HEALTH RISK INCOME, AND THE PURCHASE OF PRIVATE HEALTH INSURANCE

M. Kate Bundorf

Bradley Herring

Mark V. Pauly

ERIU Working Paper 29
http://www.umich.edu/~eriu/pdf/wp29.pdf

Economic Research Initiative on the Uninsured
University of Michigan
555 South Forest Street, 3rd Floor
Ann Arbor, MI 49104-2531

Abstract: While many believe that an individual's health plays an important role in both their willingness and ability to obtain health insurance, relatively little agreement exists on how and why health status is likely to affect coverage rates, particularly for individuals purchasing coverage in the individual market. In this paper, we examine the relationship between health risk and the purchase of private health insurance and whether that relationship differs between people purchasing coverage in the individual and large group markets and between low and high income individuals. The data source for our analysis is the panel component of the 1996-2002 Medical Expenditure Panel Survey (MEPS). We find that health risk is positively associated with obtaining private health insurance coverage. The positive relationship between health risk and coverage is stronger for individuals obtaining coverage in the large group market than for individuals obtaining coverage in the individual market. In the large group market, rates of coverage increase more quickly with health risk for low than high income individuals. We conclude that high premiums for high risks are not a significant contributor to the large uninsured population in the U.S. Among low income individuals, high premiums may represent a barrier to low risks in the large group market.

Introduction

In 2003, approximately 45 million Americans lacked health insurance coverage (DeNavas-Walt, Proctor, and Mills 2004). Although obtaining health insurance is voluntary in the U.S., surprisingly little is known about the factors that determine whether an individual obtains health insurance in private markets (McLaughlin et al. 2004). While many believe that an individual's health plays an important role in both their willingness and ability to obtain health insurance, relatively little agreement exists on how and why health status is likely to affect coverage rates, particularly for individuals purchasing coverage in the individual market.

The almost-universal belief of non-economist policy analysts is that people in worse health have greater difficulty than those in better health obtaining health insurance coverage, either because they face higher premiums that make health insurance unaffordable or because they are often denied coverage by insurers (Gauthier, Lamphere, and Barrand 1995; Hackey 2000; Simantov, Shoen, and Bruegman 2001). Economic theory of rational insurance purchasing, in contrast, predicts that, when insurance premiums reflect an individual's health risk, insurance status may be independent of medical risk (Ehrlich and Becker 1972). Finally, a large theoretical literature in economics raises the concern that, when asymmetric information exists between insurers and consumers about individual health risk, less healthy consumers purchase more health insurance at a given premium than those who are more healthy, creating instability and inefficiency in competitive health insurance markets (Rothschild and Stiglitz 1976).

Existing empirical studies provide little guidance in the face of the conflicting theory. While many studies of the employer-sponsored and Medicare markets have documented evidence of favorable risk selection into less generous relative to more generous plans (Hellinger

1995), these studies do not examine the uninsured. Similarly, studies of the individual health market have found evidence consistent with the existence of adverse selection by examining the amount of health insurance purchased by those obtaining coverage in the individual market (Browne 1992; Browne and Doerpinghaus 1993).

The few studies that explicitly compare the health status of the insured and the uninsured produce interestingly ambiguous results. Studies that find evidence of a correlation between health status and insurance status often find that the direction of the effect is not consistent across varying measures of health. Monheit and Vistnes (1994) find that uninsured workers are less likely to have a chronic condition but more likely to report fair or poor health than insured workers. Pauly and Nichols (2002) find similar conflicting effects of self-reported health status and indicators of presence of chronic conditions on the likelihood of purchasing coverage in the individual market among those without access to group coverage. In a study comparing those who decline employer-sponsored health insurance and are uninsured with those who enroll, Blumberg and Nichols (2001) find that the uninsured are healthier than the insured on some physical health measures but less healthy on others.

Other studies, in contrast, find relatively little evidence of an effect of health status on insurance status. Pauly and Herring (1999) found no statistically significant effect of relative risk (measured by expected medical expenses) on insurance purchases in the large group and individual markets, although they did find that high-risk lower income persons working for small firms were less likely to obtain insurance than otherwise similar low-risk people. Similarly, Cardon and Hendel (2001) found little evidence of adverse selection as a cause of uninsurance in their study of the health insurance coverage of single, employed workers. In particular, their

results indicate that unobserved characteristics, which included the presence of chronic conditions, were not associated with health insurance choices.

In summary, neither theory nor empirical studies provide strong evidence on the relationship between health status and health insurance coverage. There is considerable confusion, even in the health services research literature, about what we should expect. Few studies have directly examined the relationship between health status and insurance coverage, and the evidence that does exist is often contradictory.

Theoretical Framework and Hypotheses

We begin by defining the appropriate conceptual measure of health status for the purpose of examining health as a cause of uninsurance. People purchase health insurance if the utility of the expected benefits of health insurance exceeds the premium they would pay for coverage. The benefits of coverage are twofold. The first is the value of the medical care an individual ultimately consumes that is covered by insurance. Medical care is costly, however, and future spending on medical care is uncertain. As a result, health insurance also provides protection from the financial risk associated with consuming medical care. If the expected benefits of coverage, in the form of expected covered expenditures plus the value of protection from financial risk, exceed the premium, an individual will purchase health insurance.

Health status is relevant in the decision to obtain health insurance because it affects an individual's likely future consumption of medical care, and individuals vary in their expectations of health care utilization. (We consider the effect of health status on premiums below.) The likelihood of a health shock may vary systematically with observable characteristics, such as age. For example, an older person has a much higher probability of being diagnosed with cancer than

a younger person, all else equal. Expected health expenditures may also vary based on realizations of earlier health shocks, such as developing a chronic condition, or information about family history for particular diseases. People in poor health or those that anticipate future poor health are likely, but not certain, to consume more medical care in the future than those in better health.

We define health risk as an individual's expected health expenditures in the next period based on their health in the current period. Health risk is hypothesized to influence rates of coverage because consumers compare the expected benefits of coverage, in the form of covered health expenditures, with the premium they pay for coverage. Risk averse consumers will be willing to pay a premium that exceeds their expected benefits, reflecting the value they obtain from protection from financial risk associated with medical spending. However, for a given premium, the expected benefits of coverage will be greater for high risk consumers than for low risk consumers. This is because they are likely to consume more medical care and obtain more benefits from their health coverage, independent of their level of risk aversion.

Health Insurance Coverage when Premiums vary with Health Risk

While high risk consumers are more likely to purchase coverage for a given premium than low risk consumers, health insurance premiums in competitive markets will vary with health risk (Arrow 1963). Insurance premiums are "actuarially fair" when they are equal to expected health expenditures. In the simplest economic theory, premiums that are actuarially fair and therefore vary precisely with risk should result in universal purchase of insurance by risk averse individuals—regardless of risk level. In a more realistic model in which an administrative "loading" is added to insurance premium, the simple theory (assuming negligible income effects)

says that the decision to purchase insurance, and the amount of coverage, should depend only on the loading, and not on the probability of loss or on the relative sizes of expected losses across potential purchasers. As Ehrlich and Becker (1972) write in their classic article, "if (the loading proportion) were independent of (the loss probability), so also would be the real price of insurance, and (the loss probability) would then have no effect on the incentive to insure".

In contrast, the widespread policy view is that, when premiums increase with risk, high risk consumers are less likely to purchase insurance than low risk consumers facing lower premiums. The economic and policy views can potentially be reconciled by relaxing the assumption of negligible income effects. In particular, at low income levels, the assumption of minor income effects may not apply. For a person with low income, even in economic theory, the effect of high premiums on expected real income may be substantial, and this may affect insurance purchase decisions. If the risk premium—the excess over the fair premium a person is willing to pay—falls beyond some point as income falls, higher risk, low income people would be less likely to buy. But most convenient specifications of risk averse utility functions (e.g., CRRA, CARA), do not necessarily display this property. Moreover, if the demand for medical care is normal, and if the response to user price is greater at higher levels of total copayment, then moral hazard will be greater at lower (real) income levels than at high income levels and so the demand for generous coverage will be lower amongst lower income, higher risk people. So it is possible, but by no means assured, that lower income, higher risk people may demand less coverage or be less likely to demand any coverage.

The Effect of Imperfect Risk Rating of Premiums on Coverage

In the case discussed above, we assumed that premiums were actuarially fair or that the premium for an individual was proportional to his expected health expenditure, possibly with a loading. Premiums, however, may not vary perfectly with individual risk. Insurers may not be able to observe all the information necessary to calculate premiums that reflect all variation across the population in expected expenditures (Rothschild and Stiglitz 1976). Alternatively, either institutional features of the health insurance market, costly underwriting, or regulation may prevent insurers from using information they have or could potentially have in setting premiums. In either case, insurance premiums may deviate from expected expenditures, representing a movement away from actuarially fair premiums toward the average premium (Pauly 1970).

In what follows we consider three possible benchmark ways in which premiums might vary with the individual's risk. At one extreme, premiums might be *perfectly risk-rated.* We shall interpret this to mean that, across risks, premiums are proportional to risk. Line R in Figure 1 displays this case. At the other extreme, premiums may not increase at all with risk. This is the case of *pure community rating*; the horizontal line C shows this. If premiums increase with risk but less than proportionately, we can all this *partial community rating* (or we could just as well call it partial risk rating). The intermediate line I shows this case.

Figure 2 translates these cases into empirically observed "risk-purchase" gradients. Absent income effects, the line R is horizontal because (under the assumption that risk aversion is independent of risk) there should be no relationship between risk and purchase probability. The line R' shows what would happen if income effects, as discussed earlier, affect the purchase of coverage by high risk consumers. In this case, the proportion purchasing coverage declines at the highest risk levels. Other characteristics that are correlated with both expected expenditures and demand for insurance may also influence the slope of this line. For example, recent research

by Finklestein, and McGarry (2003) suggests that highly risk averse people may both buy insurance and avoid risky health behaviors, so the empirical line could even have a positive slope. The line C shows similarly for pure community rating that the high risks buy and the low risks do not; adding income effects (i.e., line C') reduces the probability for lower risks a little. In the case of community rating, insurance coverage is positively correlated with health risk.

We do not draw in a risk-purchase line for the partial risk/community rating case because, as is obvious, such a line can have any slope (with or without income effects) depending on just how the premium varies with risk, how risk averse people are, and how important income effects are. The most we can say in the case of partial risk rating is that the closer partial risk rating is to perfect risk rating, the closer the risk purchase lines in the imperfect case will look like the perfect case, and similarly for community rating. But it would be perfectly possible for the line to have a positive slope at low risks (as community rating effects predominate) and then turn negative at high risks as income effects from such risk rating as is present predominate. In short, in theory, anything could happen.

Hypotheses

Our objective is to examine the relationship between health risk and private insurance coverage. We will assume in our analysis (for reasons described below) that premiums for health insurance are generally not perfectly risk rated. This assumption is the basis of our first hypothesis:

H1: Increasing health risk is associated a higher likelihood of purchasing private health insurance.

However, the theoretical framework developed above implies that both an individual's income and the extent to which the premium they face for insurance coverage reflect their risk will affect the impact of risk on whether they purchase private health insurance. We first consider the extent of risk rating.

Most people in the U.S. obtain coverage through an employer, and in large groups, employee contribution to plan premiums appear to be community rated; the employee out-of-pocket contribution for coverage tends to be the same regardless of health risk (Pauly and Herring 1999). Because the employee contribution is usually only a small fraction of the total premium, averaging about 15% for single coverage and 26% for family coverage (Claxton et al. 2004), all rational employees who do not have coverage from other sources should enroll. However, it appears that some employees choose not to enroll in the coverage offered by their employer in response to a premium contribution (Blumberg, Nichols, and Banthin 2003; Chernew, Frick, and McLaughlin 1997; Cooper and Vistnes 2003; Gruber and Washington 2005) We hypothesize that these will tend to be the lower risks.

A smaller portion of the relevant U.S. population purchases health insurance in the individual market. In the individual market, new applicants are underwritten and risk rated, although guaranteed renewability virtually prohibits re-underwriting. Thus, there might be something approximating risk rating affecting many buyers. (Even if there is perfect risk-rating in the individual market, our first hypothesis will likely still hold since the individual market represents a small portion of the total private market in the U.S., and thus, the positive risk gradient in the group market would more than offset the neutral effect of risk on coverage in the individual market.) Underwriting might also lead to insurers refusing to offer coverage to high-risk applicants. However, it is also possible that risk rating is not perfect in the individual market

if insurers use less information on individual health expenditures in setting premiums than consumers use when choosing to purchase coverage or dropping a plan that they have. While we have no a priori hypothesis about the magnitude of the effect of risk on coverage in the individual market, we assume that the extent of risk rating of premiums is greater in the individual than in the group market. This leads to our second hypothesis:

H2: The risk gradient is steeper in the group market than in the individual market.

Once again, this is based on the assumption that premiums are closer to community rated in group settings than individual settings. If we find no relationship of insurance purchasing with risk in the individual market, that would be consistent with effective risk rating on average.

Finally, we assume that income effects of the type discussed informally above do exist at low income levels: when health insurance is risk rated, high risk, low income people will face premiums that are high relative to income. When insurance is community rated, low risk, low income people will face theses types of income effects. This assumption then generates our third hypothesis:

H3: The risk gradient varies by income, but the direction varies depending on the extent of risk rating. When premiums are risk rated, the risk gradient is less steep for low income than high income individuals. In contrast, when premiums are community rated or partially community rated, the risk gradient is more steep for low income individual than high income individuals.

When premiums are predominantly risk rated, lower income higher risks might even be less likely to obtain insurance than lower risks, so the gradient would potentially have a negative slope or be flat. Put the other way round, the higher likelihood of purchase by higher risks (relative to lower risks) under community rating increases as income falls since lower risks are

increasingly constrained by the pooled premium. There is even an awkward possibility under the intermediate risk/community rating case in which the income effects discussed in Hypothesis 3 cancel out the price effects in Hypothesis 1 so that the risk gradient is flat or has a negative slope.

Methods

We test empirically the relationship between an individual's health risk, as measured by their expected health expenditures in a given year conditional on their prior year health status, and their health insurance coverage. We then examine differences in this relationship between individuals likely to purchase coverage in the group and individual markets as well as individuals in low and high income families within each market.

Data Source

The data source is the Medical Expenditure Panel Survey (MEPS) produced by the Agency for Healthcare Research and Quality. The Household Component of the MEPS is a nationally representative survey of the U.S. civilian, non-institutionalized population, which collects information about medical care expenditures, medical care use, health care conditions and health insurance coverage for survey respondents as well as information on demographic and socioeconomic characteristics. The survey uses an overlapping panel design in which a new sample of households is contacted each year and households are followed over a two year period. Households are interviewed in five rounds conducted over a 2.5 year period to collect data on health care expenditures over two years. In this project, we exploit the panel structure, using information for a given reporting unit over the two year period in which the unit participates (called the reference period). We use data from six reference periods covering 1996-2002.

Measuring Health Risk

Our measure of health risk is a regression prediction of an individual's expected insured health expenditures in a given year if she were privately insured based on her prior year health status. We develop this measure by estimating a model of the relationship between current year insured medical expenditures and prior year health conditions among the privately insured. We then apply this model to the entire study population, regardless of insurance status, holding characteristics of individuals other than age, sex, and indicators of the presence of a chronic condition constant at the mean of the sample. The implicit assumption is that the risk individuals perceive is well proxied by the expenses predicted by our multivariate regression. If moral hazard is present, it will affect the absolute level of the risk measure across insured and uninsured people, but should not much affect the relative measure.

We develop two estimates of expected health expenditures, conditioning on different sets of variables. In the first, which we call "total risk", we estimate a model of current year privately insured medical expenditures that incorporates both age interacted with sex and indicators of the presence of health conditions in the prior year as well as controls for socioeconomic characteristics. We then use this model to predict medical expenditures for all individuals in the sample as a function of age, sex, and prior year health conditions, holding all other control variables constant at the sample mean.

In the second, which we call "condition risk", we focus on the effect of health conditions on expected expense, controlling for age and sex. We reestimate the model, omitting the indicators of chronic conditions, and generate predicted expenditures for the full study sample conditional on only age and sex. We call this estimate of predicted expenditures "demographic risk". We then construct our measure of "condition risk" by dividing our estimate of "total risk" by our estimate of "demographic risk". Condition risk, therefore, measures the extent to which

an individual's expected expenditures deviate from those of others with similar demographic characteristics due solely to the presence of chronic conditions. More information on the development of these measures is available in the technical appendix.

Defining health insurance markets

To test hypotheses regarding differences between the individual and group markets in the relationship between health risk and coverage, we identify people likely to purchase coverage in each of these markets based on family employment status. The MEPS defines an insurable unit as a sub-family relationship unit constructed to include adults plus those family members who would typically be eligible for coverage under the adults' private health insurance family plans. We examine the employment status of each parent in the insurable unit and assign individuals to the individual and large group markets as follows:

1. Large Group Market: individuals with either parent of the insurable unit employed full-time in a firm with greater than 50 employees.
2. Individual Market: both parents of the insurable unit either self-employed in a firm with only one worker or not employed.

In pooled analyses, we use a comprehensive set of market definitions, including identifiers of individuals with access to coverage through small firms or part-time workers, to control for the relative price of coverage across markets. We restrict our cross-market comparisons to individuals in the large group and the individual markets because we believe that these are the cleanest comparisons of differences in rating methods.

Empirical Model and Estimation

We estimate models of the probability of purchasing private health insurance as a function of individual health risk and income, controlling for other characteristics that affect

demand for health insurance. The dependent variable is an indicator of whether the respondent had private health insurance (employer-sponsored group, other group, non-group coverage or source unknown) in December of his second year in the survey. Because the estimates of total risk and condition risk are highly skewed, we use the log transformation of these variables in our empirical models. The control variables include race, ethnicity, education, marital status, family income and family income squared, family size, year, region, urban indicator and the interaction of region and urban indicator.

We include in our study sample 36,663 individuals who participated in the panel component of the MEPS from 1996-2001 and were 25-64 and not covered by public health insurance at any point during their second year in the survey. We exclude 1,604 individuals for missing data, primarily regarding employment status. We also exclude 114 individuals covered by someone outside the health insurance unit, since we cannot determine that policyholder's employment status. We also drop people reporting employer-sponsored coverage whom we initially assigned to the individual market based on family employment status yet we have reason to believe they may have a "legitimate" source of that group coverage. This includes 1,316 likely retirees reporting group insurance (i.e., those who are either over 50 years of age or have any family member who is retired); 26 people reporting COBRA coverage during the year; and 5 people who received workers compensation during the year. Table 1 present descriptive statistics for the study sample.

We estimate the models using maximum likelihood logistic regression. All results are weighted and standard errors are adjusted for clustering at the level of the health insurance unit. We estimate models on the entire sample and then separately for those in the individual and large group markets. These two market samples represent a subset of the full study sample because

not all respondents are classified in these two groups. For example, respondents may be in families with parents employed in a small firm or part-time in a large firm. We also estimate models by family income relative to poverty level, defining low income families as those with income below two times poverty level and high income families as those with family income equal to or exceeding two times poverty level.

We estimate separate models to differentiate between sources of health risk. We first estimate the model entering total risk (expected expenditures predicted by age, sex, and prior year health conditions). Because age and sex both contribute to expected expenditures, controlling for these variables may confound our estimate of total risk. However, age and sex may also be proxies for characteristics, other than expected expenditures (e.g., attitudes toward risk), that affect demand for health insurance. This introduces the possibility that the effects we observe using the measure of total risk, without controlling for age and sex, may be due to either health risk or omitted variables which affect demand for health insurance and are correlated with age or sex. We address this by estimating an alternative version of the model including condition risk (i.e., the measure of the extent to which an individual's expected expenditures deviate from those of others with similar demographic characteristics due to the presence of chronic conditions) and categorical indicators of age and sex. By using categorical indicators, we control for the effects of age and sex on coverage, independent of whether they are driven by health risk or other factors. We consider this a stronger test of the relationship between health risk and coverage because health conditions are less likely to be associated with unobserved preferences for coverage than demographic characteristics. In addition, this measure is more likely to capture private information on the part of the consumer because the conditions in our model are less easily observed by insurers than age and sex.

Results

Figure 1 summarizes our estimates of expected expense from our model of health expenditures. Although age and sex are correlated with expected expenditures, considerable variance exists within demographic groups in expected expenditures due to the presence of prior year conditions. The large dot for each group represents the expected expenditure conditional on age and sex. Expected expenditures rise continuously with age for men, but rise less rapidly with age for women. The distribution around the mean within age and sex group in total risk indicates that considerable within demographic group variation exists in expected (not just actual) expenditures; prior year conditions are important determinants of current year expenditures, independent of age and sex.

In Table 2, we examine how well our classification of the "potential" insurance market (based on family employment status) matches actual source of coverage. People classified as likely to purchase in the individual market are much more likely to do so (20%) than those classified as likely to purchase in the large group market (1%). However, 5% of people we classify as unlikely to have access to employer-sponsored group coverage report that they have this type of coverage. Thus, those classified as likely to purchase in the individual market appear not to be completely reliant on this market, although they are more likely to access coverage this way than others. We also find that those classified as likely to purchase in the individual market are on average higher risk and lower income than those likely to purchase in the group market. Average total risk is $1,848 for individuals likely to purchase in the individual market and $1,394 for individuals likely to purchase in the group market. This is due to both demographic characteristics ($1,599 compared to $1,415) and the presence of health conditions (1.11 vs.

0.98). Fifty-one percent of those in the individual market are classified as low income compared to 13% in the large group market.

In Table 3, we begin to examine the relationship between health risk and the probability that a person purchases private health insurance. In these unadjusted estimates, 73% of individuals in the bottom quintile of the total risk distribution had private health insurance compared to 85% of individuals in the top quintile. The positive relationship between health risk (as measured by total risk) and the purchase of private health insurance is consistent across income levels and health insurance markets. Although individuals in low income families (48%) are much less likely than those in high income families (87%) to have private health insurance, in both cases, the proportion with private health insurance increases with health risk for individuals in both low and high income families. Similarly, although those likely to purchase in the individual market are less likely to have private health insurance (32%) than those likely to purchase in the large group market (89%), in both markets, the proportion purchasing coverage increases with risk. For those in the individual market, the proportion purchasing private insurance increases from 21% to 38% from the lowest to the top quintile of risk. For those in the large group market, the corresponding change is from 84% to 92%.

Table 4 shows our multivariate results for insurance status, controlling for a variety of other characteristics. We continue to find that total risk is strongly associated with private insurance coverage (O.R.=1.49, $p<=0.01$). As discussed in the methods section, total risk may be capturing effects of unobserved characteristics of individuals that are correlated with age and sex, expected health expenditures, and insurance coverage. To address this, we estimate an alternative specification, using condition risk, controlling for demographic characteristics. In this specification, condition risk has a positive, statistically significant effect on the probability

of purchasing private health insurance in the pooled sample (O.R.=1.32, p<=0.01). The effect of condition risk, however, is not statistically significant for individuals likely to purchase coverage in the individual market, although it is in the large group market (O.R.=1.35, p<=0.01).

Table 5 shows the results for risk for the different subsamples split by income, presenting only the results for the variables measuring health risk from the different regressions. The positive relationship between health risk (whether measured by total risk or condition risk) and coverage exists for individuals in both low and high income families when the samples are pooled across markets. For high income families, however, the positive relationship between risk and coverage is driven primarily by those likely to purchase in the large group market. For high income families likely to purchase in the individual market, there is no statistically significant relationship between either measure of risk and coverage. For individuals in low income families, in contrast, both total risk and condition risk are positively associated with coverage in both the individual and large group markets. In summary, for low income families, we find evidence of a positive relationship between health risk and coverage rates in both the individual and large group markets. For high income families, in contrast, we find strong evidence of this positive relationship only in the large group market.

Although coverage generally increases with risk, the magnitude of the effect is relatively small. When the sample is pooled across both income and market, a large change in total risk (a movement from the 5^{th} to the 95^{th} percentile of the risk distribution) is associated with an increase in the probability of coverage of 10 percentage points. (Expected expense at the 5^{th} percentile is $427 while expected expense at the 95^{th} percentile is $3,469.) A similar increase in condition risk is associated with 5 percentage point increase in the probability of coverage. Figures 4a and 4b demonstrate graphically the magnitude of the effect within each market and

income level presented in Table 5. The figure indicates that the magnitude of the relationship is generally similar across markets and income levels, with the exception of low income individuals purchasing coverage in the large group market. In this case, the slope of the line appears to be steeper than those of the others. The change in the predicted probability of coverage associated with a change from the 5th to the 95th percentile of the total risk distribution is associated with a 17 percentage point increase in coverage (Table 5). Indeed, when we test the difference in the marginal effect for total risk between individuals in low and high income families in the group market, we find that the positive effect is significantly larger for low income families relative to high income families.[1] In contrast, we find no significant difference in the marginal effect of total risk between low and high income individuals purchasing coverage in the individual market. Figure 5 indicates that the results are similar when we use the condition risk measure of health risk. The magnitude of the effect of risk on coverage appears not to be very large, although the slope of the line is steepest among low income individuals purchasing coverage in the large group market. Once again, the difference in the marginal effects of condition risk between low and high income individuals in the large group market is statistically significant.

Conclusions

We find that, in aggregate, the likelihood of purchasing private health insurance nearly always increases with expected health expenditures. This was the case for both total health risk, considering age, sex and health conditions simultaneously, and health risk due to the presence of chronic conditions. We consider the relationship between condition risk and rates of coverage as stronger evidence of a positive relationship between health risk and coverage because it is less

[1] We test this by estimating the marginal effect, dy/dx, for the measure of risk, and testing whether the difference between the marginal effects estimated for low and high income individuals in the large group market differs from 0.

likely that the presence of chronic conditions is associated with unobserved preference for health insurance than age and sex.

The positive relationship between health risk, particularly risk related to the presence of prior period health conditions, and coverage is stronger in the large group than the individual market. When the data were pooled by income, condition risk did not have a statistically significant effect on the probability of coverage for those likely to purchase coverage in the individual market. For those likely to purchase coverage in the large group market, in contrast, the relationship between condition risk and coverage was positive and statistically significant. These findings are consistent with more risk rating of premiums in the individual than the large group market, as we hypothesized. While people in the individual market are less likely to purchase coverage than those in the group market, health risk does not appear to be the primary deterrent to coverage in the individual market.

The magnitude of the relationship between health risk and coverage (for the pooled sample), however, is not particularly large. A very large change in total health risk, a movement from the 5^{th} to the 95^{th} percentile of the risk distribution, was associated with a 10 percentage point change in the probability of coverage. Because a relatively small proportion of the population is uninsured, however, the small effect may explain a relatively large proportion of uninsurance, particularly among high income individuals.

Contrary to popular perception, our results provide no evidence that high premiums for high risks are a significant contributor to the large uninsured population in the U.S. Even in the individual market, where individual risk rating or coverage denials are more likely to take place, health risk is associated with either a positive or no relationship with coverage. While one possible explanation for this is that high premiums do not create large income effects among low

income, high risks purchasing coverage in the individual market, another possibility is that these types of income effects exist, but we are unable to identify them in our analysis. There are at least two explanations for why we may not observe these effects even if they exist.

First, our measures of the relevant market for an individual may be noisy. Our results provide some suggestions that this may be the case. (A more complete treatment would allow for the person's occupation to be endogenous, but there are no obvious instruments then for assigning people to different types of markets.) A small proportion of those we identify as likely to purchase in the individual market based on family employment status obtain some type of group coverage. Thus, another interpretation of our results is that individuals without obvious access to group coverage through an employer, particularly those who are high risk, work hard to find coverage through an alternative source. This may include state high risk pools, purchasing alliances, or enrolling in employment-based coverage for which it is not obvious to an econometrician they are eligible.

Alternatively, our measure of risk, particularly condition risk, is likely to capture information about health status that is observable to individuals but not easily observable to an insurer. Our measure includes information about a large number of health conditions that were present in the year prior to the one in which we observe coverage. It is unlikely that insurers have as detailed information about health risk. Indeed, other work, which measures health risk using a methodology that is likely more similar to that which insurers use, finds evidence of a negative effect of health risk on coverage for some consumers purchasing coverage in the individual market (Herring and Pauly 2005). Thus, our results may reflect the effect of a consumer's private information about their health risk on coverage in the individual market more so than the effect of risk rating or coverage denials.

Finally, our results suggest that financial deterrents to coverage may exist among low risks in the group market. In the large group market, where premiums are less likely to be risk rated, we find that the risk gradient is steeper among low than high income individuals. A potential explanation for this is that low income, low risk individuals facing either high out-of-pocket premiums or low wages for jobs with coverage may be less likely to obtain coverage from an employer than high income, low risk individuals. The former case could be true if the employee out-of-pocket premium were high relative to the expected benefits of coverage. In 2002, the average employee contribution for single coverage was approximately $450 – similar to the average expected expenditure for man 25-29 years old (Figure 3). Alternatively, if the wage difference between jobs without and without health insurance reflects the average premium for coverage ($3,060 for single coverage in 2002), the reduction in wages associated with coverage may generate income effects for low income workers that make jobs with coverage unattractive relative to those without coverage.

Does the relationship between risk and the purchase of coverage signal the existence of adverse selection? Although our results are consistent with insurers using less information in setting premiums than individuals use in purchasing decisions, which is a necessary condition for adverse selection, they do not provide direct evidence on whether adverse selection exists. First, standard models of adverse selection do not predict the existence of a persistently uninsured segment of the population. Rather, if an equilibrium exists, it involves positive, but suboptimal levels of coverage for low risks and complete coverage for high risks, the pooling of low and high risks into a single level of coverage, or in the absence of an equilibrium, cycling behavior – none of which is consistent with a large, continuously uninsured subset of the population. The alternative explanation is that unobserved characteristics of individuals are positively correlated

with both expected health expenditures and preferences for insurance. For example, if individuals become more risk averse as they age, risk aversion will be positively correlated with expected health expenditures, which also increase with age (Finklestein and McGarry 2003). In this case, even if premiums were perfectly risk adjusted, the relationship may still exist – high risks just have higher demand for health insurance. In our empirical work, we cannot rule this case out, although our approach to measuring health risk indicates that these types of preferences would need to be correlated with health status independent of age.

References:

Arrow, K. J. 1963. "Uncertainty of the Welfare Economics of Medical Care." *American Economic Review* 53(5):941-73.

Blumberg, L. J. and L. M. Nichols. 2001. "The Health Status of Workers who Decline Employer-Sponsored Health Insurance." *Health Affairs* 20(6):180-87.

Blumberg, L. J., L. M. Nichols, and J. S. Banthin. 2003. "Workers' Decisions to Purchase Health Insurance." *International Journal of Health Care Finance and Economics* 1(3-4):305-25.

Browne, M. J. 1992. "Evidence of Adverse Selection in the Individual Health Insurance Market." *The Journal of Risk and Insurance*:13-33.

Browne, M. J. and H. I. Doerpinghaus. 1993. "Information Asymmetries and Adverse Selection in the Market for Individual Medical Expense Insurance." *The Journal of Risk and Insurance* 60(2):300-12.

Cardon, J. and I. Hendel. 2001. "Asymmetric Information in Health Care and Health Insurance Markets: Evidence from the National Medical Expenditure Survey." *Rand Journal of Economics* 32(3):408-27.

Chernew, M., K. Frick, and C. McLaughlin. 1997. "The Demand for Health Insurance Coverage by Low-Income Workers: Can Reduced Premiums Achieve Full Coverage." *Health Services Research* 32(4):453-70.

Claxton, G., I. Gil, B. Finder, and E. Holve. 2004. "Employer Health Benefits - 2004 Annual Survey". pp. 1-164. Menlo Park, California: The Kaiser Family Foundation and Health Research and Educational Trust.

Cooper, P. F. and J. Vistnes. 2003. "Workers' Decisions to Take-Up Offered Health Insurance Coverage: Assessing the Importance of Out-of-Pocket Premium Costs." *Medical Care* 41(7):35-43.

DeNavas-Walt, C., B. Proctor, and R. J. Mills. 2004. "Income, Poverty, and Health Insurance Coverage in the United States: 2003". *Current Population Reports.* pp. 1-77. Washington D.C.: U.S. Census Bureau.

Ehrlich, I. and G. S. Becker. 1972. "Market Insurance, Self-Insurance, and Self-Protection." *Journal of Political Economy* 80(July/August):623-48.

Finklestein, A. and K. McGarry. 2003. "Private Information and its Effect on Market Equilibrium: New Evidence from Long-Term Care Insurance". *NBER Working Paper.* pp. 1-48.

Gauthier, A., J. A. Lamphere, and N. Barrand. 1995. "Risk Selection in the Health Care Market: A Workshop Overview." *Inquiry* Spring:14-22.

Gruber, J. and E. Washington. 2005. "Subsidies to Employee Health Insurance Premiums and the Health Insurance Market." *Journal of Health Economics* 24(2):253-76.

Hackey, R. B. 2000. "The Politics of Reform." *Journal of Health Politics, Policy, and Law* 25(1):211-23.

Hellinger, F. J. 1995. "Selection Bias in HMOs and PPOs: A Review of the Evidence." *Inquiry* 32(Summer (2)):135-42.

Herring, B. and M. Pauly. 2005. "The Effect of State-Level Rating Regulation on Premiums and Coverage in the Individual Health Insurance Market".

McLaughlin, C. G., S. E. Crow, M. Harrinington, and H. Kuttner. 2004. "Causes and Consequences of Lack of Insurance: Gaps in our Knowledge." In *Health Policy and the*

Uninsured, edited by C. G. McLaughlin, pp. xiii-xxv. Washington D.C.: The Urban Institute Press.
Monheit, A. C. and J. P. Vistnes. 1994. "Implicit Pooling of Workers from large and Small Firms." *Health Affairs* Spring(I):301-14.
Pauly, M. V. 1970. "The Welfare Economics of Community Rating." *The Journal of Risk and Insurance* 37(3):407-18.
Pauly, M. V. and B. Herring. 1999. Pooling Health Insurance Risks. Washington DC: AEI Press.
Pauly, M. V. and L. Nichols. 2002. "The Nongroup Health Insurance Market: Short on Facts, Long on Opinions and Policy Disputes." *Health Affairs* Web Exclusive(23October2002).
Rothschild, M. and J. Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90:629-49.
Simantov, E., C. Shoen, and S. Bruegman. 2001. "Market failure? Individual insurance markets for older Americans." *Health Affairs* 20(4):139-49.

Figure 1: The Relationship between Expected Health Expenditures and Premiums under Different Methods of Risk Rating of Premiums

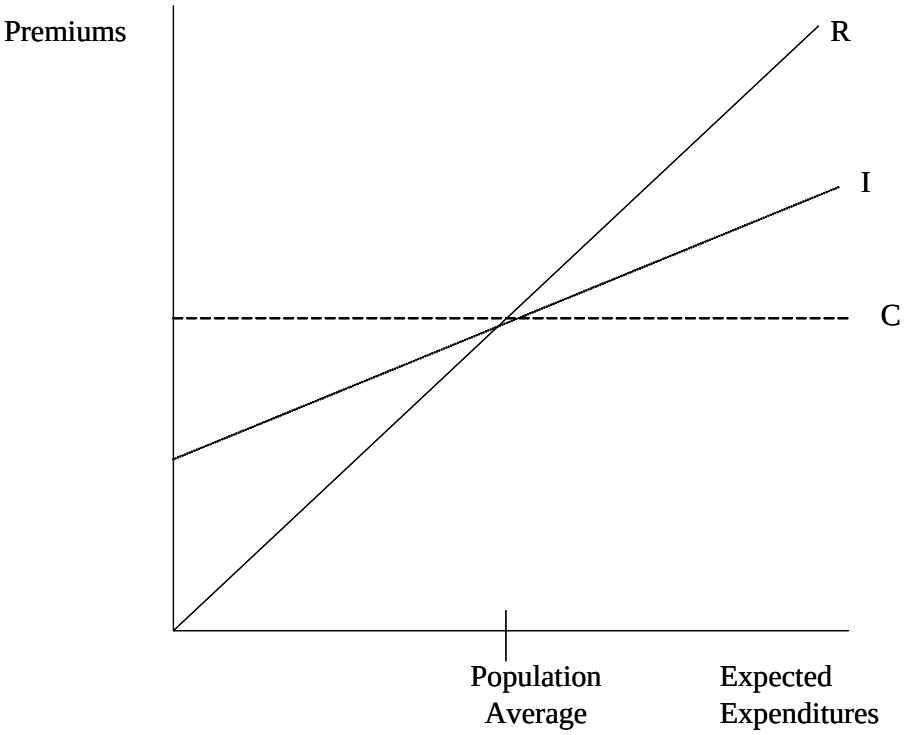

Figure 2: The Theoretical Relationship between Expected Expenditures and Coverage under Different Methods of Risk Rating of Premiums

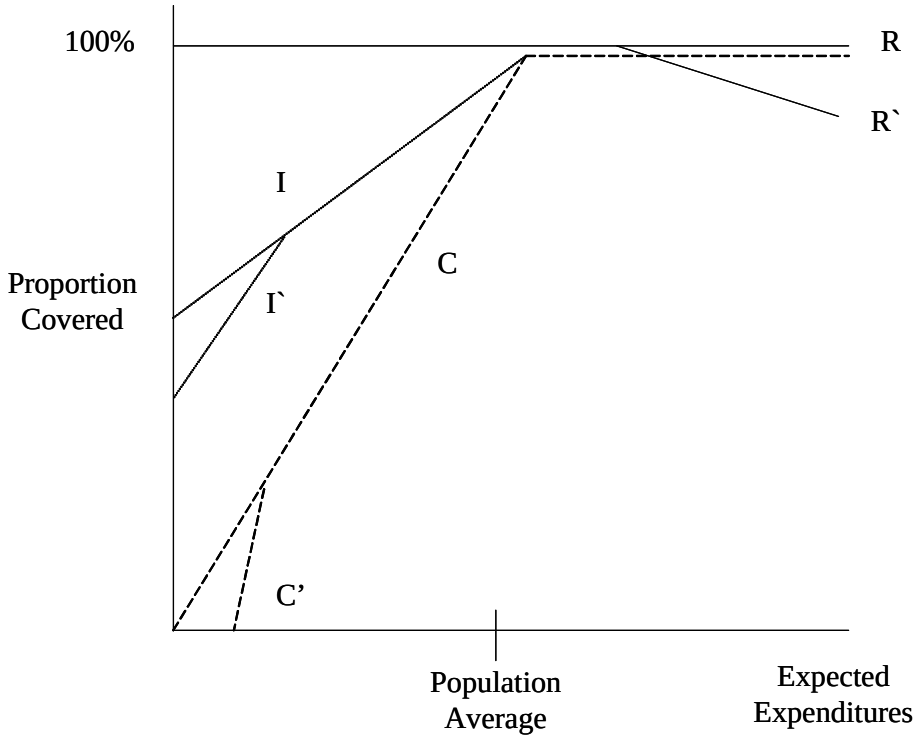

Figure 3: Distribution of Expected Expenditures by Age Group and Sex

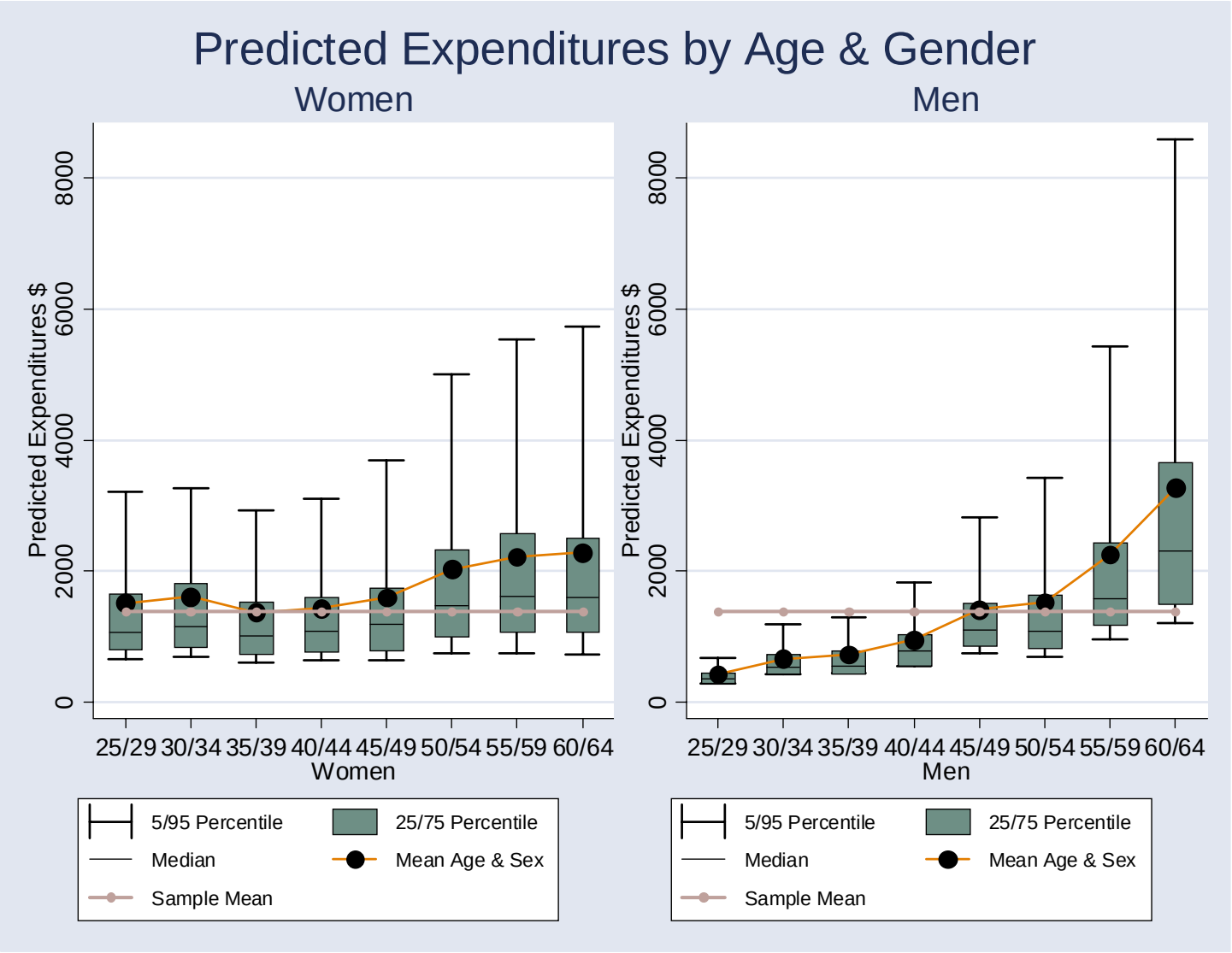

Figure 4a.

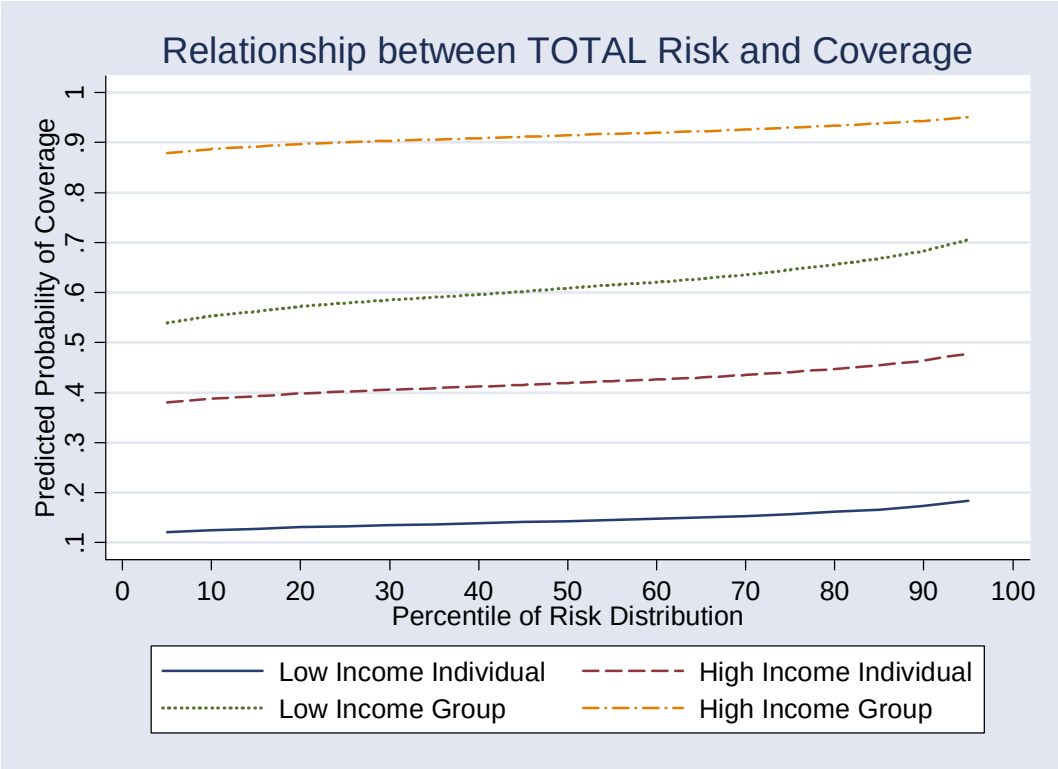

Figure 4b.

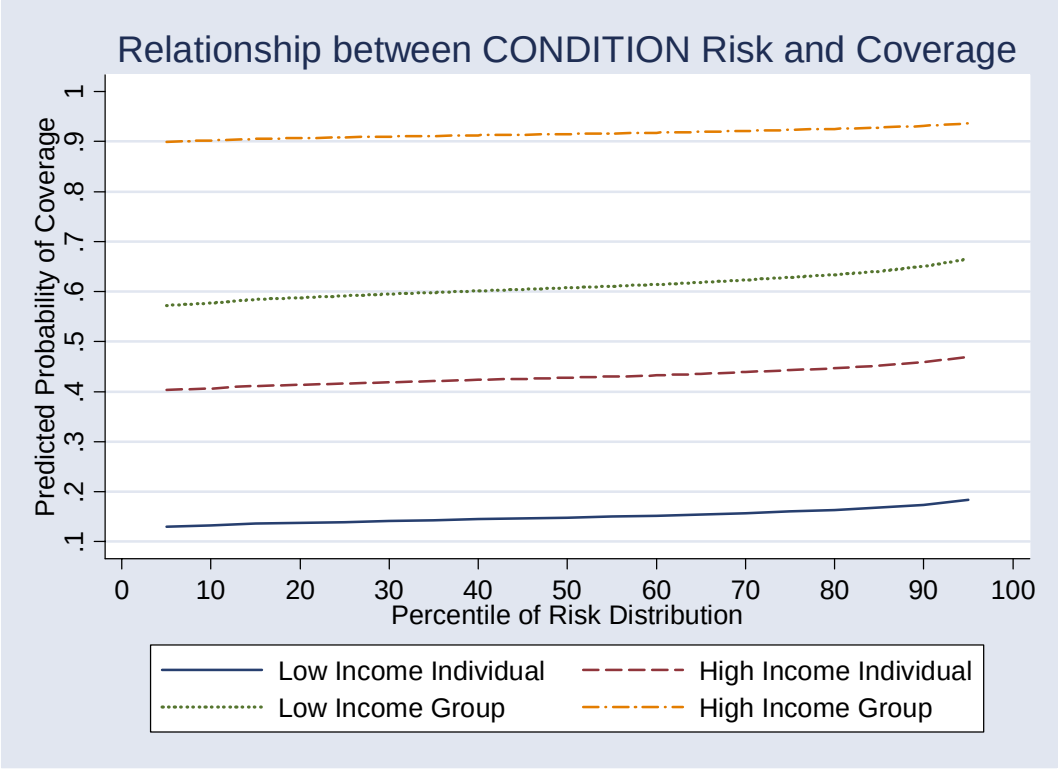

Table 1: Descriptive Statistics for Study Variables

Sample: Individuals 25-64 without Public Health Insurance
n=33,699

Variable	Mean
Demographic Risk	1413.49
Total Risk	1385.63
Condition Risk	0.98
Male 25-29	0.07
Male 30-34	0.07
Male 35-39	0.08
Male 45-49	0.07
Male 50-54	0.06
Male 55-59	0.04
Male 60-64	0.02
Female 25-29	0.06
Female 30-34	0.07
Female 35-39	0.08
Female 40-44	0.08
Female 45-49	0.07
Female 50-54	0.06
Female 55-59	0.04
Female 60-64	0.03
Black	0.10
White	0.70
Hispanic	0.11
HS Diploma or Equivalent	0.51
Bachelor's Degree	0.21
Master's Degree	0.10
Other Graduate Degree	0.08
Married	0.66
Previously Married	0.16
Family Income	34.74
Family Income^2	2024.21
Family Size	2.50
Individual Market	0.05
Uncertain Group Market	0.08
Small Group Market	0.16
Large Group Market	0.71
1997	0.16
1998	0.16
1999	0.17
2000	0.17
2001	0.17
Region Midwest	0.24
Region South	0.35
Region West	0.22
Urban	0.82

Table 2: Key Study Variables by Market

	All	By Market: Individual	By Market: Large Group	
N	33,699	1,944	23,588	
Insurance Coverage				
Any Private	0.80	0.32	0.89	**
Employer	0.76	0.05	0.88	**
Other Group	0.01	0.04	0.01	**
Non-group	0.04	0.20	0.01	**
Uninsured	0.20	0.68	0.11	**
Health Risk				
Total Risk	$ 1,386	$ 1,848	$ 1,394	**
Demographic Risk	$ 1,413	$ 1,599	$ 1,415	**
Condition Risk (Total/Demographic)	0.98	1.11	0.98	*
Income				
Low Income (<2 times poverty level)	0.18	0.51	0.13	***
High Income (> 2= times poverty level)	0.82	0.49	0.87	

Statistical significance refers to test for differences between the individual and large group markets

+ significant at 10%; * significant at 5%; ** significant at 1%

Table 3: Insurance Status by Expected Health Expenditures and Income

		Quintile of Expected Expenditures - Total Risk						
Any Private Insurance Coverage	All	1	2	3	4	5	Change	
Within Quintile Median Expenditure		$ 515	$ 750	$ 1,008	$ 1,443	$ 2,659		
All Income Levels								
All markets	0.80	0.73	0.78	0.83	0.84	0.85	0.11	**
Individual Market	0.32	0.21	0.29	0.36	0.34	0.38	0.16	**
Large Group Market	0.89	0.84	0.87	0.90	0.91	0.92	0.08	**
Low Income (Income <=2 times poverty level)								
All markets	0.48	0.44	0.45	0.51	0.51	0.52	0.09	**
Individual Market	0.18	0.09	0.12	0.21	0.23	0.23	0.15	**
Large Group Market	0.64	0.59	0.61	0.66	0.68	0.69	0.09	**
High Income (Income > 2 times poverty level)								
All markets	0.87	0.81	0.85	0.89	0.90	0.91	0.10	**
Individual Market	0.46	0.34	0.48	0.49	0.44	0.57	0.23	**
Large Group Market	0.93	0.89	0.91	0.93	0.94	0.95	0.06	**

Statistical significance refers to test of difference between the first and 5th quintile within each group

+ significant at 10%; * significant at 5%; ** significant at 1%

Table 4: Models of the Relationship between Health Risk and Coverage

Dependent Variable: Private Health Insurance Coverage
Maximum Likelihood Logistic Regression - odd ratios shown

			By Market (defined by family employment status)			
			Individual		Large Group	
	Total Risk	Components of Risk	Total Risk	Components of Risk	Total Risk	Components of Risk
Total Risk	1.4902		1.2295		1.5962	
	[0.0494]**		[0.1271]*		[0.0742]**	
Condition Risk		1.3227		1.1778		1.3491
		[0.0558]**		[0.1405]		[0.0798]**
Male 25-29		0.6383		1.0539		0.6606
		[0.0624]**		[0.4656]		[0.0798]**
Male 30-34		0.7008		0.3522		0.8348
		[0.0672]**		[0.1719]*		[0.1021]
Male 35-39		0.7564		1.0602		0.8074
		[0.0717]**		[0.4608]		[0.0996]+
Male 45-49		0.9756		0.8372		1.2357
		[0.0987]		[0.3345]		[0.1639]
Male 50-54		1.1269		1.1204		1.6171
		[0.1225]		[0.4910]		[0.2576]**
Male 55-59		1.0882		0.7527		1.5212
		[0.1389]		[0.3281]		[0.2844]*
Male 60-64		1.1163		1.424		1.425
		[0.1605]		[0.6133]		[0.3141]
Female 25-29		0.9688		2.171		1.0199
		[0.1001]		[1.0175]+		[0.1328]
Female 30-34		1.1124		1.4378		1.2272
		[0.1124]		[0.7727]		[0.1572]
Female 35-39		1.3302		1.245		1.5647
		[0.1235]**		[0.6131]		[0.1856]**
Female 40-44		1.3433		0.6782		1.4058
		[0.1217]**		[0.2735]		[0.1626]**
Female 45-49		1.8261		1.4405		2.1866
		[0.1821]**		[0.5602]		[0.2984]**
Female 50-54		1.9235		1.0766		2.2189
		[0.2091]**		[0.4502]		[0.3244]**
Female 55-59		1.7975		1.549		2.3917
		[0.2149]**		[0.6435]		[0.3961]**
Female 60-64		2.0915		1.5097		3.155
		[0.2718]**		[0.5656]		[0.7251]**
Observations	33,699	33,699	1,944	1,944	23,588	23,588
P-value of F-test of joint signficance of age and sex coefficients		0.000		0.053		0.000

Standard errors in brackets
+ significant at 10%; * significant at 5%; ** significant at 1%
Note: Models include all control variables discussed in text

Table 5: Summary of Results from Models Interacting Health Risk, Income, and Market

	All Markets		Individual Market		Large Group Market	
	Odds Ratio (Std. Error)	Magnitude[1]	Odds Ratio (Std. Error)	Magnitude[1]	Odds Ratio (Std. Error)	Magnitude[1]
All Income Levels						
Total Risk	1.4902	0.100	1.2295	0.0625	1.5962	0.092
	[0.0494]**		[0.1271]*		[0.0743]**	
Condition Risk	1.3227	0.053	1.1778	0.0372	1.3491	0.045
	[0.0558]**		[0.1405]		[0.0798]**	
Low Income						
Total Risk	1.4478	0.146	1.3162	0.063	1.4756	0.166
	[0.0733]**		[0.1727]*		[0.1043]**	
Condition Risk	1.3197	0.083	1.3728	0.054	1.3315	0.093
	[0.0851]**		[0.2289]+		[0.1226]**	
High Income						
Total Risk	1.5241	0.083	1.2954	0.098	1.6479	0.092
	[0.0665]**		[0.2042]		[0.1020]**	
Condition Risk	1.3765	0.048	1.2703	0.067	1.4043	0.037
	[0.0783]**		[0.2364]		[0.1122]**	

+ significant at 10%; * significant at 5%; ** significant at 1%

[1] Defined as the change in the predicted probability of coverage based on a movement from the 5th to the 95th percentile of the risk distribution of the entire population

Technical Appendix

Bundorf, Herring, and Pauly, "Health Risk, Income, and The Purchase of Private Health Insurance", September 2005.

Measuring Health Risk

Our measure of health risk is a regression prediction of an individual's expected privately insured health expenditures in a given year if he or she were insured based on his or her prior year health status. We develop this measure by estimating a model of the relationship between privately insured health expenditures in a given year and an individual's health status in the prior year among the privately insured. We then apply this model to the entire study population, regardless of insurance status, holding characteristics of individuals other than age, sex, and indicators of the presence of a chronic condition constant at the mean of the sample. The implicit assumption is that the risk individuals perceive is well proxied by the expenses predicted by our multivariate regression. If moral hazard is present, it will affect the absolute level of the risk measure across insured and uninsured people, but should not much affect the relative measure.

We develop two estimates of expected health expenditures, conditioning on different sets of variables. We first estimate a model of health expenditures that controls for both demographic characteristics and the presence of health conditions in year 1:

$$(1) \qquad Y_{i,2} = f\left(A_{i,1}, S_{i,1}, C_{i,1}, X_{i,2}\right)$$

where $Y_{i,2}$ is individual i's year 2 privately insured health expenditures, A and S are categorical indicators of age and sex, C is vector of year 1 health conditions for individual i, and X is a group of year 2 control variables. We then use this model to

predict annual privately insured medical expenditures for all individuals in the sample, including the insured and the uninsured, conditioning on age, sex, and conditions, holding all other control variables constant at the sample mean:

$$(2) \qquad E_{D,C} = E\left[Y_{i,2} \mid A_{i,1}, S_{i,1}, C_{i,1}, \overline{X_{i,2}}\right]$$

We then reestimate the model omitting the indicators of chronic conditions:

$$(3) \qquad Y_{i,2} = f\left(A_{i,1}, S_{i,1}, X_{i,2}\right)$$

and generate predicted expenditures for the full study sample conditional only on age and sex:

$$(4) \qquad E_{D} = E\left[Y_{i,2} \mid A_{i,1}, S_{i,1}, \overline{X_{i,2}}\right]$$

From these estimates, we develop 2 measures of risk. The first is "total risk", expected expenditures based on demographic characteristics and chronic conditions, $E_{D,C}$. Next, we isolate risk related to health conditions:

$$(5) \qquad R_C = E_{D,C} / E_D \ .$$

R_C measures the extent to which an individual's expected expenditures deviate from those of others with similar demographic characteristics due to the presence of chronic conditions.

The demographic control variables we include in the models are the interaction of sex and age in 5-year increments. Identification of prior year health conditions is based on questions in the MEPS asking respondents if they had any physical or mental health problems, accidents, or injuries. Respondents are prompted to include all conditions regardless of whether they saw a medical provider, received treatment or took medications. Thus, the survey elicits information on conditions independent of treatment and likely picks up conditions among the uninsured. We use only conditions identified

during the respondent's first year of participation in the survey. These conditions are then mapped to ICD-9 codes which are available in the public release version of the MEPS. For our project, we map the ICD-9 codes to Diagnostic Cost Groups using the DxCG software.

The control variables include race, ethnicity, education, marital status, family income, family income squared, year, and firm size of the policy holder as a proxy for plan generosity. Summary statistics for the study sample used to estimate the empirical models of health expenditures are included in Appendix Table 1.

We estimate the expenditure model using a single stage generalized linear model, log transforming the dependent variable, and assuming the variance of $Y_{i,2}$ conditional on the independent variables in the model is proportional to the mean. The results of the expenditure models are presented in Appendix Table 2.

Appendix Table 1: Descriptive Statistics for Expenditure Prediction Model Estimation Sample

Sample: Individual 25-54 covered by private group health insurance for the entire year
n=24,663

Variable	Mean
Total Private Health Expenditures in Year 2	1,528.73
Male 25-29	0.05
Male 30-34	0.07
Male 35-39	0.08
Male 40-44	0.08
Male 45-49	0.07
Male 50-54	0.06
Male 55-59	0.05
Male 60-64	0.03
Female 25-29	0.05
Female 30-34	0.07
Female 35-39	0.08
Female 40-44	0.08
Female 45-49	0.08
Female 50-54	0.07
Female 55-59	0.05
Female 60-64	0.03
Infectious and Parasitic	0.10
Malignant Neoplasm	0.02
Benign/In Situ/Uncertain Neoplasm	0.05
Diabetes	0.03
Nutritional and Metabolic	0.12
Liver	0.01
Gastrointestinal	0.23
Musculoskeletal and Connective Tissues	0.23
Hematological	0.01
Cognitive Disorders	0.00
Substance Abuse	0.00
Mental	0.11
Developmental Disability	0.00
Neurological	0.06
Cardio-Respiratory Arrest	-
Heart	0.13
Cerebro-Vascular	0.00
Vascular	0.01
Lung	0.15
Eyes	0.07
Ears, Nose and Throat	0.40
Urinary System	0.05
Genital System	0.09
Pregnancy Related	0.03
Skin and Subcutaneous	0.10
Injury, Poisoning, Complications	0.19

Symptoms, Signs and Ill-Defined Conditions	0.19	
Neonates	-	
Transplants, Openings, Other V-Codes	0.00	
Race: White	0.71	
Race: Black	0.09	
Race: Other	0.20	
Ethnicity: Hispanic	0.07	
Education: Less than HS	0.07	
Education: HS Diploma or Equivalent	0.50	
Education: Baccalaureate	0.23	
Education: Masters	0.11	
Education: Other Advanced Degree	0.09	
Married	0.73	
Previously Married	0.14	
Never Married	0.14	
Family Income (000s)	38.66	
Family Income^2	2,353.49	
Panel 1997	0.16	
Panel 1998	0.16	
Panel 1999	0.17	
Panel 2000	0.17	
Panel 2001	0.17	
Panel 2002	0.17	1.00
Firm size 1-49	0.30	
Firm size 50-199	0.19	
Firm size 200-499	0.17	
Firm size >=500	0.24	
Firm size: Unknown	0.11	

Source: Panel Component of the 1996-2001 Medical Expenditure Panel Survey

Appendix Table 2: Results of Expenditure Models

Dependent Variable: ln (Year 2 Privately Insured Health Expenditures)
Estimation Sample: Individuals 25-64 who were privately insured all of year 2
Model Estimation: Maximum likelihood assuming the errors follow the poisson distribution

	Measures of Health	
Independent Variables	Demographics and Prior Year Health Conditions	Demographic Characteristics
Male 25-29	-0.466	-0.646
	[2.20]*	[3.00]**
Male 30-34	-0.216	-0.333
	[1.75]+	[2.65]**
Male 35-39	-0.135	-0.193
	[1.05]	[1.48]
Male 45-49	0.311	0.414
	[2.85]**	[3.45]**
Male 50-54	0.258	0.488
	[2.70]**	[4.91]**
Male 55-59	0.642	0.966
	[4.50]**	[5.28]**
Male 60-64	0.758	1.233
	[5.72]**	[9.75]**
Female 25-29	0.24	0.52
	[2.18]*	[4.95]**
Female 30-34	0.351	0.673
	[3.72]**	[7.18]**
Female 35-39	0.146	0.397
	[1.64]	[4.57]**
Female 40-44	0.252	0.51
	[2.72]**	[5.35]**
Female 45-49	0.171	0.51
	[1.99]*	[6.11]**
Female 50-54	0.328	0.774
	[3.62]**	[8.99]**
Female 55-59	0.363	0.877
	[3.56]**	[9.05]**
Female 60-64	0.417	0.976
	[2.27]*	[5.28]**
Infectious and Parasitic	0.086	
	[1.61]	
Malignant Neoplasm	0.843	
	[6.17]**	
Benign/In Situ/Uncertain Neoplasm	0.01	
	[0.16]	
Diabetes	0.499	
	[4.64]**	
Nutritional and Metabolic	0.25	

	[4.31]**	
Liver	0.824	
	[2.31]*	
Gastrointestinal	0.209	
	[4.58]**	
Musculoskeletal and Connective Tissues	0.407	
	[7.98]**	
Hematological	0.303	
	[2.16]*	
Cognitive Disorders	0.321	
	[0.98]	
Substance Abuse	-0.085	
	[0.56]	
Mental	0.412	
	[6.28]**	
Developmental Disability	0.516	
	[1.39]	
Neurological	0.254	
	[3.69]**	
Heart	0.464	
	[8.11]**	
Cerebro-Vascular	0.844	
	[2.55]*	
Vascular	0.285	
	[2.29]*	
Lung	0.205	
	[2.96]**	
Eyes	0.12	
	[1.58]	
Ears, Nose and Throat	0.187	
	[4.01]**	
Urinary System	0.284	
	[3.49]**	
Genital System	0.099	
	[1.71]+	
Pregnancy Related	0.973	
	[12.40]**	
Skin and Subcutaneous	0.074	
	[1.30]	
Injury, Poisoning, Complications	0.141	
	[2.87]**	
Symptoms, Signs and Ill-Defined Conditions	0.188	
	[3.66]**	
Transplants, Openings, Other V-Codes	2.527	
	[3.77]**	
Race: Black	-0.043	-0.243
	[0.44]	[2.48]*
Race: Other	-0.131	-0.166
	[1.57]	[2.01]*
Ethnicity: Hispanic	-0.062	-0.162

	[0.86]	[2.22]*
Education: HS Diploma or Equivalent	0.075	0.115
	[0.86]	[1.31]
Education: Baccalaureate	0.052	0.074
	[0.55]	[0.79]
Education: Masters	0.284	0.357
	[2.03]*	[2.56]*
Education: Other Advanced Degree	0.195	0.232
	[1.76]+	[2.12]*
Married	0.127	0.09
	[1.86]+	[1.20]
Previously Married	0.197	0.24
	[2.18]*	[2.46]*
Family Income	-0.004	-0.007
	[2.19]*	[3.93]**
Family Income^2	0	0
	[1.40]	[2.77]**
Panel 1997	-0.156	-0.17
	[1.72]+	[1.60]
Panel 1998	-0.388	-0.295
	[4.30]**	[2.93]**
Panel 1999	-0.413	-0.449
	[1.70]+	[3.47]**
Panel 2000	-0.09	-0.492
	[0.90]	[5.07]**
Panel 2001	-0.167	-0.315
	[4.22]**	[4.92]**
Firm size 1-49	-0.027	-0.077
	[0.38]	[0.98]
Firm size 50-199	-0.002	-0.011
	[0.04]	[0.16]
Firm size 200-499	0.014	0
	[0.20]	[0.00]
Firm size: Unknown	-0.078	-0.024
	[0.83]	[0.24]
Constant	6.42	7.149
	[38.22]**	[41.84]**
Observations	24663	24663

Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%
Standard errors adjusted for clustering within household.