

NBER WORKING PAPER SERIES

VALUE BASED COST SHARING MEETS THE THEORY OF MORAL HAZARD:
MEDICAL EFFECTIVENESS IN INSURANCE BENEFITS DESIGN

Mark V. Pauly
Fredric E. Blavin

Working Paper 13044
<http://www.nber.org/papers/w13044>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2007

The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by Mark V. Pauly and Fredric E. Blavin. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Value Based Cost Sharing Meets the Theory of Moral Hazard: Medical Effectiveness in Insurance Benefits Design

Mark V. Pauly and Fredric E. Blavin

NBER Working Paper No. 13044

April 2007

JEL No. I11

ABSTRACT

The conventional theory of optimal coinsurance rates in health insurance in the presence of moral hazard indicates that, in situations of equal risk characteristics, coinsurance should vary if the price-responsiveness or price-elasticity of demand for different medical services varies, and should be larger for the more price responsive services. An alternative theory called "value-based cost sharing" indicates that coinsurance should be lower for services with higher (marginal) benefits relative to costs. This paper reconciles the two views. It shows that, if patient demands are based on correct information on benefits and costs, the conclusion of the conventional view is identical to the conclusion from the value-based approach. If patient demands differ from correct demands, it is shown that optimal coinsurance depends both on the extent and direction of information imperfection and on price-responsiveness or price elasticity. The paper also shows, as an alternative to adjusting coinsurance to deal with information imperfection, that providing better information which affects patient demands can be superior if uninformed patient demands exceed informed patient demands, but value based cost sharing can be superior to providing information (even if the cost of information is minimal) when patient demands fall short of informed demands. An extended numerical example illustrates these points.

Mark V. Pauly

Health Care Systems Department

University of Pennsylvania

208 Colonial Penn Center

3641 Locust Walk

Philadelphia, PA 19104-6218

and NBER

pauly@wharton.upenn.edu

Fredric E. Blavin

Ph.D. Candidate

Health Care Systems Department

The Wharton School

University of Pennsylvania

Colonial Penn Center

3641 Locust Walk

fblavin@wharton.upenn.edu

Introduction.

Almost all private insurances in the United States leave some fraction, usually modest, of spending on covered services unreimbursed. While the proportion of personal health expenditures American consumers pay out of pocket has fallen slightly over time, the absolute dollar amount of such cost sharing has risen, along with unit prices and total spending. The advent of tax-subsidized high deductible insurance linked to health savings accounts has heightened interest in many observers, and concern in some, about the effects of cost sharing in voluntary insurance.

Compared to full coverage of expenses, out of pocket payments have two effects, one certain and one highly likely. Such payments expose the insured to higher financial risk (compared to full coverage) if above-average spending or costly illness occurs. They also provide incentives for lower levels of medical care use. The basic theory of insurance design when there is a tradeoff between risk protection and moral hazard is well known (Pauly, 1968; and specifically, Zeckhauser, 1970). The fundamental idea is that consumers should be willing to sacrifice a modest amount of financial protection in order to experience stronger incentives to avoid the use of care of low value relative to its price. Incentives are necessary because insurers are incapable of determining exactly what losses are because they cannot tell precisely what illness has occurred (Pauly, 2007). However, recent research and policy discussion has raised anew the question of whether and how the still-limited but growing evidence base for medical care might be used in the design of insurance. Specifically, is evidence of medical effectiveness useful for the design of benefits that are efficient? An interesting and important recent analysis suggests a positive answer. First under the label of “benefits-based cost sharing” and then under the new term “value-based insurance design,” Mark Fendrick and Michael Chernew have proposed

“tailoring copayments at the point of service to the evidence-based value of specific services for targeted groups of patients” (2006, p. 17; Fendrick et al., 2001). A more recent elucidation says that coverage design should depend on costs as well as medical benefits, that is, on the cost-quality tradeoff in a particular set of clinical interventions (Chernew, Rosen, and Fendrick, 2007).

To our knowledge, no one has spelled out the differences, if any, between the levels of cost sharing that would have been specified as (second-best) optimal under the two theories. If marginal benefit or demand curves are or should be based on evidence, will the optimal level of coinsurance indicated by the theory of moral hazard differ from that under a value-based approach? Reconciling the two approaches will also require describing more carefully what measure of “evidence-based value” might be appropriate for that theory. While the link between theory and practice is never exact, there has been some interest in and even some appeal to benefits or value based cost sharing by some employers and insurers. In this paper we explore the question of whether and when taking such value into account might make a difference for efficient insurance design.

We show that if consumers and the physicians they consult both base demand and advice on the same information about medical effectiveness, ideally as accurate as the state of knowledge and evidence allow, the two approaches (properly defined) give the same message about optimal benefit design. However, there are some informal descriptions of the value based theory which, if used, would incorrectly suggest differences in such cases. In these full information models, coinsurance levels do not depend only on evidence about medical benefits and, moreover, to the extent that they do depend on such evidence, it is not the typical single summary measure of average or marginal effectiveness often invoked by evidence-based guidelines. Rather, the key

information is full schedule of marginal benefits and, to the extent that this schedule varies across consumers based on health status, income, or health preferences, on that variation as well.

More specifically, the primary determinant of the optimal level of coinsurance is, in the words of Phelps (2003, p. 328), the person-level “*price responsiveness*” of the demand or marginal benefit curve under full information, or, more elaborately, the configuration of the curve, and not any specific single measure or concept about the marginal effectiveness of medical care per se. If the demand or marginal benefit curve is linear, the measure of price responsiveness is its slope; if the curve is constant elasticity, it is the elasticity. In the diagrammatic and numerical analysis that follows we will use linear demand curves, but the conclusions will carry over to constant elasticity curves as well.

However, if information used by consumers (or individual physicians) is less perfect than that available to planners, optimal cost sharing will be affected. Optimal coinsurance will then depend on the extent and nature of the imperfect information, but it will still depend on demand or marginal benefit price responsiveness. Specifically, if actual demand curves fall short of true curves, coinsurance should be lower than under fully informed demand, but should still vary with price responsiveness. Optimal coinsurance levels relative to full information reverse if actual demand curves exceed fully informed curves, and price responsiveness still matters. However, the rationale for altering benefits in all cases is based on the optimal level of risk protection; in a paradoxical sense, when the dust clears, medical benefits per se ultimately need not matter in benefit-based cost sharing.

Clearing the underbrush.

We want to focus on insurance design in situations where the primary effect is purely medical, in the sense of better health per se. Consequently, the question of optimal cost sharing when one type of care produces cost offsets, which has been treated at length elsewhere (Pauly and Held, 1990), will be ignored. Of course, virtually any care that improves subsequent health probably also impacts subsequent medical spending, but that spending effect will not be considered. This also means that the empirical experience of some benefits managers who have emphasized the alleged ability of selectively lower cost sharing for certain services to lower overall spending (Newhouse, 2006) is not directly relevant to the contrast between theories that we want to draw.

We will also assume that “optimal insurance” means insurance that maximizes consumer welfare (generally in the sense of expected utility) so we will not be directly concerned about apparently large effects on physical measures of health except as they affect expected utility, nor will we attach a value to “containing cost growth” or cost levels per se. Finally, and more substantively, we will interpret the Chernew-Rosen-Fendrick definition of “the cost-quality tradeoff” (2007) as equivalent to basing (relative) value on cost-effectiveness ratios. One implication of this is that what is potentially relevant is not a comparison of “interventions with the highest value or effectiveness” with interventions with “little or no proven health benefit.” Cost needs to be considered also. If an intervention of the first sort is very expensive and one of the second sort is very cheap but has a little health benefit, coinsurance might be the same for both. Moreover, if costs change (e.g., if an effective medical device becomes more expensive), then cost-sharing should probably change as well. Finally, we will interpret the value-based view as not limiting coverage only to services whose marginal benefits always and everywhere

exceed their cost; (second best) optimal insurance in the presence of moral hazard may involve accepting use of services whose value is less than cost.

We will model cost sharing only as proportional coinsurance, not considering copayments, deductibles, upper limits, and stop-loss features often present in real world insurance plans. Partly this is for simplicity, but also because it corresponds with the classic treatment of moral hazard.

Most of the Zeckhauser Model in a picture.

The classic statement of optimal coinsurance in health insurance was provided by Richard Zeckhauser in 1970, in a paper that emphasized the desirability of departing from uniform cost sharing (ignoring the administrative costs of doing so) if care for different illnesses provided different configurations of marginal benefit. Figure 1 shows two demand or marginal benefit curves for two different illnesses, A and B, for a representative consumer. Assume that the unit price of care is normalized to \$1, and also assume that initially the coinsurance is set at the level C_1 . Now suppose we consider lowering the coinsurance rate to C_2 . Because this reduction triggers a larger increase in quantity for illness A, it is clear that the addition to welfare cost is larger for illness A (at MWC_A) than for illness B (at MWC_B).

Lowering the coinsurance rate will also increase the risk protection provided by the insurance, and the consumer will place a positive value on this. Assume that the variance of medical expenses is the same for both illnesses, so that the marginal risk premium associated with the increase in coverage is the same for both. The initial coinsurance rate will be the optimal rate if the increase in welfare cost just equals the marginal risk premium. It is clear that if this equality was satisfied for illness B, it will not be satisfied for illness A. The optimal

coinsurance rate for illness A, the more price-responsive demand or marginal benefit curve, will be higher than that for illness B. The reason why this is so is because the higher marginal welfare cost for A compromises the consumer's desire for risk protection; alternatively, we could say that B can be "better covered" in terms of risk protection because doing so causes less of an increase in use of care whose marginal benefit is less than its marginal cost.

This is the standard economic model of optimal coinsurance. If these demand curves also represent marginal benefit curves, how would they be characterized in the value-based approach? We first need to face a watershed difference in perspective. The moral hazard approach is intended to specify coinsurance that will lead to the optimal quantity of care. Simple value based approaches do not view coinsurance as determining different amounts of some service that a person might use, but rather as determining whether the person uses some service (for example, filling a prescription for an anti-hypertension medication) at all (with the quantity assumed to be predetermined, once the person has decided to use). That is, in the simple value based approach, there is no role for more or less; the only question is whether or not. So to have a fair comparison, we need to specify a value based approach in which marginal values vary with quantity of a given service.

Here is why. In the value based approach sometimes people are thought to differ in their benefit from the predetermined quantity (rather than all being average), but more typically services are either characterized as "high value" (for all, at the predetermined use protocol) or low value. In such a world, variation in coinsurance plays no role: coinsurance should either be zero or 100%, depending on the cost effectiveness of the intervention; there is no moral hazard in this simple value based view, because moral hazard only occurs if some service is high value for some patients at some quantities but low value for other patients at other (larger) quantities.

Rather, in the simple view coverage should be full for interventions whose cost effectiveness value is less than a target threshold, and zero for all others.

To generate discussion, we therefore first re-interpret the value based approach from the quantity-variation perspective, recognizing that this may not be how it is usually conceptualized in the informal literature. The most obvious point then is that, when there is diminishing marginal benefit associated with quantity variation, there is no simple way to rank care for each illness as having higher or lower “proven effectiveness.” For either illness in the diagram, the marginal effectiveness of care is high at low quantities and low at high quantities. There are no universally high value or low value services; value depends on quantity.

The best way to deal with this problem is to think of a standard cost effectiveness model with multiple possible interventions. For example, one might compare screening programs for cancer that screen at different intervals. The marginal benefit then is the change in benefit from a (small) reduction in the interval between screenings, and one can calculate the incremental cost effectiveness ratio from going from a given program to the next more frequent program.

This incremental or marginal formulation challenges some of the typical language. For example, at uniform coinsurance C_1 and the quantity associated with it, the marginal effectiveness of care (per dollar of cost) is the *same* for both illnesses when the consumer is in well-informed equilibrium; cost-effectiveness will not differ. Care may not have the same average effectiveness (per dollar of cost) across illnesses under uniform coinsurance, but it should have the same marginal effectiveness. At coinsurance rates which are not trivially low, informed patients will not be consuming care of “little or no proven health benefit.” (They might be consuming care of high potential health benefit that is currently unproven beyond a shadow of a doubt by evidence, if the expected benefit is high enough, however.)

Even when we use the same language, however, the two approaches still differ. If the coinsurance rates are set optimally for these two illnesses, the *marginal* cost-effectiveness ratios will differ. With higher coinsurance rates for illness A, its marginal cost effectiveness ratio will be higher than that for care for illness B. But this divergence in cost effectiveness ratios will be optimal as far as consumers are concerned, since lowering the coinsurance rate for A would cause a larger welfare cost for that illness; the improved health that would result would not be worth the additional resource cost that would have to be incorporated into the insurance premium and the out of pocket payment. In other words, cost effectiveness information or measures alone are not sufficient to specify optimal insurance coverage.

One might also look at the average cost effectiveness ratios for illnesses A and B; with the demand curves as drawn it is higher for illness B than for A. So it would in this example be correct to say that the optimal coinsurance rate should be lower for the illness with the higher average ratio.

However, this correlation need not occur. Consider demand curve for illness C in Figure 2; this curve is drawn as having the same slope as the demand curve for illness B. But note that marginal benefit is higher at every quantity for illness C than for illness B. However, as can easily be seen, the marginal welfare cost that would be associated with this illness at C_1 is the same as that for illness B. If the marginal risk premium was also the same, the optimal coinsurance rate would be the same for both illnesses even though the average benefit at any quantity is not. Of course, at the same coinsurance rate the marginal benefits per dollar would be equalized; optimality requires using a larger quantity of the more beneficial service—but uniform coinsurance in this case would bring that about.

To sum up: if consumer demand curves coincide with true (or, if evidence is sufficient, evidence-based) marginal benefit curves, the standard economic theory will still be correct even if value-based considerations are incorporated. It will, however, not be identical with a model that says that coinsurance should be determined solely by marginal (or average) evidence-based medical benefit.

Adding imperfect information.

Now we assume that patient-consumers do not initially have demand curves that map the marginal benefit curves generated by evidence. Patients might not have been fully informed by their physicians, or they might not understand the information they have been given. Such imperfection in physician behavior might lead patients to use demand or marginal benefit curves that either fall short of or exceed the true demand curves. There are many ways in which imperfect information can affect demand or perceived marginal value curves. Let us initially assume that the uninformed curves have the same slopes as the true curves, but are shifted to the right or left.

So in figure 3 assume that the illness A and B demand curves are shifted to the left as shown, to D'_A and D'_B ; patients underestimate marginal benefits at every quantity. At the full information optimal level of coinsurance, but with the new imperfect information demand curve, the marginal welfare cost is now less than before, at MWC'_B . If we assume that the marginal risk premium remains the same, it is clear what ought to happen: the optimal level of coinsurance for either illness A or illness B should fall. However, because demand shortfall leads to smaller quantities demanded (and therefore lower losses, relative to the perfect health state), the marginal risk premium will also be lower. It is plausible (though not guaranteed) that the marginal

welfare cost will fall by more than the marginal risk premium. We will assume here that this is so.

The reason why coinsurance can be lowered is that doing so now causes less marginal welfare cost (e.g., MWC'_B) at every quantity than in the full information case. That is, the reason to lower coinsurance is not to move use to a particular cost-effectiveness level, but rather because lower welfare cost at the margin allows insurance to do a better job at its main function, spreading risk. The positive impact on access and health is an incidental benefit of taking advantage of a reduction in cost of risk protection. Of course, the higher marginal benefit from the units whose use is now induced is what produces that cost reduction, so one could also give a cost effectiveness explanation.

How far should coinsurance fall? Not by as much as would push medical care use up to the level that would have been demanded with informed demand and optimal coinsurance, or level C3, but somewhere in between C3 and C1, depending on what the marginal risk reduction curve looks like.

For illness A, although the results are not shown, it is easy to conclude that its optimal coinsurance rate should probably also move down. It is likely that the new coinsurance rate for illness A will remain higher than the new rate for illness B as long as imperfect information shifts both marginal benefit curves to the same extent. So, as before, marginal medical effectiveness relative to cost only provides part of the information needed to determine the appropriate value based coinsurance rate. That rate also depends on the slope or elasticity of the marginal benefit curves, and the form of the person's utility function in disposable income, which determines the marginal risk reduction curve.

The argument in situations in which information imperfection leads to a demand in excess of the perfectly informed demand is symmetric. Coinsurance rates probably should rise to choke off welfare cost, but should still depend on slopes and utility functions as well as on the extent of information imperfection.

Pay to improve information?

Given the optimality and equivalence properties associated with the fully informed demand or marginal benefit curve, it would seem that it would be desirable to provide information to consumers to move closer to those curves if the cost of doing so is not too high. But this conclusion may not follow. Suppose that, absent costly information, patient demand falls short of the truth as in the “primed” curves. Then coinsurance rates as noted can be lowered, and more protection against risk provided. Yet the quantity can still be driven to a more efficient level, and without spending any real resources informing patients. That is, for illnesses or types of care where people tend to underuse because of imperfect information, that imperfect information itself reduces the effects of moral hazard and so should not be removed, certainly not at positive resource cost. It is salutary ignorance.

But now the case of misinformed overuse is not symmetrical. In the case of overuse, if information is not very costly, it may be better to provide information to shrink demand rather than deter overuse by exposing people to higher coinsurance and more risk. That is, it is wasteful to tell people about underuse as long as demand shortfalls relative to the truth are not so large that they cannot be moved to the optimal level by zero or higher cost sharing. (Negative cost sharing also adds risk.) In contrast, it may be worthwhile to counsel patients away from

overuse rather than force them to put more skin in the game as an alternative. Leave underusers in the dark, but inform overusers.

A classic numerical example.

To illustrate the effect of patient misperception on the optimal coinsurance rate, we present a numerical example based on the extended illustration of the optimal coinsurance rate in Phelps' textbook, *Health Economics* (2003). Phelps discusses the impact of insurance coverage with an assumed 20% coinsurance rate for a hospitalization insurance that provides coverage for the demand curves and loss probabilities indicated in Table 1 (p. 362). He also assumes that consumers have a risk aversion parameter $r = -0.0003$. He shows the total welfare cost of this level of coverage (assuming linear demand curves), compared to no coverage, and the net risk premium for that level of coverage, assuming that insurance has a 10% loading. The main point he makes is that the total welfare cost, at \$80, is less than the net risk premium, of \$136, so the person would be better off buying insurance with 20% coinsurance than remaining uninsured. But 20% is not the optimal coinsurance rate in this example. (While Phelps provides a formula for calculating the optimal coinsurance rate, his formula assumes a constant elasticity—rather than a linear—demand curve, and so is not applied to this problem.) The unit price is \$500.

To illustrate the influences that determine the optimal coinsurance rate in the linear demand curve example, we first assume that the demand curves are correct reflections of the marginal benefit from inpatient care, and show what the optimal coinsurance rate is. Then we assume that patient information imperfection leads to a particular backward shift of the evidence-based demand curve, and show how the coinsurance rate should be altered to compensate. Table 2 shows the total and marginal welfare cost, and the total and marginal risk premiums, for the

insurance policy and demand curves in Phelps' example. Beginning at 100% coinsurance, the marginal welfare cost is greater than the marginal risk premium. As coinsurance is reduced, the former rises and the latter falls, and the two become equal at a coinsurance rate of about 0.46. An assumption here is that "efficient" physicians will not recommend quantities in excess of that which would be demanded by an informed patient at the optimal coinsurance rate, since at that rate the marginal benefit from additional use falls well short of the marginal cost. There should thus be no noncompliance if physicians provide recommendations based on perfect knowledge and (identical ex ante) patients make their decisions based on that perfect knowledge, the optimal coinsurance rate, and their own values for health and financial risk.

Table 3 shows the comparable risk premia and welfare costs for a demand curve shifted backwards by imperfect information by 0.5 days in both illness events. The welfare cost at any coinsurance rate c^* along the new demand curve is now the welfare cost from the original demand curve at $c + 0.4$. For example, the welfare cost at $c^* = 0.2$ is the welfare cost at $c = 0.6$, or 20, and the marginal welfare cost is now 4.7, not the 9.7 as in the full information case. In this example, it turns out that the optimal coinsurance rate is between 0.3 and 0.35, substantially less than the perfect information optimal coinsurance rate of 0.46.

Note that imperfect information shifts down the marginal welfare cost curve, but it also shifts down the marginal risk premium curve. In the example, and in most plausible cases, the downward shift in the former is greater than the downward shift in the latter, so the optimal coinsurance rate falls. (It would be possible to think of utility functions where this need not occur, but they would be unusual.) Note also that the new optimal coinsurance rate falls between the original value of 0.46 and the coinsurance rate that would produce the same level of use as at the no-ignorance optimal rate, given imperfect information, which is about 0.1. The reason why

the coinsurance rate does not fall this far is because the total and marginal financial risk reduction is lower along the lower demand curve.

There is, however, a potential problem in simply saying that the employer or insurer should offer the plan with the optimal coinsurance rate. Misinformed consumers who judge insurance relative to their perceived marginal benefit curve will not find a policy with a coinsurance rate of 0.33 to be the one they find most attractive. Rather, if they judge coverage by using their own curves (and the insurance premiums associated with those curves), they would prefer a coinsurance rate of 0.55. It could even happen (though it does not happen in this example) that misinformed consumers could think themselves better off with no insurance at all than with insurance that tricks them into doing what will turn out to be the right thing, but is ex ante perceived to be of less value relative to its cost. The employer or insurer could offer only the optimal plan, and refuse to offer plans consumers think they like better, but that might be awkward. If consumers could be convinced that their own marginal benefit curves are wrong, they might go along with this paternalism—but they might not be easy to convince or, if convinced, might then change their demand curves back closer to the true ones with a higher optimal coinsurance rate.

One final illustration.

We now make demand more price responsive, setting quantities at 20% coinsurance at 5 and 10 respectively. As might be expected, the optimal coinsurance rate (calculations not shown) now is higher, at 0.64, compared to the 0.46 for demand curves with less steep slope. Next we impose the same ignorance-induced backward shift in the demand curve of 0.5 days at every coinsurance rate. The optimal rate now becomes 0.55 (versus 0.33), making the point that even

with identical magnitudes of consumer misconception or noncompliance, the optimal coinsurance rate depends on the price responsiveness of demand. However, the change in the optimal coinsurance rate generated by the same amount of consumer ignorance is smaller in this more responsive case, because the size of the coinsurance reduction needed to push the quantity of care a given amount back closer to the optimum is smaller. But the optimal coinsurance rate from the consumer's point of view is now about 0.7, so the divergence between the socially optimal and privately optimal rates grew.

People differ.

Thus far we have implicitly been using a model in which consumers are identical in terms of the marginal benefit schedules for health and their levels of risk aversion. If this uniformity is not present (in either dimension), optimal coinsurance levels should differ. To the extent that people attach different values to given health outcomes and different values to financial risk, they may not prefer coinsurance rates that reflect only evidence based medicine. What is relevant in terms of health values is not, however, everything that affects clinical outcomes. If new evidence caused a uniform shift in a given marginal benefit or marginal value curve, holding its slope constant, assuming that the risk premiums do not change, coinsurance rates should stay the same. That is, people who are sicker, or who attach a higher value to health, should face the same coinsurance rates as others. At those rates they will use more care, and that additional use should be reflected in higher premiums. But if people with different preferences can be identified, coinsurance need not be affected. It is a change in preferences that affects the slope or elasticity of marginal health benefits—the extent to which marginal benefits decline with more health and more use—that matters.

We believe we know very little about how the shape of demand curves varies across people, if it does vary. Consideration of variations in financial risk aversion also matter. The person with low levels of risk aversion who is highly responsive to out of pocket payments should have high values of coinsurance regardless of the level of medical benefits.

Manipulating demand for care and the demand for insurance.

The final question, as suggested in the numerical example, is whether consumers would prefer and choose insurances which used value based cost sharing to correct overuse and underuse. Making these corrections increases what we might call their ex post expected utility. But if consumers are making decisions about insurance having their misinformed demand curves in mind (and believing them to represent true marginal benefits) they may not prefer insurance that makes them override their current best judgment, wrong as it may be. Of course, if the reason for underuse is known and acknowledged reluctance to do what one ought (take a medicine with side effects to prevent an asymptomatic disease), the split-personality consumer might prefer this kind of insurance. Seeing how people would feel about such insurances—policies that generally would lead to better health relative to spending after the fact, but may not seem so before the fact—is part of our ongoing research.

Conclusion.

The only reason to use nonzero coinsurance is to control moral hazard. Moral hazard only occurs if the insurer cannot precisely tell what illness (loss) state the person is in. When and if the evidence base for the marginal benefit of medical treatment associated with particular patients at different illness levels ever becomes complete, cost sharing can be dispensed with

entirely; insurance will pay in full for efficient care and pay nothing for care whose benefit falls short of cost. But with today's imperfect medical knowledge, coinsurance has a role to play.

Measures of medical professional judgments on effectiveness are part of the information needed to develop more targeted cost sharing (assuming more variation is administratively feasible), other influences will matter, in addition to such evidence. The shape of the marginal benefit curve, the consumer's risk aversion, and the effectiveness of provision of more information, and the cost of the intervention all will be important.

References

- Chernew, M. E., A. Rosen, and A. M. Fendrick, 2007, "Value-based Insurance Design," *Health Affairs* 26.2, w195-w203.
- Fendrick, A. M., and M. E. Chernew, 2006, "Value-Based Insurance Design: A 'Clinically Sensitive' Approach to Preserve Quality and Contain Costs," *American Journal of Managed Care* 12.1, 18-20.
- Fendrick, A. M., D.G. Smith, M. E. Chernew, and S. N. Shahet, 2001, "A Benefit-Based Copay for Prescription Drugs: Patient Contribution Based on Total Benefits, Not Drug Acquisition Cost," *American Journal of Managed Care* 7.9, 861-869.
- Newhouse, J. P., 2006, "Reconsidering the Moral Hazard-Risk Avoidance Tradeoff," *Journal of Health Economics* 25.5, 1005-1014.
- Pauly, M. V., 1968, "The Economics of Moral Hazard," *American Economic Review* 58.3, 531-537.
- Pauly, M. V., 2007, "Adverse Selection and Moral Hazard: Implications for Health Insurance Markets," working paper.
- Pauly, M. V., and P. J. Held, 1990, "Benign Moral Hazard and the Cost-Effectiveness Analysis of Insurance Coverage," *Journal of Health Economics* 9.4, 447-461.
- Phelps, C. E., 2003, *Health Economics*, 3rd ed. Boston: Addison Wesley.
- Zeckhauser, R. J., 1970, "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory* 2.1, 10-26.

FIGURE 1

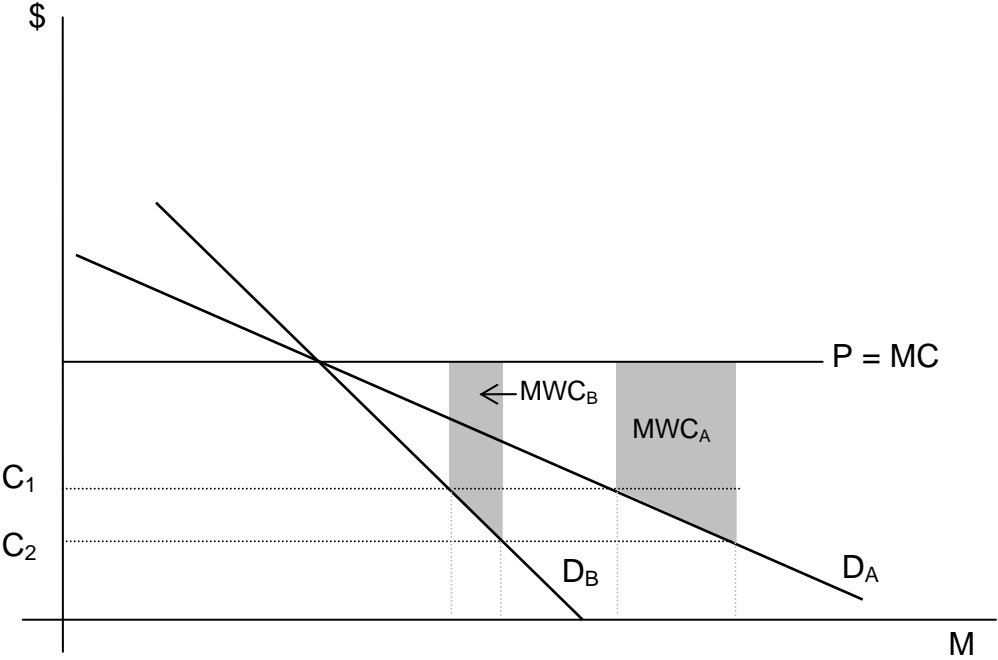


FIGURE 2

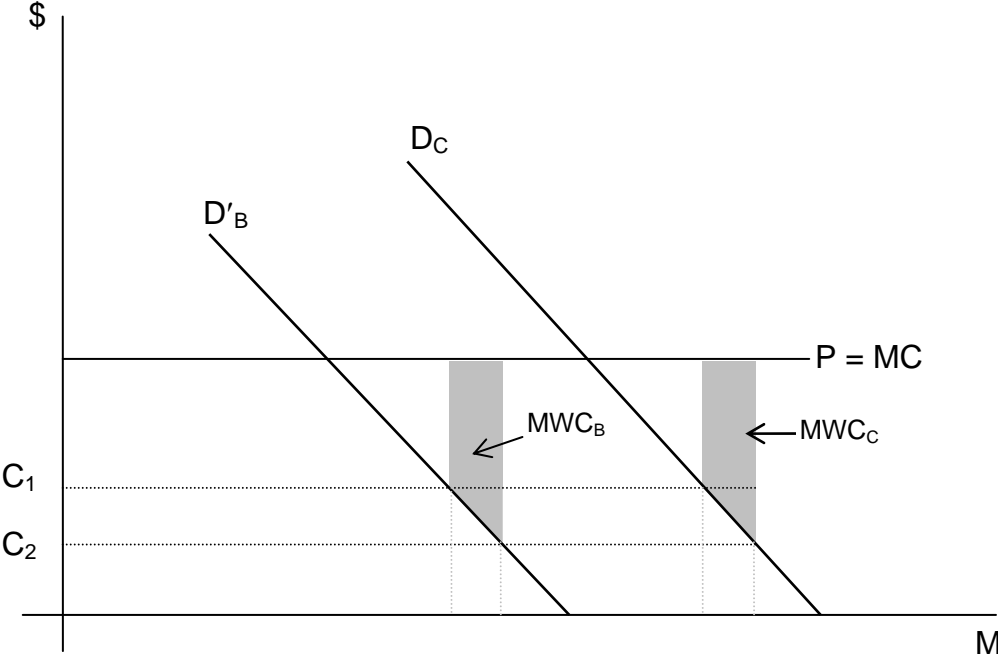


FIGURE 3

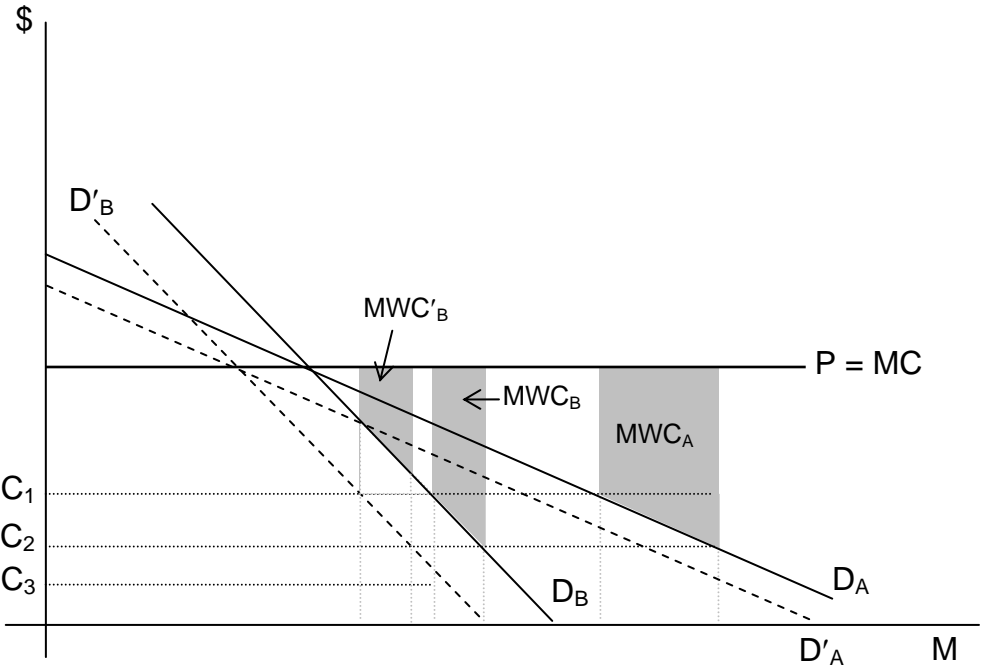


TABLE 1

A Detailed Calculation of Welfare loss

Probability	Uninsured Demand	Insured Demand
0.3	3.0	4.0
0.1	8.0	9.0

TABLE 2

Replication of Phelps Example with 10% Loading
Marginal Welfare Cost, Marginal Risk Premium, and Optimal Coinsurance Rate

C	Moral Hazard Welfare Loss	Marginal Welfare Cost	Risk Premium	Marginal Risk Premium	Net Gain
1	0.0	0	0.0	0.0	0.0
0.95	0.3	0.3	14.0	14.0	13.7
0.9	1.3	0.9	27.6	13.6	26.4
0.85	2.8	1.6	40.7	13.1	37.9
0.8	5.0	2.2	53.2	12.5	48.2
0.75	7.8	2.8	65.1	11.9	57.3
0.7	11.3	3.4	76.3	11.2	65.1
0.65	15.3	4.1	86.8	10.5	71.5
0.6	20.0	4.7	96.5	9.7	76.5
0.55	25.3	5.3	105.2	8.8	79.9
0.5	31.3	5.9	113.1	7.8	81.8
0.45	37.8	6.6	119.9	6.8	82.1
0.4	45.0	7.2	125.6	5.7	80.6
0.35	52.8	7.8	130.2	4.6	77.4
0.3	61.3	8.4	133.5	3.3	72.3
0.25	70.3	9.1	135.5	2.0	65.2
0.2	80.0	9.7	136.1	0.6	56.1
0.15	90.3	10.3	135.3	-0.9	45.0
0.1	101.3	10.9	132.8	-2.4	31.6
0.05	112.8	11.6	128.7	-4.1	15.9
0	125.0	12.2	122.9	-5.8	-2.1

Net gain is maximized at C=.46

TABLE 3

**Re-Calculation of Optimal Coinsurance Rate, Backwards Shift in Demand by .5 days
Marginal Welfare Cost, Marginal Risk Premium, and Optimal Coinsurance Rate**

With 10% Loading Fee					
C	Moral Hazard Welfare Loss	Marginal Welfare Cost	Risk Premium Gain	Marginal Risk Premium	Net Gain
1	N/A	N/A	0.0	0.0	N/A
0.95	N/A	N/A	11.5	11.5	N/A
0.9	N/A	N/A	22.7	11.2	N/A
0.85	N/A	N/A	33.4	10.7	N/A
0.8	N/A	N/A	43.7	10.3	N/A
0.75	N/A	N/A	53.5	9.7	N/A
0.7	N/A	N/A	62.6	9.2	N/A
0.65	N/A	N/A	71.2	8.5	N/A
0.6	0.0	0.0	79.0	7.9	79.0
0.55	0.3	0.3	86.2	7.1	85.9
0.5	1.3	0.9	92.5	6.3	91.2
0.45	2.8	1.6	97.9	5.5	95.1
0.4	5.0	2.2	102.5	4.5	97.5
0.35	7.8	2.8	106.0	3.5	98.2
0.3	11.3	3.4	108.5	2.5	97.2
0.25	15.3	4.1	109.8	1.3	94.5
0.2	20.0	4.7	109.9	0.1	89.9
0.15	25.3	5.3	108.8	-1.1	83.5
0.1	31.3	5.9	106.3	-2.5	75.0
0.05	37.8	6.6	102.3	-3.9	64.5
0	45.0	7.2	96.9	-5.5	51.9

Net gain is maximized at C=0.33