Volume Title: The Role of Health Insurance in the Health Services Sector

Volume Author/Editor: Richard N. Rosett

Volume Publisher: NBER

Volume ISBN: 0-87014-272-0

Volume URL: http://www.nber.org/books/rose76-1

Publication Date: 1976

Chapter Title: Welfare Analysis of Changes in Health Coinsurance Rates

Chapter Author: Kenneth Arrow

Chapter URL: http://www.nber.org/chapters/c3815

Chapter pages in book: (p. 1 - 34)

PART ONE The Market for Health Insurance

1

KENNETH ARROW
Harvard University

Welfare Analysis of Changes in Health Coinsurance Rates

SUMMARY

This is a study of the welfare implications of changes in the coinsurance rate of health insurance policies. Only efficiency aspects are studied; distributional problems are ignored.

The basic function of health insurance is the reduction of uncertainty; other things being equal, individuals prefer and are willing to pay for a reduction in their financial risks. To that extent, a reduction in the coinsurance rate would represent a welfare gain. However, given that illness has occurred, insurance constitutes a subsidy to one form of consumption and therefore implies an efficiency loss whose magnitude depends on supply as well as demand conditions.

To get a precise expression for the net welfare change associated with a change in the coinsurance rate, it is necessary to formulate the problem as a miniature general equilibrium model, with both supply and demand considerations made explicit. Account must be taken of the random factors in demand, the financing of health insurance (here assumed to be by lump-sum taxation), the elasticity

Part of Rand Corporation study for the Office of Economic Opportunity (R-1281-OEO). The author is indebted to Charles Phelps and John Stein for many helpful comments, especially with regard to implicit assumptions and to exposition. All opinions are those of the author and not necessarily of the Rand Corporation or the United States government.

of supply, and the determination of medical prices through supply and demand. For each coinsurance rate, there is an equilibrium price for medical services. Each individual, given his income net of the taxes needed to pay for health insurance, has a demand for medical services in each state of nature and therefore an expected utility, taking into account uncertainty about health and medical costs. In the study I evaluate the change in expected utility as the coinsurance rate changes (see especially theorems 1 and 3).

If the supply of medical services is totally inelastic, then a change in coinsurance rates has no efficiency effect whatsoever. The price of medical services charged by the seller changes just enough so that the coinsurance payment (the price to the buyer) remains constant; hence, there is no effect on demand or on financial risk. When supply is totally inelastic, changes in coinsurance rates affect only the distribution of income between suppliers of medical services and the rest of the population.

1. INTRODUCTION

In this paper I make the following assumptions:

1. The health status of an individual is a random variable whose distribution is independent of prices and income;[1] the individual aims to maximize his expected utility.
2. The utility depends on the amount of goods other than health care, the amount of health care, and the state of health.
3. The insurance offered reimburses the expenditure on health care by a fixed proportion.
4. Given this coinsurance rate, the individual freely chooses the amount of medical care he wants after knowing his health status.
5. The health insurance payments are financed by lump-sum taxes.

I ignore distributional considerations and assume a single person in the economy. The interaction between distribution and insurance needs separate analysis.

Involved in this study is an investigation of the gain or loss of welfare associated with a small change in the coinsurance rate. For this purpose, it is clearly necessary to consider both supply and demand. The rapid increase in the prices of medical services since the introduction of Medicare and Medicaid can possibly be interpreted as the response of a market with relatively inelastic supply to a sudden increase of demand; if medical supply were highly elastic, the consequences for demand and therefore for efficiency could well have been very different.

I have constructed a miniature general equilibrium model of the economy disaggregated only into medical and nonmedical service markets. On the supply side, the main issue is the transformation between medical and nonmedical services. The hypothesis of perfect elasticity of supply has been implicit in most previous work. The general case is treated here. The particular case in which the elasticity of supply is zero turns out to have a property that may be surprising at first glance, although it is not hard to see that it is true: welfare, and indeed the allocation of resources as a whole, is totally independent of the coinsurance rate. All a change in the rate does is to transfer purchasing power between the medical and other sectors of the economy.

The welfare gains and losses have been treated in a paper by Feldstein (1973), although the theoretical basis of the calculation is not set forth too exactly. He also considers the nonoptimal behavior of nonprofit institutions. I shall analyze Feldstein's results in another paper.

2. EXPLICIT FORMULATION OF THE MODEL

Let

s = state of health
x_{1s} = demand for goods other than health care by individual in state s
x_{2s} = demand for medical care in state s
x_1 = supply of goods other than medical care
x_2 = supply of medical care
$U(x_{1s}, x_{2s}, s)$ = utility in health state s if x_{1s}, x_{2s} is consumed in that state
p = price of medical care received by seller
q = price of medical care paid by buyer
T = lump-sum taxes needed to finance health insurance
y = income after taxes

Prices, taxes, and income are measured, with "other goods" as numeraire. To avoid distributional considerations, I assume that all individuals have identical endowments and identical utility functions. I further assume a very large population, with the state of health varying independently from individual to individual. Measure all quantity variables on a per capita basis. Then the aggregate demand for commodity i is $E(x_{is})$, which is nonstochastic and is, in equilibrium, equal to the supply, x_i.

(1) $$E(x_{is}) = x_i \ (i = 1, 2)$$

I assume that the insurance policies specify

$r = q/p$ = coinsurance rate

but that the actual values of p and q are determined by market forces, to satisfy (1). The total cost of insurance (per capita) is then $(p - q)\, x_2$, or,

(2) $$T = (1 - r)px_2$$

The representative individual derives his income by selling x_1 units of goods other than medical care (more exactly, the resources that will produce x_1 units of other goods) and x_2 units of medical care, the former at a price equal to 1, the latter at price p. Hence,

(3) $$y = x_1 + px_2 - T = x_1 + rpx_2 = x_1 + qx_2$$

I assume that the individual's supply decisions are made on a purely economic basis; there are no net advantages or disadvantages to the production of medical services. If supply is inelastic, x_1 and x_2 are given, and p is determined by demand conditions. If medical services are produced perfectly elastically at price p, $x_1 + px_2$ is given to the individual and to society, p is determined by technological considerations, and the actual values of x_1 and x_2 are determined by demand conditions.

In the general imperfectly elastic supply case, the supply functions $x_1(p)$ and $x_2(p)$ are determined so as to maximize $x_1 + px_2$ subject to a transformation constraint. Since p is the marginal rate of transformation,

(4) $$\frac{dx_1}{dp} + p\,\frac{dx_2}{dp} = 0$$

Once the state s has occurred, the individual maximizes $U(x_{1s}, x_{2s}, s)$ subject to the constraint

(5) $$x_{1s} + qx_{2s} = y$$

(Strictly speaking, the budget constraint is a weak inequality, but in this case it is assumed to be binding.) The optimality conditions then are (5) and the equations,

(6) $$\frac{\partial U}{\partial x_{1s}} = \lambda_s, \ \frac{\partial U}{\partial x_{2s}} = \lambda_s q$$

where λ_s is the marginal utility of income in state s.

The optimization defines, for each s, the demand functions $x_{1s}(q, y)$ and $x_{2s}(q, y)$. For given r, q and y are in turn functions of p, x_1, and x_2. The aggregate demands, $E(x_{is})$ $(i = 1, 2)$, are therefore also functions of these variables. The two equations (1) are not independent; if the equation for $i = 2$ is multiplied by q and added to that for $i = 1$,

$$E(x_{1s} + qx_{2s}) = x_1 + qx_2$$

but from (3) and (5), this reduces to the tautology, $E(y) = y$ (recall that y is not dependent on s). Hence, the equilibrium values of p, x_1, and x_2 are defined by one of the equations (1) together with the supply conditions. There is thus an equilibrium allocation of resources between medical care and other goods for each value of the coinsurance rate r, and I wish to evaluate these alternative equilibria. Ideally, I would like to optimize on r; as a minimum, I would like to determine whether an increase or a decrease in the coinsurance rate would increase welfare.[2]

The criterion of welfare is taken to be the expected utility of the representative individual. This respects his attitude to risk aversion; it also has the implication of respecting his tradeoff between medical care and other goods in any given state of health, a point that might be more arguable but will not be challenged here. Let W be the individual's welfare:

$$W = E[U(x_{1s}, x_{2s}, s)] \tag{7}$$

The equilibrium magnitudes of the system are functions of r. In particular, for each s, x_{1s} and x_{2s} are functions of q and y, which in turn are determined, for fixed r, by p, x_1, and x_2. Hence, W is a function of r. I shall examine the effects of marginal changes in W; the spirit of this analysis is therefore very similar to Lesourne's (1975, Ch. 3, Sec. I).

3. THE GENERAL FORMULA FOR WELFARE EFFECTS: FIRST FORM

First, differentiate the budget equation (5) with respect to r.

$$\frac{dx_{1s}}{dr} + q\frac{dx_{2s}}{dr} = \frac{dy}{dr} - x_{2s}\frac{dq}{dr}$$

But from (3),

$$\frac{dy}{dr} = \frac{dx_1}{dr} + q\,\frac{dx_2}{dr} + x_2\,\frac{dq}{dr}$$

so that

$$\frac{dx_{1s}}{dr} + q\,\frac{dx_{2s}}{dr} = \frac{dx_1}{dr} + q\,\frac{dx_2}{dr} + (x_2 - x_{2s})\,\frac{dq}{dr} \tag{8}$$

Now differentiate the welfare criterion (7) with respect to r.

$$\frac{dW}{dr} = E\left(\frac{\partial U}{\partial x_{1s}}\,\frac{dx_{1s}}{dr} + \frac{\partial U}{\partial x_{2s}}\,\frac{dx_{2s}}{dr}\right)$$

First substitute from (6) and then from (8):

$$\frac{dW}{dr} = E\left[\lambda_s\left(\frac{dx_{1s}}{dr} + q\,\frac{dx_{2s}}{dr}\right)\right]$$

$$= E\left[\lambda_s\left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr} + (x_2 - x_{2s})\,\frac{dq}{dr}\right)\right]$$

Notice, however, that the magnitudes

$$\frac{dx_1}{dr},\ \frac{dx_2}{dr},\ \text{and}\ \frac{dq}{dr}$$

are independent of the state of health and hence are not random variables. They can therefore be factored out of any expectation. The marginal effect of coinsurance rate on welfare therefore takes the form

$$\frac{dW}{dr} = \left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr}\right) E(\lambda_s) - E\left[\lambda_s(x_{2s} - x_2)\right]\frac{dq}{dr} \tag{9}$$

The first term of (9) represents the welfare gain within each state s resulting from an increase in the coinsurance rate, which is a decrease in the subsidy to the consumption of medical services. The second term is the distinctive element that measures the welfare loss because of increased risk-bearing. With regard to the second term, notice that from (1), x_2 is the mean value of x_{2s}; hence, by the usual definition of a covariance,

$$E\left[\lambda_s(x_{2s} - x_2)\right] = \sigma_{\lambda_s\,x_{2s}} \tag{10}$$

the covariance of medical services used with the marginal utility of income. I shall return to this term in Section 5.

At the moment, I use standard methods of second-best analysis (see, for example, the discussion in Lesourne, 1975, cited above) to restate the first factor in the first term of (9). This is the conventional measure of marginal welfare effect, and I give it the symbol

$$W_0 = \frac{dx_1}{dr} + q\,\frac{dx_2}{dr} \tag{11}$$

Because of uncertainty and the fact that markets are therefore not perfect (in the sense that the full set of contingent markets is not available), there are two somewhat different expressions that can be found for W_0, the first of which is more useful econometrically and the second of which is more useful for theoretical analysis.

Since x_{2s} is a function of q and y, and with the aid of the definition of income (3),

$$\begin{aligned}\frac{dx_{2s}}{dr} &= \frac{\partial x_{2s}}{\partial q}\frac{dq}{dr} + \frac{\partial x_{2s}}{\partial y}\frac{dy}{dr}\\ &= \frac{\partial x_{2s}}{\partial q}\frac{dq}{dr} + \frac{\partial x_{2s}}{\partial y}\left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr} + x_2\,\frac{dq}{dr}\right)\\ &= \left(\frac{\partial x_{2s}}{\partial q} + x_2\,\frac{\partial x_{2s}}{\partial y}\right)\frac{dq}{dr} + \frac{\partial x_{2s}}{\partial y}\left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr}\right)\end{aligned} \tag{12}$$

Take expectations of both sides of (12). From (1),

$$E\left(\frac{dx_{2s}}{dr}\right) = \frac{dE(x_{2s})}{dr} = \frac{dx_2}{dr}$$

and therefore,

$$\frac{dx_2}{dr} = \frac{dq}{dr}\left[E\left(\frac{\partial x_{2s}}{\partial q}\right) + x_2\,E\left(\frac{\partial x_{2s}}{\partial y}\right)\right] + \left(\frac{dx_1}{dr} + q\,\frac{dx_2}{dr}\right)E\left(\frac{\partial x_{2s}}{\partial y}\right)$$

Now, from (4) and (11),

$$\begin{aligned}W_0 &= \frac{dx_1}{dr} + q\,\frac{dx_2}{dr} = \frac{dx_1}{dr} + p\,\frac{dx_2}{dr} + (q-p)\,\frac{dx_2}{dr}\\ &= (q-p)\,\frac{dx_2}{dr}\\ &= (q-p)\,\frac{dq}{dr}\left[E\left(\frac{\partial x_{2s}}{\partial q}\right) + x_2\,E\left(\frac{\partial x_{2s}}{\partial y}\right)\right]\\ &\quad + (q-p)\,E\left(\frac{\partial x_{2s}}{\partial y}\right)W_0\end{aligned} \tag{13}$$

The expression

$$(14) \qquad \bar{S}_{22} = E\left(\frac{\partial x_{2s}}{\partial q}\right) + x_2 E\left(\frac{\partial x_{2s}}{\partial y}\right)$$

resembles a Slutsky compensated derivative but, in fact, is not, nor is it the expectation of one. However, it can also be written

$$\bar{S}_{22} = \frac{\partial E(x_{2s})}{\partial q} + E(x_{2s})\frac{\partial E(x_{2s})}{\partial y}$$

If a demand curve for medical services is fitted to time series in the usual way, the dependent variable is $E(x_{2s})$, and therefore all the terms in the expression can be calculated from the econometric analysis.

Solve in (13) for W_0, using the abbreviation (14).

$$(15) \qquad W_0 = \frac{(q-p)\bar{S}_{22}}{1+(p-q)E\left(\frac{\partial x_{2s}}{\partial y}\right)}\frac{dq}{dr}$$

When there is some insurance, $q < p$; since medical services are a normal good, the denominator is certainly positive, and the sign of W_0 is opposite to that of $\bar{S}_{22}$ (if $dq/dr > 0$).

An alternative expression for $\bar{S}_{22}$ will strongly suggest that it must be negative. Let S_{22s} be the compensated effect of a change in the price of medical services for a given state s—that is, the derivative of x_{2s} with respect to q when the consumer remains on an indifference curve for that state.

$$S_{22s} = \frac{\partial x_{2s}}{\partial q} + x_{2s}\frac{\partial x_{2s}}{\partial y}$$

Then

$$\frac{\partial x_{2s}}{\partial q} + x_2\frac{\partial x_{2s}}{\partial y} = S_{22s} - (x_{2s} - x_2)\frac{\partial x_{2s}}{\partial y}$$

Taking expectations,

$$\bar{S}_{22} = E(S_{22s}) - E\left[(x_{2s} - x_2)\frac{\partial x_{2s}}{\partial y}\right]$$

Of course, for each s, $S_{22s} < 0$, so the first term is negative. The second term is the covariance between medical services and the marginal propensity to consume them (remember that this is the covariance across states of health for a given individual). This term can also, and perhaps more illuminatingly, be rewritten as follows:

Since

$$(x_{2s} - x_2)\frac{\partial x_{2s}}{\partial y} = \tfrac{1}{2}\,\frac{\partial (x_{2s} - x_2)^2}{\partial y}$$

$$E\left[(x_{2s} - x_2)\,\frac{\partial x_{2s}}{\partial y}\right] = \tfrac{1}{2}E\left[\frac{\partial (x_{2s} - x_2)^2}{\partial y}\right] = \tfrac{1}{2}\,\frac{\partial E[(x_{2s} - x_2)^2]}{\partial y} \tag{16}$$

$$= \tfrac{1}{2}\,\frac{\partial \sigma^2_{x\,2s}}{\partial y}$$

$$\overline{S}_{22} = E(S_{22s}) - \tfrac{1}{2}\,\frac{\partial \sigma^2_{x\,2s}}{\partial y}$$

The second term is rather unexpected; it represents the effect of income on the variance of health expenditure. It seems reasonable to assume that a higher income permits higher medical expenditures in more serious illnesses; hence, one would expect that the variance of medical expenditures would increase with income. Therefore, it is to be presumed *a fortiori* that $\overline{S}_{22}$ is negative and that $W_0 > 0$ when $q > p$. Phelps has taken the absolute values of the residuals from a regression of physician visits on a number of variables including income (see Newhouse and Phelps, 1973) and shown that the correlation with income is slightly negative, in contrast to this argument. However, the effect is small compared with the first term in (16), so that the negativity of $\overline{S}_{22}$ is not in question.

Remark 1 Clearly, when $q = p$ (no insurance at all), then W_0 vanishes completely.

Remark 2 If $dq/dr = 0$, then again $W_0 = 0$. Then by (11), the first term in (9) is 0 and so is the second term, so that $dW/dr = 0$. To bring this out more clearly, notice that the specific definition of r played no role in the analysis; any parameter of the insurance contract would have yielded the same formulas. In particular, r might have been replaced by q everywhere. That is, one could imagine an insurance system in which the government chose the buyer's price, rather than a coinsurance rate, and then let the forces of the market determine seller's price and from that the needed lump-sum taxes. The analysis would have proceeded along the same lines, except that dq/dr would have been replaced by the number 1. If

$$W'_0 = \frac{dx_1}{dq} + q\,\frac{dx_2}{dq} \tag{17}$$

then the analogue of (15) is

$$W_0' = \frac{(q-p)\bar{S}_{22}}{V}, \text{ where } V = 1 + (p-q)E\left(\frac{\partial x_{2s}}{\partial y}\right) \tag{18}$$

and (15) itself becomes

$$W_0 = W_0' \frac{dq}{dr} \tag{19}$$

From (9), (10), (11), (18), and (19), one can write Theorem 1.

Theorem 1 A general formula for the marginal effect of an increase in the coinsurance rate on expected welfare is

$$\frac{dW}{dr} = \left[W_0' E(\lambda_s) - \sigma_{\lambda_s x_{2s}}\right] \frac{dq}{dr}$$

where

$$W_0' = \frac{(q-p)\bar{S}_{22}}{V}, V = 1 + (p-q)\,E\left(\frac{\partial x_{2s}}{\partial y}\right)$$

and $\bar{S}_{22}$ is given by either of the following formulas:

$$\bar{S}_{22} = \frac{\partial E(x_{2s})}{\partial q} + E(x_{2s})\,\frac{\partial E(x_{2s})}{\partial y}$$

$$= E(S_{22s}) - \tfrac{1}{2}\,\frac{\partial \sigma^2_{x_{2s}}}{\partial y}$$

and S_{22s} is the compensated effect of a change in the price of medical services within a given state. I shall argue in Section 5 that under plausible assumptions the covariance between marginal utility of income and medical services is positive.

It should be pointed out that the welfare effect in Theorem 1 is measured in utility terms and therefore in arbitrary units. Usually, welfare losses are measured in some convenient numeraire. In this case, there is no completely obvious numeraire. Goods other than medical services appear to be the obvious choice, but under conditions of uncertainty this is not a well-defined commodity; one has to distinguish among other goods in different states of health. The simplest numeraire appears to be the composite good consisting of one unit of "other goods" in every state of health. The marginal utility of this composite is $E(\lambda_s)$, and therefore the welfare loss measured in other goods is obtained by dividing dW/dr by $E(\lambda_s)$.

4. THE GENERAL FORMULA FOR WELFARE EFFECTS: TAKING EXPLICIT ACCOUNT OF SUPPLY FACTORS

In the formulas in Theorem 1 I ignored the fact that r was to be interpreted as the coinsurance rate; the formulas are valid for any shift in the insurance scheme or indeed in any other parameter. The specific effect of coinsurance rates is confined to the expression dq/dr, the effect of the coinsurance rate on the buyer's price of medical services. The evaluation requires supply considerations.

In the case of perfectly elastic supply, the seller's price, p, is given by the technology. Since $q = rp$,

$$\frac{dq}{dr} = p \tag{20}$$

when supply is perfectly elastic. Then the expressions in Theorem 1 can be evaluated from demand considerations alone.

Theorem 2 When supply of medical services is perfectly elastic, production of a unit of medical services requires giving up p units of other goods. Then the marginal welfare effect of an increase in the coinsurance rates is

$$\frac{dW}{dr} = [W_0' E(\lambda_s) - \sigma_{\lambda_s x_{2s}}]p$$

where W_0' is defined in the statement of Theorem 1.

The welfare evaluations of medical insurance that have been made (for example, Feldstein, 1973; Pauly, 1968) have assumed perfect elasticity.

The imperfectly elastic case, at least in the short run, is much more realistic. Even in the long run, the production of both physicians and hospital services occurs under such special circumstances that perfect elasticity cannot be taken for granted.

Assume that the supplies of the two types of goods are functions of $p = q/r$; in particular, x_1 and x_2 are functions of p. Let

$$x_i' = \frac{dx_i}{dp} \ (i = 1, 2)$$

Differentiate the equation (1)

$$E(x_{2s}) = x_2(p)$$

with respect to r to solve for dq/dr.

(21) $$E\left(\frac{\partial x_{2s}}{\partial q}\frac{dq}{dr}+\frac{\partial x_{2s}}{\partial y}\frac{dy}{dr}\right)=x_2'\frac{dp}{dr}$$

From (3)

$$\frac{dy}{dr}=(x_1'+qx_2')\frac{dp}{dr}+x_2\frac{dq}{dr}$$

Any shift in supplies induced by a change in p must lie on the transformation surface; by (4),

$$x_1'+px_2'=0$$

Hence, as before,

$$x_1'+qx_2'=(q-p)x_2'$$

From (21)

$$E\left(\frac{\partial x_{2s}}{\partial q}+x_2\frac{\partial x_{2s}}{\partial y}\right)\frac{dq}{dr}=x_2'\left[1+(p-q)\,E\left(\frac{\partial x_{2s}}{\partial y}\right)\right]\frac{dp}{dr}$$

or, with the notation defined in the statement of Theorem 1,

(22) $$\bar{S}_{22}\frac{dq}{dr}=x_2'V\frac{dp}{dr}$$

From the identity $p=q/r$,

$$\frac{dp}{dr}=\frac{1}{r}\frac{dq}{dr}-\frac{q}{r^2}$$

Substitute in (22) and solve for dq/dr.

(23) $$\frac{dq}{dr}=\frac{pVx_2'}{x_2'V-r\bar{S}_{22}}$$

From (23) the following observations emerge:

1. since $\bar{S}_{22}<0$, $dq/dr<p$, if x_2' is finite;
2. if x_2' is infinite (perfect elasticity), $dq/dr=p$;
3. if $x_2'=0$ (perfect inelasticity), $dq/dr=0$;
4. as r approaches 0, dq/dr approaches p (this assumes that there is satiation in demand and some elasticity of supply, so that there is a finite equilibrium value of p).

Substituting (23) into Theorem 1, after some simplification, yields Theorem 3.

Theorem 3. Let $x_2(p)$ be the supply function of medical services. Then,

$$\frac{dW}{dr}=\frac{\left[(q-p)\bar{S}_{22}E(\lambda_s)-V\sigma_{\lambda_s r_{2s}}\right]px_2'}{Vx_2'-r\bar{S}_{22}}$$

where V and $\bar{S}_{22}$ are defined as in the statement of Theorem 1. The evaluation of this expression depends on econometric estimation of the demand and supply curves and on the evaluation of the covariance term.

5. COVARIANCE BETWEEN MARGINAL UTILITY OF INCOME AND MEDICAL SERVICES

To show that this covariance is positive, it suffices to indicate that both are increasing functions of the state of health. This presupposes that the states of health are measured in order of increasing severity of illness (poor health has a high index number). The desired result is derived from the following assumptions:

A.1. For fixed levels of medical services and other goods, the marginal rate of substitution of other goods for medical services is an increasing function of the state of health.

A.2. For fixed levels of medical services and other goods, the marginal utility of other goods does not decrease with the state of health.

A.3. For a fixed state of health, the utility is jointly concave in medical services and other goods and twice continuously differentiable.

A.4. For a fixed state of health and a fixed level of medical services, the marginal rate of substitution of other goods for medical services increases with an increase in the amount of other goods.

A.1 amounts to saying that the states of health are ordered in such a way as to make it true; the assumption is not tautological because it does assert that if the marginal rate of substitution increases from one state to another for one pair of values of medical services and other goods, it does so for all.

A.2 means that if an individual is initially given a fixed level of other goods and of medical services for all states of health, he would prefer to switch, if at all, to having the level of other goods rise with illness, if the switch can be made on an actuarially fair basis and if the level of medical services in any state is not subject to change. As Joseph Newhouse has pointed out to me, A.2 is not expected to hold for all states of health (see also Arrow, 1974, p. 5). Certainly, some states of health sharply reduce the value of other goods to the patient; he is too ill to enjoy the consumption and *ex ante* would have preferred to have shifted consumption of other goods to states of better health. In many states of ill health, however, other goods (such as domestic servants and other forms of service, or travel to less demanding climates) may be valued very highly. The result contained in Theorem 4 remains valid if A.2 holds only on the average or even if the marginal utility of other goods falls, but not too rapidly, as the state of health deteriorates.

A.3 is a usual statement of risk aversion. It means that given any two possible pairs (x_1^0, x_2^0) and (x_1^1, x_2^1) of other goods and medical services, the individual would prefer their average to an even chance of getting one or the other.

A.4 is self-explanatory.

I first derive expressions for the rates of change of medical services and of the marginal utility of income with respect to state of health and then show that, under the above assumptions, both are positive.

Differentiate the optimality conditions (6) and the budget equation (5) with respect to s, the state of health.

$$U_{11}\frac{dx_{1s}}{ds} + U_{12}\frac{dx_{2s}}{ds} + \left(-\frac{d\lambda_s}{ds}\right) = -U_{1s}$$

$$U_{21}\frac{dx_{1s}}{ds} + U_{22}\frac{dx_{2s}}{ds} + q\left(-\frac{d\lambda_s}{ds}\right) = -U_{2s}$$

$$\frac{dx_{1s}}{ds} + q\frac{dx_{2s}}{ds} = 0$$

where

$$U_{ij} = \frac{\partial^2 U}{\partial x_{is}\,\partial x_{js}}\ (i, j = 1, 2),\ U_{is} = \frac{\partial^2 U}{\partial x_{ix}\,\partial s}(i = 1, 2)$$

We can treat this system in the usual way as linear in the derivatives, dx_{1s}/ds, dx_{2s}/ds, and $-d\lambda_s/ds$. From the second-order conditions for a constrained optimum, the determinant, D, of the above system must be positive. Straightforward use of Cramer's rule yields

(24) $$\frac{dx_{2s}}{ds} = \frac{U_{2s} - qU_{1s}}{D}$$

(25) $$\frac{d\lambda_s}{ds} = \frac{U_{1s}(qU_{21} - U_{22}) + U_{2s}(U_{12} - qU_{11})}{D}$$

Since $q = U_2/U_1$, the numerator of (24) can be written

$$U_2\left(\frac{U_{2s}}{U_2}-\frac{U_{1s}}{U_1}\right)=U_2\frac{\partial\log(U_2/U_1)}{\partial s}$$

Since $U_2 > 0$ and, from A.1, U_2/U_1 is increasing in s,

$$\frac{dx_{2s}}{ds} > 0$$

Equivalently,

$$U_{2s} > qU_{1s}$$

From A.4,

$$U_{12} - qU_{11} = U_2\left(\frac{U_{21}}{U_2}-\frac{U_{11}}{U_1}\right) = U_2\frac{\partial\log(U_2/U_1)}{\partial s} > 0$$

From the last two relations, the numerator of (25) satisfies the inequality

$$U_{1s}(qU_{21} - U_{22}) + U_{2s}(U_{12} - qU_{11}) > U_{1s}(qU_{21} - U_{22}) + qU_{1s}(U_{12} - qU_{11})$$
$$= - U_{1s}(U_{11}q^2 - 2U_{12}q + U_{22}) \geqslant 0$$

since $U_{1s} \geqslant 0$ by A.2, and

$$U_{11}q^2 - 2U_{12}q + U_{22} \leqslant 0$$

by A.3.

Theorem 4 According to assumptions A.1–A.4, the marginal utility of income is positively correlated with medical expenditures. The covariance constitutes an offsetting risk adjustment to the marginal welfare change with respect to an increase in coinsurance.

In particular, Theorem 4 establishes that some insurance is better than no insurance. From Theorem 1 and from the fact that $V = 1$ when $r = 1$ (that is, $q = p$),

$$\frac{dW}{dr} = -\sigma_{\lambda_s x_{2s}}\frac{dq}{dr} < 0$$

if $dq/dr > 0$.

It appears that nothing further can be estimated on a theoretical basis except for the special case of inelastic supply discussed in the next section. It is not even excluded, so far as I can see, that complete insurance be optimal, although it is unlikely. In that case, $q = 0$; hence, the budget constraint tells us that $x_{1s} = y$, and x_{2s} is determined by the condition

$$\frac{\partial U_s}{\partial x_{2s}} = 0$$

Since medical care is always costly in terms of discomfort and time, we can suppose that the demand will be satiable. The solution to the last equation, when $x_{1s} = y$, will be denoted by x_{2s}^0. When $q = 0$, it is easy to calculate that $D = -U_{22}$. Hence (24) and (25) become

$$\frac{dx_{2s}^0}{ds} = -\frac{U_{2s}}{U_{22}}, \quad \frac{d\lambda_s}{ds} = U_{1s} - \frac{U_{2s}U_{12}}{U_{22}} = U_{1s} + U_{12}\frac{dx_{2s}^0}{ds}$$

The consumption of free medical services is certainly increasing with the state of health (measured to increase with increasing illness); indeed, since the state of health has so far appeared only ordinally, it is reasonable to identify x_{2s}^0 with s, so that $dx_{2s}^0/ds = 1$. Hence,

$$\frac{d\lambda_s}{ds} = U_{1s} + U_{12}, \quad \sigma_{\lambda_s x_{2s}} = \sigma_{\lambda_s s}$$

The relation between marginal utility of income and state of health when medical care is free depends on the cross-effects of state of health and of medical services on the marginal utility of other goods. It can be shown (see Appendix 1) that the covariance in question equals the variance of free medical services multiplied by an average value of the derivative $d\lambda_s/ds = U_{1s} + U_{12}$; in symbols,

$$\sigma_{\lambda_s s} = U\sigma_s^2$$

where U is a weighted average of the values of $U_{1s} + U_{12}$ for varying s.

The risk-aversion term may not vanish even for zero coinsurance. (Remember that this is a term in the *marginal* welfare effect; the risk-aversion welfare gain is of the first order in the coinsurance rate.) It is therefore conceivable that it outweighs the allocation term. In general, the values of U_{1s} and U_{12} should be small, so perfect insurance should not be optimal.

The calculation of $\sigma_{\lambda_s x_{2s}}$ can be made only by assuming specific forms for the utility function and the distribution of states of health. A specific example is developed in Appendix 2.

6. THE CASE OF PERFECTLY INELASTIC SUPPLY

In welfare economics we are accustomed to the argument that when supply is totally inelastic, changes in prices have no welfare effects. The argument may need to be reexamined here because of the presence of uncertainty in demand and the absence of contingent markets; but the conclusion remains valid. It does not seem to have been adequately emphasized in the literature on health insurance that *if the supply of medical care is perfectly inelastic, then there is no welfare effect at all from a change in the coinsurance rate*. There is, however, a rise in the price of medical services paid to the supplier; in a multiperson world this amounts to a redistribution of income to the suppliers of medical services.

This conclusion follows immediately by setting $x_2' = 0$ in Theorem 3. Notice that if x_1 and x_2 are both given, either of the equations (1) has only a single unknown, q; r, the coinsurance rate, does not enter, and income, y, is determined by q, from (3). The equilibrium buyer's price and income are the same for all values of r; in particular, the demands for medical services and for other goods in each state s is the same for all r, and therefore expected utility is independent of r.

The only variable that does change with changing r is p, since $p = q/r$. That is, the price of medical services rises as coinsurance rates fall. The pre-tax income of society is increasingly directed to medical services. To the extent that taxes to pay for medical services do not fall on medical income, there is a transfer of income to the suppliers of medical services. To illustrate, suppose that the cost of medical insurance is paid for by a proportional income tax at a rate t.

$$t = \frac{(p - q)x_2}{x_1 + px_2} = \frac{(1 - r)qx_2}{rx_1 + qx_2}$$

Then, the ratio of post-tax nonmedical incomes to their level with no insurance is

$$\frac{(1 - t)x_1}{x_1} = 1 - t = \frac{r(x_1 + qx_2)}{rx_1 + qx_2}$$

which decreases from 1 toward 0 as r decreases from 1 to 0. Correspondingly, the ratio of post-tax medical incomes to their no-insurance level is

$$\frac{(1 - t)px_2}{qx_2} = \frac{1 - t}{r} = \frac{x_1 + qx_2}{rx_1 + qx_2}$$

which rises from 1 as r decreases.

APPENDIX 1

Covariance of Marginal Utility of Income and Health

Theorem If X is a random variable and $f(X)$ is a function, then $\sigma_{f(X)X} = u\,\sigma_X^2$, where u is a weighted average of $f'(X)$.

Proof Let $g(x)$ be the density of X, a and b the limits of the range of X, $G(x)$ the cumulative distribution of X, and

$$H(x) = \int_a^x y g(y)dy$$

$$\sigma_{f(X)X} = E\left[f(X)X\right] - E(X)E\left[f(X)\right]$$

Integrating by parts,

$$E\left[f(X)X\right] = \int_a^b f(x)xg(x)dx = f(b)H(b) - f(a)H(a) - \int_a^b f'(x)H(x)dx$$

$$E\left[f(X)\right] = \int_a^b f(x)g(x)dx = f(b)G(b) - f(a)G(a) - \int_a^b f'(x)G(x)dx$$

By definition,

$$H(b) = E(X),\ H(a) = 0,\ G(b) = 1,\ G(a) = 0$$

Hence,

$$\sigma_{f(X)X} = \int_a^b f'(x)\ \left[E(X)G(x) - H(x)\right]dx$$

Let

$$W(X) = E(X)G(x) - H(x)$$

Then,

$$\sigma_{f(X)X} = \int_a^b f'(x)W(x)dx$$

This holds for any function $f(X)$. In particular, let $f(X) = X$, so that $f' = 1$.

$$\sigma_x^2 = \int_a^b W(x)dx$$

so that

$$\sigma_{f(X)X}/\sigma_x^2 = \int_a^b f'(x)w(x)dx$$

where

$$w(x) = W(x)\Big/\int_a^b W(x)dx$$

By construction,

$$\int_a^b w(x)dx = 1$$

To show that $u = \sigma_{f(X)X}/\sigma_X^2$ is a weighted average of $f'(X)$, it suffices to show that $w(x)$ is nonnegative, or, equivalently, that $W(x)$ is nonnegative.

Differentiating,

$$W'(x) = E(X)g(x) - xg(x) = g(x)[E(X) - x]$$

Hence, $W(x)$ is increasing for $x < E(X)$ and decreasing for larger values of x. It has a maximum at $x = E(X)$ and minimums at the extremes, $x = a$ and $x = b$. But $W(a) = 0$, $W(b) = E(X) - H(b) = E(X) - E(X) = 0$, so $W(x) > 0$ for all x, $a < x < b$.

In the text, X is interpreted as the state of health s (as measured by the consumption of medical services, x_{2s}^0, when free) and $f(X)$ is interpreted as the marginal utility of income, λ_s, with the derivative $d\lambda_s/ds = U_{1s} + U_{12}$.

APPENDIX 2

Covariance of Marginal Utility of Income and Medical Services for a Specific Utility Function and Distribution of States of Health

I seek here to illustrate how expressions might be found for the covariance term of Section 5 if assumptions are made about the nature of the utility function and the distribution of medical services for a given coinsurance rate.

Assume that

$$U(x_{1s}, x_{2s}, s) = -(1/c)e^{-cx_{1s}} + U_2(x_{2s}, s)$$

That is, I assume (1) that utility is additive in other goods and in medical services, and (2) the utility function for other goods has constant absolute risk aversion (this assumption is made by Feldstein, 1973). Assume further that the distribution of medical services for a given income and a given coinsurance rate is described by a gamma distribution (see Friedman, 1971),

$$\frac{a^b}{\Gamma(b)} e^{-ax_{2s}}(x_{2s})^{b-1}$$

Then, for any numbers m, n,

$$E(e^{-mX_{2s}}X_{2s}^{n}) = \frac{a^b}{\Gamma(b)} \int_0^{+\infty} e^{-mx_{2s}}x_{2s}^{n}e^{-ax_{2s}}x_{2s}^{b-1}dx_{2s}$$

$$= \frac{a^b}{\Gamma(b)} \int_0^{+\infty} e^{-(a+m)x_{2s}}x_{2s}^{n+b-1}dx_{2s}$$

Let

$$y = (a+m)x_{2s}$$

(2.1) $$E(e^{-mX_{2s}}X_{2s}^{n}) = \frac{a^b}{\Gamma(b)} \frac{1}{(a+m)^{n+b}} \int_0^{+\infty} e^{-y}y^{n+b-1}dy$$

$$= \frac{a^b\Gamma(n+b)}{\Gamma(b)\,(a+m)^{n+b}}$$

If $m = 0$, $n = 1$, we have

$$E(X_{2s}) = \frac{\Gamma(b+1)a^b}{\Gamma(b)a^{b+1}} = \frac{b}{a}$$

Since

$$\lambda_s = \frac{\partial U}{\partial x_{1s}} = e^{-cx_{1s}} = e^{-c(y-qx_{2s})} = e^{-cy}e^{cqx_{2s}}$$

with the aid of the budget constraint,

$$E(\lambda_s) = e^{-cy}E(e^{cqx_{2s}}) = \frac{a^b\Gamma(b)e^{-cy}}{\Gamma(b)(a-cq)^b} = \frac{a^be^{-cy}}{(a-cq)^b}$$

from (2.1), with $m = -cq$ and $n = 0$.
Then,

$$\sigma_{\lambda_s x_{2s}} = E(\lambda_s X_{2s}) - E(\lambda_s)E(X_{2s})$$

$$= E(e^{-cX_{1s}}X_{2s}) - \frac{ba^{b-1}e^{-cy}}{(a-cq)^b}$$

$$= e^{-cy}E(e^{cqX_{2s}}X_{2s}) - \frac{ba^{b-1}e^{-cy}}{(a-cq)^b}$$

$$= e^{-cy}\left[\frac{a^b\Gamma(b+1)}{\Gamma(b)(a-cq)^{b+1}} - \frac{ba^{b-1}}{(a-cq)^b}\right].$$

$$= \frac{e^{-cy}ba^{b-1}cq}{(a-cq)^{b+1}}$$

where use is made of (2.1) with $m = -cq$ and $n = 1$.

This calculation is designed to show merely that manageable formulas are not impossible, even though at the cost of strong assumptions. The parameters a and b of the distribution of medical services demanded are in principle observable. The absolute risk aversion cannot be inferred from data on the demand for medical services but is at least inferrable for observed behavior in the presence of uncertainty—for example, the choice of stock portfolios. The assumption of constant absolute risk aversion is uncomfortable; it implies, for example, that the demand for risky assets does not increase with wealth. However, alternative assumptions, such as constant relative risk aversion, do not lead to simple formulas, though in any case they always lead to expressions that can be evaluated numerically.

NOTES

1. This assumption may be false if higher income leads to better living conditions or to preventive medicine, which decreases the probability of serious illnesses. In the present context, the income effects are those arising from changes in the coinsurance rate and could not reasonably loom large.
2. Of course, this is within the context of insurance schemes that are purely linear in medical costs. Indeed, actual insurance plans, with their deductibles followed by coinsurance, are nonlinear and need closer investigation.

REFERENCES

1. Arrow, K. J., "Optimal Insurance and Generalized Deductibles," *Scandinavian Actuarial Journal*, 1974, pp. 1–42.
2. Feldstein, M., "The Welfare Loss of Excess Health Insurance," *Journal of Political Economy*, 81 (March–April 1973), pp. 251–280.
3. Friedman, B. S., "A Study of Uncertainty and Health Insurance," Ph.D. dissertation, Massachusetts Institute of Technology, 1971.
4. Lesourne, J., *Cost-Benefit Analysis and Economic Theory* (Amsterdam: North-Holland, 1975).
5. Pauly, M., "The Economics of Moral Hazard," *American Economic Review*, 68 (June 1968), pp. 531–537.

1 COMMENTS

William A. Brock
University of Chicago

Kenneth J. Arrow has constructed a miniature general equilibrium model that is, on the surface, especially tailored to answer welfare questions about coinsurance rates. Actually, however, a completely different interpretation can be given to the model that will yield additional insights.

Let $f(s)$ be the density function of states of health. Instead of thinking of

$$W = E[u(x_{1s}, x_{2s}, s)] = \int u(x_{1s}, x_{2s}, s) f(s) ds \tag{1}$$

as the expected utility of the representative individual, think of it as a welfare function for a society composed of heterogeneous individuals with no uncertainty in which $f(s)$ equals the "*proportion*" of individuals of type s. W is a "natural" welfare function in the sense that it gives equal weight to equal numbers—that is, it is an equal treatment property. (See Arrow and Kurz [1970] for a discussion of welfare functions in growth theory in which the utility of future generations is weighted by their numbers in the social welfare function.) Given this interpretation, Arrow's work looks more like a standard welfare analysis of distortions. If there is only one type s, it *is* a standard distortions analysis and the results of Foster and Sonnenschein (1970) and Kawamata (1972) can be applied to yield theorems on the welfare effects of increased "coinsurance" rates for general n goods models. In most n goods cases, welfare decreases as distortion increases if there is some substitutability in production, and welfare remains the same if supply is perfectly inelastic.

The same can be said for the insurance interpretation of Arrow's model. If there is only one state of health in the world, there is nothing to insure—thus there is nothing to compensate the economy for the welfare loss resulting from the wedge imposed between the selling price and the buying price of medical services.

According to the principle of continuity, this result tells us that we need substantial variations (1) in the utility functions $u(x_{1s}, x_{2s}, s)$ and/or (2) in the density $f(s)$ before insurance has any chance of becoming worthwhile from a welfare point of view. Observations (1) and (2) are brought out fairly clearly in Arrow's Equation (9) and his Theorem 4.

Some diagrams will be useful. These diagrams depict the standard classroom presentation of the welfare analysis of distortions. OAB is the transformation frontier. For given $1 > r \equiv q/p$ = buying price/selling price, a distorted equilibrium E is depicted in Figure 1. Distorted equilibrium is a selling-price p and a point E such that indifference curve DF is tangent to the line with slope $-1/q$ through E, the line through E with slope $-1/p$ is tangent to the transformation frontier, and $r = q/p$. It is intuitive from Figure 1 that E will

FIGURE 1 One 1 = 1 state case, elastic supply

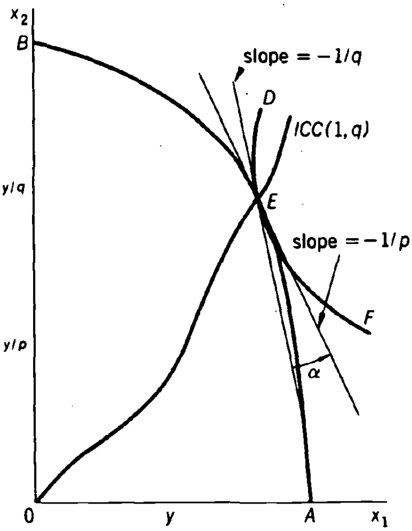
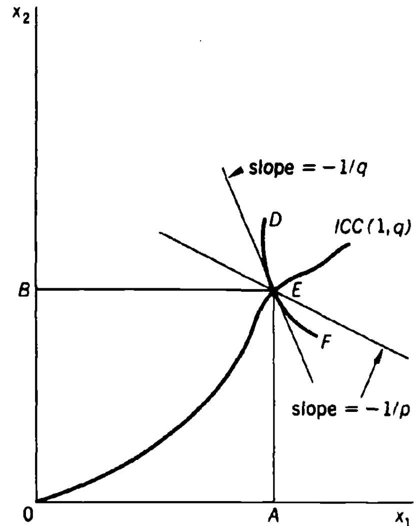


FIGURE 2 One 1 = 1 state case, perfectly inelastic supply



change when r changes and that welfare will fall when the wedge α increases—i.e., when r decreases, the subsidy to the consumption of medical care increases. It is obvious from Figure 2 that since E does not change when r changes, welfare does not change in the perfectly inelastic supply case. Given r, find distorted equilibrium in Figure 2. Merely find q so that $ICC(1,q)$ passes through the corner point E of AEB and set p so that $r = q/p$. Here $ICC(1,q)$ denotes the income consumption curve of the representative individual as a function of prices $(1,q)$.

Basically, what Arrow has done in his paper is to extend the analysis of figures 1 and 2 to the case of heterogeneous individuals. That is, $ICC(1,q)$ is replaced by

$$(2) \qquad \int_s ICC_s(1,q)f(s)ds$$

where (2) is a vector integral function of q. It is just the vector sum of income consumption curves across s—each ICC_s weighted by $f(s)$. Arrow's analysis in Section 6 points out that W in (1) does not change when r changes for the case of perfectly inelastic supply. It is obvious that the equilibrium E does not change from Figure 2. It is *not* obvious at first glance, however, that W does not change since the *distribution* of demand across individuals may change. I.e., it is not clear that there is *only* one q such that

$$(3) \qquad \int_s ICC_s(1,q)f(s) = E$$

even though, as Arrow correctly points out, q appears only in the left-hand side of Equation (3). Something needs to be assumed to guarantee uniqueness of the solution q to (3). A sufficient condition for uniqueness is

$$\frac{d}{dq}\int_s x_{2s}(1,q)\,ds < 0$$

for q satisfying (3), plus an appropriate boundary condition (see Dierker [1972]). At any rate, r has *no* impact on welfare because r does not affect the *set* of equilibria defined by (3).

It is worth pointing out that distorted equilibrium may not be unique for nonlinear transformation frontiers even though utility is normal in both goods and only one state exists. Normality is sufficient for uniqueness if the transformation frontier is *linear*. (See Foster and Sonnenschein [1970] and Kawamata [1972] for a general discussion.) If equilibrium is not unique, Arrow's differential analysis may break down.

A general uniqueness theorem may be formed by writing the equations that define distorted equilibrium as

$$G(q,r) = 0 \tag{4}$$

Then follow Brock's argument (1973, p. 555). His argument is easily extended to general n goods models with a *vector* of distortions $r \in R^k$ by assuming Dierker's uniqueness condition (1972) for a fixed r_0 and assuming

$$\frac{\partial G}{\partial q} \tag{5}$$

is non-singular on the set $\{q \mid G(q,r) = 0\}$ for each r. Just use (5) to prove that the number of equilibria is independent of r and use Dierker's condition for $r = r_0$ to obtain unique equilibrium for *all* r.

Thus, it appears that Arrow's analysis may be extended to n goods models with vectors of distortions. The concept of radial increase in distortion introduced by Foster and Sonnenschein (1970) will probably be needed here.

I think that I have said enough about the mathematics of Arrow's exercise. Let us now turn to the economics. The following points suggest natural lines for extending Arrow's analysis.

1. Arrow assumes only one coinsurance rate r for all states of health. It is natural to ask from the perspective gained by viewing Arrow's model as a model of people of type s if there are incentives for insurance companies to divide states into classes and offer different coinsurance rates for each class and be able to make a profit net of enforcement cost by so doing. Enforcement cost must be paid because a person may try to convince companies that he exists in a state of health that has a different coinsurance rate from the state that he is actually in. Since there is some sort of tradeoff between risk and distortion in Arrow's model, it is natural to perform such a subdivision of the set of states with corresponding coinsurance rates so as to maximize ex-

pected utility, taking into account the resource cost of the subdivision. Under what conditions can private enterprise solve this problem without government intervention? Extending Arrow's model to include such optimum partitioning of the set of states that I suggested above is likely to be important for policy purposes.

Let me expand on this point. Divide states of health into S_1, S_2, where S_1 denotes those states in which the consumption of medical care is very painful (involving radiation therapy, heart surgery, removal of sex organs, etc.) and where S_2 denotes those states in which consumption of medical services is less painful (corrective eyeglasses, minor dentistry, minor cosmetic surgery, treatment of baldness, nose jobs, sex therapy at Masters and Johnson, etc.). Quite clearly, one would want a higher coinsurance rate for S_2 states.

It seems to me that a very interesting general theory could be built by *explicitly* modeling the amount of overconsumption in a given class of states as a function of the particular characteristics of the medical services—a Lancaster-Becker type of approach.

In a world in which the law of large numbers applies, it seems possible that private enterprise would solve the problem without government intervention subject to solving the difficulties pointed out by Rothschild and Stiglitz (1973) and Wilson (1973). Obviously, incentives will be set up to subdivide states of health in order to capture the efficiency losses, but equilibrium may not exist if differentiation generates lumpy costs (Rothschild and Stiglitz, 1973; Wilson, 1973). It may be that some kind of equilibrium may exist if the "cost" of subdividing sets of states is explicitly modeled.

2. Is the optimal choice of r in Arrow's problem *enforceable* in the sense that it will be sustained by a competitive insurance industry? Since it is probably not enforceable in this sense, will it be enforceable at reasonable administrative cost?
3. The model needs to be disaggregated into different classes of people if it is to be part of a "project to study the financing of medical care services for the poor and near-poor" (Arrow, 1973, p. iii). But this should be a relatively routine extension of Arrow's basic theory.

REFERENCES

1. Arrow, K. J., "Welfare Analysis of Changes in Health Coinsurance Rates," R-1281-OEO, The Rand Corporation, November 1973.
2. ———, and M. Kurz, *Public Investment, the Rate of Return, and Optimal Fiscal Policy* (Baltimore: The Johns Hopkins Press, 1970).
3. Brock, W. A., "Some Results on the Uniqueness of Steady States in Multi-Sector Models of Optimum Growth when Future Utilities Are Discounted," *International Economic Review*, 14 (October 1973), pp. 535–559.
4. Dierker, E., "Two Remarks on the Number of Equilibria of an Economy," *Econometrica*, 40 (1972), pp. 951–954.

5. Foster, E., and H. Sonnenschein, "Price Distortion and Economic Welfare," *Econometrica*, 38 (March 1970), pp. 281–297.
6. Kawamata, K., "Price Distortion and Potential Welfare," Ph.D. dissertation, University of Minnesota, 1972.
7. Rothschild, M., and J. Stiglitz, "Analysis of Equilibrium in Insurance Markets," Departments of Economics, Princeton and Yale universities, 1973.
8. Wilson, C., "An Analysis of Simple Insurance Markets with Imperfect Differentiation of Consumers," University of Rochester, December 1973.

Isaac Ehrlich

University of Chicago and
National Bureau of Economic Research

Professor Arrow's paper addresses the welfare implications of changes in health coinsurance rates in a rigorous and stimulating fashion. It formally restates the ubiquitous moral hazard argument within a general equilibrium setting in which the effects of changes in coinsurance rates on the demand for and the supply of medical care are considered simultaneously. The paper also systematically considers the effects of coinsurance rate changes on the presumed utility costs of risk-bearing under some restrictive assumptions concerning the relation between preferences for goods and medical care services and various states of health. Both analyses, rigorous throughout, are developed with admirable clarity and simplicity. In the course of this investigation Arrow develops an intuitively plausible, though generally unrecognized, proposition. He shows that if the supply of medical services is perfectly inelastic, then changes in coinsurance rates—indeed, any changes in the terms under which market health insurance is provided—would have no effects on welfare. As he has already taught us in his analysis of employer or co-worker discrimination,[1] the existence of discrimination under perfectly inelastic supplies of relevant factors may affect only the distribution of income among various groups in the economy without engendering any allocative effects. In this study Arrow has developed a similar proposition under conditions of uncertainty by demonstrating that when the supply of relevant goods is perfectly inelastic, the expected utility of a representative individual is unaffected by an increase in the wedge between buyer's and seller's price of medical care brought about by a reduction in the health coinsurance rate.

Yet I have a basic reservation concerning the generality of the analysis of welfare effects associated with changes in health coinsurance rates as developed in this paper. My reservation stems from the implicit assumption from which the welfare implications have been derived that there are no close substitutes or complements to market health insurance, or, put differently, that coverage of medical care expenditures via a market insurance contract is the only means of shifting related financial risks. But surely various alternative means of "insurance" are available that may reduce vulnerability to illness, its severity, or, in general, the burden of medical care expenditures associated with poor states of health. I call these means "self-insurance" and "self-protection."[2] Things like nonprescription drugs, health foods and diets, physical exercise, frequent (uninsured) medical check-ups or, in general, the effective practice of "preventive medicine" are illustrative. Indeed, there are even more general methods of risk-shifting or self-insurance, such as dependence on family members or friends to assume financial responsibilities and to provide some health care services in poor states of health, or the choice of specific occupations and residential locations in which specific health risks are relatively low. The base of the argument I wish to develop in connection with these alternative means of health insurance is that they, like insurance-induced consumption of medical care services, draw scarce resources away from the production of other goods and services. Preventive care may be quite costly in terms of time and other resource expenditures, and raising large families as a means of shifting financial risks seems an inferior method of insurance. Since the availability of market health insurance does not eliminate expenditure of resources on self-insurance and self-protection, the relevant question from a welfare point of view is how changes in health coinsurance rates affect the *overall* allocation of resources to the general "health" sector encompassing both market- and self-insured health services and the transaction costs of insurance. In general, one may not be able to answer this question unambiguously without referring to the relative efficiency or inefficiency implicit in the provision of each form of insurance at the margin.

To illustrate the argument more pointedly I shall assume, merely for expositional convenience, that medical care outlays do not enhance utility directly, but rather are dictated as an objective medical requirement following the onset of poor states of health. Put differently, given the state of health, the demand for medical care services is perfectly inelastic.[3] Poor states of health are assumed to affect consumers' welfare mainly via their impact on the magnitude of the ensuing financial liabilities. These entail both direct costs of medical care and the opportunity costs of sick days, which are assumed to be monotonically related to expenditures on medical care (for simplicity, nonpecuniary costs are ignored). I shall also make the symmetrical assumption that preventive care and other means of self-insurance are not desired per se but only as means of shifting health risks, and that their sole impact is to reduce the extent of financial liabilities (medical care expenditures in particular) without affecting the probabilities that such states would occur.[4] That is to say, more self-insurance merely reduces the outlays on medical care services during poor states of health. Under these assumptions, an optimal allocation of resources between market-insured medical care (including transaction costs of insurance) and self-insurance services would equate the shadow price of transferring income between any two states of health via self-insurance and the corresponding market (health) insurance terms of exchange of income.

Figure 1 is an equilibrium for which a representative individual is faced with only two states of health: good health (g) with probability $1 - p$ and poor health (b) with probability p. Let us denote the person's income in state g by I^e; his loss function from poor health by $L(c)$, L being conditional on self-insurance expenditure in the amount of c "dollars"; the magnitude of market insurance purchased by s—the net income added to the insuree's income in state b via market insurance (the absolute amount of $L(c)$ covered through market insurance net of the premium paid); and the market health insurance implicit terms of exchange between net income in states b and g (defined in terms of state g's income) by π. The optimal combination of market and self-insurance occurs where [5]

(1) $$\pi = -\frac{1}{L'(c) + 1}$$

In equilibrium, the market insurance line is tangent to the transformation curve of income between states g and b as determined by the productivity of self-insurance (point A in Figure 1). If there were no transaction costs associated with the provision of health insurance and if, as in Arrow's system, the amount collected in premiums exactly matched the expected amount of medical care expenditures covered by insurance, then the price of insurance would be actuarially fair, $\pi = p/(1-p)$; the optimal coinsurance rate would be zero; and Equation (1) could be restated as $-pL'(c) = 1$, or by approximation

(2) $$p\,\Delta L = \Delta c$$

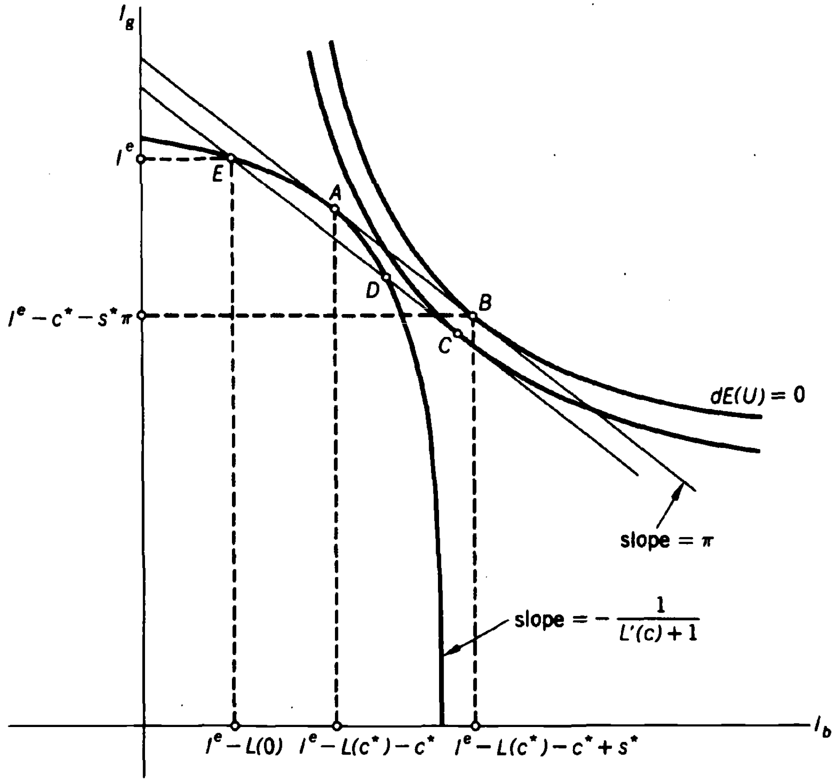
where $p\,\Delta L$ denotes the expected absolute value of the marginal reduction in the financial costs of illness (its actual value for the representative individual)[6] and Δc is the marginal expenditure on self-insurance. Since in practice the transaction costs of market insurance may be nonnegligible, let $\pi = (1+\lambda)p/(1-p)$, where λ denotes the net loading factor (the proportion by which π exceeds the fair price of insurance). Then the optimal coinsurance rate would be greater than zero, and by similar analysis it can easily be shown that in this case

(3) $$p\,\Delta L + \lambda p(\Delta L - \Delta c) = \Delta c$$

where, as before, $p\,\Delta L$ denotes the expected reduction in health-related losses and $\lambda p(\Delta L - \Delta c)$ represents the corresponding expected reduction in the transaction costs of insurance.[7] In a welfare-maximizing equilibrium position, the sum of these two terms must equal the marginal cost of self-insurance. Notice that for the representative individual, a decrease in the demand for self-insurance services and an increase in the demand for insured medical care services, for example, may involve a reduction in the price of self-insurance services and an increase in the price of medical care services if the relevant supply curves were upward sloping. Since a representative individual may be assumed to be both a producer and a user of health care and health insurance services, his expected utility would be affected only by real exchanges between self-insurance services and market-insured medical care. L, c, and s, the basic variables of the model, therefore will be defined throughout the following analysis in real rather than nominal terms.

Clearly, market health insurance, s, and self-insurance services, c, are substitutes in the sense that a decrease in the (real) price of one induces a greater demand for the other.[8] Moreover, any exogenous decrease in self-

FIGURE 1



insurance would, by assumption, expand the potential losses from ill health and, hence, would increase the demand for coinsured medical care. Similarly, an exogenous decrease in health coinsurance rates, which necessarily implies an increase in s, unambiguously would lead to reduced resource expenditures on self-insurance, an increase in the actual losses in states of poor health, and an increase in the demand for medical care services (see Note 8). As long as the supply of market-insured medical services and self-insurance services are not perfectly inelastic, an exogenous change in health coinsurance rates is thus expected to generate allocative effects. However, the net effect on welfare depends on the overall change in the actual amount of resources allocated to medical care, market insurance, and self-insurance activities.

If the initial equilibrium position were optimal, then the welfare implications of changes in health coinsurance rates would be readily evident: any *exogenous* change in these rates would bring about an increase in the total amount of resources allocated to market- and self-insured health care services and a reduction in the market value of personal income (indicated by the height of the market insurance lines in Figure 1). The expected utility of a representative individual will fall. As Figure 1 illustrates, expected utility would decrease following any change in the market and self-insurance nexus away from points A and B, corresponding to the optimal values s^* and c^*. This is seen, for example, by a shift from the initial equilibrium to, say, points E and C following an exogenous decrease in self-insurance expenditure and a consequent decrease in the health coinsurance rate (and an increase in market insurance purchased). But precisely the same reduction in expected utility would be associated with an exogenous increase in self-insurance and a resulting decrease in market insurance and the health coinsurance rate relative to their optimal magnitudes (compare points D and C with A and B).The important implication of this analysis is that in comparing situations in which coinsurance rates are *suboptimal*, one cannot determine unambiguously whether a reduction in the coinsurance rate would increase or lessen the misallocation of resources. A reduction in health coinsurance rates could be corrective if, for example, initially self-insurance were less efficient than market insurance at the margin so that Equation (3) were an inequality such that

(4) $$p\,\Delta L + \lambda p(\Delta L - \Delta c) < \Delta c$$

In this case more market insurance and medical care and less self-insurance services would be desirable.

Professor Arrow has focused attention on an important welfare issue associated with health coinsurance rates—namely, the potential resource misallocation arising from an increase in the wedge between buyers' and sellers' prices when the demand for health care services in each state of health is elastic with respect to its price. It has been argued in this comment, however, that in order to adequately analyze the direction of welfare effects resultings from (exogenous) changes in health coinsurance rates, one would wish to consider jointly the production of market-insured health services and alternative health insurance activities, including specific uninsured medical services. Since Arrow's general formula describing welfare effects does not account for resource shifts involving such alternatives to market-insured health services, empirical estimates of welfare effects relying solely on this formula would not appear to be complete.

NOTES

1. See Arrow (1973).
2. See Ehrlich and Becker (1972).
3. This assumption does not imply that the demand for insured medical care is unresponsive to changes in the coinsurance rate or the underlying price of insurance. As the following analysis demonstrates, a reduction in the coinsurance rate—an increase in the coverage ratio of ill-health-related losses—generally would enhance the demand for insured medical care because of a substitution away from self-insurance toward market insurance.
4. The analysis of the interaction between market health insurance and self-protection efforts designed to reduce the likelihood of states of ill health is more complicated than the analysis of

the interaction between market insurance and loss-reducing self-insurance, which is why the first analysis is ignored here. However, the basic implications developed in the following analysis concerning the welfare effects of changes in health coinsurance rates are general and would hold equally well when one's own efforts were expected to reduce the probabilities of poor states of health as well as the severity of illness and other related losses in such states.

5. The equilibrium condition is arrived at through maximization of the expected utility function

(1a) $E(U) = (1-p)U(I^e - c - s\pi) + pU[I^e - L(c) - c + s]$

with respect to c and s. Alternatively, let the coverage ratio (1 minus the coinsurance rate) be denoted by ∂ and let the loading factor be defined in terms of the gross amount paid in claims, $\lambda' \equiv (k-1)$. One then can proceed by maximizing the expected utility function

(1b) $E(U) = (1-p)U\left[I^e - c - kp\,\partial L(c)\right] + pU\left[I^e - c - \{1 - \partial(1-kp)\}L(c)\right]$

directly with respect to c and δ. The equilibrium condition then would be given by

$$(1c) \qquad - \frac{1 + kp\,\partial L'(c)}{1 + [1 - \partial(1-kp)]L'(c)} = \frac{kp}{1-kp}$$

Notice that since self-insurance is assumed to be effective, the term $L'(c) + 1$ in Equation (1) is less than or equal to zero. The function $L(c)$ is also assumed to be continuously differentiable and convex.

6. Following Arrow's analysis, I assume that health risks are independently distributed among a large number of individuals, so that the expected change in resources devoted to market-insured medical care would be the same as the actual resources devoted to accomodate the representative individual.

7. Alternatively, using Equation (1c) in Note 5, the equilibrium condition can be restated more simply in terms of the gross loading factor as follows

(3a) $(1 + \lambda')p\Delta L = \Delta c$

8. Strictly speaking, this theorem holds unambiguously only insofar as the relation between $_c$ and $_s$, not the coverage ratio ∂, is concerned. However, an *exogenous* increase in ∂, given L, implies an increase in the quantity of insurance, since quantity of insurance is defined by $s = \partial L\,(1 - kp)$. Furthermore, it can easily be shown through appropriate differentiations of Equation (1b) in Note 5 and utilization of the assumption that individuals are risk averse that an exogenous decrease in the coinsurance rate $r = 1 - \partial$ would necessarily reduce the optimal value of self-insurance expenditures, c. Such a decrease in r would thus raise L, hence the demand for (insured) medical care services, and the quantity of insurance purchased, s

REFERENCES

1. Arrow, Kenneth J., "The Theory of Discrimination," in Orly Ashenfelter and Albert Rees (editors), *Discrimination in Labor Markets* (Princeton: Princeton University Press, 1973).
2. Ehrlich, Isaac, and Gary Becker, "Market Insurance, Self-Insurance and Self-Protection," *Journal of Political Economy*, 80 (July–August 1972), pp. 623–648.