prevention good substantially reduces the marginal efficiency of effort (π_e). We say that an accident prevention good for which this holds is very effort-retarding. Note that this property can be consistent with the second-order conditions of the individual's maximization problem.

The sign of $(\pi_f - \frac{\pi_{ef}\pi_e}{\pi_{ee}})$ depends on the "normality" of accident-prevention equipment in the probability-of-accident-function; viz., if a decrease in the probability of accident, holding the "relative price" of effort and accident-prevention equipment fixed, is most efficiently achieved with an increase (decrease) in accident-prevention equipment, then such equipment is normal (inferior). Inferiority is necessary but not sufficient for an accident-prevention good to be very effort-retarding.[14a] Figure 1 portrays a normal and an inferior accident-prevention good.

We present the result of this subsection in:

Proposition 2: In the economy treated in this subsection (two outcomes; separable, outcome-independent utility; a single consumer good; a single accident-prevention good which can be taxed linearly; and accident-prevention effort which is untaxable), the accident-prevention good should be taxed if it is very effort-retarding and subsidized otherwise.[15]

14a. A very effort-retarding good is roughly analogous to a Giffen good.

15. The Proposition can be extended straightforwardly to an arbitrary number of outcomes, accident-prevention goods, consumer goods, and effort types.

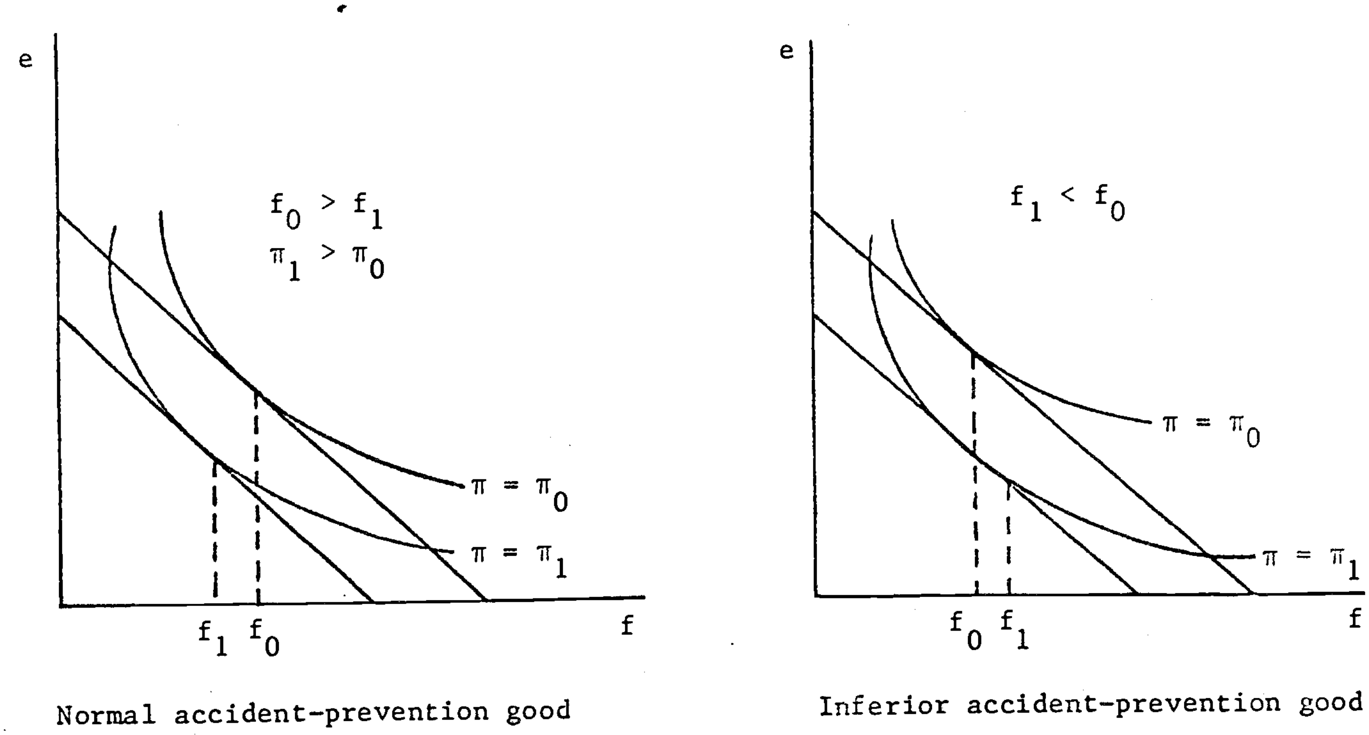

Figure 1: Normal and inferior accident-prevention goods

2.2 Accident-prevention equipment

In this subsection, we treat an economy that differs from that of the previous subsection in two respects: i) there is no accident-prevention effort; and ii) the utility function is not necessarily outcome-independent.

For this case, the symmetric information optimum, in which the planner can directly control the individual's purchases, is the solution to

$$\max_{f,c_1,c_0} \quad (1-\pi[f])u_0[c_0]+\pi[f]u_1[c_1] \tag{15}$$

$$\text{s.t. i) } (1-\pi[f])c_0+\pi[f]c_1+f = (1-\pi[f])y_0+\pi[f]y_1,$$

and is characterized by the resource constraint ((15)i), and

$$u_0' = u_1' = u', \text{ and} \tag{16a}$$

$$\pi_f((u_1-u_0)-u'((c_1-y_1)-(c_0-y_0))) = u'. \tag{16b}$$

We now examine the asymmetric information optimum, in which the planner can only indirectly influence the individual's purchase of f by setting x_0, x_1, and p. For reasons that will shortly be apparent, we first treat the case in which $u_0 > u_1$ for all c_0, c_1 such that $u_0' = u_1'$. We say that in this case, the accident is relatively-utility-decreasing.[16] As in the previous subsection, we set the producer price of the consumption and accident-prevention goods and the consumer price of the consumption good equal to unity.

(13') reduces to

$$p-1 = \pi_f(x_1-x_0). \tag{17}$$

The equation characterizing the individual's choice of how many units of the accident-prevention good to buy is

$$\pi_f(u_1-u_0)-p(\pi u_1'+(1-\pi)u_0') = 0. \tag{18}$$

16. We say that an accident is absolutely-utility-decreasing or simply utility-decreasing if $u_0[\Phi]>u_1[\Phi]$, for all $\Phi>0$ and marginal-utility-decreasing if $u_0'[\Phi]>u_1'[\Phi]$ for all $\Phi>0$.

The assumption that an accident is relatively-utility-decreasing implies, for instance, that an individual who has a limb severed and has it replaced by an artificial limb, and is compensated to the point where his marginal utility of income is the same as if his limb had not been severed, would still have preferred not to have the limb severed.

The resource constraint is

$$(p-1)f = (1-\pi[f])x_0+\pi[f]x_1. \tag{19}$$

And if one solves for the optimal values of x_0 and x_1, one obtains

$$u_0' = u_1'. \tag{20}$$

Let * denote values of variables at the symmetric information optimum. It may easily be checked that if the government sets $p^* = \frac{(u_1^*-u_0^*)\pi_f^*}{(u')^*} = \pi_f^*(x_1^*-x_0^*) + 1$ (from (17) and (18)), $x_0=x_0^*$ and $x_1=x_1^*$, then the individual will choose f^*, and the symmetric information optimum will be attained. This optimum may also be attained if the government sets only $p^*-1 = \pi_f^*(x_1^*-x_0^*)$, distributes the revenue (typically negative) collected in lump-sum fashion to insurance companies, and allows these companies to choose x_0 and x_1.[17]

---

17. An insurance company then faces the (dual of the) problem

$$\max_{x_0,x_1} \; (1-\pi)u_0+\pi u_1 \quad \text{s.t.} \quad \text{i) } f = f[p^*,x_0,x_1]$$
$$\text{ii) } (p^*-1)f = (1-\pi)x_0+\pi x_1.$$

Where β is the Lagrange multiplier on ii), the first-order conditions are:

$$x_0: \quad (1-\pi)u_0'-\lambda((\pi_f(x_1-x_0)-(p^*-1))\frac{\partial f}{\partial x_0} + (1-\pi)) = 0$$

$$x_1: \quad \pi u_1'-\lambda((\pi_f(x_1-x_0)-(p^*-1))\frac{\partial f}{\partial x_1} + \pi) = 0.$$

When $p^*-1=\pi_f^*(x_1^*-x_0^*)$, this pair of equations is solved by $x_0=x_0^*$ and $x_1=x_1^*$. Because of possible nonconvexities, there may be other (x_0,x_1) pairs which solve the pair of equations, in which case the economy could settle down at an inferior local optimum.

Upon reflection, this result is not surprising. Competitive equilibrium without government intervention fails to achieve the symmetric information optimum only because individuals, if they were provided with full insurance, would purchase an inefficient quantity of accident-prevention goods. This can be remedied directly by subsidizing or taxing these goods. If, with full insurance, the individual has an incentive both to purchase and to use the accident-prevention goods, then the government can achieve the symmetric information optimum by providing full insurance and adjusting the price of the accident-prevention goods so that the individual purchases the optimal amount. Accident-prevention effort, however, is different. Since it is untaxable, one can stimulate it only indirectly, by taxing substitutes and subsidizing complements, which causes an efficiency loss relative to the symmetric information optimum.

When the accident is relatively-utility-decreasing, the accident-prevention good will typically be subsidized. There is one circumstance in which the accident-prevention good should be taxed. If the accident is so relatively-utility-decreasing that when $u_0'=u_1'$, $x_0 > x_1$ then from (17) it follows that $p^* > 1$. Loosely, individuals are apt to be excessively careful to prevent accidents that severely reduce the ability to experience pleasure.

Now let us turn to the case of <u>relatively-utility-increasing</u> accidents; i.e. $u_0 < u_1$ for all c_0, c_1 such that $u_0' = u_1'$. In this case, if full insurance is provided and if the government sets $p = \frac{(u_1^*-u_0^*)\pi_f^*}{(u')^*}$, the price of the accident-prevention good is negative. The individual would "buy" as many units of the good as possible, since doing so, with a negative price, would increase his income. But he would not use them since his utility is higher when an accident occurs. Thus, with relatively-utility-increasing accidents, the symmetric information optimum is not attainable under conditions of asymmetric information. The symmetric information optimum could be attained if individuals were compelled to use all the accident-prevention goods they purchased.[18] The tax system, however, while a form of compulsion, cannot by itself force individuals to use what they purchase. To give individuals the incentive to use accident-prevention goods, it is necessary that $u_0 > u_1$. With relatively-utility-decreasing accidents, the asymmetric information optimum is the solution to

$$\max_{f,c_1,c_0} \quad (1-\pi[f])u_0[c_0]+\pi[f]u_1[c_1]$$

$$\begin{aligned} \text{s.t.} \quad \text{i)} \quad & (1-\pi[f])c_0+\pi[f]c_1+f \\ & = (1-\pi[f])y_0+\pi[f]y_1, \\ \text{ii)} \quad & u_0 > u_1, \end{aligned} \tag{21}$$

where constraint ii), which was absent from the symmetric information optimum problem (see (15)), is binding. The asymmetric information optimum can be decentralized by having the government set the price of the accident-prevention good arbitrarily low but positive, financing the subsidy from a lump-sum tax on firms, and allowing firms to choose what exclusive contract to offer. Where ^ denotes the value of a variable at the asymmetric information optimum, firms will choose x_0 and x_1 so that with the $\hat{p}$ set by the government, the individual will (via (18)) choose $f=\hat{f}$ and the individual's expected utility is maximized.

---

18. With enforcement costs, the most efficient method of compuslsion would be to apply an infinitely large fine for non-compliance and to inspect for non-compliance with an infinitesmal but strictly positive probability.

If utility is outcome-independent, so that $u_0 = u_1$ for all c_0, c_1 such that $u_0' = u_1'$, a different problem arises. If full insurance is provided and if the government sets $p = \frac{(u_1^* - u_0^*)\pi_f^*}{(u')^*}$, the price of the accident-prevention good is zero, the individual's purchases of the good are indeterminate (from (18)), $\lim_{p \to 0} \frac{df}{dp} = -\infty$). In this case, by setting the price of the accident-prevention good arbitrarily low but positive, the planner can move the economy arbitrarily close to the symmetric information optimum but not all the way to it. Firms will then have an incentive to provide just short of full insurance, choosing x_1 and x_0 so that individuals will choose arbitrarily close to the optimal quantity of accident-prevention goods.[18a]

We summarize the above results in

Proposition 3: In an economy with two outcomes, a single accident-prevention good which can be taxed linearly, no accident-prevention effort, and a single consumption good:

i) if the accident is relatively-utility-decreasing, the symmetric information optimum can be achieved by the government setting p optimally, and insurance companies choosing x_0 and x_1.

ii) if utility is outcome-independent, the symmetric information optimum can almost (arbitrarily closely) be achieved by the government providing the accident-prevention good at an arbitrarily low but positive price, allowing insurance companies to choose what contract to offer, and financing the subsidy via a lump-sum tax.

iii) if the accident is strongly relatively-utility-increasing, the symmetric information optimum cannot be achieved. The asymmetric information is achieved by the government providing the accident-prevention good at an arbitrarily low price, allowing insurance companies to choose what contract to offer, and financing the subsidy via a lump-sum tax.

18a. This case is examined, in a specific context, in Diamond and Mirrlees [1978].

The Proposition extends to the case where there is more than one consumer good. If the accident is relatively-utility-decreasing, the symmetric information optimum can be achieved by taxing only the accident-prevention good at the optimal rate; no tax should be applied to the consumer goods. If, however, the accident is relatively-utility-decreasing, then differential consumer good taxation is desirable. The Proposition also extends straightforwardly to the case where there are several accident-prevention goods. When the accident has multiple outcomes, the analog to whether the accident is relatively-utility-decreasing is whether the individual, when he is provided with full insurance and charged an arbitrarily small positive price for the accident-prevention good, will purchase it. When the accident has multiple outcomes and there are several accident-prevention goods, the analysis becomes more difficult since there may be multiple local optima; as fuller insurance is provided, the individual will discontinue the use of the first one accident-prevention good, then another, and for each set of accident-prevention goods employed, there will be a local optimum.

The complication pointed out in this section, that the interior first-order conditions of the two-tier maximization problem need not characterize the optimum, was first pointed out in Mirrlees [1975] and applies to all the optimal tax problems treated in this paper. To simplify, we ignored the complication in section 1, circumvented it by assumption in section 2.1, and shall assume it away for the rest of the paper.

In this subsection we framed the discussion in terms of the taxation and subsidization of accident-prevention goods. In so doing, we implicitly treated the other side of the resource constraint (19) which pertains to insurance. When the net revenue raised from taxation is positive (negative), insurance companies should be subsidized (taxed) in lump-sum fashion.

2.3 Accident-prevention effort and more than one consumer good

The economy to be treated in this subsection is like that of subsection 2.1 except that instead of one consumer good and one accident-prevention good there are two consumer goods. To simplify the analysis, we treat a special case in which one of the goods (good a) is separable from effort in the utility function, while the other (good b) is not. Having set all producer prices equal to one, we choose good a as the numéraire. Utility in state i is then

$$u_i[c_i^a]+v_i[c_i^b,e],$$

or, since $y_i+x_i = c_i^a+q^b c_i^b$ from an individual's budget constraint,

$$u_i[y_i+x_i-q^b c_i^b]+v_i[c_i^b,e].$$

Expected utility is therefore

$$\begin{aligned} EU = {} & (1-\pi[e])(u_0[y_0+x_0-q^b c_0^b]+v_0[c_0^b,e]) \\ & +\pi[e](u_1[y_1+x_1-q^b c_1^b]+v_1[c_1^b,e]), \end{aligned} \tag{22}$$

and the economy's resource constraint is

$$(1-\pi[e])x_0+\pi[e]x_1-(q^b-1)((1-\pi[e])c_0^b+\pi[e]c_1^b) = 0. \tag{23}$$

In this case, (10) becomes

$$q^b-1 = \frac{\sum_{i=0}^{1} \left(\frac{\partial \pi_i}{\partial q^b}\right)_\theta (x_i-(q^b-1)c_i^b)}{\sum_{i=0}^{1} \pi_i \left(\frac{\partial c_i^b}{\partial q^b}\right)_\theta}, \tag{24}$$

where $\pi_0 = 1-\pi$, and $\pi_1 = \pi$. Upon rearrangement and simplification, (24) may be rewritten as

$$q^b - 1 = \frac{\left(\frac{\partial\pi}{\partial q^b}\right)_\theta (x_1 - x_0)}{\left(\frac{\partial\pi}{\partial q^b}\right)_\theta (c_1^b - c_0^b) + \pi\left(\frac{\partial c_1^b}{\partial q^b}\right)_\theta + (1-\pi)\left(\frac{\partial c_0^b}{\partial q^b}\right)_\theta} . \tag{24'}$$

Letting $\tilde{c}^b$ denote the expected value of c^b, i.e.

$$\tilde{c}^b = (1-\pi)c_0^b + \pi c_1^b,$$

(24') may be further simplified to give

$$q^b - 1 = \frac{\left(\frac{\partial\pi}{\partial q^b}\right)_\theta (x_1 - x_0)}{\left(\frac{\partial\tilde{c}^b}{\partial q^b}\right)_\theta} , \tag{24"}$$

which states that q^b should be set so that the "compensated" change in tax revenue from a unit increase in q^b equals the "compensated" change in insurance payouts. Unfortunately, while this interpretation is clear, the implications of (24") concerning the determinants of the sign and magnitude of q^b-1 are far from straightforward.

Table 1 categorizes the circumstances under which good b should be taxed or subsidized, the entries in the table indicating the sign of q^b-1. To make the analysis more concrete, we treat the example of a painful car accident in which the issue is whether pain-killers should be taxed or subsidized. Suppose, for example that: i) effort and pain-killers are substitutes

Table 1

	$x_1 > x_0$		$x_1 < x_0$	
	$(\frac{\partial\pi}{\partial q^b})_\theta > 0$	$(\frac{\partial\pi}{\partial q^b})_\theta < 0$	$(\frac{\partial\pi}{\partial q^b})_\theta > 0$	$(\frac{\partial\pi}{\partial q^b})_\theta < 0$
$(\frac{\partial\tilde{c}^b}{\partial q^b})_\theta > 0$	+	-	-	+
$(\frac{\partial\tilde{c}^b}{\partial q^b})_\theta < 0$	-	+	+	-

in the sense that $(\frac{\partial\pi}{\partial q^b})_\theta = \pi_e(\frac{\partial e}{\partial q^b})_\theta < 0$; ii) the "compensated" expected demand for pain-killers is inversely related to their price, $(\frac{\partial\tilde{c}^b}{\partial q^b})_\theta < 0$ (one scenario consistent with this is $(\frac{\partial c_1^b}{\partial q^b})_\theta < 0$, $(\frac{\partial c_0^b}{\partial q^b})_\theta < 0$, and $c_1^b > c_0^b$ -- the "compensated" taxation of pain-killers reduces their consumption in both events, and more pain-killers are consumed in the event of accident; these conditions combined with $(\frac{\partial\pi}{\partial q^b})_\theta < 0$ imply $(\frac{\partial\tilde{c}^b}{\partial q^b})_\theta < 0$); and iii) the accident is not so severe as to destroy the individual's capacity for pleasure, so that $x_1 > x_0$. In this case ($x_1 > x_0$, $(\frac{\partial\pi}{\partial q^b})_\theta < 0$, and $(\frac{\partial\tilde{c}^b}{\partial q^b})_\theta < 0$) pain-killers should be taxed.

The signs of $x_1 - x_0$, $(\frac{\partial\pi}{\partial q^b})_\theta$, $\frac{\partial\tilde{c}^b}{\partial q^b}$ are not primitive characteristics of the utility or probability-of-accident functions. To convert (24") into a more primitive form, we employ the expected utility function (22), in which case (24") may be written as

$$q^b - 1 = \frac{-\pi_e(u_0'AE + u_1'BD)(x_1 - x_0)}{-\pi_e(u_0'AE+u_1'BD)(c_1^b - c_0^b) + \pi(u_1'(FB-A^2)+u_0'AD) + (1-\pi)(u_0'(FE-D^2)+u_1'AD)}, \qquad (24''')$$

where $A = (1-\pi)\dfrac{\partial^2 v_0}{\partial e \partial c_0^b}$ $D = \pi \dfrac{\partial^2 v_1}{\partial e \partial c_1^b}$

$B = (q^b)^2 u_0'' + \dfrac{\partial^2 v_0}{\partial (c_0^b)^2}$ $E = (q^b)^2 u_1'' + \dfrac{\partial^2 v_1}{\partial (c_1^b)^2}$

$$F = -\pi_{ee}((u_0+v_0)-(u_1+v_1))-2\pi_e(\frac{\partial v_0}{\partial e} - \frac{\partial v_1}{\partial e})+(1-\pi)\frac{\partial^2 v_0}{\partial e^2} + \pi\frac{\partial^2 v_1}{\partial e^2} .$$

This equation has one interesting implication: when effort is separable from good b as well as good a in the utility function, good b should be neither taxed nor subsidized. Thus, the substitutability or complementarity between effort and good b is an essential determinant of how good b should be taxed. Unfortunately, it appears that other neat results are not to be had.[19] Thus, with moral hazard, the determinants of optimal tax rates in realistic situations will be complex, and the determination of these rates will require numerical solution. Nevertheless, the insight that the optimal tax systems reflects two factors - Ramsey-pricing to reduce the deadweight loss associated with prices not reflecting production costs, and pricing to reduce accident probabilities and hence the deadweight loss associated with moral hazard - is useful.

In interpreting the models of this subsection we have treated outcome 0 as "no accident" and outcome 1 as "accident". We could just as well, however, have assumed that an accident always occurs, and that the amount of effort affects the damage caused by the accident. In this case, instead of

---

19. The difficulty stems from the fact that (24) contains terms that reflect global rather than local properties of the utility function.

speaking of accident-prevention goods and effort, one would want to speak of damage-reducing goods and effort. Fire extinguishers and seat belts are two obvious examples of damage-reducing goods. Thus, one wants to subsidize fire extinguishers as long as doing so does not make the individual much more careless.

To say that a good should be taxed is equivalent to saying that the shadow price of the good exceeds its market price. Put this way, however, this result points to an interesting implication of our analysis: In an economy with moral hazard, pecuniary externalities "matter";[20] by "matter" we mean that government intervention is justified to internalize the externality. In a classic competitive economy, pecuniary externalities do not matter since the social benefit associated with the marginal unit equals the social cost. The purchase of the marginal unit causes prices to change throughout the economy, which induces a string of marginal reallocations. But since the social benefit of each of these marginal reallocations equals the social cost, the pecuniary externality creates no deadweight loss. In an economy with moral hazard, however, since the social benefit of the marginal reallocations is not in general equal to the corresponding social cost, pecuniary externalities alter the aggregate deadweight loss in the economy and thus matter. This result implies that the welfare properties of economies with moral hazard are markedly different from those of Arrow-Debreu economies. For one, the Invisible Hand Theorem usually fails to hold in economies with moral hazard; other differences are examined at some length in Arnott and

20. This statement needs to be qualified. When the allocation in an economy with moral hazard coincides with the corresponding symmetric information optimum, pecuniary externalities do not matter.

Stiglitz [1983c].

In the previous analysis, we assumed that an individual's total purchases of consumption and accident-prevention goods were unobservable. As a result, firms could not make (and enforce) the terms of the insurance they offer contingent on individuals' total purchases of these goods. Furthermore, competitive firms had no choice but to price goods and insurance at cost; any firm which tried to do what the tax system does, taxing some goods and subsidizing others, would make a loss since individuals would purchase the subsidized but not the taxed goods.[21] In addition, the government could apply only linear taxation. What happens if, instead, individuals' total purchases of consumption and accident-prevention goods are observable? Firms could then sell total packages to individuals, tieing their purchases of insurance, accident-prevention goods, and consumption goods, in which case government intervention would be unnecessary since there is no allocation governments could achieve which firms could not also. If this possibility is ruled out, then observability of total purchases allows the government to apply non-linear taxation. How does this alter the previous analysis? First, the problem identified in subsection 2.2, that taxation cannot by itself coerce an individual to use the accident-prevention goods he purchases (nor, for that matter, can packaging by firms), arises whether taxation is linear or non-linear. Second, linear taxation along with lump-sum transfers, which is what we treated, can be used to sustain any allocation that is sustainable using general non-linear taxation. And third, non-

21. Firms might, however, be able to sell bundles in which the ratios of the various goods in each bundle are those characterizing the asymmetric information optimum.

linear taxation can be designed to circumvent the possible non-uniqueness of competitive equilibrium (the tax system fixed) while linear taxation with lump-sum transfers cannot. Thus, the use of non-linear taxation by governments or packaging by firms, where possible, may improve social welfare, but does not eliminate the deadweight loss caused by the unobservability of effort.

3. Policy Applications

We start out with a fairly lengthy discussion of one application, optimal taxation vis-à-vis automobile accidents, and then touch on a variety of others.

In this paper, we have assumed that individuals are identical in order to circumvent problems associated with adverse selection. In fact, of course, almost all insurance is characterized by both moral hazard and adverse selection. The presence of adverse selection will generally affect the optimal tax structure and in interesting ways. The taxation of commodities will alter the self-selection constraints and therefore the deadweight loss associated with adverse selection.[22] The reader should keep in mind

---

22. This is analogous to commodity taxation in an economy with an income tax. See Atkinson and Stiglitz [1976] and Stiglitz [1982b].
    Consider, for instance, a simple farming economy in which everyone owns his own plot of land. All plots are identical, output is either zero or one, depending on farming ability and the state of nature. The insurer can observe a farmer's output, but neither his ability nor the state of nature he experiences. There are high and low ability farmers and farmers know their own ability. The insurer will offer two packages of insurance, one low-price, low-quantity for high-ability farmers, the other higher-price, higher-quantity for low-ability farmers. High ability farmers would like to purchase more insurance, but if the insurer offers a low-price policy with more insurance, low ability farmers will

that the discussion which follows, by treating an economy of identical individuals, ignores these considerations.

3.1 Automobile accidents

Because of moral hazard, the provision of automobile accident insurance will increase both the probability of accident and the average size of damage, conditional on an accident occurring. People will tend to drive too much and too carelessly, and to be insufficiently careful in the maintenance of their cars (all of which cause the probability of accident to be too high). And they will tend to drive too fast, to purchase cars that are insufficiently safe and excessively expensive, to take inadequate precaution to be alert, and to neglect wearing seat belts when they should (all of which cause average damage, conditional on an accident occurring, to be too high).

We have seen that the structure of optimal taxation is complicated by interdependencies between and within the probability-of-outcome function and the outcome-contingent utility functions. In the discussion which follows, we ignore these complications. Then the optimal tax structure will involve: taxing gasoline (to encourage people to drive less), taxing alcohol (to encourage people to drive more carefully), imposing penalties on careless driving, subsidizing maintenance, taxing cars, with a higher tax on less safe and

purchase it in preference to the policy designed for them. Thus, the presence of low-ability farmers constrains the set of (separating) contracts that can be offered high-ability farmers. This adverse selection externality entails a deadweight loss. Now suppose that the government taxes those goods whose consumption makes people less risk-averse (alcohol, cigarettes?). This will cause farmers to become more risk-averse and will alter the pair of equilibrium insurance contracts, thereby altering the deadweight loss associated with the adverse selection externality.

more expensive cars, subsidizing coffee breaks (to encourage people to be more alert when driving),[23] and imposing penalties for not wearing seat belts. One might also want to tax complements and subsidize substitutes. This could involve subsidizing alternative modes of transport. Accidents on these alternative modes are insured against, and this insurance too is characterized by moral hazard. But one suspects that, per passenger-mile, the deadweight loss associated with moral hazard is higher for private cars than for other forms of transportation.

We now turn to some other issues. First, we have not yet stated what normalization we have made; we have adopted the most intuitive, and normalized on those goods the level of consumption of which has no effect on the probability of accident. Second, if an insurance company can effectively regulate its clients' total purchases of insurance,[24] the government does not necessarily want to tax automobile insurance. If it does, it reduces the probability of accident, but it also exposes the individual to more risk and there is a welfare cost associated with this. The government through its taxation wants to increase the level of precaution at every level of insurance, rather than alter the level of insurance *per se*. With the normalization we have made, whether the government taxes or subsidizes insurance is a residual. It imposes the appropriate taxes and subsidies on goods; if doing so raises negative revenue, it imposes lump-sum taxes on insurance companies and otherwise it pro-

---

23. This has been done on some highways in Pennsylvania.

24. Insurance companies do this by stating that they will pay a client in the event of an accident only if he claims from no other company.

vides them with lump-sum subsidies. Third, there is an important qualification to this line of reasoning. We have treated automobile accidents as if they involve only the insured. But many car accidents involve at least two cars. A private insurance company, in deciding on the contract to offer a client, will neglect to consider that by offering its client more insurance and hence increasing his probability of accident, it also increases the accident probability of other firms' clients and hence these firms' profitability. This reciprocal negative externality, which results in over-insurance, would be internalized if there were only one insurance firm providing insurance against car accidents. This provides an argument in favor of the socialization of automobile insurance, or alternatively a higher tax rate on insurance. Fourth, any outcome the government can achieve via the taxation (not necessarily linear) of body-work or, more generally, accident damage repair, it can also achieve via the differential taxation of payouts. Both are essentially outcome-contingent taxes. Thus, arguments in favor of the taxation of accident damage repair are neither stronger nor weaker than arguments in favor of the taxation of insurance.[25] Fifth, it has been noted elsewhere in the literature that experience-rating is a means of reducing the social loss attributable to moral hazard.

3.2 Other policy applications

We shall not attempt an exhaustive survey of policy applications,

25. There is another important moral hazard phenomenon in this context. Both the accident victim and the repair company have an incentive to inflate the repair bill (splitting the excess of claimed over actual repair costs).

but rather choose a few, each of which illustrates a point of interest.

i) social security

Diamond and Mirrlees [1978] consider an economy in which individuals are *ex ante* identical, but some, as they age, develop disabilities which make early retirement more attractive. The problem faced by the government is how to design social security when it cannot ascertain the extent to which a particular individual is disabled. The more attractive are provisions for early retirement, the more insurance is provided the disabled, but at the same time, the more attractive it becomes for able individuals to retire "too early".

The probability distribution of retirement ages is analogous to the probability distribution of accident damages in our model (with earlier retirement corresponding to higher damage). Consumption in different periods is analogous to different goods in our model, which are, at the same time, consumer goods and accident-prevention goods. Consumption in earlier periods is damage-reducing since it causes the individual to save less, which makes retiring less attractive; and similarly, consumption in later periods is damage-increasing. Relative to the symmetric information optimum, the provision of social security causes too many people to retire at too early an age. The deadweight loss associated with this can be reduced by subsidizing consumption when young and taxing consumption when old, or more straightforwardly by taxing savings. This discourages saving, which makes an individual less inclined to retire. Depending on how increasing disability affects the utility function, since there is no analog to accident-prevention effort, the symmetric information optimum may or may not be achievable.

One could enrich the D-M model to allow for different types of goods. In such an economy one might want to subsidize those goods that encourage later retirement of the able, and tax those that encourage earlier retirement. One might therefore want to subsidize companies' attempts to improve working conditions, and tax goods that are complementary to retirement interests. To reduce the number of early retirees, one might also want to subsidize medical expenses.

ii) <u>informal insurance</u>

In Arnott and Stiglitz [1982c], we argued that the provision of informal insurance - mutual assistance among friends and family members in times of need - that is supplementary to market insurance may result in inefficiency, relative to the case where all insurance is provided through the market. For example, unemployment is likely to be a less unhappy state if an individual is helped out by family and friends; such support dampens the individual's incentive to search for a new job. The amount of informal insurance provided can be influenced by the government through the taxation of complements and subsidization of substitutes, through regulation, or by direct prohibition. In many contexts, however, the costs of such government intervention, on both economic and moral grounds, are likely to exceed the benefits. In these cases there is potential but not actual market failure. The "market" (i.e., spontaneous economic activity, including markets and spontaneous non-market institutions) is efficient in the sense that it cannot be improved upon by government intervention, though not in the sense employed by the naive functionalists who argue that where markets are deficient, the non-market institutions which arise will correct all potential "market" failures.

iii) unemployment insurance

The optimal provision of unemployment insurance trades off the deadweight losses associated with search and matching externalities, the incomplete provision of insurance, and moral hazard. The deadweight loss associated with this form of moral hazard (on the one hand, the provision of unemployment insurance encourages quitting, decreases the costs of poor job performance resulting in dismissal, reduces search intensity, and increases the reservation wage; on the other, it may give rise to better matching) can be reduced by taxing goods that makes unemployment more enjoyable, and subsidizing those goods that make search and employment more attractive. To make employment more pleasant, the government could subsidize firms' attempts to improve working conditions. And to make search more pleasant, the government could subsidize the operation of employment agencies. One could also argue that beer should be taxed if it is found that the unemployed drink it in disproportionately large quantities; one suspects, however, that this too is a case of potential rather than actual market failure.

iv) medical insurance

Here the policy prescription is obvious. The government should tax those goods and activities that are conducive to ill health, and subsidize those associated with good health. Some such measures are already in effect in most developed countries. The taxation of alcohol and cigarettes is an example, though it is doubtful whether the principal aim of these taxes was the reduction in the deadweight loss from the moral hazard associated with the provision of medical insurance.

It is not clear whether moral hazard considerations support the taxation or subsidization of non-preventive medical treatment (this is analogous to body work in the case of automobile accidents).

v) <u>sharecropping and other principal-agent problems</u>

Most economists have a good understanding of what moral hazard is, but it is our feeling that the importance and pervasiveness of moral hazard tends to be underestimated. For almost all risks, the probability of accident or the size of damage conditional on the accident occurring is affected by the tim's actions, which are only imperfectly observable. As a result, insurance markets against virtually every risk will be incomplete or absent. And when an insurance market is incomplete, institutions will develop that provide non-market insurance, perhaps explicit, perhaps implicit, perhaps formal, perhaps informal. Thus, moral hazard is present not only in insurance markets, but in these numerous social institutions as well.

We have given examples of market insurance, informal insurance, and government-provided insurance. There is another large class of institutions which provide (implicit) insurance--principal-agent relationships. An example which has been extensively discussed is the landlord-tenant contract (Stiglitz [1974], e.g.). The basic problem is that the provision of implicit insurance to the agricultural laborer discourages effort. The deadweight loss associated with this can be reduced by subsidizing complements to effort. Thus, if a happy worker is a good worker, it might benefit the landlord to construct entertainment facilities. And if there is a problem with workers spending too much time at the local café, he should raise its prices. Since indebtedness and purchases of inputs may both affect effort, they too should be taxed or subsidized. The attempt to internalize these "externalities" leads to interlinkage of land, labor, credit, and commodity markets (Braverman and Stiglitz [1982] and Mitra [1982]). Other familiar principal-agent relationships are those between employer and employee (e.g. Arnott and Stiglitz [1981], and Arnott, Hosios, and Stiglitz [1983]), borrower and lender (e.g. Stiglitz and Weiss [1981]), and physician and patient (e.g. Arrow [1965]).

3.3 Some comments

The reader will no doubt have asked himself: i) how significant is the deadweight loss associated with the various forms of moral hazard we have discussed; ii) what is the approximate magnitude of taxes and subsidies based on these considerations; iii) how are the optimal tax and subsidy rates to be computed; iv) and may not the administrative and other costs of imposing a complex system of taxes and subsidies to reduce the deadweight loss associated with moral hazard exceed the benefits? The simple answer is that neither the theoretical nor the empirical work that would be required to answer these questions has been done. Insurance actuaries probably have a good idea of how responsive the probability of a particular accident is to the parameters of the insurance contract. But this datum reflects not only moral hazard, but also adverse selection; as the contract is modified, not only may each client alter his accident-prevention behavior, but also the client population may change. Since the welfare properties of economies with both moral hazard and adverse selection have not yet been investigated, we do not know what actuarial data would be necessary to compute the deadweight loss associated with moral hazard *cum* adverse selection nor how to undertake the computation.

Our intuition is that the deadweight loss due to moral hazard is important enough in some contexts, most notably health, theft, fire, automobile accidents, social security, and unemployment insurance, to warrant corrective taxation.

4. Concluding Comments

This paper complements three others we have written dealing with the pure theory of moral hazard. While each treated different topics, there was a common theme: The properties of economies with asymmetric information are very different from those of the classic, competitive, Arrow-Debreu economy. Within this theme, there were several sub-themes:

1. In the analysis of economies with asymmetric information, one must guard against making assumptions that are not primitive, but are rather characteristics of the classic, competitive economy. For example, in Arnott and Stiglitz [1982a] we showed that moral hazard may give rise to intrinsic non-convexities in the relevant parameter space; it would be a mistake to blindly assume convexity. And in Arnott and Stiglitz [1982b], we demonstrated that the form of the insurance contract depends on who has what information - contracts may specify not only price, but also the quantity of insurance to be purchased, and may furthermore be characterized by random premia and payouts; thus, it would be a mistake to assume the form of the insurance contract.
2. The basic theorems of classic welfare economics no longer hold in economies with asymmetric information. Efficiency is not in general consistent with decentralized, atomistic behavior, and may entail large firms (even with constant returns in the production technology) and/or extensive government intervention. One of the sources of potential market failure, that shadow prices generally differ from market prices, was the focus of this paper. Other

sources of market failure are discussed in Arnott and Stiglitz [1983].

3. Relatedly, considerable care must be taken to ensure that the appropriate efficiency concept is employed. Efficiency should be defined not only contingent on the information acquisition technology, or more restrictively on what information is and is not available to whom, but also treating market structure as endogenous and taking into account the technology of government intervention. The latter are particularly difficult. In the paper, we pointed out potential market failures on the assumption that the private sector price goods at production cost. This assumption is appropriate if, as we assumed, the individual's total purchases of each commodity are unobservable. But if all an individuals' purchases could be costlessly monitored, there is nothing in our model per se to rule out the emergence of a super-firm which produces all goods and sells all insurance. If the market were contestable, this super-firm would act efficiently; there is nothing the goverment could do which it could not. This result in turn is predicated on the assumption that individuals are perfectly informed concerning all contracts and prices. If this rather unpalatable assumption were relaxed, one would expect a monopolistically competitive equilibrium to emerge.

The above discussion points to the importance of giving careful consideration to the information available to firms and to how this affects the feasible scope of their actions, and to relating government intervention to those characteristics of governments, notably their coercive powers, that permit them a broader range of actions than firms.

A final problem is that the costs of government intervention should include not only information acquisition and administration costs, but also the efficiency losses associated with internal incentive problems. Efficiency should be defined contingent on all these, and other, considerations.

In this paper, we have made a simple yet, we believe, important point. In particular, in competitive equilibrium (with no government intervention) in an economy with asymmetric information, shadow prices will generally deviate from market prices. There is a welfare loss associated with the informational asymmetries, the size of which depends on individuals' consumption patterns. If increased consumption of a good reduces this welfare loss, the good should be subsidized whenever the reduction in welfare loss exceeds the costs of government intervention.

We investigated a particular case of this general proposition.[26] We considered an economy with identical individuals, in which the presence of asymmetric information gave rise to moral hazard. And we considered the determinants of the optimal tax structure in such an economy. Broadly speaking, our records accorded with intuition. The moral hazard arising from the provision of insurance causes people to self-protect too little, so those goods whose consumption encourage an individual to self-protect more/less should be subsidized/taxed. Thus, fire extinguishers should be subsidized if having a fire extinguisher reduces accident damage, while alcohol should be taxed if its consumption causes individuals to drive more recklessly. Some of our results, however, are not immediately obvious. First, we demonstrated that there are cir-

26. The general proposition itself is examined in greater generality in Greenwald and Stiglitz [1982].

cumstances, albeit improbable, in which accident-prevention goods should be taxed. Second, we showed that, as in the optimal commodity tax problem, the determination of the optimal tax structure is far from straightforward. And third, we argued that if the probability of accident is dependent on an individual's (anonymous) purchases of accident-prevention goods and not at all on his accident-prevention effort, that optimal taxation may lead the economy to the symmetric information optimum.

There are two obviously worthwhile extensions to our analysis. We treated information as if it were either costless or infinitely costly to acquire; as a result, something was either perfectly observable or unobservable. One would like to treat explicitly the costs of acquiring insurance-relevant information,[27] particularly since one could then determine the circumstances under which such information acquisition should be subsidized. The analysis should also be extended to treat adverse selection and moral hazard simultaneously. Not only would this alter the optimal tax structure in interesting ways, but also the development of such a model is a necessary condition for both sound, qualitative policy advice, and for the empirical estimation necessary to compute optimal tax rates.[28]

---

27. The collection of information has been studied by Holmstrom [1979] and Shavell [1979a] but for simple economies with moral hazard in which competitive equilibrium is constrained efficient.
    The transmission of information, in particular whether firms have an incentive to share information on their common clients is the focus of current work by Hellwig [1982].

28. We assumed that it is prohibitively costly to monitor an individual's purchases of commodities. With this assumption, one cannot improve on linear taxation. There are some commodities, however, for which individual consumption can be monitored. It would be interesting to investigate the optimal non-linear taxation (or linear taxation <u>cum</u> rationing) of these commodities.

The belief that an unregulated market (and spontaneous non-market institutions) would provide many forms of insurance in an inefficient manner is widespread, at least among policy makers and the lay public, and has given rise to extensive government intervention vis-à-vis insurance. This paper and Arnott and Stiglitz [1983] have provided a theoretical basis for this belief, and have indicated at least some of the factors that should be considered in determining appropriate corrective action by the government.

Appendix

Derivation of Equation (10)

Rewrite (9) as

$$\max_{\{q^k\},\{p^\ell\},\{x_i\}} \mathcal{L} = \sum_i \pi_i V_i(q,p,y+x) - \lambda(\sum_i \pi_i (x_i - \sum_k (q^k-1)c_i^k - \sum_\ell (p^\ell - 1) f^\ell). \qquad \text{(i)}$$

$$\frac{\partial \mathcal{L}}{\partial x_m} = \sum_i \pi_i \frac{\partial V_i}{\partial x_m} + \sum_i \frac{\partial \pi_i}{\partial x_m} V_i - \lambda\pi_m + \lambda\pi_m(\sum_k (q^k-1)\frac{\partial c_m^k}{\partial x_m} + \sum_\ell (p^\ell-1)\frac{\partial f^\ell}{\partial x_m}) - \lambda(\sum_i \frac{\partial \pi_i}{\partial x_m}(x_i - \sum_k (q^k-1)c_i^k - \sum_\ell (p^\ell-1)f^\ell)) = 0. \qquad \text{(ii)}$$

From the individual's maximization problem, $\sum_i \pi_i \frac{\partial V_i}{\partial x_m} + \sum_i \frac{\partial \pi_i}{\partial x_m} V_i = \pi_m \alpha_m$. Thus,

$$\sum_m c_m^j \frac{\partial \mathcal{L}}{\partial x_m} = \sum_m \pi_m c_m^j \alpha_m - \lambda(\sum_m \pi_m c_m^j) + \lambda(\sum_m \pi_m c_m^j(\sum_k (q^k-1)\frac{\partial c_m^k}{\partial x_m} + \sum_\ell (p^\ell-1)\frac{\partial f^\ell}{\partial x_m})) - \lambda(\sum_m c_m^j \sum_i \frac{\partial \pi_i}{\partial x_m}(x_i - \sum_k (q^k-1)c_i^k - \sum_\ell (p^\ell-1)f^\ell)) = 0. \qquad \text{(iii)}$$

Also,

$$\frac{\partial \mathcal{L}}{\partial q^j} = \sum_m \pi_m \frac{\partial V_m}{\partial q^j} + \sum_m \frac{\partial \pi_m}{\partial q^j} V_m + \lambda(\sum_m \pi_m c_m^j) + \lambda(\sum_m \pi_m(\sum_k (q^k-1)\frac{\partial c_m^k}{\partial q^j} + \sum_\ell (p^\ell-1)\frac{\partial f^\ell}{\partial q^j})) - \lambda(\sum_i \frac{\partial \pi_i}{\partial q^j}(x_i - \sum_k (q^k-1)c_i^k - \sum_\ell (p^\ell-1)f^\ell)) = 0. \qquad \text{(iv)}$$

Again from the individual's maximization problem, $\sum_m \pi_m \frac{\partial V_m}{\partial q^j} + \sum_m \frac{\partial \pi_m}{\partial q^j} V_m =$ $- \sum_m \pi_m c_m^j \alpha_m$. Thus, we obtain from (iii) and (iv) that

$$\sum_m c_m^j \frac{\partial \mathcal{L}}{\partial x_m} + \frac{\partial \mathcal{L}}{\partial q^j} = \lambda(\sum_m \pi_m (\sum_k (q^k-1)(c_m^j \frac{\partial c_m^k}{\partial x_m} + \frac{\partial c_m^k}{\partial q^j})$$

$$+ \sum_\ell (p^\ell-1)(\sum_m \pi_m c_m^j \frac{\partial f^\ell}{\partial x_m} + \frac{\partial f^\ell}{\partial q^j})) \tag{v}$$

$$- \lambda(\sum_i ((\sum_m \frac{\partial \pi_i}{\partial x_m} c_m^j + \frac{\partial \pi_i}{\partial q^j})(x_i - \sum_k (q^k-1)c_i^k - \sum_\ell (p^\ell-1)f^\ell))) = 0.$$

Applying the notation that

$$(\frac{\partial c_m^k}{\partial q^j})_\theta = \frac{\partial c_m^k}{\partial q^j} + c_m^j \frac{\partial c_m^k}{\partial x_m}, \tag{vi}$$

$$(\frac{\partial f^\ell}{\partial q^j})_\theta = \frac{\partial f^\ell}{\partial q^j} + \sum_m \pi_m c_m^j \frac{\partial f^\ell}{\partial x_m}, \text{ and} \tag{vii}$$

$$(\frac{\partial \pi_i}{\partial q^j})_\theta = \frac{\partial \pi_i}{\partial q_j} + \sum_m \frac{\partial \pi_i}{\partial x_m} c_m^j \tag{viii}$$

to (v) gives

$$\sum_m c_m^j \frac{\partial \mathcal{L}}{\partial x_m} + \frac{\partial \mathcal{L}}{\partial q^j} = \lambda\{\sum_m \pi_m (\sum_k (q^k-1)(\frac{\partial c_m^k}{\partial q^j})_\theta + \sum_\ell (p^\ell-1)(\frac{\partial f^\ell}{\partial q^j})_\theta)$$

$$- \sum_i (\frac{\partial \pi_i}{\partial q^j})_\theta (x_i - \sum_k (q^k-1)c_i^k - \sum_\ell (p^\ell-1)f^\ell)\} = 0. \tag{ix}$$

Thus,

$$\sum_i \pi_i (\sum_k (q^k-1)(\frac{\partial c_i^k}{\partial q^j})_\theta + \sum_\ell (p^\ell-1)(\frac{\partial f^\ell}{\partial q^j})_\theta)$$

$$- \sum_i (\frac{\partial \pi_i}{\partial q^j})_\theta (x_i - \sum_k (q^k-1)c_i^k - \sum_\ell (p^\ell-1)f^\ell) = 0, \tag{x}$$

which is equivalent to (10) for $\chi = q^j$. The proof for $\chi = f^\ell$ is analogous.

BIBLIOGRAPHY

Arnott, R.J., and J.E. Stiglitz, "Labor Turnover, Wage Structures and Moral Hazard: The Inefficiency of Competitive Markets," September 1981, Econometric Research Program, Princeton University, research memorandum 289.

Arnott, R.J., and J.E. Stiglitz, "Equilibium in Competitive Insurance Markets: The Welfare Economics of Moral Hazard Part I: Basic Analytics and Part II: Existence and Nature of Equilibrium," Institute for Economic Research, Queen's University, discussion papers 465 (1982a) and 483 (1982b).

Arnott, R.J., and J.E. Stiglitz, "Equilibrium in Competitive Insurance Markets: The Welfare Economics of Moral Hazard. Part III. Some Welfare Implications of the Analysis," mimeo, 1983.

Arnott, R.J., A. Hosios, and J.E. Stiglitz, "Implicit Contracts, Labor Mobility, and Unemployment," June 1983, mimeo.

Atkinson, A.B., and J.E. Stiglitz, "The Design of Tax Structure," Journal of Public Economics, 6, 1976, 55-75.

Arrow, K., Aspects of the Theory of Risk-Bearing (Helsinki: Yrjo Jahnssönin Säätio, 1965).

Braverman, A., and J.E. Stiglitz, "Sharecropping and the Interlinking of Agrarian Markets", American Economic Review, 72, 1982, 695-715.

Debreu, G., Theory of Value (New York: John Wiley and Sons, Inc., 1959).

Diamond, P.A., and J.A. Mirrlees, "Optimal Taxation and Public Production I: Production Efficiency, and II: Tax Rules," American Economic Review, 61, 1971, 8-27 and 261-278.

Diamond, P.A., and J.A. Mirrlees, "A Model of Social Insurance with Variable Retirement," Journal of Public Economics, 10, 1978, 295-336.

Diamond, P., and M. Rothschild, comments on readings, Uncertainty in Economics (New York: Academic Press, Inc., 1978).

Dionne, G., "Moral Hazard and State-Dependent Utility Function", Journal of Risk and Insurance, XLIX, 1982, 405-422.

Greenwald, B., and J.E. Stiglitz, "The Welfare Economics of Imperfect Information and Incomplete Markets", Bell Labs., mimeo., 1982.

Grossman, S. and O. Hart, "An Analysis of the Principal-Agent Problem," CARESS Working Paper #80-17, University of Pennsylvania.

Hellwig, M., "Moral Hazard and Monopolistically Competitive Insurance Markets," 1982, mimeo.

Helpman, E., and J.-J. Laffont, "On Moral Hazard in General Equilibrium," Journal of Economic Theory, 10, 1975, 8-23.

Holmström, B., "Moral Hazard and Observability," Bell Journal of Economics, 10, 1979, 74-91.

Marshall, J.M., "Moral Hazard," American Economic Review, 66, 1976, 880-890.

Mirrlees, J.A., "An Exploration in the Theory of Optimum Income Taxation," Review of Economic Studies, 38, 1971, 175-208.

Mirrlees, J.A., "The Theory of Moral Hazard with Unobservable Behaviour, Part I," October 1975, mimeo.

Mirrlees, J.A., "The Implications of Moral Hazard for Optimal Insurance", April 1979, mimeo.

Mirrlees, J.A., "Lecture Notes on the Economics of Uncertainty," August 1980, mimeo.

Mitra, P., "A Theory of Interlinked Rural Transactions", Journal of Public Economics, 20, 1983, 167-192.

Pauly, M., "Overprovision and Public Provision of Insurance," Quarterly Journal of Economics, 88, 1974, 44-62.

Rothschild, M., and J. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," Quarterly Journal of Economics, 90, 1976, 629-649.

Shavell, S., "On Moral Hazard and Insurance," Quarterly Journal of Economics, 93, 1979a, 541-562.

Shavell, S., "Risk Sharing and Incentives in the Principal and Agent Relationship," Bell Journal of Economics, 10, 1979b, 55-73.

Spence, M., "Product Differentiation and Performance in Insurance Markets," Journal of Public Economics, 10, 1978, 427-447.

Spence, M., and R. Zeckhauser, "Insurance, Information, and Individual Action," American Economic Review, Papers and Proceedings, 61, 1971, 380-387.

Stiglitz, J.E., "Incentives and Risk Sharing in Sharecropping," Review of Economic Studies, 1974, 219-255.

Stiglitz, J.E., "Utilitarianism and Horizontal Equity: The Case for Random Taxation," Journal of Public Economics, 18, 1982, 1-34.

Stiglitz, J.E., "Self-Selection and Pareto Efficient Taxation," Journal of Public Economics, 17, 1982, 213-240.

Stiglitz, J.E., and A. Weiss, "Credit Rationing in Markets with Imperfect Information," American Economic Review, 71, 1981, 393-410.

Weiss, L., "The Desirability of Cheating, Incentives and Randomness in the Optimal Income Tax," Journal of Political Economy, 1976, 1343-1352.