Journal of Health Economics 57 (2018) 168–178

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Decomposition of moral hazard

John A. Nyman[a,*], Cagatay Koc[b], Bryan E. Dowd[a], Ellen McCreedy[c], Helen Markelova Trenz[d]

[a] *Division of Health Policy and Management, School of Public Health, University of Minnesota, 420 Delaware St. SE, Box 729, Minneapolis, MN, 55455-0392, United States*
[b] *Cornerstone Research, 1919 Pennsylvania Avenue, N.W., Suite 600, Washington, D.C., 20006-3420, United States*
[c] *Center for Gerontology and Healthcare Research, Brown University, School of Public Health, 121 South Main Street, Suite 6, Providence, RI, 02903, United States*
[d] *Optum Labs, 11000 Optum Circle, Eden Prairie, MN, 55344, United States*

ARTICLE INFO

Article history:
Received 23 January 2017
Received in revised form
25 September 2017
Accepted 7 December 2017
Available online 11 December 2017

Keywords:
Moral hazard
Health insurance income effects
Efficient moral hazard

ABSTRACT

This study seeks to simulate the portion of moral hazard that is due to the income transfer contained in the coinsurance price reduction. Healthcare spending of uninsured individuals from the MEPS with a priority health condition is compared with the predicted counterfactual spending of those same individuals if they were insured with either (1) a conventional policy that paid off with a coinsurance rate or (2) a contingent claims policy that paid off by a lump sum payment upon becoming ill. The lump sum payment is set to be equal to the insurer's predicted spending under the coinsurance policy. The proportion of moral hazard that is efficient is calculated as the proportion of total moral hazard that is generated by this lump sum payment. We find that the efficient proportion of moral hazard varies from disease to disease, but is the highest for those with diabetes and cancer.

Introduction

The concept of moral hazard has played a central role in U.S. health policy since Pauly introduced it into the health economics literature 50 years ago (Pauly, 1968). It debuted as a counter-argument to Arrow's famous 1963 evaluation of the welfare effects of health insurance derived from Friedman and Savage's risk avoidance model (Arrow, 1963; Friedman and Savage, 1948). Instead of health insurance being overwhelmingly welfare-increasing (under the right circumstances) as Arrow had suggested, Pauly warned that the welfare loss from moral hazard could be so large as to render health insurance welfare-reducing (Pauly, 1968). Analysts who adopted Pauly's model concluded that moral hazard generated so much inefficiency that coinsurance rates should be raised dramatically (to 67% or greater according to Feldstein, 1973, or to about 45% across-the-board according to Manning and Marquis, 1996) in order to ensure that insurance increases welfare. Apparent empirical support of this theory and policy came from the influential RAND Health Insurance Experiment that found that large increases in cost-sharing could be implemented causing substantial decreases in healthcare spending and utilization, but with no important impact on health (Newhouse, 1993).[1]

In Pauly's model, the additional health care consumed when insured was viewed solely as the product of a price distortion (Pauly, 1968). His model did not recognize that the price reduction (that is, the change from the market price to the coinsurance rate stipulated in the health insurance contract) alternatively could represent a vehicle for transferring income from those who purchase insurance and remain healthy, to those who purchase insurance and become ill. It also did not recognize that this income transfer could cause the purchase of additional health care, resulting in a welfare gain. While Marshall (1976) and de Meza (1983) observed that such income effects were possible if insurance paid off with a direct income transfer, neither recognized that an insurance price reduction effectively represented a similar income transfer and could produce a similar income effect.

[1] RAND findings have been challenged by those who point out that attrition rates were 16 times larger in the cost-sharing plans than in the free plan (Nyman, 2007; Aron-Dine et al., 2013). Differential attrition rates suggest that participants who became ill dropped out of the cost-sharing arms in order to revert to their more complete pre-RAND insurance coverage. A lack of ill patients in the cost-sharing arms would help explain the finding that the remaining participants in those arms could receive substantially less health care, especially fewer hospitalizations, but with no important effects on health.

* Corresponding author.
E-mail address: Nyman001@umn.edu (J.A. Nyman).

https://doi.org/10.1016/j.jhealeco.2017.12.003
0167-6296/

In this paper, the theory is adopted that, rather than to avoid risk, at least a portion of health insurance is purchased in order to receive an income transfer when ill, and that the insurance price reduction (represented by the coinsurance rate) is the vehicle by which income is transferred from the healthy to the ill (Nyman, 1999a,b, 2003). Because the literature on the welfare effects of moral hazard did not originally recognize an income effect and has only belatedly attempted to redefine moral hazard in healthcare as the pure price effect (Cutler and Zeckhauser, 2000), we refer in this paper to the entire effect of insurance on spending as moral hazard. Accordingly, this paper seeks to decompose ex post moral hazard—the additional healthcare expenditures caused by becoming insured—into an efficient income-generated portion and an inefficient portion that arises because a price reduction is used to transfer income.

Chetty (2008) uses the term "liquidity" to refer to the subject of this transfer. In some respects, "liquidity" is more accurate than "income" because it captures the fact that the transfer in insurance typically does not have a periodic time dimension, but instead represents the transfer of a stock of resources or wealth that can be captured at an instant of time. On the other hand, if the transfer occurs by means of a price reduction (as it does with health insurance), the transfer is far from "liquid" in the sense that it would be difficult, if not impossible, to resell most health care services (e.g., an appendectomy) on the market and thereby convert them into cash. Accordingly, we have chosen to continue to refer to this notion as an "income transfer" because it matches the simple theoretical model better and because this language was used in the analyses that originally presented this theory (Nyman 1999a,b, 2003).

In the next section, the theory is summarized. It is shown how an income transfer is contained in the structure of a standard coinsurance contract. Then, a 7-step empirical approach is presented. In the following section, the results of the decomposition of moral hazard are presented for the various health conditions. In general, we find evidence of a substantial efficient moral hazard effect, but one that varies in size depending on the type of illness and assumptions of the empirical model. In the final section, the implications of the results are discussed and the limitations acknowledged.

Theory

The theoretical model has been presented elsewhere (Nyman, 1999a,b, 2003), so it is summarized here. First, consider the *ex post* case of a consumer who has become ill, with and without insurance. The consumer maximizes utility when ill, U_s, derived from medical care, M, and other goods and services, Y. If uninsured, the consumerspends his or her income (or wealth), Y_o, on M and Y, assuming that the price of a unit of medical care is unity. As a result, the consumer's problem is written:

$$\max U_s(M, Y)$$

$$s.t. Y_o = M + Y$$

The consumer solves the problem at (M_u, Y_u), where the marginal rate of substitution equals 1, the (negative of the) price ratio.

If the consumer is insured, he or she has purchased insurance with an actuarially fair premium payment, R, out of income, but in return, the price of medical care has dropped from 1 to the coinsurance rate, c. The consumer's problem is now:

$$\max U_s(M, Y)$$

$$s.t. Y_o - R = cM + Y$$

The consumer solves the problem at (M_i, Y_i) where the marginal rate of substitution equals c, the new effective price ratio.

Because the price of M has dropped from 1 to c, the consumer purchases $(M_i - M_u)$ more medical care, which constitutes moral hazard spending using the traditional definition. Assume that an actuarial study has previously estimated M_i and the insurer sets the premium, R, to be actuarially fair: $R = \pi(1-c)M_i$, where π is the probability of illness. Thus, the insurer spends $(1-c)M_i$ on this consumer, of which $\pi(1-c)M_i$ is financed by the consumer's own premium contribution and the rest by $(1-\pi)(1-c)M_i$ worth of income transferred from the insurance pool.

Alternatively, the consumer could have spent the same premium payment, R, on a contingent claims insurance contract that would pay out the same amount, $(1-c)M_i$, as was spent by the insurer under the coinsurance contract, but in this case as a lump-sum payoff, P, paid directly to the insured consumer upon becoming ill. The ill consumer's problem under this contract would be written:

$$\max U_s(M, Y)$$

$$s.t. Y_o - R + P = M + Y$$

The payoff, P, would consist of the consumer's own contribution to the insurance pool, $R = \pi(1-c)M_i$, plus an amount contributed by others and transferred to the ill consumer, $T = (1-\pi)(1-c)M_i$. Thus, the ill consumer with this insurance both pays out R *ex ante* and receives $(T+R)$ *ex post* as a result of this contract, so T, the income transfers, is the net increase in income. The consumer solves this problem at (M_c, Y_c), where the marginal rate of substitution equals 1, the original price ratio without distortion.

This model can be used to decompose the moral hazard generated by the coinsurance contract. Assuming that healthcare is a normal good and that $M_u < M_c < M_i$, then $(M_c - M_u)$ would represent the income transfer effect produced by a coinsurance contract and $(M_i - M_c)$ would represent the price effect. Thus, $(M_c - M_u)/(M_i - M_u)$ is the percentage of moral hazard that is income-generated, and so, efficient.

The *ex ante* portion of the model captures the motivation for purchasing insurance. In the *ex ante* period, the (healthy) consumer does not know whether she will become ill during the contract period. If the consumer is uninsured and becomes ill, the consumer is assumed to spend M_u on medical care and the rest of income on other goods and services. If the uninsured consumer remains healthy, she is assumed to derive no utility from medical care, so $U_h(0, Y_o)$. Thus, this model applies best to the type of healthcare that only those who are seriously ill would consume, the type of healthcare that is likely to represent a large portion, if not the majority, of total healthcare spending. Accordingly, if uninsured, the consumer's *ex ante* expected utility is:

$$EU_u = \pi U_s(M_u, Y_o - M_u) + (1-\pi)U_h(0, Y_o).$$

If the consumer is insured with a coinsurance policy, she would have paid the fair premium, R (worth $\pi(1-c)M_i$ of income), even if healthy. The coinsurance rate payoff would include a net transfer, T (worth $(1-\pi)(1-c)M_i$ of income), if ill, which could be spent on other goods and services after the M_i healthcare spending is accounted for. Again, it is assumed that the healthy consumer derives no utility from, and therefore does not purchase, any of this type of health care. For example, what *healthy* consumer would choose to receive a course of chemotherapy, an amputation, or an organ transplant, even if insured? So, an insured consumer has *ex ante* expected utility of:

$$EU_i = \pi U_s(M_i, Y_o + T - M_i) + (1-\pi)U_h(0, Y_o - R).$$

Thus, even though the insurance is actuarially fair, in that the expected premium payment if healthy, $(1-\pi)R = (1-\pi)\pi(1-c)M_i$, equals the expected transfer if ill,

$\pi T = \pi(1-\pi)(1-c)M_i$, the assumption of state dependent utility generates a utility gain. That is, insurance is purchased because the additional healthcare and other consumption spending permitted by the income transfer in the ill state, U_s, produces a greater utility gain than the utility lost from paying the premium in the healthy state, U_h.

Even though an overall welfare gain would exist for those who purchase a coinsurance health insurance contract for the purpose of transferring income across states of health, using a price reduction to transfer this income still contributes a loss of some welfare to the overall welfare calculation. This welfare loss, however, may be regarded simply as a transactions cost of the transfer. A contingent claims insurance contract could accomplish this transfer directly and avoid this loss, but it would entail transactions costs associated with writing complex contingency contracts (specifying the various payments for each illness and complication, and each sequela from the various treatments) and monitoring for fraud (confirming the existence of illness before making the insurance payment). Because coinsurance contracts require the consumer to receive healthcare before the insurer would pay for it, such contracts are able to avoid a large portion of these transactions costs. Indeed, the use of the standard coinsurance instead of contingent claims contracts may minimize the welfare loss and thus be optimal under the theory of the second best. That such contracts dominate health insurance markets in practice suggests this is the case.

It should be noted that this model of the decomposition of moral hazard is not the Hickisan (compensated) decomposition of an exogenous reduction of the market price. This is because the price effect in insurance is a consequence of a *quid pro quo* contract: in order for the price of healthcare to fall to the coinsurance rate level, it is necessary to purchase the price reduction through paying an insurance premium. Moreover, the purchase of increasingly smaller coinsurance rates generally requires larger upfront premium payments and, assuming that healthcare is a normal good, these reductions in *ex ante* disposable income would have larger implications for the amounts of *ex post* healthcare consumed at each coinsurance rate. For example, the increase in healthcare generated by the purchase of, say, a 20% coinsurance rate for a \$20,000 premium would be smaller than the increase generated by exogenous fall in the market price of the same magnitude, and therefore would produce a different theoretical decomposition (Nyman, 1999a, 2003).[2]

Also, it should be noted that in this model, the arguments of the state-dependent utility function when ill include spending on medical care. Thus, it assumes that income transferred to the ill state can be used for both additional medical care spending and spending on other consumption goods and services. This differs from Viscusi and Evans (1990) and the others who have investigated the effect of health on the marginal utility of income in the purchase of consumption goods and services alone.

Finally, our implied hand-to-mouth consumption model assumes that it is sometimes not possible to save or borrow an amount of income equal to the amount of the income transfer in the insurance contract when ill. This assumption reflects the reality that consumers who become seriously ill may not have the time or ability to work and save enough to pay for a costly procedure or their chronic care. Similarly, a loan (collateralized or not, depending on the cost of the medical procedure being considered) may not be available (at a manageable interest rate) to the seriously ill because of the lender's concerns about repayment. Thus, according to this theory, having purchased income-transferring insurance may be the only marketed source of additional income to gain access to care when ill. This constraint is central to understanding the theory of an income-transfer demand for health insurance. However, the various ways in which society and uninsured individuals have sought to circumvent this constraint in practice–from making free or reduced-price healthcare available to the ill to saving in anticipation of illness–make identifying and accurately measuring this effect an empirical challenge.

[2] In Nyman (1999b), a Hicksian decomposition is shown in the diagrams that allegedly describe this decomposition of the coinsurance effect. This error was made at the insistence of the paper's editor who could not be convinced that the welfare analysis of a purchased insurance price decrease was different from the welfare analysis of an exogenous market price decrease and who was prepared to reject the paper if the Hicksian (compensated demand) effect was not shown. The correct diagrams of this decomposition appear in Nyman (2003).

Methods

Overview

The goal of this study is to simulate as closely as possible differences in healthcare spending between (1) a consumer who becomes ill when uninsured and (2) the same consumer who becomes ill when insured under either (a) a standard coinsurance insurance policy, where the insurer pays off by reducing the price of the ill consumer's care, or (b) a contingent claims policy, where the insurer pays directly to the consumer at the time of illness an amount of income equal to what the insurer would have spent on the insured consumer under the coinsurance policy. These spending amounts are used to find the proportion of moral hazard spending that is efficient for various ill health states.

The empirical analysis required a number of technical decisions. We adopted an approach where an initial set of assumptions is defined, but where the results with alternative assumptions are also presented as a sensitivity analysis. Although all the empirical findings are presented in the results section, for clarity of exposition we also indicate the tabular location of the alternative results when the issue is raised in the methods section.

Data

Data for the simulation analysis come primarily from the Medical Expenditure Panel Survey for the pre-Affordable Care Act years 1996–2010. The MEPS contains information on the insurance status, total healthcare expenditures, out-of-pocket expenditures, income and financial resources, a series of important chronic health conditions and other demographic variables for a large nationally representative sample of non-institutionalized Americans. The theoretical approach suggests that the decomposition should focus on those who are or become ill, thus we calculate a series of estimates that vary with the illnesses of the MEPS respondents. By constraining the analysis to those with illnesses, the empirical issue of a possible mass of cases with zero expenditures is avoided and so is the need for a 2-part estimation procedure.

The MEPS data are self-reported and have been shown to understate expenditures because of recall bias. An adjustment developed at the Agency for Healthcare Quality and Research (AHRQ), the parent governmental agency for the MEPS, was applied to both insured and uninsured total expenditures to reflect this underreporting (Zuvekas et al., 2005). Issues also arise regarding whether the out-of-pocket spending by the uninsured accurately captures the healthcare consumed in a way that is consistent with national income account expenditure levels. Hadley and Holahan (2003) estimate that total spending on the uninsured is about 2.88 times greater than the level of out-of-pocket spending captured in the MEPS, reflecting the availability of uncompensated healthcare provided pro bono by hospitals and physicians, and the programmatic healthcare provided by charities and government agencies that is not linked to any specific individual by administrative records. Uninsured healthcare expenditures are adjusted for this as well.

Our analysis seeks in part to simulate the spending behavior of an uninsured consumer who experiences an income gain, as if by a direct transfer from a contingent claims insurer at the time of illness. Such a simulation requires that the transfer be plausibly exogenous to the spending decision. Because of limitations in the MEPS, a second data set with a more distinctly exogenous relationship between healthcare spending and additional financial resources was introduced into the analysis: the Health and Retirement Survey. The HRS was used to estimate the out-of-pocket healthcare spending response that is generated by the increase in liquid assets occurring 2 years before the period of the spending. Because this asset increase variable was temporally distant from the spending variable and likely generated by market price changes that are beyond the individual's control, it would thus be sufficiently exogenous and better capture the spending behavior generated by the payoff of a contingent claims insurance contract. At the same time, the temporal distance between the increase in liquid assets and the measure of healthcare spending may serve to reduce the magnitude of the asset elasticity of healthcare spending. Moreover, because such a payoff is not earmarked for spending on healthcare, as an insurance payoff would be, the responsiveness of spending may be further reduced. Both these elasticity-reducing factors may combine to make conservative the estimates of the proportion of moral hazard spending that is efficient.

Methodological steps

The intent of the empirical work is to simulate the additional spending that an uninsured consumer would have made if she had been insured under a contract that paid off by either (1) reducing the price of care with a coinsurance rate or (2) paying a lump sum amount equal to the amount paid out by the insurer under the coinsurance contract. To reiterate, the difference between uninsured spending and counterfactual spending by the uninsured if insured with a coinsurance contract, $M_i - M_u$, is the estimate of total moral hazard. The difference between uninsured spending and counterfactual spending by the uninsured if insured with a lump sum payoff contract, $M_c - M_u$, is efficient moral hazard, and the proportion of total moral hazard that is efficient is $(M_c - M_u)/(M_i - M_u)$. To accomplish this decomposition of moral hazard, a 7-step empirical approach was employed.

Step 1: estimating spending levels

Because the decision to become insured is endogenous, a 2-stage instrumental variable procedure was used to estimate the effect of being insured on spending on healthcare services. The first stage used a probit equation to predict being privately insured versus uninsured, among people not publicly insured. The desired instruments should predict insurance status but have no effect on healthcare spending. Studies facing a similar issue have used instruments based on (1) firm size, (2) unionization, and (3) employment, (see Dranove et al., 1998; Baker and Brown, 1999; Johnson and Crystal, 2000; Bundorf, 2002; Olson, 2002; Bhattacharya et al., 2003; Deb and Trivedi, 2006; Munkin and Trivedi, 2003; Cawley and Simon, 2005; Koc, 2005; Pauly, 2005; Deb et al., 2006a,b; Pauly and Herring, 2007; Town et al., 2007; Bates and Santerre, 2008; Bates et al., 2010; Koc, 2011). We used respondent-level versions of these instruments: (1) the size of the respondent's firm, measured by the number of full-time equivalent [FTE] employees, with firms with over 500 FTEs coded at a maximum of 500, (2) the respondent's union status, and (3) his or her self-employment status. These instruments affect the offer of insurance for the primary policy holder. Because our sample included people who received employment-based insurance through a spouse, but might not be employed themselves, an alternative analysis was conducted in which the study sample was limited to primary policyholders only (column B of Table 6).

An alternative analysis also was conducted in which the state- and county-level versions of these instruments were used: (1) average firm size in the respondent's state, (2) unionization rate of employees in the respondent's county, and (3) unemployment rate in the respondent's state. State-level average firm sizes come from the U.S. Department of Commerce, Bureau of the Census, Statistics of U.S. Businesses. County-level unemployment rates were taken from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics Series. State-level unionization rates came from the Union Membership and Coverage Database in the Current Population Survey (CPS), first constructed by Hirsch and Macpherson (2003). To preserve comparability, we constrained the data to the same observations as were used with the original respondent-level analysis (column C of Table 6).

The variable for private insurance was defined in the MEPS as an individual who had private insurance during any portion of the year, and was either uninsured or publicly insured for the rest of the year. The uninsured variable, in contrast, was defined as an individual who was uninsured during the entire year. Although only 12% of the insured were insured for less than 12 months in our 2010 data, we were concerned that their behavior might differ from the behavior of those who were insured for the entire year. Therefore, in an alternative analysis, we ran the regression using only those who were privately insured for the entire 12-month period (column D of Table 6).

In stage 2, the predicted probabilities of being privately insured or uninsured from the stage 1 are used to estimate expenditures using Mullahy's (1997, 1998); generalized method of moment's (GMM) approach to deal with the skew in the expenditure distribution. The advantages of this method are that it provides unbiased coefficient estimates and heteroskedasticity-robust standard errors with which to calculate valid standard errors for the coefficients.

The stage 2 equation was specified using a linear relationship between spending and current earned income. In an alternative analysis, a double-log specification was used so that the regression coefficient would represent the earned income elasticity of spending. This facilitated comparison with the liquid assets elasticity of spending calculated from the HRS (Step 2 below), which was estimated using a similar specification (columns E and F of Table 2).

Our theoretical analysis focused on the spending produced by an insurance income transfer to those who are or become ill. Accordingly, in both the elasticity estimates using the HRS (step 2 below) and the 2-stage equations using the MEPS, separate analyses were conducted for those with different illnesses. Illness categories were: (1) 1 or more of the "priority" health conditions, (2) cancer, (3) diabetes, (4) mental illness (depression or anxiety), (5) heart condition, (6) arthritis, and (7) hypertension (Table 5). "Priority health conditions" are a set of illnesses identified by the Agency for Healthcare Research and Quality as being worthy of tracking because of their incidence, expense and policy relevance. Thus, the first illness category captures whether the respondent has been diagnosed with 1 or more of the following illnesses: hypertension, heart disease, high cholesterol, emphysema, chronic bronchitis, diabetes, cancer, arthritis, asthma, attention deficit/hyperactivity disorder, and stroke.

Step 2: estimating insurance payout elasticities of spending

The MEPS earned income variable used in the expenditure equations data did not adequately mimic the income transfer that would occur in a contingent claims insurance contract. This is because the effect of differences in annual earned income on healthcare spending by the ill would be constrained by anticipated spending commitments—mortgages, car payments, rent, payments for utilities, transportation to and from employment, and the purchase of other necessities—that would reduce the ability to spend at the discretion of the individual who is ill. This, in turn, would act to limit the response of out-of-pocket healthcare spending to differences in income, thus tending to bias the income coefficient downward as a measure of the insurance payout effect on healthcare spending.

In order to reduce this bias, we sought to estimate (what we would interpret as) the insurance payout elasticity of spending using a more exogenous resource shock. As described above, a plausibly exogenous variable was found with the liquid financial assets variable in the HRS. Specifically, the HRS collects data on the change in all liquid financial assets for the respondent in the 2 years since the last HRS wave. The liquid assets variable is calculated as the sum of the net values of: (a) real estate (besides the primary residence), (b) vehicles, (c) businesses, (d) IRA and Keogh accounts, (e) stocks, mutual funds, and investment trusts, (f) checking, savings, or money market accounts, (g) certificates of deposit, government savings bonds, and T-bills, (h) bonds and bond funds, and (i) all other savings, minus the value of debt. The change in liquid assets variable measures the difference between the value of liquid assets in the current interview, and the same variable in the interview 2 years previous.

To estimate the asset elasticity of spending, the amount of out-of-pocket spending in the HRS was first multiplied by 2.88, the correction for the value of uncompensated and programmatic care received by the uninsured that was applied to the MEPS out-of-pocket spending. This change should not affect the calculation because the elasticities are derived from a function that is homogeneous of degree zero. Then, this spending by the uninsured was regressed on the increase in liquid assets from the 2 years previous to the current wave. Separate elasticities were calculated for those with each of the different diseases.

There was wide variation in the gain in liquid assets, from 0 to many millions of dollars. We were concerned that the marginal increase in spending by those situated in the tails of the distribution may not represent a realistic simulation of the marginal increase in spending generated by those who would receive a check that is commensurate with their healthcare needs. Therefore, in the initial analysis, we limited the sample to those who had at least a $100 gain. In alternative analyses, we also constrained the increase to those with $25,000 at most (columns C of Table 2).

The HRS equation that was used to estimate the asset elasticities of spending employed a specification as similar as possible to the MEPS expenditure equation used in step 1, given the different sets of explanatory variables available. The HRS equation was not corrected for selection because we sought HRS elasticities that represented the spending behavior of those who were (and chose to be) uninsured. That is, because in this step we were not also interested in determining the counterfactual spending of the uninsured if insured, it was not necessary to employ a 2-stage analysis.

In both our HRS and MEPS analyses, we restricted the respondents to those under age 65. The HRS respondents are restricted to those aged 50 years or greater, but the MEPS respondents are aged 18 years or greater. To alleviate our concern that the asset elasticities of spending found in the HRS would not apply to the subset of younger respondents of the MEPS, we conducted an alternative analysis using only those MEPS respondents who were 50 years of age or older (column E of Table 6).

Step 3: estimating expenditures of the insurer

The goal in this step is to predict the amount of health expenditures, $(1-c)M_i$, that are paid for by the insurer for those who are insured. To determine this amount, a Terza et al., 2008 residual inclusion model was employed that used the same instruments and right-hand-side specification as were used in step 1. The dependent variable in this step was expenditures net of any coinsurance payments by the insured consumers, and the insured respondents in the MEPS were used to estimate the model.

Step 4: predicting counterfactual insurance expenditures on behalf of the uninsured if insured

The counterfactual health expenditures that would be paid for by the insurer, if the uninsured were insured, were determined by entering the characteristics of the uninsured from the MEPS and the corresponding Terza residual into the equation from Step 3 and predicting insurer expenditures. This prediction represents the value of the payoff, P, which would be paid out in a contingent claims insurance policy.

Step 5: calculating the percentage change in spending generated by the counterfactual insurance payoff

The goal in this step is to calculate the percentage change in spending that would be generated by the counterfactual insurance payoff, P. The percentage change in spending, $(M_c-M_u)/M_u$, is determined by multiplying the liquid asset elasticity of healthcare spending, ε, by the payoff, P, as a percentage of disposable income. Disposable income for an uninsured person who has become insured is baseline income, Y_o, net of the insurance premium, R. So,

$$(M_c-\mathrm{M_u})/\mathrm{M_u} = \varepsilon[P/(Y_o-\mathrm{R})].$$

Data on insurance premiums come from the annual Kaiser surveys for the corresponding time periods, and represent the average national premium adjusted for the family size of the MEPS respondent (Kaiser Family Foundation and National Research and Education Trust, 2010). Because the uninsured often were represented by those with low incomes, premiums exceeded baseline income in a number of cases. Such observations were excluded from the analysis.

Step 6: predicting total moral hazard spending

Total moral hazard spending is estimated by first entering the characteristics of the uninsured into the insured version of the selectivity-corrected spending equation from step 1 and predicting counterfactual insured healthcare spending, M_i, for the uninsured. Next, a similar procedure is used to estimate M_u for the uninsured using the uninsured version of the same equation. Total moral hazard spending is found by subtracting M_u from M_i.

Step 7: calculating the percentage of total moral hazard spending that is efficient

Finally, an estimate of efficient moral hazard, M_c-M_u, is calculated. Efficient moral hazard is found by multiplying the percentage change in spending that would be generated by the insurance transfer, $(M_c-M_u)/M_u=\varepsilon(P/Y_o-R)$, by spending of the uninsured, M_u. This product, $\varepsilon(P/Y_o-R)^*M_u$, represents (M_c-M_u). Thus, the percentage of total moral hazard that is efficient is calculated as $(M_c-M_u)/(M_i-M_u)$. These steps are repeated for MEPS respondents with each of the various illnesses.

It should be noted that the analysis excluded disabled individuals because disabled people may have fewer resources available than the non-disabled, and may exhibit a greater healthcare spending response to an income (or wealth) shock. To test whether this was an important factor in determining our estimates, an alternative analysis was performed including respondents with a disability (column F in Table 6). Pregnant women were also excluded because pregnancies can be anticipated and so those who are uninsured may save in order to cover the spending required for labor and delivery. Spending out of anticipated savings differs from spending by those who experience the onset of an unanticipated disease and, if uninsured, must spend out of existing resources.

All healthcare spending was adjusted to 2010 by the medical care portion of the Consumer Price Index and all income levels were adjusted to the same year by the Consumer Price Index. Some of the empirical analyses used restricted data, and so the analyses were conducted at the Research Data Center on the campus of the University of Minnesota and the results were approved for public release by the AHRQ. The analysis used the statistical software package STATA 13 (StataCorp, 2013). The study was approved by the Independent Review Board of the University of Minnesota.

The next section defines and lists the descriptive characteristics of the analytical data samples and presents the empirical and simulation results, first for the liquid asset elasticities of spending using the HRS and then for the rest of the analysis using the MEPS.

Results

HRS analytical sample

Recall that Step 2 seeks to capture the effect of a contingent claims payoff on spending by estimating the spending responsiveness of the uninsured who have an illness to an exogenous increase in a liquid-assets variable from the HRS. The increase in liquid assets might capture, for example, payments from pensions at retirement, non-health insurance payouts, inheritances, increases in value of a stock portfolio, and so on. We required observations to have a positive change in total liquid assets from the previous wave of at least $100. We also required respondents to be less than 65 years old, be uninsured during the entire current wave, and have the health condition specified. Our inclusion criteria for the elasticity analysis were met by 2965 observations over the 1996–2010 period.

Table 1 shows the descriptive statistics for the uninsured HRS observations, with the corresponding values for the uninsured observations in the MEPS analysis presented for comparison. The mean change in non-housing assets since the previous period was in excess of $100,000, with a range of from just over $100 (the established minimum value) to almost $9 million. The median increase in assets, however, was $16,775, which appears to be more consistent with the notion that this change in assets is a reasonable proxy for a lump sum health insurance payout representing a health insurer's payment to providers. Mean annual out-of-pocket spending, after adjustments for under-reporting and the shortfall from national income accounts, was $4756 (for 1 year) for the older HRS sample, compared to $1686 for the MEPS. Median annual out-of-pocket spending in the HRS was $1676 versus $485 in the MEPS. Females and minorities were over-represented in the HRS compared to the MEPS and a smaller percentage was college educated. The HRS sample had substantially lower earned incomes than the MEPS sample, perhaps also owing to the greater likelihood that the older HRS respondent was retired. Again, because of the higher age, the HRS sample was generally sicker, with higher likelihoods of each of the health conditions.

HSR results

The income elasticities of healthcare spending for those with the various health conditions were estimated using the increase in liquid assets from the HRS as the income variable. Table 2 presents the elasticities (and standard errors in parentheses) from the HRS and what the elasticities would have been if the earned income variable from the MEPS (either the change in earned income from the previous year or lagged earned income) had been used, both derived from similar double log models. As expected, HRS elasticities calculated with an asset shock (Table 2, columns B and C) are generally larger than those calculated with the more predictable, less discretionary earned income of the MEPS (Table 2, columns D and E). Constraining the HRS asset increases to less than $25,000 (to be more consistent with what an insurer might pay for an insured person with a priority condition over the course of a year) did not produce a consistent change in magnitude. We used the estimates presented in Table 2, column B in our Step 5 calculations.

MEPS analytical sample

Again, we sought to simulate the spending response from a contingent claims payoff and compare that to the spending response from a coinsurance payoff that represented an equal amount of spending by the insurer on the insured. To do so, we modeled the spending behavior of privately insured and uninsured consumers who have various diseases using 1996–2010 data from the MEPS. The eligible observations for this analysis were less than 65 years old, not publicly insured, not pregnant and not disabled. They had incomes greater than $0 and one, or at least one, of the 8 priority health conditions.

Another issue was the degree that the spending of the uninsured was paid for by what the MEPS categorized as "other" sources. That is, the MEPS broke down the total healthcare expenditure variable by the sources of payment: (1) private insurance; (2) Medicaid; (3) Medicare; (4) out-of-pocket; (5) veteran's administration (but not CHAMPVA); (6) TRICARE; (7) *other* federal sources (such as Indian Health service, Military Treatment Facilities, and miscellaneous or unknown sources paid for by the federal government); (8) *other* state and local sources (such as community clinics, health departments and state programs other than Medicaid); (9) worker's compensation, (10) *other* miscellaneous or unknown sources;(11) *other* private expenditures (for example, having private coverage for dental care could generate private healthcare spending without being privately insured for medical coverage); and (12) *other* public expenditures. Some of the uninsured respondents listed payments from these "other" sources in addition to their out-of-pocket payments. As alluded to above, our hand-to-mouth theoretical approach seeks to isolate the behavior of those with no sources of payment for healthcare except for their own resources, compared to those with private insurance. To the extent that some uninsured respondents can access "other" sources of payment for healthcare when ill, the level of their out-of-pocket spending would be different from, and likely less than, what it would have been without these payments.

Approximately 9% of privately insured and 40% of uninsured respondents with a priority health condition in our sample had some sort of "other" payment for health expenditures. Table 3 shows that, among observations with "other" payments, other spending averaged $875 for privately insured and $1259 for uninsured observations. Observations with such payments tended to have higher private insurance, out-of-pocket and total expenditures as well. To maintain adequate sample size for condition-specific analyses, we did not drop observations with other payments. However, to reduce the effect of other payments on the analysis, we limited expenditures by the uninsured to out-of-pocket expenditures only. To be consistent, we also limited expenditures by the privately insured to private insurance payments and out-of-pocket expenditures.

The analytical sample from the MEPS used in our analysis of Steps 1 and 3–7 consisted of 7749 uninsured observations and 41,143 privately insured observations, all of whom met the inclusion criteria. The descriptive statistics for these two groups are compared in Table 4. As expected, out-of-pocket spending is greater for the uninsured than for the insured. Mean annual income among the uninsured is about half what it is among the privately insured. This suggests that many of the uninsured would not be able to afford

Table 1
Descriptive Statistics of the Uninsured from the HRS and MEPS.

	HRS (n = 2965)				MEPS (n = 7749)			
Variable	Mean	SD	Min	Max	Mean	SD	Min	Max
Change in non-housing assets ($)	105,433	383,335	106	8,758,000	NA	NA	NA	NA
Out-of-pocket spending (adjusted, $)	4756 (annual average over 2 year)	25,741	0	539,222	1686	4222	0	118,663
Family size	2.56	1.57	1	19	3.44	1.80	1	14
Male	0.35	0.48	0	1	0.42	0.49	0	1
Black	0.16	0.36	0	1	0.16	0.37	0	1
Hispanic	0.20	0.40	0	1	0.42	0.49	0	1
Married	0.63	0.48	0	1	0.53	0.50	0	1
Some college	0.30	0.46	0	1	0.26	0.44	0	1
Age	58.26	4.95	26	64	40.74	12.48	18	64
Earned income ($)	9573	17,943	0	181,374	24,988	24,459	5	225,727
Cancer	0.09	0.29	0	1	0.03	0.18	0	1
Diabetes	0.18	0.39	0	1	0.13	0.34	0	1
Mental health problem	0.23	0.42	0	1	0.23	0.42	0	1
Heart disease	0.14	0.35	0	1	0.06	0.24	0	1
Stroke	0.04	0.21	0	1	0.01	0.08	0	1
COPD/Asthma	0.10	0.30	0	1	0.20	0.40	0	1
Arthritis	0.63	0.48	0	1	0.15	0.36	0	1
High blood pressure	0.57	0.50	0	1	0.28	0.45	0	1

Table 2
Elasticities from the HRS and MEPS.

		HRS		MEPS	
	A. HRS observations	B. Increase in non-housing asset > $100	C. Increase in non-housing assets < $25 K	D. Change in earned income from previous year	E. Earned income from the previous year
Any priority Condition	2965	0.26** (0.03)***	0.29** (0.05)	0.08** (0.02)	0.10** (0.01)
Cancer	274	0.37** (0.10)	0.43* (0.19)	0.08* (0.04)	0.05 (0.05)
Diabetes	539	0.24* (0.07)	0.35* (0.12)	0.05 (0.04)	0.10** (0.02)
Mental illness	696	0.20* (0.19)	0.19 (0.11)	0.09 (0.05)	0.08** (0.03)
Heart disease	423	0.32** (0.08)	0.20 (0.15)	0.14** (0.04)	0.12** (0.04)
Stroke	133	0.15 (0.19)	****	0.17* (0.08)	0.07 (0.86)
COPD/Asthma	291	0.18 (0.12)	0.06 (0.21)	0.21** (0.04)	0.09 (0.02)
Arthritis	1864	0.23** (0.04)	0.18* (0.07)	0.04 (0.03)	0.11** (0.03)
Hypertension	1682	0.30** (0.04)	0.36** (0.07)	0.20* (0.08)	0.06** (0.02)

* $p < 0.05$.
** $p < 0.01$.
*** Standard errors are in parentheses.
**** Less than 100 observations.

Table 3
Expenditures of MEPS Observations with and without "Other Payments" .

		n	Other payments (mean)	Out-of-pocket payments (mean)	Private insurance payments (mean)	Total expen-ditures (mean)
Privately insured	No other payments	37,810	$0	$543	$1819	$2415
	Other payments	3593	$875	$590	$2268	$3865
Uninsured	No other payments	5018	$0	$388	$0	$391
	Other payments	3326	$1259	$865	$0	$2162

to purchase a private health insurance policy, especially given the other claims on their income, and that attempting to simulate their behavior as insured consumers would be problematic for that reason. The uninsured are a little younger, more likely to be black or Hispanic, and less likely to be married, have gone to college or be employed. The uninsured are slightly healthier than the insured, and they are more likely to live in the South, be employed at a small firm that is not unionized or be self-employed.

MEPS results

Table 5 presents the main point estimates from the calculation of the proportion of moral hazard that is efficient, for the various health conditions considered. Rows 1 and 2 show the number of observations that were used (Step 1) to generate spending equation coefficients for the uninsured and insured, respectively. Row 3 restates the (Step 2) HRS elasticities (from Table 2) used in the calculations. Row 4 presents the mean counterfactual predicted spending of an insurer for an uninsured consumer, *P*. This dollar amount is found by estimating insurer spending for the insured (Step 3) and entering the characteristics of the uninsured observations into that equation (Step 4).

The calculation of the percentage increase in spending (Step 5) required the calculation of the income of the uninsured after a premium was paid. Accordingly, row 5 of Table 5 presents the mean income of the uninsured, Y_o, and row 6 is the mean income after an average insurance premium for the uninsured's type of family, *R*, has been subtracted. As mentioned earlier, many uninsured had

Table 4
Descriptive Statistics for MEPS Sample by Insurance Status.

	Uninsured (n = 7749)				Privately Insured (n = 41,143)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Out-of-pocket spending (adjusted, $)	1686	4222	0	118,663	548	983	0	43,675
Private insurance spending ($)	0	0	0	0	1858	4960	0	183,821
Annual income ($)	24,988	24,459	5	225,727	57,104	43,960	5	362,292
Family size	3.44	1.80	1	14	3.00	1.46	1	15
Age (years)	40.74	12.48	18	64	44.12	11.64	18	64
Male	0.42	0.49	0	1	0.41	0.49	0	1
Black	0.16	0.37	0	1	0.13	0.34	0	1
Hispanic	0.42	0.49	0	1	0.14	0.35	0	1
Married	0.53	0.50	0	1	0.68	0.47	0	1
Some college	0.26	0.44	0	1	0.57	0.49	0	1
Employed	0.66	0.47	0	1	0.85	0.36	0	1
Sick pay	0.12	0.32	0	1	0.56	0.50	0	1
Cancer	0.03	0.18	0	1	0.06	0.25	0	1
Diabetes	0.13	0.34	0	1	0.10	0.31	0	1
Mental health problem	0.23	0.42	0	1	0.24	0.42	0	1
Hypertension	0.28	0.45	0	1	0.32	0.47	0	1
Heart disease	0.06	0.24	0	1	0.09	0.29	0	1
COPD/Asthma	0.20	0.40	0	1	0.26	0.44	0	1
Arthritis	0.15	0.36	0	1	0.16	0.37	0	1
Urban	0.79	0.41	0	1	0.82	0.38	0	1
Midwest	0.14	0.35	0	1	0.24	0.43	0	1
South	0.47	0.50	0	1	0.35	0.48	0	1
West	0.28	0.45	0	1	0.23	0.42	0	1
Firm size (employees)	37.59	101.07	0	500	136.46	182.89	0	500
Self employed	0.16	0.37	0	1	0.09	0.28	0	1
Unionized	0.02	0.13	0	1	0.13	0.34	0	1

Table 5
Proportion of Moral Hazard That Is Efficient by Health Condition.

	Any priority condition	Cancer	Diabetes	Mental health problem	Heart disease	Arthritis	High Blood Pressure
1. Number of uninsured	7749	375	1131	2676	681	1729	2356
2. Number of insured	41,143	3476	4570	13,839	4773	9138	14,138
3. Elasticity (ε)	0.26	0.37	0.24	0.20	0.32	0.23	0.30
4. Mean spending ($) by insurer ($P$)	1,858.43	16,336.04	2,931.73	1,428.83	4,260.13	2,155.95	2,302.44
5. Mean income ($) of uninsured ($Y_o$)	24,987.56	28,569.51	24,217.54	21,746.21	23,794.73	24,719.74	25,153.46
6. Mean income ($) of uninsured minus premium ($Y_o - R$)	17,956.28	21,624.07	15,789.79	14,620.34	16,278.93	16,296.42	16,799.37
7. Mean percent increase in spending generated by the percent increase in income generated by a lump-sum insurance payoff $\{(M_c - M_u)/M_u = \varepsilon[P/(Y_o - R)]\}$	0.10	0.26	0.13	0.05	0.12	0.05	0.17
8. Mean predicted spending ($) by uninsured if insured (M_i)	3,342.00	12,168.23	4,974.15	3,349.30	10,143.26	3,302.13	4,544.04
9. Mean predicted spending ($) by uninsured ($M_u$)	1,788.00	5,135.14	4,326.17	1,625.83	4,108.88	1,521.81	2,853.05
10. Mean moral hazard spending ($) by uninsured if insured ($M_i - M_u$)	1,555.00	7,033.09	647.98	1,723.47	6,034.38	1,780.32	1,690.99
11. Mean efficient moral hazard spending ($) by uninsured if insured $\{[(M_c - M_u)/M_u]^* M_u = M_c - M_u\}$	197.00	4,307.41	461.65	101.29	436.26	124.91	429.48
12. Mean proportion of efficient moral hazard by uninsured if insured $(M_c - M_u)/(M_i - M_u)$	13%	61%	71%	6%	7%	7%	25%

such low incomes that subtracting a standard premium would leave them with 0 or negative income. Such observations were dropped from the analysis at this point in the calculation. As a result, row 6 represents the mean for those uninsured with sufficient income to pay an insurance premium. Row 7 brings in the insurance spending from row 4 to calculate the percentage increase in spending that would be generated by a lump-sum insurance payoff equal to what the insurer spent on care, $(M_c - M_u)/M_u = \varepsilon[P/(Y_o - R)]$.

Table 6
Proportion of Moral Hazard That Is Efficient for Respondents with Any Priority Condition, by Assumption.

	A. Original Specification	B. Excludes Insured by Spouse	C. Insured All Year	D. Ages 50–64 Only	E. Includes People with Disabilities	F. People Likely Eligible for ACA Subsidies
1. Number of uninsured	7749	7749	7749	2188	11,766	4076
2. Number of insured	41,143	29,446	36,337	14,994	56,754	17,816
3. Elasticity (ε)	0.26	0.26	0.26	0.26	0.26	0.26
4. Mean spending ($) by insurer ($P$)	1,765.66	1,741.36	2,010.33	2,778.67	2,910.81	3,381.15
5. Mean income ($) of uninsured ($Y_o$)	24,987.56	24,987.56	24,987.56	28,672.70	23,380.94	24,693.18
6. Mean income ($) of uninsured minus premium ($Y_o - R$)	17,956.28	17,956.28	17,956.28	21,082.10	16,492.42	17,613.12
7. Mean percent increase in spending generated by the percent increase in income generated by a lump-sum insurance payoff $\{(M_c - M_u)/M_u = \varepsilon[P/(Y_o-R)]\}$	0.10	0.09	0.12	0.13	0.21	0.10
8. Mean predicted spending ($) by uninsured if insured (M_i)	3,342.00	3,061.16	3,332.59	7,411.19	5,820.80	4,742.60
9. Mean predicted spending ($) by uninsured ($M_u$)	1,788.00	1,982.96	1,893.19	4,374.13	2,790.89	3,127.22
10. Mean moral hazard spending ($) by uninsured if insured ($M_i - M_u$)	1,555.00	1,078.20	1,439.40	3,037.06	3,029.91	1,615.38
11. Mean efficient moral hazard spending ($) by uninsured if insured $\{[(M_c - M_u)/M_u]^*M_u = M_c - M_u\}$	197.00	208.49	263.62	562.18	883.56	391.76
12. Mean proportion of efficient moral hazard by uninsured if insured $(M_c - M_u)/(M_i - M_u)$	13%	19%	18%	19%	29%	24%

The calculation of moral hazard (Step 6) requires estimating counterfactual private insurance spending and uninsured spending of the uninsured. Row 8 shows the mean counterfactual private insurance spending of the uninsured, M_i, and row 9 reports the mean estimated uninsured spending, M_u, both found by entering the characteristics of the uninsured into the appropriate Step 1 spending equation. Row 10 shows the difference between row 8 and row 9 as the estimate of mean moral hazard spending of those who were originally uninsured, $(M_i - M_u)$.

The calculation of the percentage of moral hazard that is efficient (Step 7) requires the calculation of the increase in spending generated by the insurance payoff. Row 11 shows this amount as calculated by multiplying the percentage increase in spending from row 7 by baseline spending of the uninsured, $[(M_c - M_u)/M_u]^*M_u = (M_c - M_u)$. Finally, row 12 of Table 5 shows the calculation of efficient moral hazard spending as a percentage of total moral hazard spending, $(M_c - M_u)/(M_i - M_u)$.

Focusing on the Step 7 results presented in row 12 of Table 5, efficient moral hazard spending as a percentage of all moral hazard spending ranges from a low of less than 6% for those with mental health problems to a high of over 71% for those with diabetes. These amounts indicate that significant portions of moral hazard spending can be efficient and that policies that reduce the transfers across the board are likely to be counter-productive in a welfare sense for those diseases with a large proportion of efficient moral hazard.

Sensitivity analysis

Separate analyses were conducted to investigate the effect of various changes in the approach. Table 6 shows the alternative results for those MEPS respondents with any priority condition, with the columns representing the various alternative assumptions. The original results are reported again in column A for ease of comparison.

Column B shows the effect of requiring respondents to be the principal on their policy and not the spouse. Efficient moral hazard as a proportion of total moral hazard (hence, percent efficient moral hazard) increased substantially, perhaps owing to an elimination of some measurement error in the first stage analysis.

In the main analysis, the privately insured variable captured those who were privately insured at any time during the year. In column C, the analysis restricts this variable to only those respondents who were insured over the entire year and this change resulted in a higher estimate of the mean percent efficient moral hazard.

Column D shows the analysis if the ages of respondents in the MEPS are restricted to those 50 years of age or older (but less than 65 years of age) to match the age range of the HRS from which the spending elasticities were derived. As a result of this change, the percent of efficient moral hazard estimate increased.

Column E now includes respondents in the MEPS with a disability. Originally we had dropped respondents with any IADL, ADL, functional, activity, or sensory limitation on the assumption that they represented a group that is structurally different from the typical person with the diseases in question. Adding the disabled respondents back into the analysis substantially increased the percent efficient moral hazard for those with at least 1 of the priority health conditions.

Finally, we sought to estimate the effect of becoming privately insured on those with a priority health condition who were also eligible for premium subsidies under the Affordable Care Act (ACA) of 2010. Constraining our MEPS analysis to only those qualified respondents, the percentage of efficient moral hazard rose to 24%, suggesting that the ACA premium subsidies would work to increase access to care.

Discussion

Implications: theoretical, methodological and policy

This study found that between 13 and 29% (depending on assumptions) of the additional healthcare that those with at least 1 priority condition purchase if insured would have also been purchased if they had received a commensurate increase in income (liquid assets) directly from the insurer. Because these individuals could have spent this income transfer on anything of their choosing but are predicted in our simulation to spend a portion of it on medical care, this additional spending represents an increase in welfare to the extent that it results in an increase in the consumer surplus. For those with diabetes or cancer, over half of moral hazard spending was found to be efficient using our initial assumptions. These findings provide empirical support that one of the important motivations for purchasing health insurance is the additional healthcare made possible by a transfer of income from those who purchase insurance and remain healthy to those who purchase insurance and become ill.

This motivation to purchase health insurance differs from the motivation conventionally assumed in theory: to avoid or reduce risk of loss (Friedman and Savage, 1948). Those who favor conventional theory, however, must find a way to accomodate the many empirical studies showing that consumers actually prefer uncertain losses to certan ones of the same magnitude (Kahneman and Tversky, 1979). Moreover, they must also acknowledge that there is a fundamental difference between insurance against a loss of property, such as an automobile accident or house fire where something of value is simply lost, and insurance against a “loss” represented by the quid pro quo purchase of medical care (Nyman, 2003). That is, while the loss of income or wealth as medical expenditures is captured by conventional theory, the other side of the transaction is not, and the access to medical care can represent an extremely valuable gain. This suggests that conventional theory may be inappropriately applied to health insurance and that an alternative theory is needed to explain better the demand for this important commodity.

It was not until relatively recently that the literature explicitly recognized that the price reduction in a health insurance contract represented the vehicle for transferring income from those who purchased insurance and remained healthy, to those who purchased insurance and became ill (Nyman 1999a,b, 2003). Until that time, the ascendant theory held that all moral hazard was welfare decreasing, because it was due entirely to a price distortion, and therefore cost more than it was worth to consumers (Pauly, 1968). This theory was seemingly supported by the most influential empirical study of the day, the RAND Health Insurance Experiment, which found that cost-sharing reduced healthcare utilization substantially, but with no important effects on health (Newhouse, 1993). Policy recommendations to reduce healthcare costs therefore centered on increasing coinsurance rates for all healthcare in order to reduce the seemingly inefficient additional healthcare consumed by the insured (Feldstein, 1973; Manning and Marquis, 1996). Value-based insurance design (Chernew et al., 2007) questioned the across-the-board application of cost sharing, but attempted to re-orient copayments away from those healthcare interventions where the health benefits clearly exceeded costs, on the basis of cost-effectiveness analysis alone. The recognition of an insurance-generated income effect, however, would further re-orient copayments away from healthcare interventions that would otherwise be unaffordable, bringing the financial and disease characteristics of the individual care recipient into the picture as well. Of the diseases studied, we found that diabetes and cancer have the largest proportion of efficient moral hazard, but these diseases may also be those where previous healthcare expenditures have depleted available budgets the most. More directly, we found that the efficient proportion of moral hazard was also higher for ill uninsured individuals at the lower end of the income distribution, those who would be eligible for subsidies under the Affordable Care Act of 2010.

This paper also presented a methodological approach for simulating the behavior of uninsured consumers under two counterfactual insurance regimes: (1) having purchased health insurance that pays off with a coinsurance rate and (2) having purchased health insurance that pays off using a lump-sum amount equal to the insurance spending under the coinsurance rate regime. From this methodology, the additional healthcare expenditures (moral hazard) from the coinsurance policy were predicted, as was the proportion of these expenditures attributable to the income payoff contained within the coinsurance rate price reduction. This simulation indicated that the proportion of moral hazard attributable to income varies across diseases, but can be substantial for those with certain illnesses, notably cancer and diabetes. This study represents the first attempt to estimate the proportion of moral hazard generated by income transfers and that is therefore efficient. Because of the following limitations, however, its primary contribution to the literature may lie more in the development of methods than in the actual findings.

Limitations

Our analytical approach approximated but did not precisely capture the theoretical behavior we sought to describe. For example, we would have liked to have used a representative cross-section of the population and to have assigned them at random to being uninsured or purchasing insurance privately, but instead we used the behavior of those who are presently uninsured, a group of individuals with fewer resources. We would have liked to have used expenditure data that was verified by a third party, but instead we used self-reported expenditure data from the MEPS and HRS. We would have liked to have assigned those who became ill an income/asset shock that was equal to the subsequent spending that would have been done under a coinsurance payoff policy, but instead we used an asset shock that was unrelated to the insurance expenditures. We would have liked to have eliminated any supply-induced demand stemming from becoming insured, but instead we used the entire difference between privately-insured spending and uninsured spending as the measure of moral hazard. Unfortunately, many of these issues would still be unresolved, even with data from a randomized controlled trial. For example, for those trial participants assigned to a contingent claim policy arm, it would be impossible to know exactly how much insured spending actually would have been incurred if they been insured under a coinsurance policy, given the idiosyncratic nature of diseases and variations in physician practice styles.

Perhaps, the greatest departure from theory is caused by using data from the U.S. healthcare system, where it is possible to gain access to healthcare without having insurance or paying for it out of pocket. The theory assumes a hand-to-mouth economy where insurance is the only way to avoid paying for care out-of-pocket, yet uninsured Americans can and do take advantage of various free sources of healthcare—charity and pro bono care; bad debts and other uncompensated care; federal, state and local government and non-profit organization direct-care programs; emergency room care; and so on–where this assumption is violated. It is likely that because of these realities, the efficient proportion of moral hazard as measured here is smaller than the proportion that would have occurred under a more draconian system. As a result, the access value of insurance, as represented in this paper by estimates of the proportion of efficient moral hazard, is likely to be understated.

Acknowledgements

The authors are grateful to the Agency for Healthcare Policy and Research and the Institute of Aging of the National Institutes of Health for providing funding for this study (grant numbers US DHHS AHRQ/5R01-HS018328-02 REV and NIH/5R01-AG041753-02, respectively).

References

Aron-Dine, Aviva, Einav, Liran, Finkelstein, Amy, 2013. The RAND health insurance experiment, three decades later. J. Econ. Perspect. 27 (1), 197–222.
Arrow, Kenneth J., 1963. Uncertainty and the welfare economics of medical care. Am. Econ. Rev. 53, 941–973.
Baker, L.C., Brown, M.L., 1999. Managed care, consolidation among health care providers and health care: evidence from mammography. Rand J. Econ. 30, 351–374.
Bates, L.J., Santerre, Rexford E., 2008. Do health insurers possess monopsony power in the hospital services industry. Int. J. Health Care Finance Econ. 8, 1–11.
Bates, L.J., Mukherjee, K., Santerre, Rexford E., 2010. Medical insurance coverage and health production efficiency. J. Risk Insur. 77, 211–229.
Bhattacharya, Jay, Goldman, D., Sood, N., 2003. The link between public and private insurance and HIV-related mortality. J. Health Econ. 22, 1105–1122.
Bundorf, M.K., 2002. Employee demand for health insurance and employer health plan choices. J. Health Econ. 21, 65–88.
Cawley, John, Simon, Kosali. I., 2005. Health insurance coverage and the macroeconomy. J. Health Econ. 24, 299–315.
Chernew, Michael E., Rosen, Allison B., Mark, Fendrick A., 2007. Value-based insurance design. Health Affairs 26, w195–w203.
Chetty, Raj, 2008. Moral hazard versus liquidity and optimal unemployment insurance. J. Political Econ. 116 (2), 173–234.
Cutler, David M., Zeckhauser, Richard J., 2000. The anatomy of health insurance. In: Culyer, J.A., Newhouse, J.P. (Eds.), Handbook of Health Economics. Elsevier, Amsterdam, Chapter 11.
Deb, Partha, Trivedi, P.K., 2006. Specification and simulated likelihood estimation of a non-normal treatment-outcome model with selection: application to health care utilization. Econom. J. 9, 307–331.
Deb, Partha, Li, C., Trivedi, P.K., Zimmer, D.M., 2006a. The effect of managed care on use of health care services: results from two contemporaneous household surveys. Health Econ. 15, 743–760.
Deb, Partha, Munkin, M.K., Trivedi, P.K., 2006b. Private insurance, selection, and health care use: a Bayesian analysis of a roy-type model. J. Bus. Econ. Stat. 24, 403–415.
Dranove, David C., Simon, J., White, W.D., 1998. Determinants of managed care penetration. J. Health Econ. 17, 729–745.
Feldstein, Martin S., 1973. The welfare loss of excess health insurance. J. Political Econ. 81, 251–280.
Friedman, Milton, Savage, L.J., 1948. 'The utility analysis of choices involving risk. J. Political Econ. 56 (4), 279–304.
Hadley, Jack, Holahan, John, 2003. How Much Medical Care Do the Uninsured Use, and Who Pays for It? Health Affairs Web Exclusive, 12 February, 2003 W 3-66-3-81.
Hirsch, Barry T., Macpherson, David A., 2003. Union membership and coverage database from the current population survey: note. Ind. Labor Relat. Rev. 56 (2), 349–354.
Johnson, R.W., Crystal, S., 2000. Uninsured status and out-of-pocket costs at midlife. Health Serv. Res. 35, 911–932.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–291.
Kaiser Family Foundation, National Research and Education Trust, 2010. Employer Health Benefits: 2010 Annual Survey. Henry J. Kaiser Family Foundation, Menlo Park, California, and Health Research & Educational Trust, Chicago, IL.
Koc, Cagatay, 2005. Health-specific moral hazard effects. South. Econ. J. 72, 98–118.
Koc, Cagatay, 2011. Disease-Specific moral hazard and optimal health insurance design for physician services. J. Risk Insur. 78 (2), 413–446.
Manning, Willard G., Marquis, Susan M., 1996. Health insurance: the tradeoff between risk pooling and moral hazard. J. Health Econ. 15 (5), 609–640.
Marshall, John M., 1976. Moral hazard. Am. Econ. Rev. 66 (5), 880–890.
Mullahy, John, 1997. Instrumental variable estimation of count data models: applications to models of cigarette smoking behavior. Rev. Econ. Stat. 79, 586–593.
Mullahy, John, 1998. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. J. Health Econ. 17, 247–281.
Munkin, M.K., Trivedi, P.K., 2003. Bayesian analysis of a self-selection model with multiple outcomes using simulation-based estimation: an application to the demand for health care. J. Econom. 114, 197–220.
Newhouse, Joseph P., 1993. Free For All? Lessons from the RAND Health Insurance Experiment. Harvard University Press, Cambridge, MA.
Nyman, John A., 1999a. The value of health insurance: the access motive. J. Health Econ. 18 (2), 141–152.
Nyman, John A., 1999b. The economics of moral hazard revisited. J. Health Econ. 18 (6), 811–824.
Nyman, John A., 2003. The Theory of Demand for Health Insurance. Stanford University Press, Stanford, CA.
Nyman, John A., 2007. American health policy: cracks in the foundation. J. Health Polit. Policy Law 32 (5), 759–783.
Olson, C.A., 2002. Do workers accept lower wages in exchange for health benefits. J. Labor Econ. 20, S91–S114.
Pauly, Mark V., Herring, Bradley, 2007. The demand for health insurance in the group setting: can you always get what you want? J. Risk Insur. 74, 115–140.
Pauly, Mark V., 1968. The economics of moral hazard: comment. Am. Econ. Rev. 58 (3), 531–537.
Pauly, Mark V., 2005. Effects of insurance coverage on use of care and health outcomes for nonpoor young women. Am. Econ. Rev. 95, 219–223.
StataCorp, 2013. Stata Statistical Software: Release 13. StataCorp LP, College Station, TX.
Terza, Joseph V., Basu, Anirban, Rathouz, Paul J., 2008. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. J. Health Econ. 27, 531–543.
Town, Robert J, Wholey, Douglas, Feldman, Roger, Burns, L.R., 2007. Revisiting the relationship between managed care and hospital consolidation. Health Serv. Res. 42, 219–239.
Viscusi, Kip W., Evans, William N., 1990. Utility functions that depend on health status: estimates and economic implications. Am. Econ. Rev. 80 (3), 353–374.
Zuvekas, Samuel, Cohen, Joel, Pylypchuk, Yuriy, 2005. Comparison of MEPS and MarketScan Health Care Expenditure Estimates. Internal Memorandum. Agency for Healthcare Research and Quality, Rockville, MD.
de Meza, David, 1983. Health insurance and the demand for medical care. J. Health Econ. 2 (1), 47–54.