

**The Theory of the  
Demand for Health Insurance**

**John A. Nyman  
University of Minnesota**

January 4, 2001

John A. Nyman, Ph.D.  
University of Minnesota  
420 Delaware St. SE, Mail Route 729  
Minneapolis, MN 55455-0392  
USA  
telephone: 612.626.4425  
fax: 612.624.2196  
email: nyman001@tc.umn.edu

**Keywords:** health insurance, moral hazard, income transfer effect

This paper is not to be cited or quoted without the author's written permission. Copyright, 2001, John A. Nyman.

## **Abstract**

Conventional theory holds that moral hazard--the additional health care purchased as a result of becoming insured--is an opportunistic price response and is welfare-decreasing because the value of the additional health care purchased is less than its costs. The theory of the demand for health insurance presented here suggests that moral hazard is primarily an income transfer effect. In an estimation based on parameters from the literature, the value of moral hazard consumption is found to be 3 times greater than its costs, suggesting that income transfer effects dominate price effects and that moral hazard is welfare-increasing.

## 1. Introduction

The conventional theory of health insurance has held that becoming insured acts like a reduction in the price of health care, just as if the price reduction had occurred exogenously in the market. Newhouse writes:

“For the purpose of studying the relationship between health insurance and demand, the important point is that insurance is like a subsidy to purchase medical care; that is, it lowers the per-unit price of care. Although there is an income effect caused by premiums or taxes paid to finance the insurance benefits, these income effects can be shown to be empirically negligible in their effect on the demand for care...”(Newhouse, 1978, p. 9).

According to this theory then, the mechanism by which insurance is financed can be ignored because the effect of premiums on the demand for medical care--an income effect--is empirically negligible.

A central implication of this theory is that any additional health care consumed as a result of becoming insured--that is, any moral hazard--is welfare-decreasing. This welfare-loss argument, first made by Pauly (1968), is illustrated in Figure 1. The consumer's Marshallian demand for medical care is represented by  $D$  and the firms' supply by  $MC$ , the marginal cost of medical care. Without insurance, the market price is  $P=1$  and  $M^u$  medical care is consumed. With insurance, the price drops to  $P=0$  and  $M^i$  is consumed. The value of the additional medical care is measured by area  $aM^iM^u$ , and the cost by area  $abM^iM^u$ , so the welfare loss from the additional care is  $abM^i$ .

Because of this theory, many health economists have focused on policies that would reduce consumption at the margin. For example, Feldstein (1973) argues that the tax subsidy for

employer-based health insurance has resulted in American families spending too much on health care. He concludes that raising the coinsurance rate from 33% to 67% would increase society's welfare. Others (e.g., Manning and Marquis, 1996; Feldman and Dowd, 1991) have drawn similar conclusions.

In viewing insurance as a price effect, however, the origins of the insurance contract as a vehicle for transferring income to the ill have been overlooked. The prototypical insurance contract is a voluntary *quid pro quo* exchange where many consumers pay a premium in exchange for a claim on the pooled premiums, contingent on becoming ill. The smaller the probability of illness, the smaller is the premium that each purchaser of insurance must pay for any given payoff if ill. The difference between the payoff and the premium is a transfer of income from those who remain healthy to the person who becomes ill. Health insurance is purchased to obtain this income transfer when ill.<sup>1</sup>

Because of this income transfer, those who become ill purchase more health care (and other goods and services) than they would without insurance. For example, they may purchase an extra day in the hospital to recuperate, or they may purchase an expensive life-saving procedure that would otherwise be unaffordable. This additional health care is the income transfer effect of insurance. But because of the problems with verifying illness, fraud, and the complexity of writing contingent-claims contracts, the payoffs in actual private health insurance contracts occur through a reduction in the price of health care. Thus, of the additional health care purchased, that is, of the moral hazard, a portion is an opportunistic response to the reduced

---

<sup>1</sup>For example, for a payoff of \$10,000 if ill, the fair premium would be \$10 if the probability of illness were 1/1000, but \$5,000 if the probability of illness were 1/2. The income transfer is \$9,990 in the former case, and \$5,000 in the latter.

price, but a portion remains the original intended response to the income transfers.

The remainder of this paper is organized as follows. The price and income transfer effects from insurance that pays off by reducing price are modeled in the next section. Section 3 then describes the corresponding graphical analysis and section 4 presents an expected utility model of the decision to purchase insurance. Following that, contingent-claims insurance contracts, adverse selection, the requirements for determining optimal coinsurance rates are addressed in separate sections. In section 8, the welfare gain from moral hazard is estimated, and the paper concludes with the implications for public policy.

## 2. Model

Assume that all consumers have the same preferences and endowments. Without insurance, the consumer who becomes ill solves the standard problem:

$$\begin{aligned} \max U^s(M, Y) \\ \text{s.t. } Y^o = M + Y, \end{aligned} \tag{1}$$

where  $U^s$  is utility when ill,  $M$  is medical care,  $Y$  is income spent on other goods and services, and  $Y^o$  is the endowment. The price of  $M$  is assumed to be normalized at 1, so the first order conditions are:

$$U^s_M / U^s_Y = -1 \tag{2}$$

$$Y^o - M - Y = 0, \tag{3}$$

and demand for medical care without insurance is written  $M^u = M(1, Y^o)$ . With insurance that pays off by reducing the price of  $M$  from 1 to  $c$ , the ill consumer solves:

$$\begin{aligned} \max U^s(M, Y) \\ \text{s.t. } Y^o - R = cM + Y, \end{aligned} \quad (4)$$

where  $R$  is the fair premium, taken as a given, and  $c$  is the coinsurance rate. First order conditions are:

$$U_M^s / U_Y^s = -c \quad (5)$$

$$Y^o - R - cM - Y = 0, \quad (6)$$

and the demand for medical care with insurance is written  $M^i = M(c, R, Y^o)$ .

The premium is assumed to be actuarially fair, such that,

$$R = \pi(1-c)M^i, \quad (7)$$

where  $\pi$  is the probability of illness. For any given payoff--that is, for any given health care expenditures paid for by the insurer,  $(1-c)M^i$ -- the size the premium,  $\pi(1-c)M^i$ , depends on  $\pi$ . If the illness is rare and  $\pi$  is small, then the premium is small. In the limit, as  $\pi$  approaches 0, the premium becomes negligible and it appears as if the consumer faces the same first order conditions as would characterize an exogenous decrease in price to  $c$ :

$$U_M^s / U_Y^s = -c \text{ and} \quad (8)$$

$$Y^o \approx cM + Y. \quad (9)$$

In reality, however, the difference between the premium and the payout in expenditures is paid for by the many purchasers of insurance who remain healthy and transfer their premium through the insurer to those consumers who becomes ill. That is, the net payoff to each ill consumer is an income transfer from the  $(1-\pi)/\pi$  consumers who pay their premium into the pool but, because they remain healthy, have no claims on the pool. For example, say that a liver transplant cost \$300,000, but the probability of illness were only 1/75,000 during the contract period. The fair

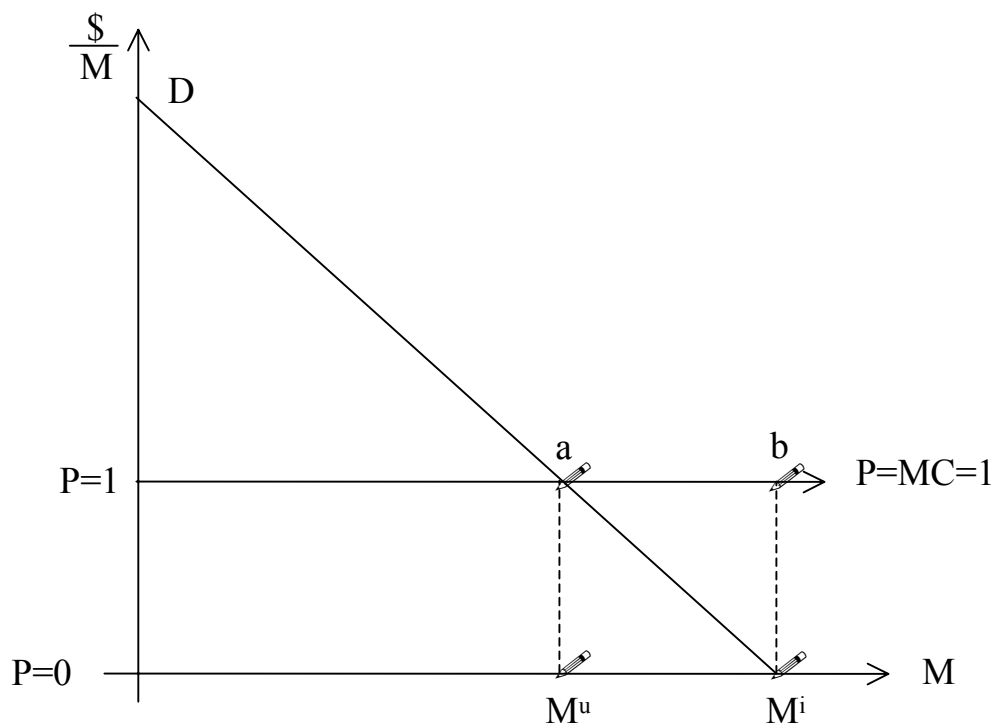


Figure 1

Conventional Moral Hazard  
Welfare Loss

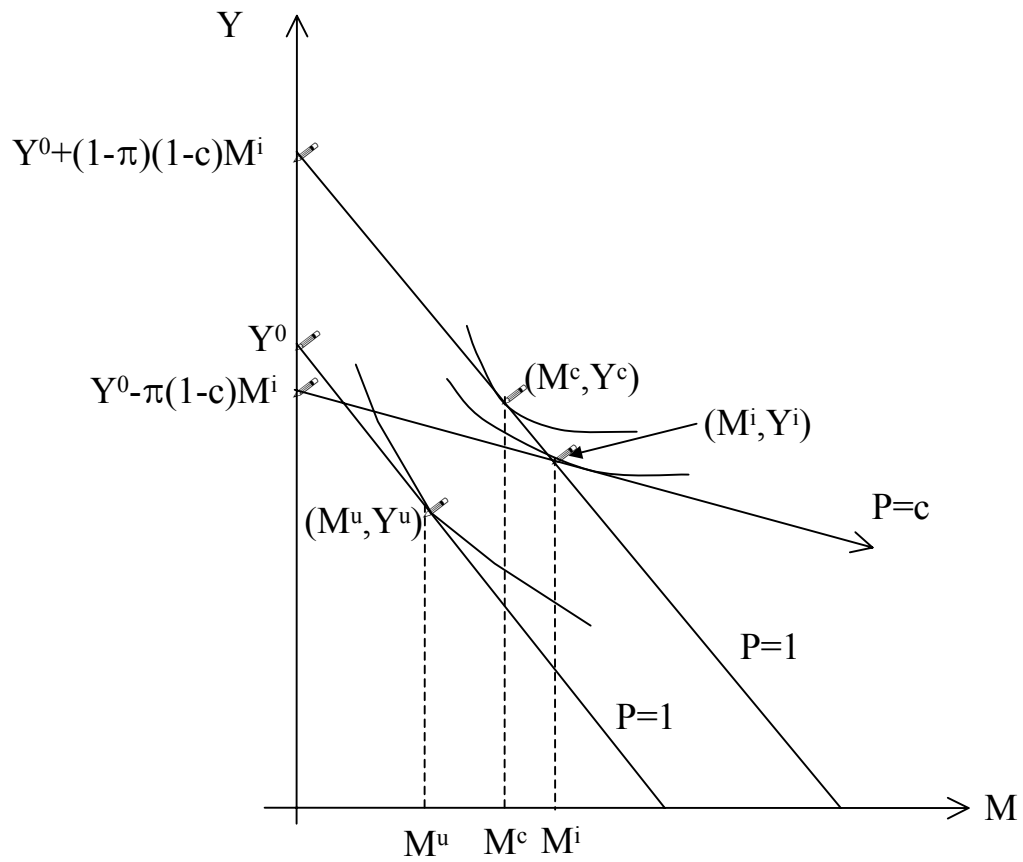


Figure 2

Income Transfer Effect of Insurance

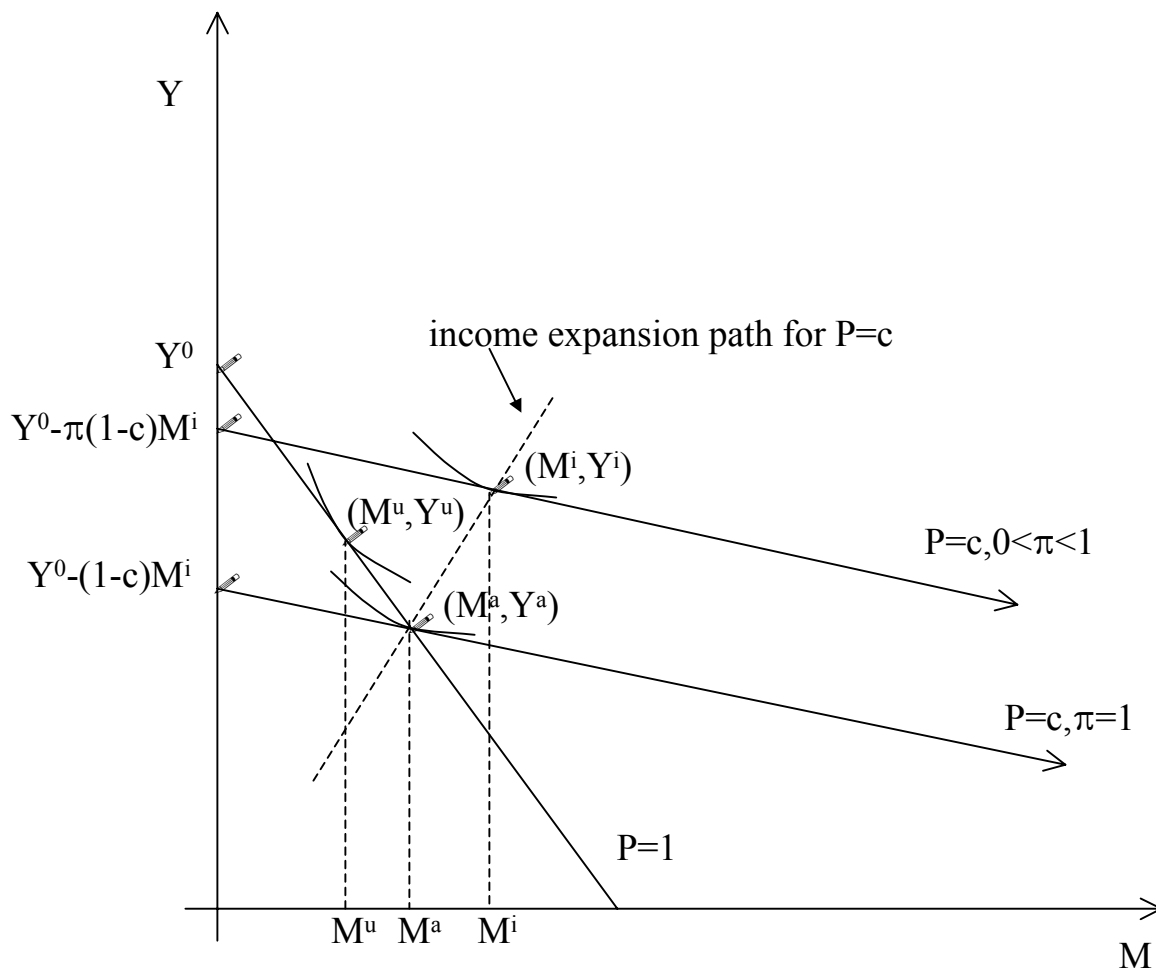
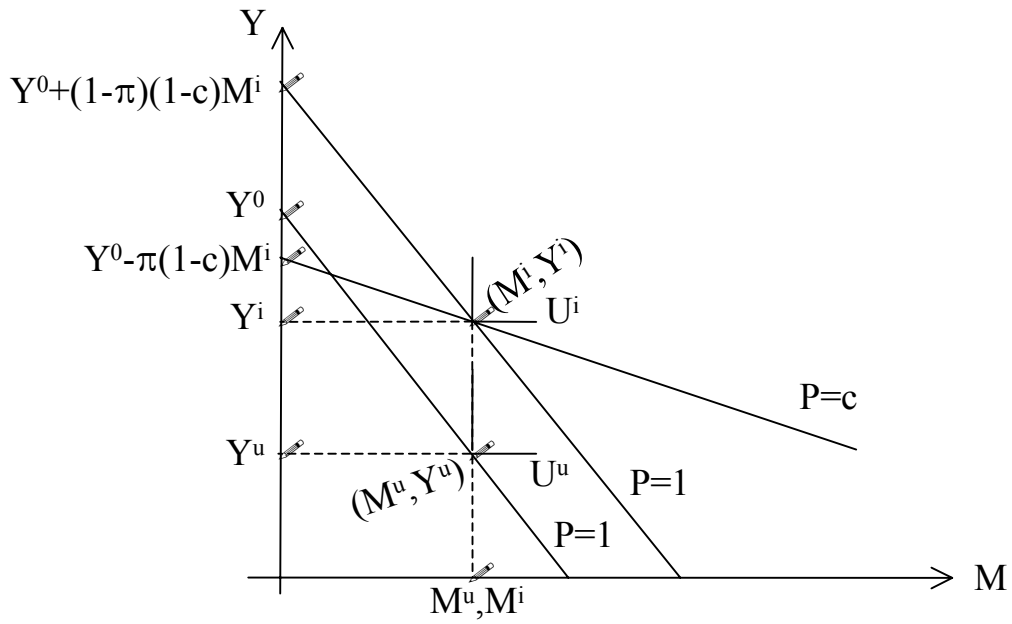
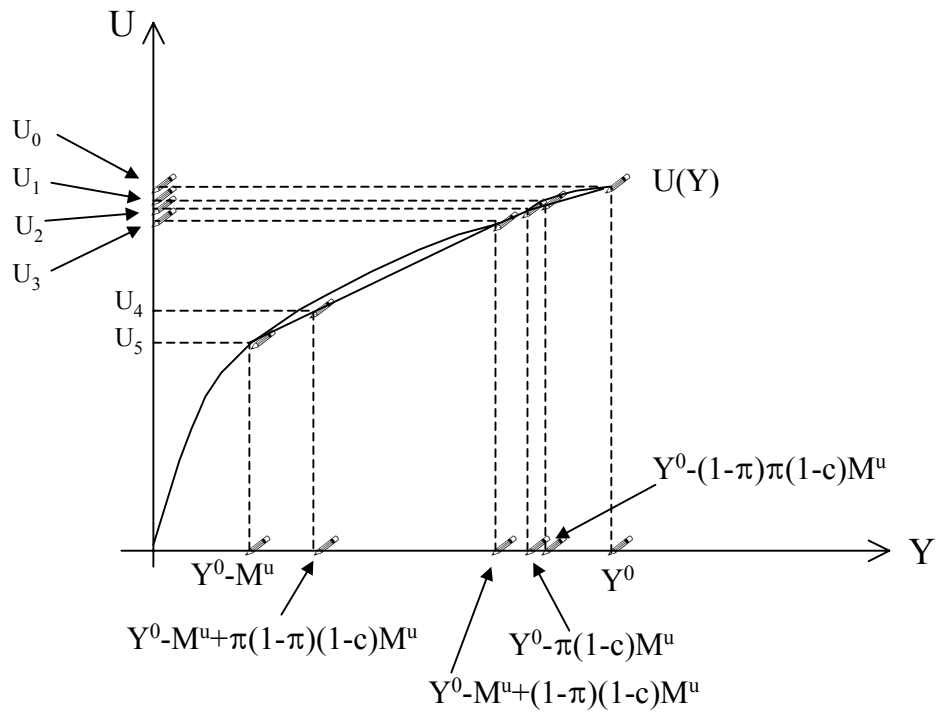


Figure 3

Income Transfer Effect: Compensating Variation Approach

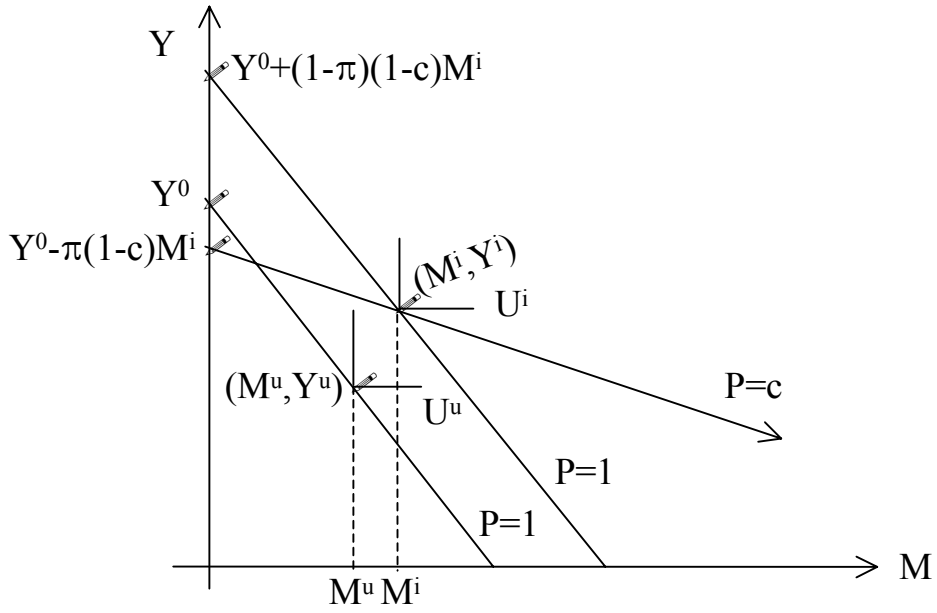


Panel A

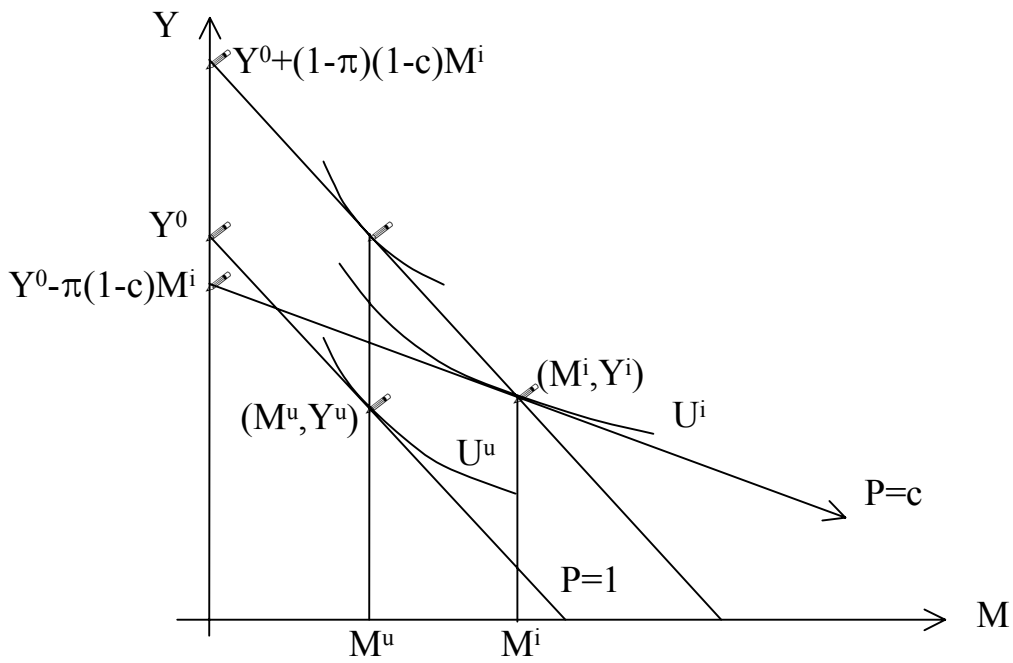


Panel B

Figure 4



Panel A



Panel B

Figure 5

premium for 0% coinsurance rate coverage of a liver transplant would be  $(\$300,000/75,000 =)$  \$4, so that the aggregate income transfer to the ill consumer from the 74,999 others who do not become ill during the contract period is  $(74,999*\$4 =)$  \$299,996.

If the illness is common, a larger percentage of the payoff is paid for in the premium by each consumer who becomes ill. In the limit, as  $\pi$  approaches 1, the first order conditions approach:

$$U_M^s/U_Y^s = -c \tag{10}$$

$$Y^o - (1-c)M \approx cM + Y,$$

or rearranging terms

$$Y^o \approx M + Y, \tag{11}$$

and the consumer pays a premium that covers the entire cost of the price reduction so that no income transfers occur. For example, if all insured consumers required one \$50 office visit to a physician during the contract year, the insurer's portion of the cost of that visit--\$50 at a 0% coinsurance rate--would be totally included in each consumer's fair premium and the income transfer would be \$0. Thus, if  $\pi=1$ , even though the price is reduced from 1 to  $c$ , the consumer is constrained to consume within his original budget,  $Y^o = M + Y$ .

The probability of illness, therefore, has an important effect on consumption with insurance. If insurance were an exogenous price decrease from 1 to  $c$ , the quantity response would simply reflect the observed price elasticity of demand for medical care, and the probability of illness would not be a factor. In fact, however, the private purchase of insurance is not an exogenous price decrease. The price reduction must be purchased by someone in order to exist. The extent to which the ill consumer pays for this price reduction and the extent to which those

who remain healthy pay for it depends on  $\pi$ , the probability of illness (or more realistically, the probabilities of the various diseases that the insurance contract covers). Thus, the probability of illness determines the quantity response to insurance that pays off with any given coinsurance rate  $c$ . This is analogous to the prototypical insurance contract for a given lump-sum payoff, where smaller probabilities of illness are associated with larger income transfers, and larger income transfers result in commensurately larger quantity responses. Similarly, in a price-payoff contract for a given expenditure payoff  $(1-c)M^i$ , smaller probabilities of illness are associated with larger portions of the payoff that are paid for by income transfers, and larger income transfers result in larger quantity responses.

The insurer conducts an actuarial study to determine insured demand,  $M^i|_{c,R,Y^0}$ . Once  $M^i$  is determined, income transfers become apparent by rewriting equation (6), the ill consumer's budget constraint after insurance, as

$$Y^0 - \pi(1-c)M^i = cM^i + Y^i, \quad (12)$$

adding  $(1-c)M^i$  to both sides of the equation, and rearranging terms:

$$Y^0 + (1-\pi)(1-c)M^i = M^i + Y^i. \quad (13)$$

So, compared to the ill consumer's budget constraint before insurance,

$$Y^0 = M^u + Y^u, \quad (14)$$

spending with insurance,  $(M^i + Y^i)$ , is larger than spending without,  $(M^u + Y^u)$ , by  $(1-\pi)(1-c)M^i$ , the income transfers. That is,  $(1-\pi)(1-c)M^i$  is the additional income that is provided by those who remain healthy and that is transferred to those who become ill, who in turn, spend this income transfer on additional medical care and other goods and services.

Empirically, the effect of this income transfer on medical care spending--that is, the

income transfer effect--can be estimated by:

$$\% \Delta M = \varepsilon(\% \Delta Y), \quad (15)$$

where  $\% \Delta M$  is the percentage increase in medical spending due to income transfers,  $\varepsilon$  is the appropriate income elasticity of demand for health care, and  $\% \Delta Y$  is the percentage increase in income due to income transfers. From equation (13),

$$\% \Delta Y = (M^i + Y^i - Y^o) / Y^o = (1 - \pi)(1 - c)M^i / Y^o. \quad (16)$$

The price effect can then be estimated as a residual if the total change in consumption is known.

### 3. Graphical Model

Figure 2 shows health care spending of the ill consumer both without insurance and with insurance that pays off by reducing the price of care from 1 to  $c$ , with a probability of illness,  $\pi$ , that lies between 0 and 1. Without insurance, optimal consumption by the ill consumer is  $(M^u, Y^u)$ , and with insurance, optimal consumption is  $(M^i, Y^i)$ . Suppose the ill consumer instead purchased a contingent-claims insurance contract where the lump-sum payoffs are exactly the same as the various expenditures by the insurer--the  $(1 - c)M^i$ 's--under the price-payoff contract. If the contract were perfectly designed, a fair premium for this contract would reflect the same expected total expenditures as under the price-payoff contract, and the difference between the aggregated payoffs and the fair premium to the ill consumer would constitute the same income transfers.<sup>2</sup> After these income transfers, the corresponding budget constraint would contain

---

<sup>2</sup>That is, if the payoff is  $(1 - c)M^i$  in a lump-sum income payment and the premium is  $-\pi(1 - c)M^i$ , then income transfers are  $(1 - \pi)(1 - c)M^i$ .

point  $(M^i, Y^i)$  but reflect the original price of medical care,  $P=1$ . Optimal consumption under this contingent-claims contract would occur at  $(M^c, Y^c)$  in Figure 2. Thus, the pure price effect of insurance would be the increase in medical care consumption caused by using a price reduction rather than lump-sum income payments as the payoff mechanism, or  $(M^i - M^c)$ . The income transfer effect of the price-payoff insurance would be the increase in medical care consumption caused by the contingent-claims (lump-sum payoff) insurance compared with consumption with no insurance, or  $(M^c - M^u)$ .

The foregoing discussion represents an “equivalent variation” type approach, where the income transfer effect is shown as the result of an increase in income transfers, evaluated at the uninsured price. Alternatively, a “compensating variation” approach would isolate the price effect of insurance by eliminating the income transfer effect after the price decrease has occurred. Figure 3 shows the income transfer and price effects using this approach. Again, optimal consumption is  $(M^u, Y^u)$  without insurance, and  $(M^i, Y^i)$  with insurance that pays off by reducing the price from 1 to  $c$  with a probability of  $\pi$ , where  $0 < \pi < 1$ . Because the probability is less than 1, income transfers have occurred and as a result, some portion of this increased consumption of medical care  $(M^i - M^u)$  is due to an income transfer effect. In order to eliminate the income transfer effect, it is necessary to eliminate income transfers, which would only be eliminated if  $\pi$  were 1. Therefore, the optimal consumption bundle without income transfers is the bundle that maximizes utility when the consumer is compelled to purchase a contract for a reduction of price from 1 to  $c$  *when already ill*. That is, the consumer faces insured prices but cannot consume beyond his budget constraint, reflecting equations (10) and (11). Such a bundle would occur at  $(M^a, Y^a)$ , the intersection of the original budget constraint and the income expansion path for the

insured price  $c$ . Thus, the price effect is  $(M^a - M^u)$  and the income transfer effect is  $(M^i - M^a)$ .

#### 4. The Decision to Purchase Insurance

The decision to purchase insurance depends to a certain extent on which of these effects tends to dominate. If  $U^h$  describes utility in the healthy state where  $M$  is 0, the consumer's ex ante decision to purchase insurance therefore compares expected utility without insurance,

$$EU_u = \pi U^s(M^u, Y^u) + (1-\pi)U^h(0, Y^o) \quad (17)$$

$$= \pi U^s(M^u, Y^o - M^u) + (1-\pi)U^h(0, Y^o) \quad (18)$$

to expected utility with insurance,

$$EU_i = \pi U^s(M^i, Y^i) + (1-\pi)U^h[0, Y^o - \pi(1-c)M^i] \quad (19)$$

$$= \pi U^s[M^i, Y^o + (1-\pi)(1-c)M^i - M^i] \\ + (1-\pi)U^h[0, Y^o - \pi(1-c)M^i] \quad (20)$$

$$= \pi U^s[M^i, Y^o - \pi(1-c)M^i + (1-c)M^i - M^i] \\ + (1-\pi)U^h[0, Y^o - \pi(1-c)M^i]. \quad (21)$$

In equation (21), income spent on other goods and services if ill is equal to the original income endowment ( $Y^o$ ), minus the premium  $[\pi(1-c)M^i]$ , plus the payoff  $[(1-c)M^i]$ , minus the income spent on medical care ( $M^i$ ). The payoff minus the premium is, of course, the income transfers, or  $[(1-\pi)(1-c)M^i]$  in equation (20). Such insurance is voluntarily purchased if  $EU_i - EU_u > 0$ .

The response to the income transfer and the price reduction will vary with the different illnesses, with different consumers and consumers' agents, even at different points in time for the same consumer and same illness. Therefore, to understand the demand for insurance, it is

necessary to consider the various possible responses to becoming insured.

**4.1. No additional medical care.** First, consider the case where  $M^i = M^u$ , so that no additional health care is consumed as a result of becoming insured. For example, the illness may be so predictable in its course and the procedure for curing it so standardized that there is no difference between the treatment with insurance and without. This case is illustrated in Panel A of Figure 4. Under this case, equation (20) becomes

$$EU_i = \pi U^s[M^u, Y^o + (1-\pi)(1-c)M^u - M^u] + (1-\pi)U^h[0, Y^o - \pi(1-c)M^u] \quad (22)$$

and the expected utility gain from being insured is

$$\begin{aligned} EU_i - EU_u &= \pi U^s[M^u, Y^o + (1-\pi)(1-c)M^u - M^u] + (1-\pi)U^h[0, Y^o - \pi(1-c)M^u] \\ &\quad - \pi U^s(M^u, Y^o - M^u) - (1-\pi)U^h(0, Y^o) \\ &= \pi \{U^s[M^u, Y^o + (1-\pi)(1-c)M^u - M^u] - U^s(M^u, Y^o - M^u)\} \\ &\quad + (1-\pi)\{U^h[0, Y^o - \pi(1-c)M^u] - U^h(0, Y^o)\}. \end{aligned} \quad (23)$$

Whether insurance is purchased or not depends entirely on the income consequences of this insurance. This is the conventional “risk avoidance” benefit from insurance, only in this specification, the benefit is expressed as a gain from the income transfer in the ill state net of the cost of the premium in the healthy state. Under this quid pro quo specification, the consumer weighs the payment of a premium that moves him from  $U^h(0, Y^o)$  “down” the utility function to  $U^h[0, Y^o - \pi(1-c)M^u]$  if healthy, against an income transfer that moves him from  $U^s(M^u, Y^o - M^u)$  “up” the utility function to  $U^s[M^u, Y^o + (1-\pi)(1-c)M^u - M^u]$  if ill. If the person voluntarily purchases this insurance, it can be assumed that the expected utility gain from the income transfer if ill exceeds the expected utility loss from paying the premium if healthy.

If the consumer faces the same utility function for  $Y$  when healthy or when ill, and if it

exhibits the standard “risk averse” functional form over income where  $U_Y > 0$  and  $U_{YY} < 0$ , insurance would be purchased. Panel B of Figure 4 illustrates this result. The consumer is originally at  $Y^o$  income and without insurance, would remain at  $Y^o$  if healthy. If he became ill without insurance, the consumer would spend  $M^u$  on medical care and have  $Y^o - M^u$  to spend on other goods and services.

If the consumer purchases insurance, he loses income when healthy by paying the actuarially fair insurance premium,  $-\pi(1-c)M^u$ , so that his expected income loss is from  $Y^o$  to  $(Y^o - \pi(1-c)M^u)$ , with a probability of  $(1-\pi)$ , or an expected dollar loss of  $-(1-\pi)[\pi(1-c)M^u]$  from income level  $Y^o$ . The corresponding expected utility loss from purchasing insurance and remaining healthy is  $(1-\pi)[U(Y^o) - U(Y^o - \pi(1-c)M^u)]|_{U(Y^o)}$ , or a loss of utility from  $U_0$  to  $U_1$  in Panel B of Figure 4, evaluated at the uninsured level of utility when healthy,  $U_0 = U(Y^o)$ . However, if the consumer purchases insurance and becomes ill, he gains income equal to the income transfers,  $(1-\pi)(1-c)M^u$ , and the expected gain in income is  $\pi[(1-\pi)(1-c)M^u]$ , from income level  $(Y^o - M^u)$ , the uninsured income level when ill. That is, the expected dollar gain from income transfers,  $\pi[(1-\pi)(1-c)M^u]$ , equals the expected dollar loss from the fair premium,  $-(1-\pi)[\pi(1-c)M^u]$ , so if the gain is evaluated on a steeper portion of a “risk averse” utility function than is the loss, the expected utility gained will exceed the expected utility lost. If ill then, the consumer’s expected utility gain from purchasing insurance is  $\pi\{U[Y^o + (1-\pi)(1-c)M^u - M^u] - U(Y^o - M^u)\}|_{U(Y^o - M^u)}$  or from  $U_5$  to  $U_4$  in Panel B of Figure 4, evaluated at the uninsured level utility if ill,  $U_5 = U(Y^o - M^u)$ . Because the expected gain in utility from the income transfer,  $U_4 - U_5$ , exceeds the expected loss in utility from the payment of the premium,  $U_0 - U_1$ , insurance is purchased.

**4.2. Additional affordable medical care.** Next, consider the case where  $M^u < M^i < Y^o$ .

That is, insurance results in an increase in consumption of medical care caused by either a price or income transfer effect, or both, but that without insurance, the ill consumer would still be able to afford to purchase same amount of the medical care as is purchased under insurance. In other words,  $M^i$  is feasible given the consumer's original endowment,  $Y^o$ . Under these circumstances, the voluntary purchase of insurance implies that

$$\begin{aligned}
EU_i - EU_u &= \pi U^s[M^i, Y^o + (1-\pi)(1-c)M^i - M^i] + (1-\pi)U^h[0, Y^o - \pi(1-c)M^i] \\
&\quad - \pi U^s(M^u, Y^o - M^u) - (1-\pi)U^h(0, Y^o). \\
&= \pi \{U^s[M^i, Y^o + (1-\pi)(1-c)M^i - M^i] - U^s(M^u, Y^o - M^u)\} + \\
&\quad (1-\pi) \{U^h[0, Y^o - \pi(1-c)M^i] - U^h(0, Y^o)\} \\
&= \pi \{U^s(M^i, Y^i) - U^s(M^u, Y^o)\} + \\
&\quad (1-\pi) \{U^h[0, Y^o - \pi(1-c)M^i] - U^h(0, Y^o)\} > 0. \tag{24}
\end{aligned}$$

In general, both an income transfer and a price effect are assumed to be present. If such an insurance contract is voluntarily purchased, equation (24) suggests that the expected net gain in utility if ill--the income transfer utility gain net of the price effect loss--is greater than the expected utility loss from paying the premium and remaining healthy. This case was described in Figure 2.

At one extreme, the entire increase in consumption ( $M^i - M^u$ ) could be due to income transfers, with no price effect. For example, the treatment protocol for a chronic disease might be so standardized that there is a negligible level of substitutability between the medical care required to treat this illness and other goods and services. This case is shown in Panel A of Figure 5. The voluntary purchase of insurance for this case would imply that the income

transfers are welfare-increasing because of the additional medical care and other goods and services that the consumer is able to purchase as a result of the income transfer. That is, the expected benefits from the income transfer if ill are positive and exceed the expected costs of paying the insurance premium if healthy.

At the other extreme, the entire increase in consumption could be caused by an opportunistic price effect, with no income transfer effect. This extreme case represents the assumptions (but not the analysis) of the conventional health insurance model. The benefit derives from the effect of the income transfer on discretionary income--purchases of other goods and services--alone, and the entire increase in consumption of medical care is a price-effect loss--“the moral hazard welfare loss.” This case is illustrated in Panel B of Figure 5. If such insurance were purchased voluntarily, the expected net gain in utility if ill would still be greater than the expected loss in utility from paying the premium and remaining healthy. However, insurance for such illness/treatment pairs is less likely to be purchased than insurance for illness/treatment pairs where the income effect dominates.

**4.3. Additional otherwise unaffordable medical care.** Finally, consider the case where  $Y^o < M^i$ . In this case, the medical care spending  $M^i$  can only occur with the income transfers from insurance, because borrowing or saving  $M^i$  dollars is assumed not to be feasible. Further, assume that the procedure is sufficiently lumpy so that without insurance,  $M^u = 0$ . For example, this case may be represented by the end-stage disease of certain internal organs, where an expensive transplant procedure is the only viable therapy for staying alive. Insurance coverage for such care would be purchased if:

$$EU_i - EU_u = \pi U^s[M^i, Y^o + (1-\pi)(1-c)M^i - M^i] + (1-\pi)U^h[0, Y^o - \pi(1-c)M^i]$$

$$\begin{aligned}
& -\pi U^s(0, Y^o) - (1-\pi)U^h(0, Y^o) \\
= & \pi\{U^s[M^i, Y^o + (1-\pi)(1-c)M^i - M^i] - U^s(0, Y^o)\} + \\
& (1-\pi)\{U^h[0, Y^o - \pi(1-c)M^i] - U^h(0, Y^o)\} \\
= & \pi\{U^s(M^i, Y^i) - U^s(0, Y^o)\} + \\
& (1-\pi)\{U^h[0, Y^o - \pi(1-c)M^i] - U^h(0, Y^o)\} > 0. \tag{25}
\end{aligned}$$

If  $c = 0\%$ , equation (25) implies that the value of insurance for these services is the expected consumer surplus of the medical procedure itself (see Nyman, 1999a).

Because of the lumpiness of the procedure, there is little if any substitutability between the procedure and other goods and services at the expenditure margin. For example, the expenditure options might be either a \$300,000 liver transplant procedure or \$0 for no care at all. Because of the lack of substitutability at the expenditure margin, the additional procedure that is consumed with insurance (but that would not be consumed without insurance) is entirely the result of an income transfer. Because the procedure is an income transfer effect and does not have a price-related welfare loss associated with it, its purchase can only be welfare-increasing.<sup>3</sup>

## 5. Contingent-Claims v. Price-Payoff Contracts

**5.1. Administrative costs.** As Figure 2 indicates, if a price effect is present, a price-

---

<sup>3</sup>Substitutability at the 0,1 procedure margin could be measured empirically by the proportion of ill persons who purchase a procedure if paid off with a lump sum. In the liver transplant example, substitutability would be measured by paying off those who were sufficiently ill to warrant a liver transplant with a check for \$300,000 and observing the percentage who then obtain the procedure. This percentage is assumed to be very high, although it has never been determined empirically.

payoff contract results in a reduction of utility compared with a contingent-claims contract that pays off with the same income transfers as does the price-payoff contract. If contingent-claims contracts yield higher utility, why do price-payoff contracts dominate the health insurance market?

The result shown in Figure 2 is based on the assumption that the premiums are actuarially fair with both types of policies. That is, the premium equals only the expected medical care expenditures in the price-payoff contract, or only the expected lump-sum income transfers in the contingent-claims contract. If the expected lump-sum transfers are equal to the expected medical care expenditures by the insurer, then the fair premiums of the two types of policies will also be equal. In practice, however, a premium would need to exceed these expected payoffs by an amount that covers the administrative and other costs of insurance. Under the price-payoff contract, these costs would include the costs of claims processing, marketing, underwriting, overhead, and normal profits. Under a contingent-claims contract, the administrative costs would include all of the above, plus additional costs for verifying illnesses, policing against fraud, and writing complex contingency contracts. Thus, it is likely that the administrative costs of the contingent-claims contract would exceed those of the price-payoff contract.

The administrative costs are lower with price-payoff contracts in part because the insurer is able to use the policyholder's physician to avoid some of the costs. For example, under a contingent-claims contract, the insurer would need to employ a physician to verify that the policyholder has the disease that he claims to have. Under a price-payoff contract, the physician implicitly verifies that the policyholder has the disease by actually performing the treatment for the claimed disease on the policyholder before the payoff is made. Also, under a contingent-

claims contract, the policyholder alone might fraudulently claim that he has a disease that warrants a \$100,000 payoff, but under a price-payoff contract, this policyholder would need to enlist the cooperation of the physician as a co-conspirator in the fraud, a much harder task. Under contingent-claims insurance, the contract would not only need to specify a schedule of payoffs for each disease but the payoffs for all the possible complications, sequellae, and courses that each disease might take. Under a price-payoff contract, the physician (and other providers) would be paid for those services that the physician (as the policyholder's disease treatment manager) deems medically necessary for treating each idiosyncratic patient/disease combination.

While using a price-payoff mechanism saves some administrative costs compared to using a contingent-claims payoff mechanism, it increases others. Foremost of these is the additional medical care that the consumer demands because he is responding opportunistically to the reduced price, that is, the price effect. Thus, the costs associated with the price effect in a price-payoff contract can be viewed as an offset to the additional administrative costs that would have been incurred under a contingent-claims contract. It may further be inferred that because price-payoff contracts dominate the market, the price effect costs are less than the additional administrative costs of the contingent-claims contract.

**5.2. Payoffs, treatment costs, and full coverage.** It is also important to note that in price-payoff contracts, the payoffs and income transfers are automatically correlated with the cost of treating the various diseases. Likewise, a contingent-claims contract with payoffs, premiums, and income transfers corresponding to the payoffs, premiums, and income transfers of a specific price-payoff contract, will also have payoffs and income transfers that are correlated with medical care expenditures for the various diseases. Under an actual contingent-claims contract,

however, any schedule of payments would need to be developed explicitly and any correlation between payoffs and treatment costs for the various diseases would need to be intentional. A perfect correlation between the cost of treating the disease and the payoff may not necessarily be desirable. For example, if it is difficult to verify the presence of certain diseases, it may be in the interest of the consumer to set payoffs for some diseases at or near zero to avoid the costs of monitoring for fraud, even though the disease has positive treatment costs. On the other hand, consumers may want payoffs set well above treatment costs for diseases with long convalescent periods and extended periods without income to cover some of the indirect costs of the disease.

Under the conventional model, the degree to which the contract provides “full coverage” is determined by the coinsurance rate alone: a coinsurance rate of 0% would indicate full coverage and any lesser rate would be less than full coverage. Under the present model, “full coverage” would be determined by both the chosen coinsurance rate and the largely exogenous probability of illness. That is, the treatment of a disease is only fully covered by the healthy if both  $c$  and  $\pi$  are (approach) zero, a different definition. If  $c$  were 0% but  $\pi$  were 1, there would be no “coverage” at all since the entire cost of the care would be paid for by the consumer up front in the premium.

## **6. Adverse Selection and Moral Hazard**

An implication of this model is that the behavior that has conventionally been identified as moral hazard may actually be adverse selection instead. That is, moral hazard has conventionally been conceptualized as an opportunistic price effect that occurs after becoming

insured. If an exogenous price change were the only mechanism for increasing consumption when insured, then a person with a chronic disease who becomes insured would purchase more care only because of the price effect. For example, suppose a consumer with myopia does not purchase designer prescription sunglasses--regarded by some as a quintessential example of a frivolous moral hazard purchase--when uninsured, but purchases them when insured at a 0% coinsurance rate. For the consumer who purchases insurance with the preexisting condition of myopia, this consumption effect is conventionally attributed to an opportunistic price effect and moral hazard under the conventional model, not adverse selection.

Under the new model, when purchasing insurance, the consumer compares the premium if healthy to the income transfers if ill. Therefore, it is clear that part of additional consumption of those who are ill and become insured is an adverse selection demand effect that is due to this comparison. Without insurance, again the consumer does not purchase designer sunglasses, but in considering whether to purchase insurance, the consumer recognizes that not all policyholders will have the condition of myopia during the contract period. If the coinsurance rate were 0% and only about 10% of the population have myopia (all of whom are assumed to purchase designer sunglasses if insured), then the consumer could purchase insurance for a fair premium representing only 10% of the cost of designer sunglasses; the remaining 90% would be paid for by income transfers from those without myopia. Thus, the consumer with myopia purchases insurance in part because of this income transfer and the additional medical care consumption that he would purchase with insurance. In contrast, if the coinsurance rate were again 0% but 100% of those in the insurance pool had myopia and purchased designer sunglasses, then 100% of the cost of designer sunglasses would be included in the myopic consumer's premium and the

consumer would no longer have a selection effect reason to purchase insurance, even though the cost at the point of purchase is \$0.

Thus, part of the increased consumption by those with insurance (compared to those without) is due to the fact that those who know ahead of time that they will benefit from the income transfers in insurance are more likely to both purchase insurance and have more medical care purchases, which is an adverse selection story. That is, even though without insurance no designer sunglasses would be purchased, the consumer with myopia self selects to purchase insurance in part because of the income transfers and the designer sunglasses he intends to purchase if insured. This effect is due to the convention that insurers sell contracts that cover some preexisting conditions, not moral hazard.

## **7. Welfare Consequences at the Margin**

Another implication of this model is that a change in the coinsurance rate, holding  $\pi$  constant, will change the size of both the price effect on medical care, and the income transfers, which will affect purchases of both medical care and other goods and services. Unless the utility gains from the income transfers can be evaluated, it will be difficult to determine the welfare implications of any coinsurance rate change or to find an optimal coinsurance rate.

In the case of an income transfer that pays for otherwise unaffordable but life-preserving care, any price effect would be negligible, and only one gain would obtain from the income transfers: the otherwise unaffordable medical procedure. As a result, the utility gain is related to the willingness to pay for the procedure, but such information on expensive procedures may not

be available. In lieu of the willingness to pay for the procedure, a measure of the willingness to pay for the *outcome* of the procedure can be used. For example, if the medical procedure for a certain illness results in an average increase of 7 additional quality adjusted life years (QALYs) and if the willingness to pay for a QALY is \$100,000, then the willingness to pay for the procedure can be measured at approximately \$700,000.

The value of an increase in the income transfer used to purchase additional *affordable* medical care and other goods and services in the ill state would be more difficult to ascertain. For those affordable medical procedures for which there is a positive consumer surplus, insurance coverage that resulted in their purchase would be welfare-increasing. Willingness to pay for these procedures could be estimated from demand analyses. These income related gains would then need to be netted against the welfare losses from the price effect and from paying the premium in the healthy state, in order to determine the net marginal gain.

## **8. The Welfare Gain from Moral Hazard**

Becoming insured increases consumption of medical care through a moral hazard effect (or an adverse selection demand effect for those who purchase insurance with preexisting conditions). According to the theory, some of the increase is due to the price effect, some to the income transfer effect, but in all cases, the additional medical care consumed is assumed to have a positive effect on the welfare of the consumer. An overall welfare loss from moral hazard would only be realized if the cost of the additional care exceeded the value of the gains because of the dominance of the price effect.

Evidence from the literature confirms that being insured does increase consumption of medical care (see Brown et al., 1998, for comprehensive review of these studies). For example, persons without insurance report fewer physician visits (Kleinman et al., 1981; Newacheck, 1988; Freeman et al., 1990; Hafner-Eaton, 1993; Newacheck et al., 1997) and may delay or forgo treatment if ill (Aday, Anderson, 1984; Freeman et al., 1987; Haywood et al., 1988; Weissman et al., 1991; Himmelstein and Woolhandler, 1995; Overpech and Kotch, 1995). The Rand Health Insurance Experiment, although comparing participants with different levels of insurance coverage rather than those with and without insurance, found that greater insurance coverage resulted in greater medical care expenditures (Newhouse, 1993). Empirical evidence also confirms that being insured reduces morbidity and mortality. Those who are uninsured arrive sicker at the hospital (Billings, 1990; Weisman et al., 1992; Weissman and Epstein, 1989; Wilson and Sharma, 1995) and are more likely to die from serious illnesses (Yergan et al., 1988; Hadley et al., 1991; Young and Cohen, 1991; Foster et al., 1992; Greenberg et al., 1988).

In particular, Franks, Clancy and Gold (1993) found being uninsured had a significant mortality effect using a large representative data set, the National Health and Nutrition Examination Survey Epidemiologic Follow-up Study (NHEFS). The NHEFS collected baseline characteristics of a sample of 6,913 adults aged 25 to 74 years and followed them for up to 16 years to determine mortality. In their analysis, the authors excluded those with Medicaid, Veterans Administration, or Medicare insurance at baseline, leaving 5,218 who reported that they were either uninsured or had private insurance. The exclusion of the Medicare participants imply that their results probably apply best to adults aged 25 through 64. The authors found that those who were reported to be uninsured at baseline survey had an increased risk of mortality

represented by a 1.25 hazard ratio (95% confidence interval of 1.55 to 1.00) compared with those who reported being insured, holding constant the other baseline characteristics.<sup>4</sup>

Using the 1.25 hazard ratio from the Franks, Clancy and Gold (1993) study, the moral hazard welfare gain from becoming insured for a year can be estimated for an arbitrary group of 40 million uninsured consumers. Assuming that these 40 million uninsured are uniformly distributed across age categories so that 1 million uninsured adults are allocated to each of the 40 age categories between age 25 and age 64, the added years of life expectancy can be estimated using the life tables. The 1996 Interpolated Abridged Life Table (Public Health Service, 2000) was used in this estimate. This Life Table reflects the experience of US residents adjusted from the 1990 census to 1996. The death rates in the table were assumed reflect the death rates of those who are continually insured, because about 85 percent of the population was insured in 1996.

For ease of computation, the differential effect of one year of insurance coverage on mortality was found by estimating the number of people who would die at each age level due to the 25% greater mortality without insurance, and multiplying this number by the insured life-expectancy at that age level. As a result of one year of being uninsured, a total of 1,225,380 life years is lost for these 40 million people, or a loss of about 9/10,000 of the total life expectancy of

---

<sup>4</sup>The Franks, Clancy and Gold (1993) study suffers from the use of a data set where insurance status is only observed at baseline. Migration in and out of being insured would likely have occurred for this population over the 16 year follow-up period. This migration, however, would interject a conservative bias into the insurance effect. That is, if the observed persons were continually uninsured or continually insured over the entire 16 year follow-up, the mortality effect would probably be greater. An improved study would observe insurance status at regular intervals, and would likely find a larger insurance effect on mortality. Moreover, if adverse selection is present, it suggests that the insured will be sicker, and this would again interject a conservative bias into the results.

the 40 million.<sup>5</sup> Thus, if it were assumed that a life year is conservatively valued at \$100,000 (Hirth et al., 2000), the value of moral hazard from being insured for 1 year is about \$123 billion.<sup>6</sup>

According to the 1996 Medical Expenditure Panel Survey, the average expenditures of those less than 65 years of age with insurance were \$1,918, and for those in the same age group without insurance were \$942, for a cost of moral hazard of about \$976 per insured person (Cohen et al., 2000). The cost of moral hazard--that is the cost of the marginal care due to becoming insured from all sources, including an adverse selection demand effect--is estimated at \$39 billion for the same 40 million. Thus, the net welfare gain from moral hazard for this arbitrary group of 40 million Americans is \$84 billion, or a welfare gain from moral hazard that is more than 3 times its cost.

---

<sup>5</sup>This analysis assumes an instantaneous change in the death rate, without any ramping up to the 25% differential, which is consistent with the 16 years of follow-up represented by the data.

<sup>6</sup>This number would be smaller if adjusted for quality of life. For example, assuming a quality of life weight of 0.66--a weight commensurate with metastatic breast cancer while undergoing chemotherapy or a migraine headache (Tengs and Wallace, 2000)--the gain would still be about 81 million quality adjusted life years (QALYs), about twice the cost. This QALY adjustment to the mortality gain, however, would be offset to a large extent by a QALY morbidity gain, attributable to moral hazard but not captured in the mortality figures.

## 9. Conclusions

To understand the relationship between health insurance and demand for medical care, it is important to recognize that private health insurance is primarily a transfer of income from those who remain healthy to those who become ill. Although there may be a price effect caused by using price reductions as the payoff mechanism, the welfare consequences of the price effects are small relative to the welfare consequences of income transfer effects, especially the role of income transfers in making accessible those expensive, life-saving procedures that would otherwise be unaffordable.

This understanding of health insurance is almost diametrically opposed to the conventional view that the relationship between health insurance and demand--that is, moral hazard--is exclusively a price effect.<sup>7</sup> Because of the emphasis on price effects in conventional theory, health economists over the last 30 or so years have tended to characterize health insurance as a source of incentives to consume too much health care, care whose value was less than the cost of producing it (Pauly, 1968). As a result of this theory, the high cost of health care in the U.S. has been generally perceived as a problem of excessive quantity. Thus, in order to reduce health care costs, U.S. economists have promoted policies--cost sharing, capitation, utilization review, managed care, etc.--that were designed to reduce inefficient consumption at

---

<sup>7</sup>One explanation for the early preoccupation with price is that the original expected utility theory of the demand for health insurance did not allow for the possibility that a lump-sum payoff could result in a larger "loss" with insurance than without (for example, see Newhouse, 1978). That is, there was the implicit assumption in expected utility theory that income elasticities for medical care were 0. If payoffs simply covered fixed losses, as they did with conventional expected utility theory, then by process of elimination, any increase in consumption due to insurance had to be a price effect.

the margin. Few U.S. health economists have focused on the plight of the uninsured, perhaps because under conventional theory, to insure them would lead to additional inefficient consumption, higher costs, and welfare losses.

The results from this analysis suggest that the conventional perception of the situation is completely backwards. The problem is not that care at the insurance/no-insurance margin is less valuable than it costs, but generally *more valuable* than it costs. Indeed, because of the high value that consumers place on their lives and health, the value of the marginal health care made available through insurance generally *far exceeds* the cost of producing that care. What are the broad-brush policy implications?

If health care costs are too high, policies should be directed at reducing the price mark-ups, rather than reducing quantity. Consider again the liver transplant example. If a liver transplant cost \$300,000 but saved 7 years of life at \$100,000 per year, its value is \$700,000. Given the 1/75,000 annual probability of needing a liver transplant, if the price of the procedure doubled to, say, \$600,000, the actuarially fair price of insurance coverage would increase from \$4.00 to \$8.00. Thus, the price of the procedure could double and the cost of insurance coverage of that procedure would still be small, generally affordable, and would represent an expected net gain to the consumer, who would not change his behavior. Even though behavior would not change, there may be welfare implications of this redistribution beyond those that have already been identified (Pauly, 1995).

Appropriate policies to counteract this problem are those that place downward pressure on prices. For example, Pfaff (1990) suggests that countries with organized buyers have lower health care costs. If so, policies where the government or a purchasing cooperative negotiates

prices with providers and health plans would hold promise. Such policies would encourage countervailing monopsony power to offset the monopoly power of providers (and health plans) in setting prices, thus reducing price mark-ups and keeping health care costs lower than they otherwise would be.

Most of all, this theory suggests that policy should be directed at insuring the uninsured. The uninsured represent about 17% of the U.S. population, or close to 40 million persons (McKay et al., 2000). If these 40 million were distributed by age approximately along the ages described in the simulation above, this analysis suggests that the benefit from moral hazard's effect on reducing mortality would likely exceed the entire cost of the moral hazard. Not counted in the above estimate are the value of the additional benefits derived from reductions in morbidity and from the truly frivolous purchases that have conventionally been associated with moral hazard by its critics. Finally, the analysis does not include the value of the additional expected benefit that would derive from the other goods and services purchased when ill that would not be available without insurance. All these gains would result in a net gain to society from policies that led to insuring the uninsured.

The policies that could accomplish this would not necessarily be expensive. For example, the regressiveness of the existing employer/employee tax subsidy represents a costless opportunity to rearrange the existing tax spending for greater gain. Instead of structuring the tax subsidy so that it increases with income, the tax subsidy could be redesigned with income limits. These limits would not reduce the number of uninsured, but they might reduce the richness of the insurance package for those large firms with high income employees. If so, the tax spending saved could then be used to encourage smaller firms that do not currently offer insurance to offer

it as a fringe benefit. Such a reworking of the tax code could be tax-spending-neutral but welfare-increasing to the extent that more hitherto uninsured workers purchased health insurance.

Finally, it should be noted that this theory and its policy implications are more difficult because they give rise to a number of equity issues. Economists have tended to focus on efficiency issues more than equity ones because they have better tools for evaluating efficiency. Equity is more difficult because it often involves interpersonal utility comparisons. But even though the light is better under the efficiency lamppost, the keys to understanding health insurance have been lost somewhere in the darker, equity sector of the analytical pavement. These equity issues--such as how to evaluate the income transfers from insurance to determine optimal coinsurance rates, how to evaluate policies that reduce the price mark-ups, and how to evaluate policies that encourage the uninsured to become insured--may require new analyses that are not yet developed. This is both the challenge and the opportunity of this new theory.

## References

- Aday, L.A., R.M. Andersen. The National Profile of Access to Medical Care: Where do we Stand? *American Journal of Public Health*. 1984;74:1331-1339.
- Billings, J., N. Teicholz. Uninsured Patients in District of Columbia Hospitals. *Health Affairs*. 1990;9:158-165.
- Brown, Margaret E., A.B. Bindman, N. Lurie. Monitoring the Consequences of Uninsurance: A Review of Methodologies. *Medical Care Research and Review*, Vol. 55 No.2, (June 1998) 177-210.
- Cohen, Joel W., S. R. Machlin, S. H. Zuvekas, et al. Health care expenses in the United States, 1996. Rockville (MD): Agency for Healthcare Research and Quality, 2000.
- Feldman, R. and B. Dowd. 1991, A New estimate of the welfare loss of excess health insurance, *American Economic Review* 81, 297-301.
- Foster, D.C., D.S. Guzick, and R.P. Pulliam. 1992. The Impact of Prenatal Care on Fetal and Neonatal Death Rates for Uninsured Patients: A "Natural Experiment" in West Virginia. *Obstetrics & Gynecology* 79 (January):40-45.

Franks, Peter, C.M. Clancy, M.R. Gold. Health Insurance and Mortality: Evidence From a National Cohort. *J of the American Medical Association*. Vol. 270(6) (August 11, 1993):737-741.

Freeman, H.E., L.H. Aiken, R.J. Blendon, C.R. Corey. Uninsured working-age Adults: Characteristics and Consequences. *Health Services Research*. 1990;24:8111-823.

Greenberg, E.R., C.G. Chute, T. Stukel, et al. Social and Economic Factors in the Choice of Lung Cancer Treatment: A Population-Based Study in Two Rural States. *New England Journal of Medicine*. 1988;318:612-617.

Hadley, J. E.P. Steinberg, J. Feder. Comparison of Uninsured and Privately Insured Hospital Patients: Condition on Admission, Resource Use, and Outcome. *JAMA* 1991;265:374-379.

Hafner-Eaton, C. Physician Utilization Disparities between the Uninsured and insured: Comparisons of the Chronically Ill, Acutely Ill, and Well Non-elderly Populations. *JAMA*. 1993;269:787-792.

Hayward, R.A., M.F. Shapiro, H.E. Freeman, C.R. Corey. Inequities in Health Services Among insured Americans: Do working-Age Adults Have Less Access to Medical Care than the Elderly? *New England Journal of Medicine*. 1988;318:1507-1512.

Himmelstein, D.U., and S. Woolhandler. 1995. Care Denied: U.S. Residents Who are Unable to Obtain Needed Medical Services. *American Journal of Public Health* 85(March):341-44.

Hirth, Richard A., Michael E. Chernew, Edward Miller, Mark Fendrick, William G. Weissert. Willingness to pay for a quality-adjusted life year: In search of a standard. *Health Economics* vol. 20, no. 3, Jul-Sep 2000, pp. 332-342.

Kleinman, J.C., M. Gold, D. Makuc. Use of Ambulatory Medical Care by the Poor: Another Look at Equity. *Medical Care*. 1981;19:1011-1029.

Manning, W. and M. Marquis. 1996, Health insurance: The tradeoff between risk pooling and moral hazard, *Journal of Health Economics* 15, 609-640.

McKay, A.P., L.A. Fingerhut, C.R. Duran. Adolescent Health Chartbook. Health United States, 2000. Hyattsville, MD: National Center for Health Statistics, 2000.

Newacheck, P.W., D.C. Hughes, and J.J. Stoddard. 1996. Children's Access to Primary Care: Differences by Race, Income, and Insurance Status. *Pediatrics* 97 (January):26-32.

Newacheck, P.W. Access to Ambulatory Care for Poor Persons. *Health Service Research*. 1988;23:401-419.

Newhouse, J. P. 1978, *The Economics of Medical Care*. Reading, MA: Addison-Wesley.

Newhouse, J.P. *Free for All? Lessons from the RAND Health insurance Experiment*.  
Cambridge, MA: Harvard, 1993.

Nyman, J. A., 1999a. The value of health insurance: The access motive. *Journal of Health Economics* 18, 141-152.

Nyman, J. A., 1999b. The economics of moral hazard revisited. *Journal of Health Economics* 18, 811-24.

Overpeck, M.D., and J.B. Kotch. 1995. The Effect of U.S. Children's Access to Care on Medical Attention to Injuries. *American Journal of Public Health* 85 (March):402-4.

Pauly, M. V., 1968, The economics of moral hazard: Comment, *American Economic Review* 58, 531-537.

Pauly, M. V., 1995, When does curbing health care costs really help the economy? *Health Affairs* vol. 14, no. 2, pp. 68-82.

Pfaff, M. Differences in Health Care Spending Across Countries: Statistical Evidence. *Journal*

*of Health Politics, Policy & Law.* 1990;15(1):1-67.

Weissman, J.S., C. Gatsonis, A.M. Epstein. Rates of Avoidable Hospitalization by Insurance Status in Massachusetts and Maryland. *JAMA.* 1992;268:2388-2394.

Weissman, J.S., R. Stern, S.L. Fielding, A.M. Epstein. Delayed Access to Health Care: Risk Factors, Reasons, and Consequences. *Ann Intern Medicine.* 1991;114:325-331.

Weissman, J.S., A.M. Epstein. Case Mix and Resource Utilization by Uninsured hospital Patients in the Boston Metropolitan Area. *JAMA.* 1989;261:3572-3576.

Wilson, B.E., and A. Sharma. 1995. Public Cost and Access to Primary Care for Hyperglycemic Emergencies, Clark County, Nevada. *Journal of Community Health* 20 (June):249-56.

Yergan J., A.B. Flood, P. Diehr, J.P. LoGerfo. Relationship between patient source of Payment and the Intensity of Hospital Services. *Medical Care.* 1988;26:1111-1114.

Young, G.J., B.B. Cohen. Inequities in Hospital Care, The Massachusetts Experience. *Inquiry.* 1991;28:255-262.

Zweifel, P. and F. Breyer, 1997. Health economics. New York: Oxford.