NBER WORKING PAPER SERIES

# HEALTH RISK, INCOME, AND EMPLOYMENT-BASED HEALTH INSURANCE

M. Kate Bundorf
Bradley Herring
Mark Pauly

Working Paper 11677
http://www.nber.org/papers/w11677

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2005

This project was funded by a grant from the Economic Research Initiative on the Uninsured (ERIU). An earlier version of this manuscript was presented at the ERIU 2004 Consumer Preferences and Coverage Choices Research Meeting, the 2005 AcademyHealth Annual Research Meeting, and the 2005 International Health Economics Association Meeting.

Health Risk, Income, and Employment-Based Health Insurance
M. Kate Bundorf, Bradley Herring, and Mark Pauly
NBER Working Paper No. 11677
October 2005, Revised December 2008
JEL No. I1


## ABSTRACT

While many believe that an individual's health plays an important role in both their willingness and ability to obtain health insurance in the employment-based setting, relatively little agreement exists on the extent to which health status affects coverage rates, particularly for those with lower incomes. In this paper, we examine the relationship between health risk and the purchase of group health insurance and whether that relationship differs by a person's income and whether they obtain coverage in the small, medium, or large group market. Using the panel component of the 1996-2002 Medical Expenditure Panel Survey (MEPS), we find that health risk is positively associated with private health insurance across the different markets, and that this positive relationship is stronger for low and middle income people, particularly in the large group market. Our results are consistent with the existence of adverse selection in the group market in the form of low rates of coverage among low risks due to an absence of risk rating of premiums. We conclude that pooled premiums for low risks, particularly those with low incomes, may represent a more important financial barrier to coverage in voluntary group insurance than high premiums for high risks.



M. Kate Bundorf
Health Research and Policy
Stanford University
HRP T108
Stanford, CA 94305-5405
and NBER
bundorf@stanford.edu

Bradley Herring
Department of Health Policy & Management
Bloomberg School of Public Health
Johns Hopkins University
624 North Broadway, HH #408
Baltimore, MD 21215
bherring@jhsph.edu

Mark Pauly
Health Care Systems Department
University of Pennsylvania
208 Colonial Penn Center
3641 Locust Walk
Philadelphia, PA 19104-6218
and NBER
pauly@wharton.upenn.edu

## I. Introduction

In 2007, approximately 45.7 million Americans were uninsured (DeNavas-Walt, Proctor et al. 2008). 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, Crow et al. 2004). While many believe that an individual's health plays an important role in both their willingness and their ability to obtain health insurance, relatively little agreement exists on how and why health status is likely to affect coverage rates, particularly for people with low incomes. Other things equal, a person in poorer health will attach more value to a given health insurance policy than a person in better health, but the person in poorer health may be charged a higher premium or find it harder to obtain coverage. There is a demand-side effect (stronger demand) and a supply-side effect (higher premiums); which effect will predominate is ambiguous in theory and can only be settled by empirical estimation.

In the group health insurance market, people rarely pay premiums that vary explicitly with their health status, and many identify risk pooling, or the ability to maintain heterogeneous risk pools in the absence of individual risk rating of premiums, as an important advantage of employer-sponsored coverage (e.g. Schoen, Davis et al. (2008) and Furman (2008)). Yet, a large theoretical literature in economics points to potential problems with pooled premiums. When asymmetric information exists between insurers and consumers about a person's risk, high risk consumers will purchase more coverage than low risk consumers at a given premium, creating instability and inefficiency in competitive health insurance markets (Rothschild and Stiglitz 1976).[1] Indeed, policy analysts have long been concerned that younger, healthier workers will

[1] Newhouse (1996) demonstrates that the introduction of contracting costs into the Rothschild and Stiglitz (1976) model may change the equilbirum outcome. If it is costly for insurers to develop products that generate self selection of low risks into less generous coverage, these costs effectively increase the load on the less generous plan. If the load is large enough, it is possible to generate a pooling equilibrium in which coverage is less than full and

forego coverage – and have increasingly considered the use of individual mandates to address the problem of uninsurance in voluntary markets.

In the small group market, policy concern generally revolves around the limitations of pooling relative to the large group market. Due to their small size, small groups are characterized by greater variation in their expected expenditures than large groups, leading to greater variation in risk-rated premiums. While economic theory of rational insurance purchasing predicts that, when insurance premiums reflect an individual's risk, insurance status may be independent of risk (Ehrlich and Becker 1972), the nearly universal policy concern over risk rating in the small group market is that it makes health insurance unaffordable for high risk individuals or groups. Differences in risk across small employers also generate concern on the part of insurers that their plans will experience unfavorable selection. The actions that plans take in response to these concerns, such as adopting pre-existing conditions exclusions or denying coverage to particular types of groups, may make coverage less accessible or less valuable to high risks.

In addition, recent research shows that high risks with small group coverage are much more likely than those with individual insurance to lose coverage, likely because they lose their coverage when they lose or change their jobs (Pauly and Lieberthal 2008). More generally, employers may be reluctant to hire workers who are high risks or who have high risk dependents; however protective group insurance may be for people who can get it, in order to limit adverse selection, employers or unions may have made it difficult for high risks to obtain coverage.

While there is concern about adverse selection (when high risks get coverage but low risks do not) and about problems of affordability when premiums are risk rated (when low risks

---

represents the efficient level for low risks in response to the higher price due to the higher load but represents an inefficiently low level of coverage for high risks.

get coverage but high risks do not), existing empirical studies of group health insurance markets provide little guidance in the face of these conflicting potential problems. While a large body of research documents the existence of risk-based selection in community-rated settings in private health insurance markets, almost all studies focus on differences among the insured with regard to the plan or level of coverage chosen, rather than whether or not the person has any insurance at all (Browne and Doerpinghaus 1993; Hellinger 1995; Glied 2000). The few studies that explicitly compare the health status of people with and without insurance produce interestingly ambiguous results. Some have found relatively little evidence of an effect of health status on insurance status. Pauly and Herring (1999) did not find a statistically significant relationship between risk (measured by expected medical expenses) and insurance coverage in the large group market, but they did find that high-risk, lower income people 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 among single, employed workers.

Other studies, which document a correlation between health status and employment – related insurance status, find that the direction of the effect is not consistent across varying measures of health. Studies comparing uninsured and insured people in the employer-sponsored market have found that, while people with chronic conditions are more likely to obtain coverage, those with worse self-reported health are less likely to obtain coverage (Monheit and Vistnes 1994; Bernard and Selden 2006). Studies examining people offered health insurance from an employer have found that who decline coverage and are uninsured are healthier than those who enroll on some physical health measures but less healthy on others (Blumberg and Nichols

2001). Similarly, those who decline coverage are more likely to report poor health, yet less likely to have a high cost chronic condition than those who enroll (Bernard and Selden 2006).

In summary, neither theory nor empirical evidence provides a clear picture of the relationship between health status and 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.

In what follows, we examine the relationship between having private coverage and risk for households in which at least one member is employed full time, and we estimate separate parameters for employment settings involving large, medium, and small firms. We also estimate separate models by family income to determine if the relationship between risk and coverage varies by family income. Our dependent variable measures whether or not the person has any private insurance coverage, whether obtained in the group market or the individual market. Among people likely to purchase in the large or medium group markets, the proportion in our sample with private individual insurance is one percent or less, much too small to treat as a separate category. Among people likely to purchase coverage in the small group market, the proportion is larger - 5 percent. The data suggest that, for this subset of the population, insurance status depends both on how group and individual markets function. If higher risk workers lose group coverage and then fail to obtain individual coverage, that is an overall effect of the employment-based system. Similarly, the relative pricing between the individual and small group markets may influence coverage decisions. In either case, the relationship between coverage and risk is not properly attributed to either group or individual markets. Thus, while we focus on coverage from either source for the bulk of our analyses, for people likely to

purchase coverage in the small group market, we also test the sensitivity of the results by estimating models in which we allow the relationship between risk and coverage to vary between group and individual coverage.

## II. Theoretical Framework

We begin by outlining a theoretical framework, based on expected utility theory, for analyzing how health status should affect obtaining health insurance in an individual setting. (We will later link these individual demands to models of insurance purchase for groups.) If the expected benefits of coverage, including both the value of the expected covered expenditures and the value of protection from financial risk, exceed the premium the person would pay to obtain coverage, he or she is likely to purchase health insurance.[2] Holding the premium constant, the benefits of coverage are greater for people who are likely to use greater amounts of insured medical care, and current health status provides a signal of likely future utilization. The likelihood of a health shock and associated health spending 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 at any point in time may also vary based on realizations of earlier health shocks, such as the earlier development of a chronic condition, on family history for particular diseases, or on other factors that determine the demand for care, such as income and tastes.

[2] This development assumes that a person without health insurance pays out-of-pocket for any medical care he consumes. However, the availability of uncompensated care for the uninsured may function as informal insurance by reducing the out-of-pocket spending associated with medical care consumed while uninsured and thus reduce the benefits of obtaining coverage (Coate 1995; Herring 2005).

People in poor health with a given insurance policy who anticipate future poor health are likely, but not certain, to consume more medical care in the future than those in better health.[3]

*II.A. Demand for Coverage and Risk Rating of Premiums*

The extent to which premiums vary with risk will affect demand for coverage. If health insurance premiums are proportional to expected health expenditures, the decision to purchase insurance, and the amount of coverage, should depend only on the administrative "loading," and not on risk, in the absence of net income effects (Ehrlich and Becker 1972).[4] (We will discuss income effects below.) In practice, premiums may not vary perfectly with what individual buyers know about their 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 (Pauly 1970). A movement from a situation in which premiums are risk rated toward one in which they are community rated creates a positive correlation between risk and coverage if consumers use information about their expected expenditures when purchasing coverage and insurers do not know or are not allowed to use the information buyers have.

[3] Health status may also affect the value of health insurance through an effect on variation in expected medical expenditures, which increases the uncertainty associated with future spending on medical care. In theory, poor health could be either positively or negatively associated with variation in expected health expenditures. Although little explicit analysis of the relationship between health status and expenditure variation exists to guide the development of hypotheses, the results of our expenditure prediction models indicate that the variance of expected expenditures is positively correlated with the mean.

[4] When people vary in both their risk and their risk aversion, the correlation between the two will determine whether risk is positively, or negatively correlated or even uncorrelated with coverage. Cutler, Finkelstein, and McGarry (2008) examine the correlation between risk and risk aversion in health insurance (as well as other insurance markets) and find that measures of risk aversion are positively correlated with some and negatively correlated with other measures of medical care utilization.

*II.B. Income Effects, Affordability, and Moral Hazard*

These relationships may not hold, however, if the assumption of minor income effects does not apply. There are two income effects that can arise in health insurance markets. One (which we will focus on here) arises from the size of the premium relative to household income. The other, which we will not consider here, arises from any (net) income effects on the demand for medical care (Nyman 1999).

For a low income buyer, the proportionate effect of high premiums on the amount of income remaining for non-health consumption may be substantial, and this may affect insurance purchase decisions. If premiums bear some (positive) relationship with risk, then the proportion purchasing coverage may decline at the highest risk levels due to budget constraints. The important consideration here is how the buyer's income available for other consumption if insurance would be purchased at a (high) premium compares to what consumption would be if the person runs the risk of paying for care if illness strikes versus owing nothing if the person stays perfectly healthy. In contrast, when premiums are community rated, the proportion purchasing coverage may decline at the lowest risk levels; for low risk, low income families, community rated premiums may represent a substantial portion of income.[5]

The extent of moral hazard associated with a given policy may vary by both risk and income. If low risks are more responsive to a reduction in the user price created by insurance, the value of the expected benefits of coverage (relative to either a risk rated premium or a community rated premium) will be lower for low risks than for high risks. Moreover, if the

[5] This relationship between coverage and income may also be explained by risk aversion. If the risk premium—the excess over the fair premium a person is willing to pay—falls beyond some point as income falls, low income people facing high premiums may be less likely to buy. However, most convenient specifications of risk averse utility functions (e.g., CRRA, CARA), do not necessarily display this property.

demand for medical care displays greater responsiveness to higher user prices when copayments are high relative to one's level of income, then moral hazard will be greater at lower income levels than at high income levels and so the demand for generous coverage will be lower amongst lower income people. So it is possible, but by no means assured, that lower income, higher risk people may demand relatively more coverage or be relatively more likely to demand any coverage than lower income, low risks, even when premiums are risk rated.

*II.C. Conceptual Framework for Employment-Based Insurance*

How might employment-based insurance impact these relationships? Because there is no complete and generally accepted model of equilibrium in labor markets with both employer and employee benefits choice, we do not propose a single conceptual framework. Instead, we ask what frameworks are consistent with each of three possible outcomes for people employed by firms that might offer group insurance: lower risks are more likely to be uninsured, higher risks are more likely to be uninsured, or the likelihood of being uninsured is independent of risk, other things equal. The answer to this question is bound to be complex since both the total premium per insured person and the explicit premium charged to employees can vary with the risk composition of the labor force and with employers' knowledge and goals for the premium-risk relationship.

A model that is easy to analyze (though not literally realistic) is one in which workers are sorted into firms whose work forces are homogeneous with regard both to risk and risk aversion. This "Tiebout type" model reproduces the equilibrium that would occur if individual workers could buy insurance at the (net-of-tax-subsidy) group premium (Goldstein and Pauly 1976) and groups could exclude as well as include members. Absent state regulation of group premium

rates, the equilibrium premium will vary with the form and generosity of the insurance offered and with the risk level of the (homogeneous) workers in each firm. The form and generosity of insurance in turn is equal to what the firm's workers would have demanded had they faced the premium schedule associated with their group's size and risk level. In this model, all workers in a group offering coverage will take that coverage, since there is no reason to experience the wage offset for this premium and yet not take the benefit it pays for. Some firms would not offer insurance, would pay higher money wages than those firms which do offer, and would attract workers whose value of insurance was below the wage differential. As noted above, the administrative loading and risk aversion would affect the demand for insurance, but the risk level would not. To make the risk level matter for purchase of insurance, one might assume that there are positive income effects on the demand for insurance (even though this implies a special form of the utility function) and for insured care. Low income, high risk consumers might not be willing to make the money wage sacrifice needed to obtain coverage, both because the less attractive but less costly alternative of charity care is available, and because they have lower willingness to pay for insurer administrative efforts. In this model, risk and insurance purchase would not be related for higher income people, but might well be related for a lower income subset.

An alternative model is one in which firm workforces become heterogeneous with respect to risk as workers age but the costs of changing jobs keep heterogeneous pools together to some extent. In such a model, Bhattacharya and Vogt (2006) show that higher risk workers are more likely than lower risk workers to have coverage, since lower risk workers are likely to see a larger gain from moving from a firm that does not offer insurance, where they would be paid a higher money wage. Some low risk workers may remain in some pools, however, depending on

the makeup of the different groups and variations in risk aversion. In effect, this model reproduces one of the stages of the Rothschild-Stiglitz story, but uses friction costs to make that stage a stable equilibrium. In this model, the low risks who go without coverage are all in firms that attract low risks and do not offer coverage.[6]

It is also possible (though hard to explain in theory) that firms which levy relatively high explicit employee premiums (perhaps in an attempt to shift the source of coverage to a working spouse's employer) may cause some very low risk workers to judge taking the offered coverage to be worth less than the explicit premium. This requires that the explicit premium exceed the expected value of their benefits for the lowest cost plan, and by enough to offset risk aversion. Models of within-firm adverse selection have been offered (e.g., Cutler and Reber (1998)) but they usually do not predict zero coverage, only excessive choice of the lowest coverage option.

In each of the situations above, when risk was correlated with coverage, it was low risks, rather than high risks, who were uninsured. An important caveat to these results is that it is, in theory, possible for low risks to obtain coverage in the individual market. Even if adverse selection results in non-coverage in the group market, low risks could obtain insurance in a risk-rated individual market. Thus, for adverse selection in the group market to be a cause of uninsurance, low risks must find either coverage too costly (the load too high) or search too costly in the individual market.

Is it possible to generate uninsurance among high risks, rather than low risks, in an employment-based setting? In cases in which individual premiums are perfectly risk rated through either cross firm sorting of workers based on risk or within firm variation in wage offsets

[6] This model generates uninsurance through the inability of employers to observe the health status of workers and, correspondingly, the non-existence of differential wage offsets within a firm. Thus, the degree to which wage offsets reflect individual risk is important for determining the extent of uninsurance; the greater the degree to which premiums are risk rated through differential wage offsets, the less uninsurance through adverse selection this model will generate.

for coverage based on risk, the particular scenarios we discussed earlier in which income effects among high risks lead to non-coverage may apply. It is also possible that workers in particular types of firms, especially small firms, may not be able to obtain coverage because they are high risk. Insurers may collect information on the health status of workers in a small firm seeking insurance through underwriting (and insurers have compiled information on health status for workers in a small firm renewing coverage), and thus may be more likely not to offer coverage or quote a premium that is very high when observed risk is disproportionately high. Moreover, insurers may not sell coverage in particular industries if the possibility of unfavorable selection based on unobserved risk is high.

*II.D. Resulting Hypotheses*

Taken together, this discussion suggests that the relationship between health risk and employment-based coverage may vary considerably, depending on how the premium varies with risk, how insurers and employers use information on health status when determining eligibility for coverage, how workers sort across jobs, how risk averse people are, and how important income effects are. While the relationships are potentially complex, our analyses are based on the following four hypotheses which we believe incorporate the predominant features of demand for coverage in the group market.

*H1: Because the explicit premiums faced by workers in the employment-based insurance market are not perfectly risk rated, increasing health risk is associated a higher likelihood of obtaining private health insurance.*[7]

[7] In principle, premiums in the employment-based market may be risk rated through variation by risk in either the explicit contributions that employees make to enroll in a plan offered by their employer or the implicit amount the

*H2: Because the use of medical underwriting and risk-rating may be more prevalent in the small group market, the positive relationship between risk and coverage is stronger in the large group market than in the small group market. (Because of the lower loading, people at all risk levels will be more likely to have insurance in large group markets than in smaller group markets.)*

*H3: Because high risk, low income people in the small group market may face premiums that are high relative to income (due to the relatively higher prevalence of risk rating), the positive relationship between risk and coverage may be weaker for low-income people than for high-income people in the small group market.*

*H4: Because low risk, low income people in the large group market may face premiums that are high relative to their income (due to the relatively higher prevalence of community rating), the positive relationship between risk and coverage may be stronger for low income people than for high income individuals.*

## III. Methods

We test empirically the relationship between health insurance coverage and an individual's health risk, as measured by their expected health expenditures in a given year conditional on their prior year health status. We examine differences across the market in which

---

employee pays for coverage in the form of lower wages. In practice, employee contributions tend to be the same regardless of health risk (Pauly and Herring, 1999; Keenan, Buntin et al., 2001), although employee contributions do generally vary based on family size. The extent to which employees of varying risk pay different premiums for coverage through differential wage offsets is less clear. While there is some evidence that wage offsets for health insurance vary across workers with different demographic characteristics (Gruber 1994) or even indicators of health status such as obesity (Bhattacharya and Bundorf 2005), many observers are skeptical that extensive risk rating of premiums is feasible through this mechanism which would require either differential wage offsets among insured workers within a firm or extensive sorting of workers across firms based on risk.

an individual is likely to purchase coverage (where the market is defined by firm size) in the relationship between the likelihood of purchasing coverage and risk. We also examine differences by income within each market to assess the importance of income effects in this relationship.

*III.A. 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 for a sample of adults in wage-earning households between 25 and 64, excluding those covered by public insurance, from six reference periods covering 1996-2002.

*III.B. Measuring Health Risk*

Our main measure of health risk is a regression-based 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's predictions 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 people 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.

Our primary estimate of risk, which we call "total risk", is expected insured expenditures conditioning on age, sex, and prior year health conditions. We obtain this prediction by estimating a model of current year, privately insured medical expenditures on age (in 5-year increments) interacted with sex and indicators of the presence of health conditions in the prior year as well as controls for a variety of socioeconomic characteristics (listed below) on the subsample of individuals who were continuously privately insured throughout the reference period. Identification of prior year health conditions is based on questions in the MEPS in which surveyors ask respondents if they had any physical or mental health problems, accidents, or injuries. Respondents are prompted to include all conditions regardless of whether or not they saw a medical provider, received treatment, or took medications. Thus, the survey is designed to elicit information on conditions independent of treatment. We limit to prior year conditions by using only conditions identified by the respondent during the first year of participation in the survey relating these conditions to expenditures during the second year. In the public release version of the MEPS, these conditions are mapped to ICD-9 codes. We map the ICD-9 codes to Diagnostic Cost Groups (DCGs) based on Aggregated Condition Categories (ACCs) using the

DxCG algorithm. We exclude a subset of ACCs that are either non-existent or extremely rare (cardio-respiratory arrest; neonates; and transplants, openings and other V-codes), that are unlikely to represent chronic conditions (pregnancy-related; injury, poisonings and complications), and that the reporting of which is likely to be influenced by insurance status (screening/history). We then use this model to predict covered 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.

We estimate the model on the subset of the study sample that was continuously covered by employer-sponsored health insurance for two years, and we include the following control variables in the models: race (black, other), Hispanic ethnicity, family income relative to poverty level (poor, near-poor, low income, middle income, high income), education (< high school, high school degree, bachelor degree, masters or doctorate, other degree, and education missing), marital status (married, formerly married, unmarried), employer size as a control for coverage generosity, region (northeast, midwest, south, west), urban indicator, and indicators of panel years. After exclusions for missing data for control variables, the number of observations in the estimation sample is 24,372.

We estimate a two-stage model of health expenditure. In the first stage, we estimate the probability of any insured expenditure using logistic regression. In the second stage, we estimate the level of insured spending among those for whom insured expenditures were greater than zero, using a generalized linear model assuming a log link for the expectation of expenditures and a Poisson distribution for the variance of expenditure conditional on the covariates (Manning and Mullahy 2001).[8]

[8] We chose this model following the tests outlined in Manning and Mullahy (2001). We chose GLM over OLS on either raw or log transformed expenditures based on the results of tests of bias in the predictions at particular points

The estimate of total expected expense – based on age, gender, and health status – is our primary measure of risk. For sensitivity analyses, we isolate expected expenditures based on the prevalence of health conditions from expected expenditures based on age and sex. We first define "demographic risk" by re-estimating the model without the indicators of chronic conditions, so that the resulting predicted expenditures are based on only on age and sex. We then define "condition risk" as the difference between total risk and "demographic risk". Condition risk measures the extent to which an individual's expected expenditures differ from people with similar demographic characteristics and is likely to be more costly for insurers to observe. We also use prior period self-reported health status as an indicator of information on health that is difficult or costly for insurers to observe.

*III.C. Defining Health Insurance Markets*

To test hypotheses regarding differences by the rating practices of insurers in the relationship between health risk and coverage, we identify people likely to purchase coverage in the large, medium, and small group markets based on family employment status. We define a family using the MEPS definition of the insurable unit - a sub-family relationship unit including 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 adult in the insurable unit and assign individuals to markets as follows:

1. Large Group Market: any adult member of the insurable unit (other than a dependent) employed full-time in a firm with 100 or more employees.

---

in the distribution. Among the GLM estimators, we chose the Poisson over the gamma distribution based on tests of precision, although, in practice, the differences between models estimated using the Poisson and gamma assumptions as well as models estimated in one and two stages were small. Predictions from the two-stage model using the Poisson assumption were less biased at the high end of the expenditure distribution than the predictions from the two-stage model assuming the gamma distribution (which were biased upward).

2. Medium Group Market: any adult member of the insurable unit employed full-time in a firm with 25 to 99 employees and no adult member of the insurable unit employed in a large firm.
3. Small Group Market: any adult member of the insurable unit employed full-time in a firm with 2 to 24 employees and no adult member of the insurable unit employed in either a medium or a large firm. This includes people who indicate that they are self-employed but work at an establishment with 2 or more workers.

We exclude people in insurable units with only part-time workers due to difficulties in assigning them to markets based on firm size. While it is inappropriate to group full- and part-time workers due to differences in their access to employer-sponsored coverage within a firm, the set of people in families with only part-time workers is not large enough to analyze independently (3.5% of the study sample).

*III.D. Study Sample*

We include in our study sample 37,820 individuals who participated in the panel component of the MEPS from 1996-2002 and who, in their second year in the survey, were 25-64 and not covered by public health insurance at any point during the year. We exclude 1,543 observations for missing data for private insurance coverage and 1,294 observations for missing data for variables used to construct the market indicators, primarily regarding employment status. We also exclude 106 individuals covered by someone outside the health insurance unit, since we cannot determine that policyholder's employment status. We exclude an additional 4,597 observations by limiting the analysis to people with a wage-earner in the insurable unit. The final study sample consists of 28,913 observations.

*III.E. Data Analysis*

Our empirical analysis is based on a series of models of the probability of purchasing private health insurance as a function of individual health risk, controlling for other characteristics that affect demand for health insurance. The dependent variable is an indicator of whether the respondent had private health insurance in December of his second year in the survey. In our main models, we measure health risk based on an individual's position in the distribution of expected health expenditures. We create categorical variables based on quintiles of the distribution, although we also identify the top and bottom $5^{th}$ percentiles to isolate very low and very high risks. Thus, total risk is defined by the following 7 categories of the distribution of expected expenditures: 0-$5^{th}$, >$5^{th}$-$20^{th}$, >$20^{th}$-$40^{th}$, >$40^{th}$-$60^{th}$, >$60^{th}$-$80^{th}$, >$80^{th}$-$95^{th}$, >$95^{th}$-$100^{th}$. This specification allows for non-linearities in the relationship between risk and coverage. As shown in Table 1, mean expected privately insured expenditures are $329, $1091, and $5,421 for the lowest, middle, and highest risk individuals, respectively.

We estimate the model separately by market (large group, medium group, and small group) and by income category. Income categories include low (<2 times poverty level), medium (2 to < 4 times poverty level), and high (>= 4 times poverty level). The control variables in the models 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.

Because the dependent variable is a binary indicator of insurance status, we estimate logistic regression models, applying the appropriate sample weights and adjusting for clustering at the level of the primary sampling unit. When presenting the results, we transform the

coefficient estimates to marginal effects to facilitate the interpretation of the magnitude of the estimates.

## IV. Results

### *IV.A. Main Results*

Table 2 compares health status between the insured and the uninsured. The uninsured report better health on many, but not all of the measures we examined. The total risk of the uninsured is lower than that of the insured. Five and three percent of the insured and uninsured, respectively, are in the highest risk category. And four and ten percent of the insured and uninsured, respectively, are in the lowest risk category. The measure of condition risk indicates that the lower total risk of the uninsured is driven not only by demographic characteristics, but also by the prevalence of conditions within demographic groups, mostly because high risks are more likely to be insured. The uninsured were less likely to report the existence of many, but not all, health conditions in the prior year. For some conditions, such as diabetes and mental conditions, the differences between the two groups are not statistically significant. In contrast to demographic characteristics and the presence of chronic conditions, the self-reported health status of the uninsured is worse than that of the insured. Twelve percent of the uninsured compared to seven percent of the insured report fair or poor health. Twenty-five percent of the uninsured compared to 31% of the insured report excellent health.

In Figure 1, we begin to examine differences across markets in the relationship between health risk and the probability that a person purchases private health insurance. The figure, which plots unadjusted rates of coverage by risk category for each market, demonstrates two important relationships. First, rates of coverage vary dramatically across markets. Coverage

rates are highest for people likely to have access to the large group market (90%) and lower for those relying primarily on the medium (80%) and small group (63%) markets (the average coverage rate by market is not in any table or figure). These differences in coverage rates by potential market are consistent across risk categories and likely reflect differences in the administrative loading across markets. Second, a positive relationship between health risk (as measured by total risk) and the purchase of private health insurance exists consistently across health insurance markets. In the large group market, rates of coverage increase from 85% for those in the bottom 5$^{th}$ percentile of the risk distribution to 94% for those in the top 5$^{th}$. In the small group market, rates of coverage range from 52% for low risks to 81% for high risks.

Table 3 presents our main results on the relationship between insurance coverage and risk by both market (large, medium and small group) and by family income. Each column in the table presents the coefficient of the indicators of risk category, transformed to the marginal effect, from models estimated separately for each market and income group. The omitted category is the middle of the risk distribution, expected expenditures in the 40$^{th}$ to 60$^{th}$ percentile of the distribution. The models also include the control variables discussed above, although the coefficients for the control variables are not shown in the table.

High risks are generally more likely than low risks to obtain group health insurance, although the magnitude of the effect varies both by income and by market. Among people in low income families, high risks are more likely than low risks to be covered in each market and the magnitude of the effect is similar across markets. The highest risks are 18, 20, and 14 percentage points more likely to have coverage than the lowest risks in the large group, medium group, and small group markets, respectively. Among people in medium income families, the difference between coverage rates for the highest and lowest risks increases from nine

percentage points in the large group market to 22 percentage points in the medium group market and to 24 percentage points in the small group market. In summary, for people in both low income and medium income families, the percentage with coverage rises consistently with risk across the market categories.

For adults in high income families, rates of coverage increase with risk for those in the small and medium group markets, but not in the large group market. For low income people purchasing coverage in the large group market, the lowest risks are 18 percentage points less likely to purchase coverage than the highest risks, and for medium income people, the lowest risks are 10 percentage points less likely to purchase coverage than the highest risks. For those with high incomes purchasing in the large group market, in contrast, the difference between lowest and highest risks is small (0.007) and not statistically significant.

Figures 2a through 2c summarize the results of the multivariate models in Table 3. In these figures, we present the estimate of the marginal effect of each risk category from these models, centered at zero. In other words, on the x-axis is the indicator of the risk category and on the y-axis is the effect of risk on coverage relative to the middle category ($40^{th}$-$60^{th}$ percentile of the risk distribution). We plot results for each market by income to identify within market differences by income in the risk gradient. Table 3 presents the standard errors and tests of statistical significance from these models.

The figures demonstrate that, in most cases, total risk is strongly positively associated with private insurance coverage in the group market. This is consistent with our first hypothesis, in which we propose that a lack of explicit risk rating of premiums in the employer-sponsored market leads to higher rates of coverage among high than low risks. In the large group market, the magnitude of the positive relationship between risk and coverage declines with income

(Figure 2a). This is consistent with our hypothesis that the risk gradient is steeper for low income than high income individuals in the group market due to the existence of either full or partial community rating in the large group market (Hypothesis 4).

In contrast, we find little evidence consistent with our hypothesis that the positive relationship between risk and coverage is stronger in the large than the small group market (Hypothesis 2) or our hypothesis that the positive (yet weaker) relationship between risk and coverage in the small group market may be weaker for low-income people than for high income people (Hypothesis 3). Both of these hypotheses were driven by an assumption of greater risk rating of premiums in the small group market (with the latter considering income constraints of high-risk, low-income people facing premiums with partial risk rating). The extent of risk rating of premiums may actually not differ much between the small group and large group markets. Moreover, our results point to the reverse of Hypothesis 2: a stronger positive relationship between risk and coverage in the small group than the large group market. This might be instead explained by differences between small and large firms in the extent of within-firm variation in the risk composition of workers. In particular, Bhattacharya and Vogt (2006) propose that both high worker switching costs and low exogenous turnover rates increase the likelihood that different risk types pool within a firm. It is possible that switching costs are greater for workers in large firms, leading low risks to be less likely to change firms in response to pooled premiums. Our results suggest that this type of effect dominates any opposing effect of differences across firms in the risk rating of premiums

In Table 4, we focus on the small group market and differentiate between coverage obtained through an employer and coverage purchased in the individual market in the dependent variable by estimating a multinomial logit model of coverage choice. As discussed earlier, we

limit this analysis to the small group market because rates of individual coverage among those we classify as having potential access to coverage through a medium or large employer are negligible. For adults in medium and high income families, we find that while high risks are more likely than low risks to have group coverage, high risks are less likely to have individual coverage. For these income groups, our results are consistent with low risks responding to pooled premiums in group coverage by substituting that group coverage with risk-rated coverage from the individual market. For adults in low income families, in contrast, we do not find evidence of this type of effect. These findings support those in Table 3 which indicate a positive relationship between risk and coverage in the employment-based system, although they indicate that the overall effect on coverage rates of pooled premiums in the group market is offset by access to coverage among low risks in the individual market for individual in medium and high income families.

*IV.B. Specification Tests*

Our main specification for risk in the results shown above is a set of seven indicator variables identifying the point in the distribution of total risk. In Tables 5 through 7, we present the results from a series of models that test the sensitivity of this main specification to alternative explanations. One possibility is that, since our measure of total risk includes variation in expected expenditures based on demographic characteristics, any effects of this measure may be driven primarily by differences across demographic groups that are correlated both with expected expenditures and preferences for insurance. For example, perhaps older people have both higher expected health expenditures and are more risk averse. In this case, the relationship between total risk and coverage may be driven by risk preferences that are correlated with age rather than

expected health expenditures. We examine this possibility by estimating models for insurance demand in which we replace the total risk variables with categorical indicators of demographic groups and the "condition risk" measure which, as described above, is the difference between total risk and expected expenditures conditioning only on demographic characteristics. We specify this measure as an indicator of whether the individual was in the bottom 10% (mean difference=-$1,296)) or the top 10% (mean difference=$2,328) of the distribution.[9] By testing whether demographic characteristics and health status have independent relationships with coverage, we provide evidence on the extent of pooling across different types of characteristics.

In these models, we also include categorical indicators of prior year self-reported health status (excellent, good, and fair or poor relatively to very good). These two measures, condition risk and self-reported health status, allow us to test whether the relationship between coverage and health status varies based on the extent to which indicators of health status are likely to be observable to insurers. We hypothesize that condition risk and self-reported health status, are more difficult and more costly for insurers to observe than are demographic characteristics. On one hand, self-reported health status may be relatively more unobservable than condition risk because it reflects asymmetric information held by the person. On the other hand, Fair/poor self-reported health status may be an indicator of the severity of condition risk, and thus be relatively more observable to the insurer (if the insurer focuses on identifying the most costly conditions).

We find that conditions risk has a positive and statistically significant effect on the probability of obtaining coverage in each market (Table 5 – Columns 1-3). This indicates that the relationship between coverage and total risk is driven at least in part by the existence of

[9] We use a three-category rather than a seven category measure simply to make the model more parsimonious, particularly since we also added the 15 indicators of demographic groups.

particular conditions, although the models indicate the demographic characteristics also play a role.

In contrast, we find that those who reported either good or fair/poor health in the prior year are less likely to have health insurance and that the negative effect is consistent across markets (Table 5 – Columns 1 - 3). While our theory predicts that self-reported health status, like other measures of poor health, would be associated with higher rates of coverage when premium are pooled, our results are consistent with findings from other studies indicating that a self-report of fair or poor health status often has a negative relationship with coverage. While our method differs from existing studies because we use prior year, rather than concurrent, self-reported health status, the negative effect of self-reported poor health contrasts the positive effect of poor health due to the existence of particular health conditions.

Since an individual's health insurance is often purchased as part of a family policy, we also test whether our findings with respect to individual health risk are influenced by the health status of family members by re-estimating the models separately for singles and families and including a measure of expected expenditures of family members (the sum of total risk of family members) in the models of individuals who are part of families (Table 6).

In column 1 of Table 6, we present the results of the model pooling the data over all markets for the purpose of comparison. Then in columns 2 and 3, we estimate the model separately by family size (based on the MEPS health insurance unit). We find that the effect of an individual's total risk is positive both for singles and for people with additional family members. While the expected health expenditures of family members is positively correlated with the probability that an individual has coverage, including this variable has little effect on the estimates of the effect of an individual's total risk (Table 6 – Columns 3 and 4).

Finally, other research has demonstrated that individual health insurance decisions are related to preferences for health insurance and medical care so we test whether our estimates of the relationship between risk and coverage are likely to be biased by these types of relationships (Monheit and Vistnes 2006). We measure preferences using questions regarding whether an individual uses a seatbelt regularly, whether the individual believes he is more likely to take risks than the average person, and whether the individual believes he can overcome illness without help from a medically trained person.[10] The first two questions are proxies for risk aversion and the last is a proxy for preferences for medical care. We code these as 0/1 binary variables where one indicates always or nearly always for the seat belt question and agree strongly or agree somewhat for the risk taking and medical preference questions. The availability of these variables, however, is limited to the three most recent years of the survey data we use, dramatically reducing the sample size. Thus, we pool the data across markets and income for this analysis.

The results for the variables measuring preferences are similar (Table 7). While their effects are in the expected directions (people who regularly wear seat belts are more likely to be insured and those who identify themselves as risk takers and as able to overcome illness without help from a medically trained person are less likely to be insured), including these variables in the models has virtually no effect on the magnitude of the estimates of the effects of total risk on coverage.

## V. Conclusions

[10] Two additional questions, which have been used in other analyses, ask whether the person believes health insurance is worth the money it costs and whether a person thinks they do not need health insurance. We do not examine these measures because they are likely to reflect a person's health risk, as well as their preferences.

We find that, in aggregate, the likelihood of obtaining employer-sponsored coverage nearly always increases with expected health expenditures. The positive relationship between insurance status and expected expenditures is generally consistent across the large group, medium group, small and group markets. In addition, both total health risk, considering age, sex and health conditions simultaneously, and health risk due to the presence of chronic conditions, controlling for demographic characteristics, are positively associated with insurance coverage.

The existence of a positive relationship between risk and coverage is consistent with a moderate amount of adverse selection. Our measure of risk includes information about health status that is relatively easily observable to consumers. The positive relationship between this measure of risk and coverage suggests that the premiums consumers face for coverage, in the form of employee contributions and potential wage offsets, incorporate less information about individual risk than consumers use when deciding whether to obtain coverage. Although our results are consistent with insurers using less information in setting premiums than individuals use in purchasing decisions, we only provide indirect evidence that this is the case. In our data, we are unable to observe the premium paid by individuals with coverage and the premium at which coverage would have been available for those who did not purchase it.

At least two potential alternative explanations for our findings exist. First, unobserved characteristics of individuals may be positively correlated with both expected health expenditures and preferences for insurance (Finkelstein and McGarry 2006; Cutler, Finkelstein et al. Forthcoming). 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. In this case, even if premiums were perfectly risk rated, the relationship may still exist – high risks just have greater demand for health insurance. We cannot rule this case out in our empirical

work, although our approach to measuring health risk indicates that these types of preferences would need to be correlated with health status independent of age. Second, while our measure of risk is likely robust to the effect of moral hazard on the utilization of services, it may be influenced by the effect of insurance on the reporting of conditions. In other words, by providing access to medical care, health insurance may cause people to report conditions that would have gone unnoticed in the absence of coverage. If this is true, our findings may be driven by people who think they are low risk rather than people who are actually low risk from the perspective of the insurer.

The positive relationship between health risk and coverage is stronger for low than for high income individuals in the large group market. This is consistent with our hypothesis that community-rated or partially community-rated premiums would have larger negative effects on rates of coverage for low income low risks than for high income low risks. This effect is relatively large. Among people with low and medium incomes purchasing coverage in the large group market, the lowest risks are 18 and 9 percentage points, respectively, less likely than the highest risks to be covered by health insurance. 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.

However, we find no evidence that the relationship between risk and coverage is stronger in the large group than the small group market. While our hypothesis was that a positive relationship between coverage and risk would be stronger in the large group market due to the greater prevalence of risk rating and medical underwriting in the small group market, our empirical results indicate that the relationship is about the same in each market or even stronger in the small group market. This finding is consistent with relatively little difference in the extent

of risk rating of premiums between markets. The results also imply that some other factor, such as higher switching costs, may make pooling arrangements somewhat more stable in large than small firms.

In conclusion, contrary to popular perception, our results provide little evidence that high premiums for high risks are a significant contributor to the large uninsured population in the U.S. Even in the small group market, where individual risk rating or coverage denials are more likely to take place, health risk is positively associated with coverage. Rather, our findings indicate that financial deterrents to coverage associated with risk may be more important among low risks, particularly in the large group market. In the individual market, in contrast, the primary barrier to the affordability of coverage is likely the high loading, which affects people at all risk levels.

We believe that our results have important implications for evaluating the extent to which risk pooling is a benefit of employment-based coverage. Among wage-earning families, we find little evidence that high premiums for high risks are a significant contributor to the large uninsured population in the U.S. Even in the small group market, where individual risk rating or coverage denials are thought to be more likely to take place, health risk is positively associated with coverage. But there is a cost associated with pooled premiums. In particular, we find that pooled premiums in the group market represent a financial deterrent leading to lower rates of coverage among low risk workers and their families, particularly those with low incomes.

**References:**


Bernard, D. and T. M. Selden (2006). "Workers who Decline Employment-Related Health Insurance." Medical Care 44(5): I12-I18.

Bhattacharya, J. and M. K. Bundorf (2005). The Incidence of the Healthcare Costs of Obesity. Cambridge, MA, National Bureau of Economic Research.

Bhattacharya, J. and W. Vogt (2006). Employment and Adverse Selecton in Health Insurance. NBER Working Paper #12430.

Blumberg, L. J. and L. M. Nichols (2001). "The Health Status of Workers who Decline Employer-Sponsored Health Insurance." Health Affairs 20(6): 180-187.

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): 300-312.

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): 408-427.

Coate, S. (1995). "Altruism, the Samaritan's Dilemma and Government Transfer Policy." American Economic Review 85(1): 46-57.

Cutler, D. M., A. Finkelstein, et al. (Forthcoming). "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance." American Economic Review, Papers and Proceedings.

Cutler, D. M. and S. Reber (1998). "Paying for health insurance: The tradeoff between competition and adverse selection." Quarterly Journal of Economics 113(2): 433-466.

DeNavas-Walt, C., B. D. Proctor, et al. (2008). Income, Poverty, and Health Insurance Coverage in the United States; 2007. Current Population Reports: Consumer Income. Washington, DC, U.S. Census Bureau: 1-71.

Ehrlich, I. and G. S. Becker (1972). "Market Insurance, Self-Insurance, and Self-Protection." Journal of Political Economy 80(July/August): 623-648.

Finkelstein, A. and K. McGarry (2006). "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market." American Economic Review 96(4): 938-958.

Furman, J. (2008). Reforming the Tax Treatment of Health Care: Right Ways and Wrong Ways. Raxes and Health Insurance: Analysis and Policy, Tax Policy Center and the American Tax Policy Institute.

Glied, S. (2000). Managed Care. Handbook of Health Economics. A. J. Culyer and J. P. Newhouse. New York, Elsevier Science B. V. 1A: 708-53.

Goldstein, G. S. and M. V. Pauly (1976). Group Health Insurance as a Local Public Good. The Role of Health Insurance in the Health Services Sector. R. N. Rosett. New York, National Bureau of Economic Research: 73-110.

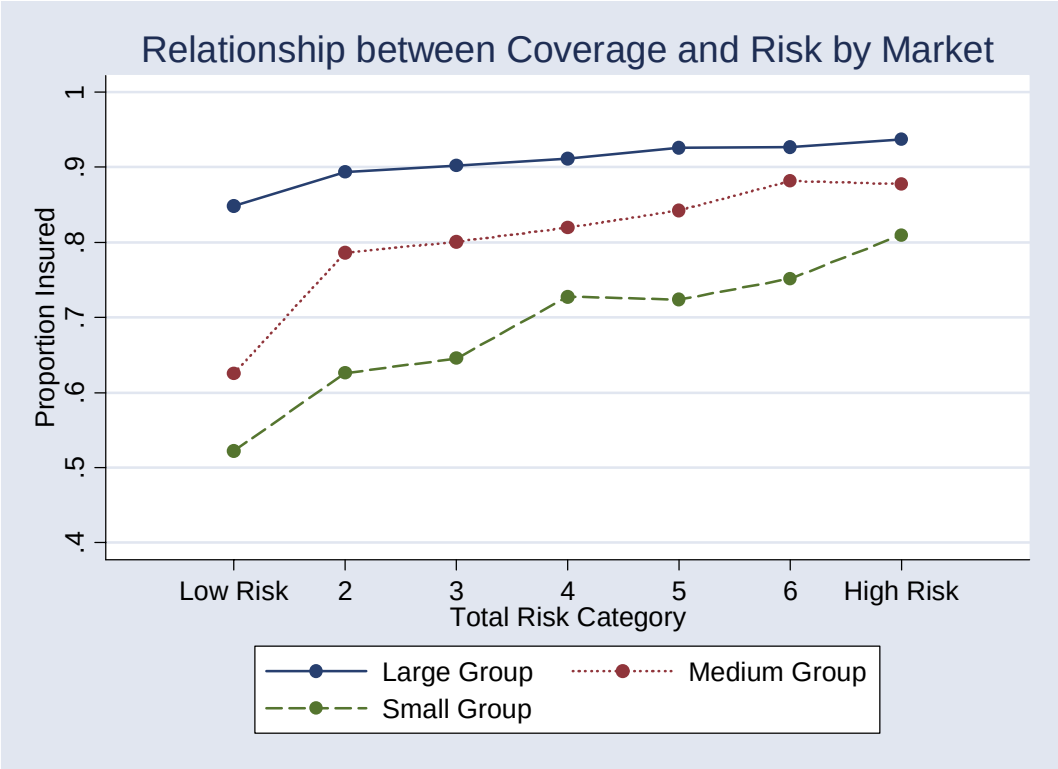
Gruber, J. (1994). "The Incidence of Mandated Maternity Benefits." American Economic Review 84(3): 622-641.

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

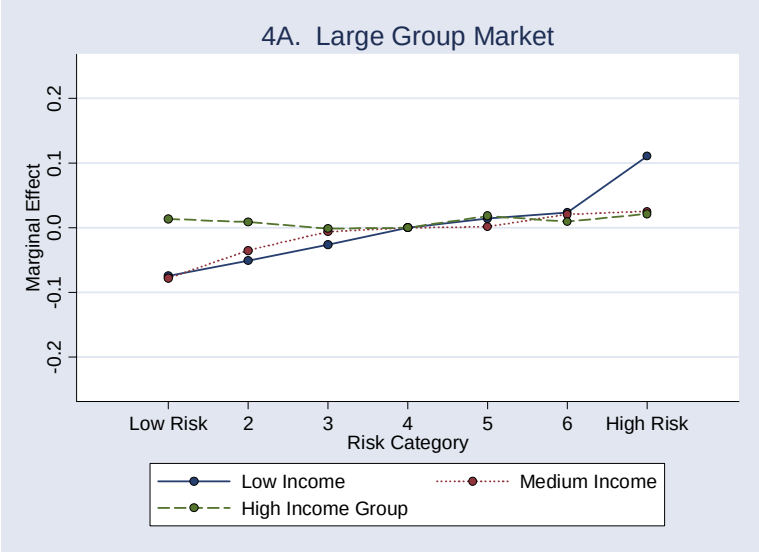
Herring, B. (2005). "The effect of availability of charity care to the uninsured on the demand for private health insurance." Journal of Health Economics 24(2): 225-252.

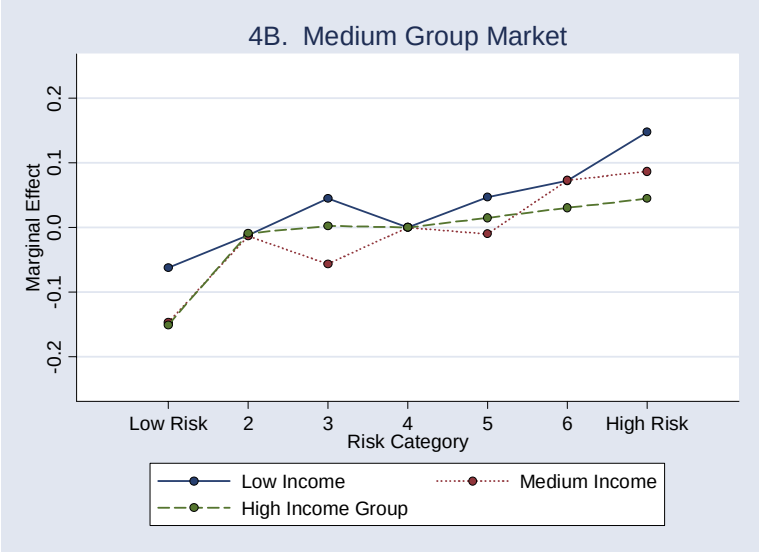
Keenan, P. S., M. J. B. Buntin, et al. (2001). "The Prevalence of Formal Risk Adjustment in Health Plan Purchasing." Inquiry 38(Fall): 245-259.
Manning, W. G. and J. Mullahy (2001). "Estimating log models: to transform or not to transform?" Journal of Health Economics 20: 461-494.
McLaughlin, C. G., S. E. Crow, et al. (2004). Causes and Consequences of Lack of Insurance: Gaps in our Knowledge. Health Policy and the Uninsured. C. G. McLaughlin. Washington D.C., The Urban Institute Press: xiii-xxv.
Monheit, A. C. and J. P. Vistnes (1994). "Implicit Pooling of Workers from large and Small Firms." Health Affairs Spring(I): 301-314.
Monheit, A. C. and J. P. Vistnes (2006). Health Insurance Enrollment Decisions: Preferences for Coverage, Worker Sorting, and Insurance Take Up. NBER Working Paper.
Nyman, J. A. (1999). "The value of health insurance: the access motive." Journal of Health Economics 18: 141-152.
Pauly, M. V. (1970). "The Welfare Economics of Community Rating." The Journal of Risk and Insurance 37(3): 407-418.
Pauly, M. V. and B. Herring (1999). Pooling Health Insurance Risks. Washington DC, AEI Press.
Pauly, M. V. and R. D. Lieberthal (2008). "How risky is individual health insurance?" Health Affairs 27(3): w242-w249.
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.
Schoen, C., K. Davis, et al. (2008). "Building Blocks for Reform: Achieving Universal Coverage with Private and PUblic Group Health Insurance." Health Affairs 27(3): 646-657.

Figure 1: Relationship between Insurance Coverage and Health Risk by Market



Figures 2A-2C: Within Market Differences by Income in the Effect of Risk on Coverage



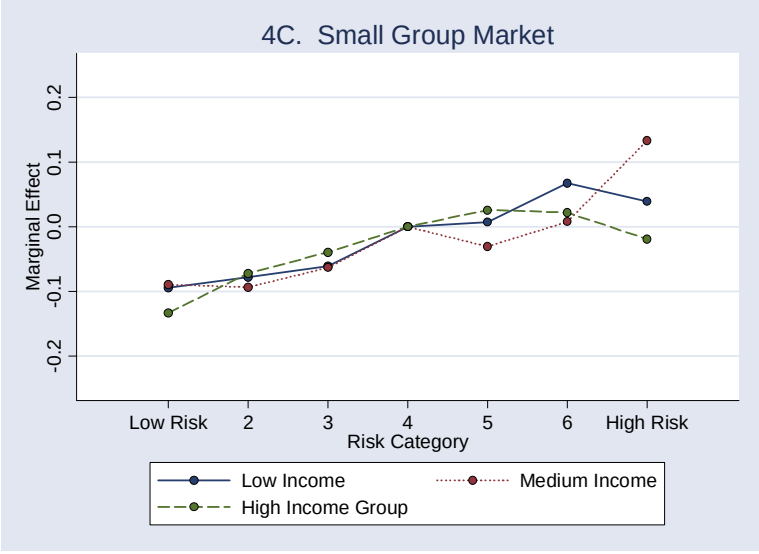



Table 1: Total Risk Measure of Expected Expenditures

| Risk Category | Percentile of the Distribution of Expected Insured Expenditures | Mean Expected Insured Expenditures |
|---|---|---|
| 1 | 0-5th | 329 |
| 2 | >5$^{th}$-20$^{th}$ | 510 |
| 3 | >20$^{th}$-40$^{th}$ | 778 |
| 4 | >40$^{th}$-60$^{th}$ | 1,091 |
| 5 | >60$^{th}$-80$^{th}$ | 1,559 |
| 6 | >80$^{th}$-95$^{th}$ | 2,493 |
| 7 | >95$^{th}$-100$^{th}$ | 5,421 |

Table 2: Comparison of Health Status between Insured and Uninsured

| N=28,913 | Year 2 Insurance Status | | |
|---|---|---|---|
| Measures of Health Status based on Year 1 Conditions | Insured | Uninsured | |
| Total Risk Category 1 (lowest) | 0.04 | 0.10 | ** |
| Total Risk Category 2 | 0.15 | 0.20 | |
| Total Risk Category 3 | 0.18 | 0.21 | |
| Total Risk Category 4 | 0.20 | 0.19 | |
| Total Risk Category 5 | 0.21 | 0.17 | |
| Total Risk Category 6 | 0.16 | 0.11 | |
| Total Risk Category 7 (highest) | 0.05 | 0.03 | |
| Condition Risk Category 1 (lowest) | 0.09 | 0.09 | ** |
| Condition Risk Category 2 | 0.81 | 0.84 | |
| Condition Risk Category 3 (highest) | 0.10 | 0.07 | |
| SRHS: Excellent | 0.31 | 0.25 | ** |
| SRHS: Very Good | 0.37 | 0.32 | |
| SRHS: Good | 0.25 | 0.30 | |
| SRHS: Fair or Poor | 0.07 | 0.12 | |
| Male: 25-29 | 0.06 | 0.13 | ** |
| Male: 30-34 | 0.07 | 0.10 | |
| Male: 35-39 | 0.08 | 0.10 | |
| Male: 40-44 | 0.08 | 0.08 | |
| Male: 45-49 | 0.07 | 0.06 | |
| Male: 50-54 | 0.06 | 0.04 | |
| Male: 55-59 | 0.04 | 0.03 | |
| Male: 60-64 | 0.02 | 0.01 | |
| Female: 25-29 | 0.07 | 0.09 | |
| Female: 30-34 | 0.08 | 0.08 | |
| Female: 35-39 | 0.09 | 0.08 | |
| Female: 40-44 | 0.08 | 0.08 | |
| Female: 45-49 | 0.08 | 0.05 | |
| Female: 50-54 | 0.06 | 0.04 | |
| Female: 55-59 | 0.04 | 0.02 | |
| Female: 60-64 | 0.02 | 0.01 | |
| Infectious and Parasitic | 0.10 | 0.07 | ** |
| Malignant Neoplasm | 0.01 | 0.01 | ** |
| Benign/In Situ/Uncertain Neoplasm | 0.05 | 0.03 | ** |
| Diabetes | 0.03 | 0.03 | |
| Nutritional and Metabolic | 0.11 | 0.05 | ** |
| Liver | 0.01 | 0.01 | |

Table 2: Comparison of Health Status between Insured and Uninsured (Continued)

| | Year 2 Insurance Status | | |
|---|---|---|---|
| Measures of Health Status based on Year 1 Conditions | Insured | Uninsured | |
| Gastrointestinal | 0.23 | 0.20 | ** |
| Musculoskeletal and Connective Tissues | 0.23 | 0.19 | ** |
| Hematological | 0.01 | 0.01 | ** |
| Cognitive Disorders | 0.00 | 0.00 | |
| Substance Abuse | 0.00 | 0.00 | |
| Mental | 0.11 | 0.11 | |
| Developmental Disability | 0.00 | 0.00 | |
| Neurological | 0.06 | 0.04 | ** |
| Heart | 0.12 | 0.08 | ** |
| Cerebro-Vascular | 0.00 | 0.00 | |
| Vascular | 0.01 | 0.01 | |
| Lung | 0.15 | 0.13 | ** |
| Eyes | 0.07 | 0.05 | ** |
| Ears, Nose and Throat | 0.40 | 0.30 | ** |
| Urinary System | 0.04 | 0.03 | ** |
| Genital System | 0.08 | 0.05 | ** |
| Skin and Subcutaneous | 0.10 | 0.06 | ** |
| Symptoms, Signs and Ill-Defined Conditions | 0.18 | 0.15 | ** |
| Proportion of Sample | 0.85 | 0.15 | |

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

Note: Estimates are weighted. Statistical significance is calculated using either a t-test or a chi-squared test depending on the variable.

Table 3: The Relationship between Insurance Coverage and Health Risk

| | Low Income | | | Medium Income | | | High Income | | |
|---|---|---|---|---|---|---|---|---|---|
| | Large Group | Medium Group | Small Group | Large Group | Medium Group | Small Group | Large Group | Medium Group | Small Group |
| Risk Category 1 (low) | -0.074 | -0.061 | -0.102 | -0.071 | -0.131 | -0.127 | 0.013 | -0.122 | -0.102 |
| | [0.037]* | [0.071] | [0.038]** | [0.023]** | [0.047]** | [0.047]** | [0.008]+ | [0.043]** | [0.037]** |
| Risk Category 2 | -0.05 | -0.011 | -0.082 | -0.036 | -0.012 | -0.079 | 0.009 | -0.009 | -0.086 |
| | [0.024]* | [0.050] | [0.032]* | [0.015]* | [0.045] | [0.034]* | [0.004]* | [0.020] | [0.040]* |
| Risk Category 3 | -0.023 | 0.045 | -0.068 | -0.006 | -0.058 | -0.075 | -0.002 | 0.003 | -0.056 |
| | [0.025] | [0.042] | [0.031]* | [0.018] | [0.032]+ | [0.048] | [0.007] | [0.018] | [0.035] |
| Risk Category 5 | 0.017 | 0.045 | -0.004 | 0.002 | -0.01 | -0.029 | 0.019 | 0.015 | 0.006 |
| | [0.039] | [0.074] | [0.042] | [0.014] | [0.022] | [0.041] | [0.005]** | [0.019] | [0.040] |
| Risk Category 6 | 0.021 | 0.073 | 0.075 | 0.02 | 0.07 | 0.001 | 0.009 | 0.03 | 0.001 |
| | [0.021] | [0.093] | [0.030]* | [0.019] | [0.022]** | [0.039] | [0.007] | [0.011]** | [0.027] |
| Risk Category 7 (high) | 0.107 | 0.138 | 0.039 | 0.022 | 0.089 | 0.119 | 0.021 | 0.041 | 0.001 |
| | [0.032]** | [0.078]+ | [0.072] | [0.018] | [0.044]* | [0.037]** | [0.007]** | [0.014]** | [0.049] |
| Risk 7 less Risk 1 | 0.181 | 0.199 | 0.141 | 0.092 | 0.219 | 0.246 | 0.007 | 0.163 | 0.103 |
| | [0.046]** | [0.120]+ | [0.073]+ | [0.022]** | [0.042]** | [0.045]** | [0.008] | [0.021]** | [0.050]* |
| Observations | 2,761 | 1,192 | 1,802 | 6,433 | 1,875 | 1,978 | 9,303 | 1,905 | 1,628 |

Standard errors in brackets

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

Note: Table presents the marginal effect of each risk category relative to the middle category (Category 4) from a logistic regression of health insurance coverage as a function of risk and control variables. The marginal effects are the sample average of the change in probability given a change from risk category 4 to the indicated risk category. The control variables include race, ethnicity, education, marital status, income and income squared of health insurance unit, number of people in health insurance unit, year indicators, urban indicator, and region indicators. Estimates are weighted and the standard errors allow for clustering by primary sampling unit.

Table 4: The Relationship between Risk and Coverage in the Small Group Market (Differentiating between Group and Individual Coverage

| | Low Income | | Medium Income | | High Income | |
|---|---|---|---|---|---|---|
| | Group | Individual | Group | Individual | Group | Individual |
| Risk Category 1 (low) | -0.075 | -0.02 | -0.104 | -0.018 | -0.161 | 0.062 |
| | [0.044]+ | [0.011]+ | [0.048]* | [0.018] | [0.051]** | [0.047] |
| Risk Category 2 | -0.052 | -0.021 | -0.043 | -0.029 | -0.098 | 0.011 |
| | [0.029]+ | [0.010]* | [0.033] | [0.015]* | [0.052]+ | [0.021] |
| Risk Category 3 | -0.032 | -0.028 | -0.052 | -0.017 | -0.085 | 0.027 |
| | [0.028] | [0.008]** | [0.049] | [0.011] | [0.045]+ | [0.019] |
| Risk Category 5 | 0.013 | -0.019 | 0.015 | -0.037 | 0.001 | 0.005 |
| | [0.041] | [0.013] | [0.040] | [0.008]** | [0.041] | [0.013] |
| Risk Category 6 | 0.061 | 0.023 | 0.015 | -0.007 | 0.016 | -0.014 |
| | [0.042] | [0.020] | [0.038] | [0.025] | [0.032] | [0.012] |
| Risk Category 7 (high) | 0.027 | 0.026 | 0.18 | -0.056 | 0.033 | -0.031 |
| | [0.076] | [0.044] | [0.042]** | [0.014]** | [0.053] | [0.014]* |
| Risk 7 minus Risk 1 | 0.103 | 0.046 | 0.284 | -0.038 | 0.194 | -0.093 |
| | [0.092] | [0.059] | [0.050]** | [0.012]** | [0.057]** | [0.015]** |
| Observations | 1,794 | 1,794 | 1,967 | 1,967 | 1,615 | 1,615 |

Standard errors in brackets

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

Note: Table presents the marginal effect of each risk category relative to the middle category (Category 4) from a multinomial logistic regression of health insurance coverage (group, individual or none) as a function of risk and control variables. The marginal effects are the sample average of the change in probability given a change from risk category 4 to the indicated risk category. The control variables include race, ethnicity, education, marital status, income and income squared of health insurance unit, number of people in health insurance unit, year indicators, urban indicator, and region indicators. Estimates are weighted and the standard errors allow for clustering by primary sampling unit.

Table 5: The Relationship between Coverage and Alternative Measures of Health Risk

| | (1) Large Group | (2) Medium Group | (3) Small Group |
|---|---|---|---|
| Condition Risk - Low | -0.032 | -0.033 | -0.065 |
| | [0.011]** | [0.020] | [0.023]** |
| Condition Risk - High | 0.019 | 0.053 | 0.049 |
| | [0.006]** | [0.015]** | [0.025]+ |
| Prior Year SRHS: Excellent | 0.007 | -0.02 | 0.008 |
| | [0.007] | [0.013] | [0.018] |
| Prior Year SRHS: Good | -0.011 | -0.029 | -0.006 |
| | [0.005]* | [0.011]* | [0.013] |
| Prior Year SRHS: Fair or Poor | -0.035 | -0.053 | -0.09 |
| | [0.009]** | [0.021]* | [0.022]** |
| Male: 25-29 | -0.003 | -0.074 | -0.129 |
| | [0.007] | [0.025]** | [0.029]** |
| Male: 30-34 | -0.004 | -0.024 | -0.081 |
| | [0.008] | [0.020] | [0.031]** |
| Male: 35-39 | 0.003 | -0.036 | -0.071 |
| | [0.009] | [0.021]+ | [0.031]* |
| Male: 45-49 | 0.01 | 0.065 | -0.097 |
| | [0.007] | [0.016]** | [0.032]** |
| Male: 50-54 | 0.033 | 0.034 | -0.016 |
| | [0.010]** | [0.029] | [0.036] |
| Male: 55-59 | 0.028 | 0.072 | 0.047 |
| | [0.011]* | [0.021]** | [0.045] |
| Male: 60-64 | 0.034 | 0.082 | 0.011 |
| | [0.016]* | [0.030]** | [0.043] |
| Female: 25-29 | -0.004 | -0.015 | -0.011 |
| | [0.009] | [0.018] | [0.030] |
| Female: 30-34 | 0.006 | -0.007 | -0.054 |
| | [0.008] | [0.031] | [0.030]+ |
| Female: 35-39 | 0.014 | 0.02 | -0.03 |
| | [0.008]+ | [0.022] | [0.026] |
| Female: 40-44 | 0.015 | 0.023 | -0.023 |
| | [0.010] | [0.020] | [0.025] |
| Female: 45-49 | 0.041 | 0.034 | 0.025 |
| | [0.006]** | [0.020]+ | [0.030] |
| Female: 50-54 | 0.038 | 0.054 | 0.08 |
| | [0.006]** | [0.019]** | [0.032]* |

Table 5: The Relationship between Coverage and Alternative Measures of Health Risk (Continued)

| | | | |
|---|---|---|---|
| Female: 55-59 | 0.043 | 0.085 | 0.096 |
| | [0.006]** | [0.026]** | [0.038]* |
| Female: 60-64 | 0.062 | 0.093 | 0.157 |
| | [0.009]** | [0.022]** | [0.029]** |
| Observations | 18,528 | 4,977 | 5,408 |

Standard errors in brackets

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

[1] Condition risk is the ratio of the prediction based on age, sex and health conditions to the prediction based only on age and sex.

Note: Table presents the marginal effect of each risk category relative to the middle category (Category 4) from a logistic regression of health insurance coverage as a function of risk and control variables. The marginal effects are the sample average of the change in probability given a change from risk category 4 to the indicated risk category. The control variables include race, ethnicity, education, marital status, income and income squared of health insurance unit, number of people in health insurance unit, year indicators, urban indicator, and region indicators. Estimates are weighted the standard errors allow for clustering by primary sampling unit.

Table 6: The Effects of Family Member Health and Individual Preferences on the Relationship between Risk and Coverage

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | | | Multiple Person Households | |
| | Pooled | Single Person Households | No control for family member health | Control for family member health |
| Risk Category 1 (low) | -0.053 | -0.09 | -0.032 | -0.032 |
| | [0.009]** | [0.020]** | [0.013]* | [0.013]* |
| Risk Category 2 | -0.023 | -0.051 | -0.011 | -0.013 |
| | [0.007]** | [0.013]** | [0.007] | [0.007]+ |
| Risk Category 3 | -0.016 | -0.031 | -0.01 | -0.01 |
| | [0.006]** | [0.014]* | [0.006]+ | [0.006]+ |
| Risk Category 5 | 0.01 | 0.004 | 0.012 | 0.012 |
| | [0.004]* | [0.012] | [0.005]* | [0.005]* |
| Risk Category 6 | 0.021 | 0.044 | 0.014 | 0.01 |
| | [0.005]** | [0.011]** | [0.005]** | [0.005]* |
| Risk Category 7 (high) | 0.042 | 0.036 | 0.043 | 0.038 |
| | [0.009]** | [0.023] | [0.010]** | [0.011]** |
| ln(Incremental Family Expected Expenditures) | | | | 0.02 |
| | | | | [0.004]** |
| Risk 7 minus Risk 1 | 0.095 | 0.126 | 0.075 | 0.070 |
| | [0.010]** | [0.024]** | [0.012]** | [0.012]** |
| Observations | 28,913 | 7,065 | 21,848 | 21,848 |

Standard errors in brackets

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

Note: Table presents the marginal effect of each risk category relative to the middle category (Category 4) from a logistic regression of health insurance coverage as a function of risk and control variables. The marginal effects are the sample average of the change in probability given a change from risk category 4 to the indicated risk category. The control variables include race, ethnicity, education, marital status, income and income squared of health insurance unit, number of people in health insurance unit, year indicators, urban indicator, and region indicators. Estimates are weighted using the MEPS panel weights and estimation of the standard errors allows for clustering by primary sampling unit.

Table 7: The Effects of Individual Preferences on the Relationship between Risk and Coverage

| Subsample with preference measures | (1) | (2) |
|---|---|---|
| | No preference controls | Preference controls |
| Risk Category 1 (low) | -0.033 | -0.021 |
| | [0.013]* | [0.013]+ |
| Risk Category 2 | -0.025 | -0.016 |
| | [0.010]** | [0.009]+ |
| Risk Category 3 | -0.017 | -0.014 |
| | [0.010] | [0.010] |
| Risk Category 5 | 0.018 | 0.017 |
| | [0.008]* | [0.008]* |
| Risk Category 6 | 0.025 | 0.022 |
| | [0.010]** | [0.010]* |
| Risk Category 7 (high) | 0.046 | 0.041 |
| | [0.016]** | [0.017]* |
| Regularly wears seatbelt | | 0.032 |
| | | [0.008]** |
| More likely than average person to take risks | | -0.042 |
| | | [0.009]** |
| Can overcome illness without medical help | | -0.023 |
| | | [0.008]** |
| Risk 7 minus Risk 1 | 0.080 | 0.062 |
| | [0.017]** | [0.018]** |
| Observations | 13,697 | 13,697 |

Standard errors in brackets

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

Note: Table presents the marginal effect of each risk category relative to the middle category (Category 4) from a logistic regression of health insurance coverage as a function of risk and control variables. The marginal effects are the sample average of the change in probability given a change from risk category 4 to the indicated risk category. The control variables include race, ethnicity, education, marital status, income and income squared of health insurance unit, number of people in health insurance unit, year indicators, urban indicator, and region indicators. Estimates are weighted the standard errors allow for clustering by primary sampling unit.