**Economic Research Initiative on the Uninsured**
**CONFERENCE DRAFT**

# THE UNINSURED: RISK, INCOME, AND "AFFORDABILITY" OF COVERAGE

Draft: Please do not cite or quote without permission.

M. Kate Bundorf

Mark V. Pauly

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

July, 2004

Abstract: The nearly universal belief among policy makers is that risk rating of health insurance premiums makes health insurance unaffordable to high risks due to the high premiums for coverage they face in competitive health insurance markets. If this is true, the affordability of health insurance among high risks may be an important contributor to the large number of uninsured in the U.S. In this paper, we propose that higher premiums do not necessarily make high risks less likely to purchase health insurance and that the relationship between health risk and coverage may vary by income as well as the extent to which premiums are risk rated in a market. We test these hypotheses by examining empirically the relationship between health risk and the purchase of private health insurance. We find that the purchase of private health insurance consistently increases with health risk across both income and health insurance market. Our results provide little evidence that the affordability of coverage for high risks is a significant contributor to the large number of uninsured in the U.S.

## I. Introduction

Private health insurance, whether obtained through work or purchased individually, is costly. This high cost is generally thought to be a large part of the reason why 16 percent of the US population does not obtain any insurance coverage: relative to their income and relative to their estimate of the value of having insurance, the premium looks high. However, not only are premiums high on average, they are thought to vary substantially across individuals at differing risks for medical care use for the same nominal coverage. Especially in the individual insurance market (but perhaps also in some group markets), higher risk people are thought to face premiums that are substantially higher than average, and this is thought to cause a serious problem of uninsurability (Pollitz and Sorian 2002).

As will be discussed further below, this premium variation is not the same economic phenomenon as variation in price for a textbook homogenous economic good. Nevertheless, the variation is an object of serious conceptual and policy interest. Specifically, there is concern that higher premiums for higher risks may in part account for both individual decisions to remain uninsured and the proportion uninsured among a population of varying risks. The impact of premium variation is also thought to be stronger at lower income levels than at higher income levels, other things equal: higher premiums for lower income higher risks may discourage coverage not only because they are "high" but also because they are "unaffordable."

Simple economic models of rational insurance purchasing and insurance market equilibrium stand in stark contrast to this virtually universal backdrop of policy concerns. There is considerable confusion, even in the health services research literature, about what we should

expect. 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) say 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." (Ehrlich and Becker 1972)

However, at low income levels, the assumption of minor income effects may not apply. For a person with low money income, even in economic theory, the effect of high risk on expected real income may be substantial, and this effect 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. Still if the demand for medical care is normal and if the response to user price is greater at higher levels of cost sharing than at lower levels, then moral hazard will be greater at lower (real) income levels and so the demand for generous coverage will be lower. 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 obvious point is that the effect of higher risk on the demand for insurance, and any interaction of that effect with money income, is ambiguous in theory. Our conclusion therefore is that the best way to determine how risk variation affects the demand for health insurance is empirical. Are higher risk people in fact less likely to demand insurance of a given type with a given level of loading than lower risk people? And does any reaction attributable to risk change as income changes?

The desirability of an empirical approach is strengthened when we recognize that, in reality, the extent to which premiums actually do vary with risk is both unknown and hard to measure. One problem is that risk is hard to measure. Another problem is that it is hard to know what the premium a person might have paid actually is—especially for those who chose not to buy. An even more serious problem is that in the group insurance market the way in which people pay for insurance and the perceptions they have of how and what they are paying is murky. Finally, in individual markets the premium often pays not only for current-period expected expenses but also (under the provisions of guaranteed renewability now present in virtually all policies) for assurances about what future premiums will be if the purchaser's risk changes (Herring and Pauly 2003). These reasons also suggest the wisdom of using a reduced form approach rather than trying (at least initially) the daunting task of including measures of risk, premiums, and policy provisions in the analysis.

## II. Developing Hypotheses.

In what follows we consider two possible ways in which premiums might vary (or be perceived to vary) with the individual's risk. On the one hand, premiums might be *risk-rated.* We shall interpret this to mean that, across risks, premiums are proportional to risk. The proportion can be equal to one (actuarially fair), greater than one (actuarially unfair) or less than one (proportionally subsidized). On the other hand, premiums may not increase proportionately with risk. If they do not increase at all with risk, we have the case of *pure community rating.* If they increase with risk but less than proportionately, we can all this partial community rating.

The extent to which premiums (hypothetically) vary with risk potentially depends on whether the set of potential purchasers obtains insurance in one of three types of markets: individual, small group, and large group. We initially assume that individuals are exogenously assigned to one of the three markets (based on their type of employment, whether at a large firm, a small firm, or a single-person or non-employment setting).

Under conventional models of insurance purchasing, higher risk people will be more likely to buy insurance (other things equal) if premiums are at least partially community rated. This leads to our first hypothesis: H1: **Increasing health risk results in a higher probability of obtaining some insurance, other things equal.** If we find no relationship of insurance purchasing with risk, that would be consistent with risk rating.

We assume that income effects of the type discussed informally above do exist at low income levels: that is, that people with higher expected expenses (higher risks) and a given money income have lower real incomes, and this lower income makes them less likely to obtain

insurance. This assumption then generates hypothesis H2: **The risk gradient increases with income.** Put the other way round, the higher relative likelihood of purchase by higher risks diminishes as income falls under partial risk rating. If in contrast premiums are risk rated, lower income higher risks might even be less likely to obtain insurance than lower risks. There is an awkward possibility under imperfect community rating in which the income effects discussed in H2 cancel out the price effects in H1 so that the risk gradient is flat or has a negative slope.

Finally, we assume that premiums are closer to community rated in group settings than individual settings. Given this assumption, we obtain hypothesis H3: **the risk gradient is steeper in the group market than in the individual market.** We look at small and large group markets separately to allow for the possibility that group size also affects the slope of the risk gradient, but we have no prior hypothesis on this.

## III. Methods

The basic design of our study is to examine the relationship between an individual's health status in a given year (year one) and their health insurance choice in the subsequent year (year two). We use the average relationship between year one health status and year two medical expenditures, conditioning on different sets of correlates of health expenditures, among individuals with private health insurance to construct a measure of health risk. We then use this measure to examine the relationship between health risk and the purchase of health insurance among individuals at different levels of income and individuals likely to be purchasing coverage in different markets including the individual, small group, and large group markets.

A. Data

The data source for this project is the Medical Expenditure Panel Survey (MEPS) produced by the Agency for Healthcare Research and Quality (AHRQ). The Household Component (HC) of the MEPS is a nationally representative survey of the U.S. civilian noninstitutionalized population. The survey collects information about medical care expenditures, medical care use, health care conditions and health insurance coverage as well as detailed information on demographic and socioeconomic characteristics. The HC 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. Although the panels may be combined to produce nationally representative annual estimates, in this project we exploit the panel structure, incorporating information for a given reporting unit over the two year period in which the unit participates (called the reference period) into our analyses.

Our study uses data from four reference periods: 1996 to 1997, 1997-1998, 1998-1999, and 1999-2000. The study sample for our primary analyses includes respondents aged 25-64 who were not covered by public health insurance at any point during their second year in the survey. Table 1 present descriptive statistics for the study sample.

*Measuring Health Risk*

Our measure of health risk is a regression prediction of a privately-insured individual's expected health expenditures based on their age, sex, and prior year health conditions if they were enrolled in a private health insurance plan. This measure is ideal for our study for two reasons. First,

expressing health risk in monetary terms (expected medical expenses) allows us to collapse multiple dimensions of health into a scalar measure. Second, based on our theoretical framework in which individuals compare their expected expenditure to the premium for coverage they face in the market, expected health expenditures are the appropriate unit to measure health risk for purposes of insurance decisions. 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 calculate expected health expenditures for an individual by developing an empirical model of the relationship between age, sex, and health conditions in a given year and privately insured health expenditures in the subsequent year.

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

where $Y_{i,2}$ is individual i's year 2 privately insured health expenditures, A is a categorical indicator of age and sex, C is vector of year 1 health conditions for individual i, and X is a group of control variables including an indicator of whether the insurance coverage with group or non-group insurance, the firm size of policy holder for those with group coverage, and categorical indicators of education, marital status, race, ethnicity, and survey year. We estimate the model on the subset of individuals in our study sample who were continuously enrolled in either group or non-group health insurance for the entire year, following Pauly and Herring (1999).

In the MEPS-HC, survey respondents are asked to identify specific physical and mental conditions they experienced, whether or not they were treated, during the reference period. This

information is used to create a list of conditions and health problems present during the first year of participation in the survey, recorded as ICD-9 codes, for each family member. We used the clinical classification system (CCS) developed by AHRQ which aggregates all diagnosis codes into 260 mutually exclusive clinically homogeneous categories (see MEPS Data Documentation HC-03). The condition indicators we ultimately include in the empirical models are these CGS indicators.

Our objective in estimating the model is to obtain a consistent estimate of $E[Y_{i,2}]$ conditional on an individual's age, sex, and year 1 conditions. The empirical issues in estimating models of health expenditures – the large number of zeros and the highly skewed distribution of the raw data for observations greater than zero – are well known, and a number of methods to deal with these issues have been proposed in the literature (Jones 2000). However, little consensus exists on the appropriate method, at least in part because the best method appears to vary by application (Buntin and Zaslavsky 2004). As a result, we experimented with a variety of methods in preliminary analyses[1], and ultimately chose to use a two stage model for estimation. In the first stage, we estimate $\Pr(Y_{i,2}) > 0$ using a maximum likelihood logit model. In the second stage, we use GLM to estimate $\ln(E(Y_{i,2}))| Y_{i,2} > 0$ assuming the variance of $Y_{i,2}$ conditional on the independent variables in the model is proportional to the mean. We then use this model, which

[1] In preliminary analyses, we estimated both one and two stage models using both least squares and GLM. We estimated OLS models using raw scale expenditures as the dependent variable. For GLM models, we used log transformed expenditures as the dependent variable and tested 3 different assumptions of the distribution of the errors (normal, poisson, gamma). We used the Park test to choose among the different assumptions of the distribution of the errors, and we compared the mean of the predictions from each model by decile with the mean of actual expenditures. Based on these types of comparisons, we chose the two-stage GLM model based on the results of the Park test and the extent to which the means of the predicted values were unbiased predictions of the means and actual expenditures in each decile.

is estimated on the subset of individual with private insurance for the entire year, to predict annual privately insured medical expenditures for all individuals in the sample as follows:

(2) $$E_{C,A}\left[Y_{i,2}\right] = f\left(A_{i,1}, S_{i,1}, C_{i,1}, \overline{X_{i,2}}\right)$$

Our measure is essentially the average expected covered medical expenditure of a privately insured person with similar age, sex and underlying health status. Our assumption is that these characteristics of individuals which affect their anticipated health spending could be incorporated by insurers into setting premiums for coverage. However, we construct two additional measures, based on this model, in order to isolate the effects of easily observable characteristics such as age and sex from less easily observed characteristics such as the presence of individual-specific health conditions. We do this by re-estimating the expenditure model on the subset of individuals with continuous private health insurance in year 2, omitting the indicators of medical conditions:

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

We use the simplified model to generate predicted expenditures adjusted only by age and sex for the full study sample:

(4) $$E_{A}\left[Y_{i,2}\right] = f\left(A_{i,1}, S_{i,1}, \overline{X_{i,2}}\right)$$

We then calculate the difference between the two predictions to develop a measure of the extent to which an individual's expected expenditures differ from the average of those of her age and sex cohort due to the presence of medical conditions:

(5) $$E_{C}\left[Y_{i,2}\right] = E_{C,A}\left[Y_{i,2}\right] - E_{A}\left[Y_{i,2}\right]$$

The intent of this measure is to capture variation in expected expenditures which may be more difficult for insurers or employers to observe when setting premiums. We want to see whether people who are higher risks for these reasons are especially likely to fail to obtain coverage.

Finally, we calculate deciles of each measure of expected expenditures and classify each survey respondent according to their decile of the expenditure distribution.

*Defining health insurance status*

Characterizing health insurance coverage on an annual basis is complicated by the fact that individuals may transition in and out of coverage as well as between the group and individual market during the year. To address this, we use two definitions of health insurance status in our empirical models. The first is a binary indicator of whether the individual had any type of private coverage at any point during the year (including insurance through employer/union, insurance through other group, non-group insurance, insurance through self-employment (firm of size one) and unknown private insurance) or was uninsured all year. The advantage of this definition is that it encompasses anyone who purchases a private plan during the year, including those who held any type of coverage for only part of the year. However, it does not allow us to effectively analyze differences in the type of coverage purchased. Thus, for the second definition, we create a mutually exclusive set of coverage types by excluding people who were covered only part year and those who changed type of coverage between the group and individual markets during the year. Thus, the sample includes only those who were continuously either covered by a group plan, covered by an individual plan, or uninsured all year.

*Defining health insurance markets*

Our study hypotheses are based on the assumption that the extent of risk rating of premiums varies across the individual, small group and large group markets. In our empirical work, we identify people likely to purchase coverage in each of these markets based on the employment status of each member of the insurable unit, which is defined in MEPS as sub family relationship units constructed to include adults plus those family members who would typically be eligible for coverage under the adults' private health insurance family plans. They include adults, their spouses and their unmarried natural/adoptive children age 18 and under as well as children under age 24 who are full time students. These assignments are intended to capture individuals who are likely to purchase in each market and are independent of the type of coverage and whether the individual actually purchased. We assign individuals to markets as follows:

1. Large Group: individuals with any adult member of the insurable unit employed in a firm with greater than 50 employees.
2. Small Group: individuals with any adult member of the insurable unit employed in a firm with 50 or fewer employees and no adult member of the insurable unit employed in a large firm.
3. Individual: all adult members of the insurable unit either self-employed or not employed.

B. Empirical Models

We estimate models of the following basic form:

$$\Pr(Y=1) = f(R, I, X) \quad (6)$$

The probability of purchasing private health insurance is a function of individual health risk (R), income (I), and other characteristics that affect demand for health insurance (X). For models of

any private coverage, we estimate linear probability models using least squares. Because people we classify as in the large group market do obtain individual insurance coverage (usually if group insurance is not offered at their current job) and, somewhat surprisingly, some of the people in the individual market report that they somehow obtained group coverage, we also estimate models in which we identify the type of coverage the individual obtained (group or individual). For these models of mutually exclusive categories, we use multinomial logistic regression, reporting the marginal effects of study variables with other variables set at the mean of the study sample. All results are weighted and standard errors are adjusted for the complex sampling design of the MEPS. We estimate models on the entire sample and separately by market (individual, small group and large group) and family income (<2 times poverty, 2-3 times poverty, >3 times poverty level) to examine differences in the effects of health risk across these variables.

We enter each type of measure of health risk into separate models. In models estimating insurance coverage as a function of expected expenditures conditional on age, sex and condition ($E_{C,A}$) and age and sex ($E_A$), we do not include categorical indicators of age and sex. The reason for this is that they are likely to be highly correlated with our measure of risk. However, these characteristics may also be proxies for demand for health insurance for which we would like to control in our empirical models. This introduces the possibility that the effects we observe may be due to either health risk as measured by expected expenditures or omitted variables which affect demand for health insurance and are correlated with age or sex. We address this in models that include the measure of health risk as expected expenditures attributable to year 1 conditions ($E_C$) by including these categorical measures of age and sex. The interpretation of the effects of

this variable is the relationship between risk attributable to health conditions, controlling for the effects of age and sex both on expected expenditures and other characteristics which affect demand for health insurance.

Finally, we conduct a series of tests of the statistical significance of the coefficients on the indicators of expected expenditure decile, with the objective of determining the statistical significance of the risk gradient. In each table, we report the statistical significance of the test that the coefficient differs from zero for each decile indicator, with the lowest decile the omitted category. At the bottom of each table, we report the results of three tests of the joint significance of the effects of the decile indicators. In the first, we test the hypothesis that the coefficients are jointly zero. In the second, we test the hypothesis that the coefficients on the decile indicators included in the model are equal. Finally, we reestimate the models recoding the decile indicators as a continuous variable. The coefficient on this variable tests the hypothesis that the probability of private insurance increases linearly by decile indicator.

## IV. Results

Although age and sex are correlated with expected expenditures, considerable variance exists within demographic groups based on prior year health conditions. Figures 1 and 2 present the distribution of predicted year two expenditures conditional on age, sex and chronic conditions from the sample with private health insurance by age and sex categories. The predictions are log transformed for the purpose of display. For men, the average of expected expenditures rises continuously with age from $398 for men 25 to 29 to $3120 for men 60-64. For women, in contrast, expected expenditures average between $1200 and $1500 between the ages of 25 and

49, peak between the ages of 50 and 54 at $2,607, and slightly decline to approximately $2,300 ages 55 to 59 and 60 to 64. The distribution around the mean, however, 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.

The distribution of expected expenditures based on age, sex, and prior period condition is highly skewed (Table 2 – Panel A), suggesting that the extent to which premiums are adjusted for risk may be important in purchasing decisions, particularly in the context of an annual choice of coverage. For example, although the average actuarially fair premium for the study population is $1,389, the average of expected expenditures is $309 in the first decile of the distribution and $5,196 in the top decile. Thus, the community rated premium is high relative to expected expenditures for those at the bottom of the spending distribution and low relative to expected expenditures for those at the top.

Of course, whether this variation in expected expenditures matters in purchasing decisions depends both on whether individuals use this information in forming expectations of future health care utilization and the extent to which insurers use this information in setting premiums for coverage. Panels 2 and 3 provide some insight into the potential effects of these issues. For example, assume individuals take into account only age and sex when forming expectations of future health expenditures. Even this minimal amount of information would make the community rated premium attractive to those at the high end of the expenditure distribution and unattractive to those at the low end (Panel 2). However, age and sex are easily observable to

insurers (and employers) and potentially easily incorporated in premium setting. Panel 3 measures the extent to which an individual's health status in year one results in their year 2 predicted expenditures deviating from those in their age-sex cohort. These results indicate that significant within demographic group variation in expected expenditures exists. In deciles 1-4 of this panel, individual expected expenditures are considerably lower than those predicted by their age and sex only. In deciles 5-8, individual expected expenditures are approximately the same as the average based on age and sex adjustment only, and in deciles 9 and 10, expected expenditures are significantly higher than those predicted by age and sex only. Risk rating based on age and sex would not eliminate incentives for selection, assuming individuals use this information in decision making.

What is the relationship between risk and the probability that a person obtains private insurance? Overall, we find that high risks are more likely than low risks to obtain private health insurance (Tables 3a and 3b). Based on unadjusted estimates, 73% of individuals in the bottom two deciles of the expected expenditure distribution had private health insurance compared to 86% of individuals in the top decile (Table 3a). The positive relationship between health risk (as measured by expected expenditures) and the purchase of private health insurance is consistent across income levels and health insurance markets. Although individuals in low income families are much less likely than those in high income families to have private health insurance (53% vs 94%), in both cases, the proportion with private health insurance increases with health risk. Among those with family income below 2 times poverty level (low income), the proportion with any private insurance increases from 0.47 to 0.60 moving from the bottom to the top of the

expected expenditure distribution. The percentage point change is positive but smaller among those classified as high income – 0.88 to 0.96.

Our findings are similar for differences between markets (Table 3b). Those likely to purchase in the individual market are less likely to have private health insurance than those likely to purchase in the large group market – 55% vs 92%. For those in the individual market, the proportion purchasing private insurance increases from 0.38 to 0.67 from the lowest to the top decile of risk. For those in the large group market, the corresponding change is from 0.87 to 0.95. The apparent risk gradient is steeper in the individual than in the group market if slope is measured in absolute terms, but it is similar if measured by the proportional reduction in the probability of being uninsured. The lowest risks are in both cases about half as likely to be uninsured as the highest risks.

The results in Table 3b also indicate how well our measure of potential markets captures individuals likely to purchase in each market. Those classified as likely to purchase in the individual market are much more likely to do so (13%) than those classified as likely to purchase in the small group or large group markets (2% and 1%, respectively). In addition, those classified as likely to purchase in the individual market make up 56% of individuals actually continuously insured in the individual market during the year. However, many people we classify as unlikely to have access to group report that they actually obtain this type of coverage, and the extent to which they obtain this coverage increases dramatically with health risk. 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.

The results of our multivariate models confirm the unadjusted results (Table 4). The likelihood of purchasing private coverage increases with health risk and the percentage point increase is greater among those in low income families than those in high income families. In each model, we reject the hypothesis that the coefficients on the decile indicators are jointly zero as well as the hypothesis that they are equal. The average effect of an increase of one decile indicator is 0.012 percentage points, ranging from 0.023 for individuals in low income families to 0.007 for individuals in high income families. We also find that our proxies for markets are consistent with a higher price of coverage in the individual and small group than the large group markets, and that this price has a larger effect on low income than high income individuals. The coefficients on other variables are also generally in the expected direction. Minorities, particularly Hispanics, are less likely to have private coverage, and higher levels of formal education and being married are positively associated with private coverage. Family income has a positive relationship with private coverage even within poverty level group.

We find that expected expenditures attributable to both individual characteristics easily observed by insurers or employers (age and sex) and those not so easily observed (prior year health conditions) are positively associated with the purchase of private insurance (Table 5). However, we begin to see a non-linearity in the likelihood of purchase in the top decile of the distribution. In particular, in the model using deciles of expected expenditure based on age and sex as the measure of health risk, the probability of purchase increases steadily from the 1$^{st}$ and 2nd to the 9$^{th}$ decile of the distribution. The incremental probability declines from 0.12 to 0.08 from the 9$^{th}$ to the 10$^{th}$ decile. This decline is consistent across income groups. This decline at the top decile,

however, is specific to expenditures conditioning on age and sex. In models with expected expenditures attributable to prior year conditions, we do not find a similar decline at the top decile that is consistent across income groups. By entering the categorical indicators of age and sex into these models, we are more effectively controlling for potential unobserved preferences for insurance that are correlated with age and sex. Thus, it is likely that the effect we observe in the top decile in the models based on age and sex is driven by these types of unobserved preferences.

When we compare the relationship between different measures of expected health expenditures in the different markets, we find some evidence of differences across markets in the effects of health risk (Table 6). Using our base model, we find that the likelihood of purchase increases relatively consistently with expected expenditures conditional on age, sex and prior year conditions. Using our constrained models, in contrast, we find a decline in the likelihood of purchase at the top decile which is constant across markets. This is consistent with the top decile of the expected expenditure distribution based only on age and sex picking up unobserved preferences for coverage that have similar effects on the probability of coverage across markets. Using the incremental measure, the probability of purchase steadily increases throughout the distribution in the small and large group markets. In the individual market, in contrast, the probability of coverage does not increase steadily above the lowest decile of health risk attributable to prior year conditions. In other words, the lowest risks in this market are less likely than others to obtain private coverage than higher risks, but the remainder of the population is about equally likely to obtain coverage. This is consistent with the results of our statistical tests. Although we reject the hypothesis that the coefficients on the decile indicators are jointly zero in

the individual market, we cannot reject the hypothesis that the coefficients on these variables are equal. In the small and large group markets, in contrast, we reject both hypotheses.

We find relatively little evidence that the relationship between the risk gradient and income varies by market (Table 7). In general, high risks are more likely to purchase coverage than low risks in each market, at each level of income, using each of the measures of expected expenditures. In addition, the risk gradient is steeper in the individual than the small and large group markets at each level of income. Finally, the slope of the gradient declines with income within each market. The one exception to this last point is in the individual market where the risk gradient initially increases from low to medium income then declines from medium to high income relatively consistently across the different measures of expected expenditures. For example, in the first panel of Table 7, the coefficient on the continuous decile indicator in the individual market increases from 0.025 to 0.039 from low to medium income then declines to 0.017. One possible interpretation of this is that low income, high risk individuals in the individual market (particularly those in the top decile of the expected expenditure distribution) are constrained in their ability to obtain private insurance relative to individuals with medium levels of income in the individual market. However, it is also possible that selection into other types of coverage, in particular state Medicaid programs, is also driving this result. In other words, low income, high risk individual who are eligible for Medicaid are more likely to take up this type of coverage than low income low risk individuals.

In Table 8, we restrict our analysis to those who were either continuously covered by a single type of coverage (group or individual) or uninsured throughout the year. This allows us to create

mutually exclusive categories of coverage in order to determine the extent to which our indicators of market are accurately capturing purchasing decisions. When accounting for the type of coverage the individual actually purchased, we find that overall, the risk gradient appears to be steeper in the group than the individual market, as hypothesized. The probability of obtaining group coverage increases continuously with decile of expected expenditure while the probability of obtaining individual coverage is slightly negative, but not statistically significant. In the analyses by market, we find that underlying these net effects of expected expenditures on the probability of coverage in each market are differing effects across markets. In particular, those who appear likely to purchase in the individual market are characterized by a steep, positive risk gradient for purchase in the *group* market, while the marginal effects of increasing risk are positive for these individuals in the individual market, they are not statistically significant. In the case of those likely to purchase in the small and large group markets, the risk gradient is actually *negative* and statistically significant, albeit the marginal effects are small. Overall, this indicates that movement exists across our market potential boundaries in the direction of selection of high risks into group coverage and low risks into individual coverage. This is consistent with less risk rating of premiums in the group than the individual market.

## V. Conclusions

We find that the likelihood of purchasing private health insurance nearly always increases with health risk. In most of the models we estimated, the likelihood of purchase continuously increased from individuals in the lowest to those in the high decile of expected expenditures. This was the case for both health risk based only on age and sex and health risk based on individual health status, independent of age and sex.

We also find that the risk gradient declines with income, rather than increases as we hypothesized, in the small and large group markets. This may be due in part to the tax treatment of group health insurance, although that effect should not alter the relative premiums at a given income level. While increasing risk lowers real income, assuming premiums are at least partially risk rated, the magnitude of the subsidy increases with risk, due to the higher premium for coverage. In the individual market, in contrast, we find some modest evidence in support of the hypothesis of the risk gradient increasing with income when comparing the relationship between risk and the purchase of coverage between low and medium income individuals. The risk gradient, however, declines between medium and high risk individuals.

Contrary to our hypothesis, we also find that the risk gradient is steeper for individuals likely to purchase in the individual than the group market. This was true when measuring expected expenditures either by easily observed characteristics such as age and sex or more difficult to observe prior year health status. A number of possible explanations for this exist. Our hypothesis was based on the assumption that the rating of premiums was closer to community rating in the group than the individual market. Either the group market may be characterized by more risk rating of premiums than is generally thought or the individual market may be characterized by less risk rating of premiums.

Alternatively, our hypothesis may be correct, but our measures of the relevant market for an individual may be noisy. Our results provide some evidence 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 significant proportion of those we identify as likely to purchase in the individual market based on family employment status obtain some type of group coverage, and the likelihood of doing so increases dramatically with risk. In fact, when we control for this, the positive risk gradient in the individual market is no longer statistically significant. In addition, we find some evidence that low risk people who appear likely to be able to obtain group coverage may seek coverage in the individual market. Both these situations are consistent with greater risk rating of premiums in the individual than the group market, with high risks seeking coverage in the group market and low risks dropping out. Albeit, these effects, particularly the low risks who appear eligible for group coverage who choose individual coverage, are not particularly large.

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. Overall, we find that risk is associated with a higher likelihood of having private health insurance. The only suggestion of issues of affordability relating to health risk is among low income, high risk individuals likely to purchase in the individual market. Although the probability of purchasing insurance increases with risk among these individuals, it increases more slowly than among similar individuals characterized by medium levels of income. However, in our current empirical specification, it is not possible to determine whether this is due to issues of affordability in the private market or the opportunity to obtain publicly funded coverage for these individuals.

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 relatively large, continuously 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.

Do our results provide evidence that the affordability of private insurance coverage is a particular problem for high risks? We find little general evidence linking health risk to behavior consistent with a lack of affordability of private health insurance to high risks. If it is a problem, it only occurs among the very high risk and low income individuals relying on the individual market.

**References:**


Buntin, M. B. and A. M. Zaslavsky (2004). "Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures." Journal of Health Economics **23**: 525-542.

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

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

Herring, B. and M. Pauly (2003). Incentive-Compatible Guarantee Renewable Health Insurance. NBER Working Paper**:** 1-43.

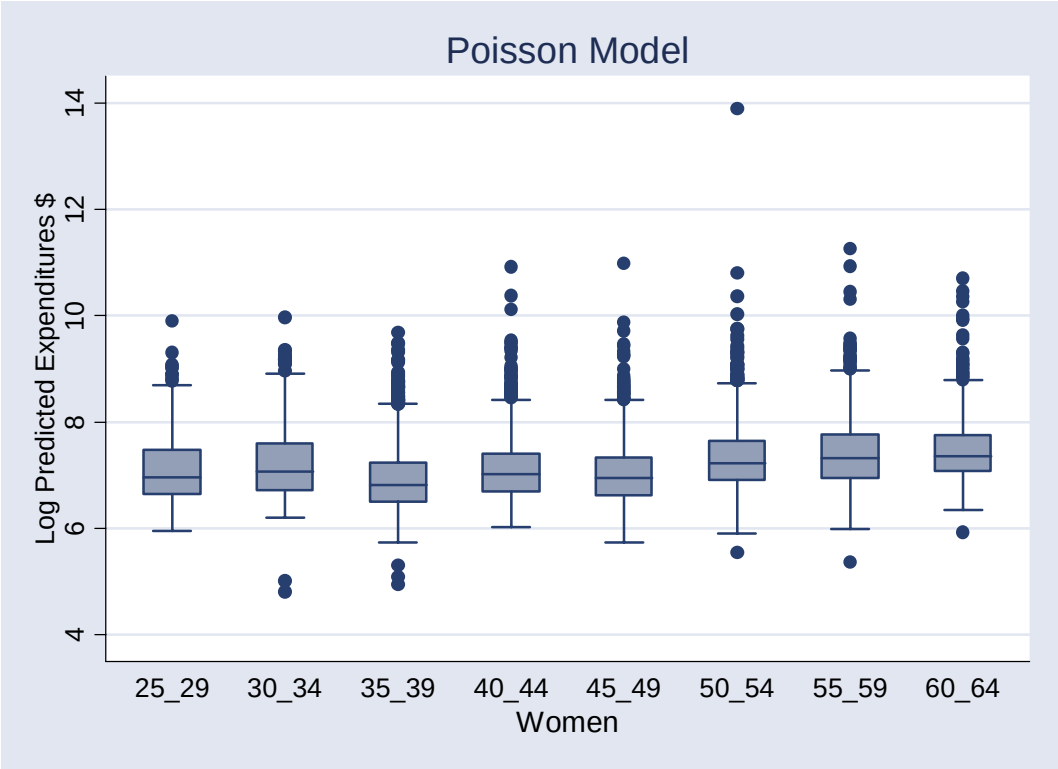
Jones, A. M. (2000). Health Econometrics. Handbook of Health Economics. A. J. Culyer and J. P. Newhouse. Amsterdam, Elsevier. **1A:** 265-344.

Pauly, M. V. and B. Herring (1999). Pooling Health Insurance Risks. Washington DC, AEI Press.

Pollitz, K. and R. Sorian (2002). "Ensuring Health Security: Is the Individual Market ready for Prime Time?" Health Affairs **Web Exclusive**(23 October 2002): W372-W376.

**Figures 1 and 2: The Distribution by Age and Sex of Predicted Year 2 Expenditures based on Age, Sex, and Year 1 Expected Expenditures**

Sample: Privately Insured Individuals 25-64



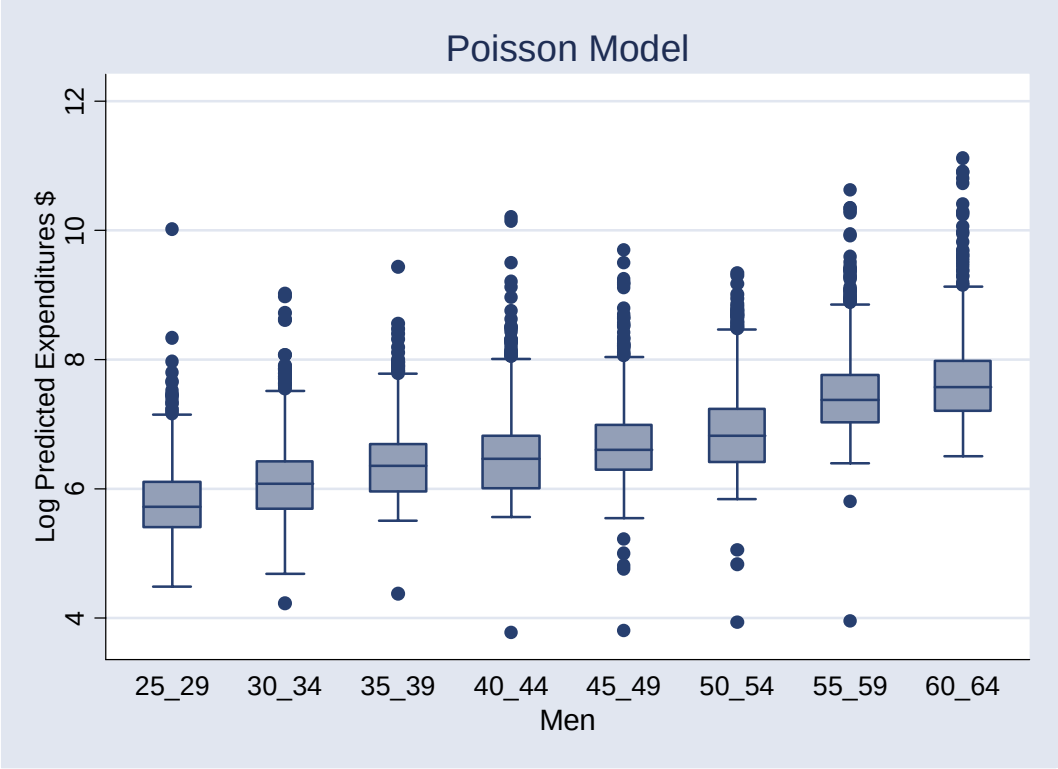



**Table 1: Descriptive Statistics for Independent Variables**
Study Sample: Individuals 25-64 without private health insurance

| | n | mean | sd | min | max |
|---|---|---|---|---|---|
| Male 25-29 | 23,343 | 0.07 | 0.25 | 0 | 1 |
| Male 30-34 | 23,343 | 0.07 | 0.26 | 0 | 1 |
| Male 35-39 | 23,343 | 0.08 | 0.27 | 0 | 1 |
| Male 40-44 | 23,343 | 0.08 | 0.27 | 0 | 1 |
| Male 45-49 | 23,343 | 0.07 | 0.25 | 0 | 1 |
| Male 50-54 | 23,343 | 0.06 | 0.23 | 0 | 1 |
| Male 55-59 | 23,343 | 0.04 | 0.20 | 0 | 1 |
| Male 60-64 | 23,343 | 0.03 | 0.17 | 0 | 1 |
| | | | | | |
| Female 25-29 | 23,343 | 0.06 | 0.24 | 0 | 1 |
| Female 30-34 | 23,343 | 0.07 | 0.25 | 0 | 1 |
| Female 35-39 | 23,343 | 0.08 | 0.27 | 0 | 1 |
| Female 40-44 | 23,343 | 0.08 | 0.27 | 0 | 1 |
| Female 45-49 | 23,343 | 0.07 | 0.25 | 0 | 1 |
| Female 50-54 | 23,343 | 0.06 | 0.24 | 0 | 1 |
| Female 55-59 | 23,343 | 0.04 | 0.20 | 0 | 1 |
| Female 60-64 | 23,343 | 0.03 | 0.18 | 0 | 1 |
| | | | | | |
| Black | 23,343 | 0.11 | 0.31 | 0 | 1 |
| White | 23,343 | 0.84 | 0.36 | 0 | 1 |
| Hispanic | 23,343 | 0.10 | 0.30 | 0 | 1 |
| Education < HS | 23,343 | 0.12 | 0.33 | 0 | 1 |
| Education HS | 23,343 | 0.32 | 0.47 | 0 | 1 |
| Education - Some College | 23,343 | 0.24 | 0.43 | 0 | 1 |
| Education College Graduate or higher | 23,343 | 0.31 | 0.46 | 0 | 1 |
| Married | 23,340 | 0.67 | 0.47 | 0 | 1 |
| Previously Married | 23,340 | 0.17 | 0.37 | 0 | 1 |
| | | | | | |
| Income | 23,343 | 33.12 | 27.24 | -75.74512 | 276.109 |
| Income^2 | 23,343 | 1,838.56 | 3,549.02 | 0 | 76236.19 |
| Child | 22,170 | 0.43 | 0.49 | 0 | 1 |
| | | | | | |
| Panel 1997 | 23,343 | 0.24 | 0.43 | 0 | 1 |
| Panel 1998 | 23,343 | 0.25 | 0.43 | 0 | 1 |
| Panel 1999 | 23,343 | 0.25 | 0.43 | 0 | 1 |
| Panel 2000 | 23,343 | 0.26 | 0.44 | 0 | 1 |
| | | | | | |
| Individual Market | 22,170 | 0.13 | 0.34 | 0 | 1 |
| Small Group Market | 22,170 | 0.31 | 0.46 | 0 | 1 |
| Large Group Market | 22,170 | 0.56 | 0.50 | 0 | 1 |

**Table 2: Summary of Three Measures of Health Risk**
Predicted Expenditures by Decile for Study Sample

| | Base Model (Age, Sex, and Year 1 Conditions) | | | |
|---|---|---|---|---|
| Decile | mean | s.d. | min | max |
| 1 | 309 | 66 | 0 | 398 |
| 2 | 463 | 42 | 399 | 520 |
| 3 | 594 | 41 | 520 | 653 |
| 4 | 716 | 32 | 654 | 778 |
| 5 | 853 | 46 | 778 | 934 |
| 6 | 1,006 | 49 | 934 | 1,097 |
| 7 | 1,211 | 71 | 1,097 | 1,343 |
| 8 | 1,511 | 107 | 1,343 | 1,708 |
| 9 | 2,030 | 204 | 1,708 | 2,452 |
| 10 | 5,196 | 22,782 | 2,452 | 1,078,937 |
| All | 1,389 | 7,330 | 0 | 1,078,937 |

| | Constrained Model (Age and Sex only) | | | |
|---|---|---|---|---|
| Decile | mean | s.d. | min | max |
| 1 | 395 | 64 | 328 | 456 |
| 2 | 630 | 0 | 630 | 630 |
| 3 | 790 | 64 | 731 | 859 |
| 4 | 1,124 | 0 | 1,124 | 1,124 |
| 5 | 1,187 | 49 | 1,141 | 1,240 |
| 6 | 1,358 | 0 | 1,358 | 1,358 |
| 7 | 1,443 | 0 | 1,443 | 1,443 |
| 8 | 1,657 | 110 | 1,551 | 1,772 |
| 9 | 2,080 | 26 | 2,049 | 2,102 |
| 10 | 2,441 | 324 | 2,158 | 2,812 |
| All | 1,234 | 594 | 328 | 2,812 |

| | Incremental (Constrained-Base) | | | |
|---|---|---|---|---|
| Decile | mean | s.d. | min | max |
| 1 | -1,078 | 274 | -2,148 | -796 |
| 2 | -676 | 54 | -796 | -597 |
| 3 | -485 | 60 | -597 | -403 |
| 4 | -326 | 39 | -403 | -252 |
| 5 | -201 | 33 | -252 | -154 |
| 6 | -99 | 28 | -154 | -47 |
| 7 | 23 | 41 | -47 | 103 |
| 8 | 214 | 71 | 103 | 353 |
| 9 | 602 | 160 | 353 | 930 |
| 10 | 3,585 | 22,760 | 930 | 1,077,165 |
| All | 155 | 7,300 | -2,148 | 1,077,165 |

**Table 3a: Insurance Status by Health Risk and Income**
**Expected Expenditures based on Age, Sex, and Year 1 Conditions**
Study Sample: Individuals 25-64 without private health insurance

| | Full Sample | | | Continuous Coverage of One Type or Continuously Uninsured | | | |
|---|---|---|---|---|---|---|---|
| Expected Expenditure Decile | N | Private Insurance at any Point During the Year | Uninsured All Year | N | Group | Non Group | Uninsured |
| **All** | | | | | | | |
| 1-2 | 4,648 | 0.73 | 0.27 | 4060 | 0.66 | 0.02 | 0.31 |
| 3-4 | 4,656 | 0.79 | 0.21 | 4122 | 0.73 | 0.03 | 0.24 |
| 5-6 | 4,658 | 0.83 | 0.17 | 4118 | 0.77 | 0.03 | 0.19 |
| 7-8 | 4,660 | 0.84 | 0.16 | 4183 | 0.79 | 0.03 | 0.18 |
| 9 | 2,331 | 0.85 | 0.15 | 2118 | 0.79 | 0.04 | 0.16 |
| 10 | 2,323 | 0.86 | 0.14 | 2092 | 0.81 | 0.04 | 0.15 |
| Total | 23,276 | 0.81 | 0.19 | 20693 | 0.75 | 0.03 | 0.21 |
| % change | | 18% | -49% | | 21% | 59% | -51% |
| **Low Income** | | | | | | | |
| 1-2 | 1,303 | 0.47 | 0.53 | 1127 | 0.36 | 0.02 | 0.61 |
| 3-4 | 1,110 | 0.50 | 0.51 | 961 | 0.39 | 0.02 | 0.58 |
| 5-6 | 1,041 | 0.56 | 0.44 | 877 | 0.43 | 0.04 | 0.53 |
| 7-8 | 967 | 0.56 | 0.44 | 830 | 0.45 | 0.04 | 0.51 |
| 9 | 512 | 0.60 | 0.40 | 445 | 0.47 | 0.05 | 0.46 |
| 10 | 523 | 0.60 | 0.40 | 455 | 0.47 | 0.06 | 0.46 |
| Total | 5,456 | 0.53 | 0.47 | 4695 | 0.42 | 0.03 | 0.54 |
| % change | | 27% | -24% | | 31% | 159% | -24% |
| **Middle Income** | | | | | | | |
| 1-2 | 1,711 | 0.78 | 0.22 | 1458 | 0.72 | 0.02 | 0.25 |
| 3-4 | 1,662 | 0.83 | 0.17 | 1438 | 0.77 | 0.03 | 0.20 |
| 5-6 | 1,556 | 0.86 | 0.14 | 1341 | 0.81 | 0.03 | 0.16 |
| 7-8 | 1,521 | 0.86 | 0.14 | 1330 | 0.80 | 0.03 | 0.17 |
| 9 | 782 | 0.88 | 0.12 | 698 | 0.82 | 0.04 | 0.13 |
| 10 | 747 | 0.91 | 0.09 | 663 | 0.87 | 0.03 | 0.10 |
| Total | 7,979 | 0.84 | 0.16 | 6928 | 0.79 | 0.03 | 0.18 |
| % change | | 16% | -59% | | 20% | 59% | -61% |
| **High Income** | | | | | | | |
| 1-2 | 1,634 | 0.88 | 0.12 | 1475 | 0.84 | 0.03 | 0.13 |
| 3-4 | 1,884 | 0.92 | 0.08 | 1723 | 0.88 | 0.02 | 0.09 |
| 5-6 | 2,061 | 0.95 | 0.06 | 1900 | 0.90 | 0.03 | 0.06 |
| 7-8 | 2,172 | 0.96 | 0.04 | 2023 | 0.92 | 0.03 | 0.04 |
| 9 | 1,037 | 0.96 | 0.04 | 975 | 0.92 | 0.03 | 0.05 |
| 10 | 1,053 | 0.96 | 0.04 | 974 | 0.92 | 0.03 | 0.05 |
| Total | 9,841 | 0.94 | 0.07 | 9070 | 0.90 | 0.03 | 0.07 |
| % change | | 9% | -65% | | 10% | 19% | -66% |

**Table 3b: Insurance Status by Health Risk and Market**

| | Full Sample | | | Continuous Coverage of One Type or Continuously Uninsured | | | |
|---|---|---|---|---|---|---|---|
| Predicted Decile | N | Private Insurance at any Point During the Year | Uninsured All Year | N | Group | Non Group | Uninsured |
| **Individual Market** | | | | | | | |
| 1-2 | 489 | 0.38 | 0.62 | 429 | 0.21 | 0.08 | 0.70 |
| 3-4 | 517 | 0.45 | 0.55 | 457 | 0.26 | 0.10 | 0.63 |
| 5-6 | 564 | 0.60 | 0.40 | 483 | 0.36 | 0.15 | 0.47 |
| 7-8 | 607 | 0.61 | 0.40 | 530 | 0.38 | 0.16 | 0.45 |
| 9 | 382 | 0.62 | 0.38 | 343 | 0.41 | 0.14 | 0.42 |
| 10 | 445 | 0.67 | 0.33 | 389 | 0.47 | 0.14 | 0.37 |
| Total | 3,004 | 0.55 | 0.45 | 2631 | 0.35 | 0.126 | 0.51 |
| % change | | 76% | -47% | | 123% | 87% | -47% |
| **Small Firm (50 or less employees) Market** | | | | | | | |
| 1-2 | 1,586 | 0.64 | 0.36 | 1369 | 0.55 | 0.03 | 0.41 |
| 3-4 | 1,439 | 0.70 | 0.30 | 1266 | 0.61 | 0.03 | 0.35 |
| 5-6 | 1,360 | 0.76 | 0.24 | 1183 | 0.68 | 0.04 | 0.28 |
| 7-8 | 1,374 | 0.78 | 0.22 | 1210 | 0.72 | 0.02 | 0.25 |
| 9 | 612 | 0.81 | 0.19 | 545 | 0.74 | 0.04 | 0.21 |
| 10 | 648 | 0.84 | 0.16 | 572 | 0.79 | 0.02 | 0.18 |
| Total | 7,019 | 0.74 | 0.26 | 6145 | 0.66 | 0.031 | 0.30 |
| % change | | 30% | -55% | | 43% | -23% | -56% |
| **Large Firm (>50 employees) Market** | | | | | | | |
| 1-2 | 2,259 | 0.87 | 0.13 | 1977 | 0.84 | 0.01 | 0.15 |
| 3-4 | 2,447 | 0.91 | 0.09 | 2171 | 0.89 | 0.01 | 0.10 |
| 5-6 | 2,509 | 0.92 | 0.08 | 2247 | 0.90 | 0.01 | 0.09 |
| 7-8 | 2,477 | 0.94 | 0.06 | 2252 | 0.92 | 0.01 | 0.07 |
| 9 | 1,236 | 0.95 | 0.06 | 1133 | 0.93 | 0.01 | 0.06 |
| 10 | 1,153 | 0.95 | 0.05 | 1058 | 0.94 | 0.01 | 0.06 |
| Total | 12,081 | 0.92 | 0.08 | 10838 | 0.90 | 0.007 | 0.09 |
| % change | | 9% | -59% | | 11% | -29% | -61% |

**Table 4: The Relationship between Health Risk and the Purchase of Private Health Insurance**
Expected Expenditures based on Age, Sex, and Year 1 Conditions (Base Model)
Study Sample: Individuals 25-64 without private health insurance

| | | Income | | |
|---|---|---|---|---|
| | **All** | **Low** | **Medium** | **High** |
| Deciles 3-4 | 0.032 | 0.051 | 0.03 | 0.028 |
| | [3.92]** | [2.21]* | [2.11]* | [2.93]** |
| Deciles 5-6 | 0.059 | 0.093 | 0.072 | 0.041 |
| | [7.48]** | [4.12]** | [5.26]** | [4.28]** |
| Deciles 7-8 | 0.073 | 0.134 | 0.079 | 0.05 |
| | [8.79]** | [5.50]** | [5.48]** | [5.54]** |
| Decile 9 | 0.094 | 0.188 | 0.106 | 0.054 |
| | [10.02]** | [6.84]** | [6.29]** | [5.16]** |
| Decile 10 | 0.107 | 0.187 | 0.145 | 0.056 |
| | [11.21]** | [6.73]** | [8.53]** | [5.32]** |
| Individual Market | -0.263 | -0.321 | -0.278 | -0.183 |
| | [24.41]** | [13.03]** | [12.26]** | [11.59]** |
| Small Group Market | -0.103 | -0.197 | -0.099 | -0.056 |
| | [15.23]** | [10.02]** | [9.47]** | [7.92]** |
| Black | -0.005 | -0.003 | 0.034 | -0.006 |
| | [0.23] | [0.06] | [1.14] | [0.29] |
| White | 0.04 | 0.021 | 0.072 | 0.017 |
| | [2.30]* | [0.45] | [2.72]** | [0.88] |
| Hispanic | -0.132 | -0.147 | -0.102 | -0.068 |
| | [12.59]** | [6.93]** | [6.28]** | [4.45]** |
| Education HS | 0.153 | 0.132 | 0.133 | 0.071 |
| | [12.86]** | [6.14]** | [7.07]** | [4.01]** |
| Education - Some College | 0.189 | 0.169 | 0.171 | 0.091 |
| | [15.95]** | [6.30]** | [9.09]** | [5.30]** |
| Education College Graduate or higher | 0.205 | 0.241 | 0.196 | 0.109 |
| | [16.75]** | [7.89]** | [10.44]** | [6.40]** |
| Married | 0.105 | 0.165 | 0.067 | 0.079 |
| | [10.28]** | [6.06]** | [4.30]** | [7.27]** |
| Previously Married | 0.008 | 0.009 | -0.007 | 0.033 |
| | [0.71] | [0.34] | [0.43] | [2.52]* |
| Income | 0.005 | 0.007 | 0.006 | 0.001 |
| | [19.51]** | [3.73]** | [6.73]** | [4.93]** |
| Income^2 | 0 | 0 | 0 | 0 |
| | [14.03]** | [1.19] | [4.42]** | [3.73]** |
| Child in Family | -0.015 | -0.03 | 0.026 | -0.011 |
| | [2.57]* | [1.48] | [2.42]* | [1.92]+ |
| Panel 1997 | -0.016 | 0.002 | 0.026 | -0.027 |
| | [2.32]* | [1.81]+ | [1.85]+ | [3.74]** |
| Panel 1998 | -0.004 | 0.045 | -0.005 | -0.008 |
| | [0.47] | [0.07] | [0.41] | [1.87]+ |
| Panel 1999 | 0.007 | 0.029 | -0.013 | -0.015 |
| | [0.90] | [1.16] | [0.79] | [0.98] |
| Constant | 0.504 | 0.359 | 0.518 | 0.751 |
| | [20.34]** | [5.94]** | [12.89]** | [24.07]** |
| Observations | 22101 | 5193 | 7559 | 9349 |
| H0: D34=D56=D78=D9=D10 | [18.37]** | [8.95]** | [13.74]** | [3.25]* |
| H0: D34=D56=D78=D9=D10=0 | [31.14]** | [15.89]** | [17.79]** | [7.87]** |
| H0: Slope=0 | 0.012** | 0.023** | 0.015** | 0.007** |
| R-squared | 0.24 | 0.20 | 0.17 | 0.12 |

Absolute value of t statistics in brackets
+ significant at 10%; * significant at 5%; ** significant at 1%

**Table 5: The Relationship between Health Risk and the Purchase of Private Health Insurance**
Expected Expenditures based on Constrained Model and the Difference between Base and Constrained Models (Incremental)
Study Sample: Individuals 25-64 without private health insurance

| | Expected Expenditures - Age and Sex | | | | Incremental Expected Expenditures | | | |
|---|---|---|---|---|---|---|---|---|
| | | Income | | | | Income | | |
| | All | Low | Medium | High | All | Low | Medium | High |
| Deciles 3-4 | 0.019 | 0.003 | 0.028 | 0.02 | 0.032 | 0.074 | 0.041 | 0.003 |
| | [2.42]* | [0.11] | [1.76]+ | [2.32]* | [4.01]** | [2.94]** | [2.67]** | [0.36] |
| Deciles 5-6 | 0.06 | 0.103 | 0.075 | 0.037 | 0.031 | 0.101 | 0.037 | 0.005 |
| | [8.74]** | [4.66]** | [5.83]** | [4.71]** | [3.17]** | [3.56]** | [2.42]* | [0.46] |
| Deciles 7-8 | 0.06 | 0.072 | 0.079 | 0.04 | 0.044 | 0.087 | 0.051 | 0.02 |
| | [8.10]** | [3.34]** | [5.80]** | [5.08]** | [4.95]** | [3.41]** | [3.03]** | [2.20]* |
| Decile 9 | 0.121 | 0.21 | 0.156 | 0.063 | 0.06 | 0.114 | 0.07 | 0.032 |
| | [11.18]** | [6.61]** | [7.09]** | [5.22]** | [6.25]** | [3.78]** | [4.06]** | [3.76]** |
| Decile 10 | 0.078 | 0.184 | 0.105 | 0.043 | 0.06 | 0.135 | 0.083 | 0.016 |
| | [6.94]** | [4.95]** | [4.31]** | [3.93]** | [6.51]** | [4.98]** | [5.22]** | [1.68]+ |
| Male 25-29 | | | | | -0.02 | -0.025 | -0.008 | -0.032 |
| | | | | | [1.32] | [0.60] | [0.36] | [1.63] |
| Male 30-34 | | | | | -0.02 | -0.037 | -0.009 | -0.021 |
| | | | | | [1.58] | [0.97] | [0.40] | [1.35] |
| Male 35-39 | | | | | -0.016 | 0.002 | -0.024 | -0.028 |
| | | | | | [1.32] | [0.04] | [0.98] | [2.17]* |
| Male 45-49 | | | | | -0.006 | -0.043 | 0.02 | -0.012 |
| | | | | | [0.52] | [0.89] | [0.88] | [0.91] |
| Male 50-54 | | | | | 0.02 | 0.042 | 0.042 | 0 |
| | | | | | [1.55] | [0.82] | [1.67]+ | [0.02] |
| Male 55-59 | | | | | 0.063 | 0.109 | 0.112 | 0.018 |
| | | | | | [4.10]** | [1.84]+ | [3.41]** | [1.23] |
| Male 60-64 | | | | | 0.11 | 0.321 | 0.124 | 0.038 |
| | | | | | [6.37]** | [5.87]** | [3.66]** | [2.31]* |
| Female 25-29 | | | | | 0.047 | 0.118 | 0.071 | 0.008 |
| | | | | | [3.58]** | [2.95]** | [3.16]** | [0.56] |
| Female 30-34 | | | | | 0.039 | 0.082 | 0.072 | 0.006 |
| | | | | | [3.36]** | [2.21]* | [3.35]** | [0.51] |
| Female 35-39 | | | | | 0.039 | 0.091 | 0.048 | 0.016 |
| | | | | | [3.53]** | [2.61]** | [2.48]* | [1.54] |
| Female 40-44 | | | | | 0.04 | 0.055 | 0.057 | 0.014 |
| | | | | | [3.90]** | [1.34] | [2.98]** | [1.23] |
| Female 45-49 | | | | | 0.055 | 0.1 | 0.079 | 0.018 |
| | | | | | [4.89]** | [2.61]** | [3.74]** | [1.52] |
| Female 50-54 | | | | | 0.085 | 0.159 | 0.115 | 0.035 |
| | | | | | [6.51]** | [3.29]** | [4.67]** | [2.84]** |
| Female 55-59 | | | | | 0.114 | 0.24 | 0.166 | 0.032 |
| | | | | | [8.03]** | [5.06]** | [6.53]** | [2.12]* |
| Female 60-64 | | | | | 0.139 | 0.252 | 0.166 | 0.069 |
| | | | | | [8.30]** | [4.76]** | [5.34]** | [3.86]** |
| Constant | 0.504 | 0.376 | 0.498 | 0.754 | 0.489 | 0.311 | 0.472 | 0.772 |
| | [20.47]** | [6.43]** | [12.02]** | [24.50]** | [17.42]** | [4.22]** | [10.27]** | [23.49]** |
| Observations | 22101 | 5193 | 7559 | 9349 | 22101 | 5193 | 7559 | 9349 |
| D34=D56=D78=D9=D10 | [29.78]** | [10.46]** | [10.96]** | [4.51]** | [5.27]** | [1.74] | [2.79]+ | [4.69]** |
| D34=D56=D78=D9=D10=0 | [37.58]** | [13.69]** | [17.08]** | [8.31]** | [11.61]** | [5.98]** | [6.14]** | [4.64]** |
| Slope>0 | 0.011** | 0.020** | 0.014** | 0.006** | 0.007** | 0.013** | 0.008** | 0.003** |
| R-squared | 0.24 | 0.2 | 0.17 | 0.12 | 0.24 | 0.21 | 0.18 | 0.12 |

Absolute value of t statistics in brackets
+ significant at 10%; * significant at 5%; ** significant at 1%
Note: Models include controls for race, education, marital status, income, income^2, panel year and (when appropriate) market

**Table 6: The effect of health risk on the purchase of health insurance by insurance market**
Varying measures of Expected Expenditures
Study Sample: Individuals 25-64 without private health insurance

| | Expected Expenditures Conditioning On Age, Sex and Conditions | | | Expected Expenditures Conditioning On Age and Sex | | | Expected Expenditures - Incremental | | |
|---|---|---|---|---|---|---|---|---|---|
| | Individual | Small Group | Large Group | Individual | Small Group | Large Group | Individual | Small Group | Large Group |
| Deciles 3-4 | 0.047 | 0.034 | 0.024 | 0.063 | 0.024 | 0.007 | 0.074 | 0.043 | 0.011 |
| | [1.37] | [2.16]* | [2.92]** | [1.90]+ | [1.55] | [0.88] | [2.39]* | [2.87]** | [1.29] |
| Deciles 5-6 | 0.129 | 0.076 | 0.031 | 0.088 | 0.092 | 0.03 | 0.065 | 0.05 | 0.013 |
| | [3.75]** | [4.69]** | [4.12]** | [2.61]** | [6.49]** | [4.47]** | [1.98]* | [2.59]** | [1.53] |
| Deciles 7-8 | 0.154 | 0.09 | 0.042 | 0.134 | 0.086 | 0.03 | 0.077 | 0.062 | 0.022 |
| | [4.23]** | [5.70]** | [5.33]** | [3.96]** | [5.62]** | [3.87]** | [2.33]* | [3.31]** | [2.62]** |
| Decile 9 | 0.191 | 0.133 | 0.044 | 0.246 | 0.136 | 0.049 | 0.123 | 0.087 | 0.025 |
| | [5.17]** | [6.86]** | [4.78]** | [6.96]** | [6.35]** | [4.72]** | [3.84]** | [4.32]** | [2.99]** |
| Decile 10 | 0.205 | 0.127 | 0.058 | 0.197 | 0.069 | 0.026 | 0.099 | 0.087 | 0.032 |
| | [6.16]** | [6.64]** | [6.63]** | [5.02]** | [3.00]** | [2.79]** | [3.43]** | [4.45]** | [3.76]** |
| Male 25-29 | | | | | | | -0.038 | -0.044 | -0.001 |
| | | | | | | | [0.56] | [1.53] | [0.08] |
| Male 30-34 | | | | | | | -0.061 | -0.016 | -0.015 |
| | | | | | | | [0.90] | [0.58] | [1.18] |
| Male 35-39 | | | | | | | -0.103 | -0.001 | -0.01 |
| | | | | | | | [1.84]+ | [0.03] | [0.78] |
| Male 45-49 | | | | | | | -0.04 | 0.017 | -0.012 |
| | | | | | | | [0.76] | [0.62] | [1.09] |
| Male 50-54 | | | | | | | 0.013 | 0.027 | 0.015 |
| | | | | | | | [0.24] | [0.96] | [1.36] |
| Male 55-59 | | | | | | | 0.107 | 0.087 | 0.02 |
| | | | | | | | [1.94]+ | [2.98]** | [1.47] |
| Male 60-64 | | | | | | | 0.214 | 0.078 | 0.041 |
| | | | | | | | [4.23]** | [2.18]* | [2.92]** |
| Female 25-29 | | | | | | | 0.088 | 0.075 | 0.018 |
| | | | | | | | [1.51] | [3.00]** | [1.30] |
| Female 30-34 | | | | | | | -0.005 | 0.063 | 0.03 |
| | | | | | | | [0.07] | [2.55]* | [2.53]* |
| Female 35-39 | | | | | | | -0.018 | 0.086 | 0.019 |
| | | | | | | | [0.35] | [3.63]** | [1.65] |
| Female 40-44 | | | | | | | 0.089 | 0.073 | 0.013 |
| | | | | | | | [1.73]+ | [3.42]** | [1.15] |
| Female 45-49 | | | | | | | 0.018 | 0.095 | 0.034 |
| | | | | | | | [0.32] | [3.89]** | [3.13]** |
| Female 50-54 | | | | | | | 0.137 | 0.135 | 0.039 |
| | | | | | | | [2.70]** | [5.45]** | [3.17]** |
| Female 55-59 | | | | | | | 0.22 | 0.144 | 0.043 |
| | | | | | | | [4.61]** | [5.06]** | [3.15]** |
| Female 60-64 | | | | | | | 0.205 | 0.156 | 0.065 |
| | | | | | | | [4.18]** | [4.75]** | [4.05]** |
| Constant | 0.008 | 0.277 | 0.69 | -0.012 | 0.27 | 0.693 | -0.005 | 0.244 | 0.686 |
| | [0.12] | [5.91]** | [28.11]** | [0.18] | [5.71]** | [28.85]** | [0.06] | [4.72]** | [25.74]** |
| Observations | 3003 | 7017 | 12081 | 3003 | 7017 | 12081 | 3003 | 7017 | 12081 |
| D34=D56=D78=D9=D10 | [6.18]** | [7.63]** | [5.99]** | [9.48]** | [10.95]** | [5.89]** | [.88] | [1.99]+ | [2.5]* |
| D34=D56=D78=D9=D10=0 | [9.48]** | [13.55]** | [10.59]** | [11.26]** | [15.95]** | [8.78]** | [3.71]* | [5.58]** | [3.98]* |
| Slope>0 | 0.024** | 0.016** | 0.006** | 0.024** | 0.013** | 0.005** | 0.011** | 0.009** | 0.003** |
| R-squared | 0.26 | 0.21 | 0.09 | 0.27 | 0.2 | 0.09 | 0.28 | 0.21 | 0.09 |

Absolute value of t statistics in brackets
+ significant at 10%; * significant at 5%; ** significant at 1%
Note: Models include controls for race, education, marital status, income, income^2, panel year and (when appropriate) market

**Table 7: The effect of health risk on the purchase of health insurance by income and insurance market**
Varying Measures of Health Risk

| | Expected Expenditures Conditioning On Age, Sex and Year 1 Conditions | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Individual | | | Small Group | | | Large Group | | |
| | Low | Medium | High | Low | Medium | High | Low | Medium | High |
| Deciles 3-4 | 0.077 | 0.066 | -0.025 | 0.014 | 0.011 | 0.076 | 0.065 | 0.031 | -0.001 |
| | [1.35] | [0.93] | [0.52] | [0.37] | [0.43] | [3.51]** | [1.93]+ | [1.94]+ | [0.19] |
| Deciles 5-6 | 0.092 | 0.212 | 0.073 | 0.109 | 0.044 | 0.094 | 0.069 | 0.057 | 0.001 |
| | [1.66]+ | [3.35]** | [1.53] | [2.60]** | [1.64] | [4.16]** | [2.33]* | [4.24]** | [0.24] |
| Deciles 7-8 | 0.188 | 0.226 | 0.058 | 0.101 | 0.07 | 0.109 | 0.123 | 0.054 | 0.007 |
| | [3.40]** | [3.56]** | [1.12] | [2.47]* | [2.46]* | [5.39]** | [3.36]** | [3.71]** | [1.12] |
| Decile 9 | 0.183 | 0.292 | 0.112 | 0.197 | 0.141 | 0.101 | 0.163 | 0.047 | 0.009 |
| | [3.06]** | [4.30]** | [2.09]* | [4.14]** | [4.42]** | [4.23]** | [4.00]** | [2.22]* | [1.44] |
| Decile 10 | 0.199 | 0.342 | 0.115 | 0.216 | 0.154 | 0.087 | 0.131 | 0.08 | 0.016 |
| | [3.58]** | [4.79]** | [2.15]* | [4.55]** | [5.04]** | [3.49]** | [2.99]** | [4.96]** | [2.67]** |
| Constant | 0.029 | -0.086 | 0.165 | 0.14 | 0.336 | 0.62 | 0.4 | 0.693 | 0.911 |
| | [0.29] | [0.74] | [1.42] | [1.24] | [4.71]** | [8.31]** | [5.50]** | [15.95]** | [43.04]** |
| Observations | 1155 | 804 | 1044 | 2057 | 2547 | 2413 | 1981 | 4208 | 5892 |
| D34=D56=D78=D9=D10 | [2.1]+ | [4.47]** | [2.06]* | [5.24]** | [8.41]** | [1.01] | [2.53]* | [3.05]+ | [3.20]* |
| D34=D56=D78=D9=D10=0 | [3.98]** | [6.79]** | [2.31]* | [6.79]** | [8.27]** | [6.02]** | [4.32]** | [6.07]** | [2.77]* |
| Slope>0 | 0.025** | 0.038** | 0.017* | 0.025** | 0.017** | 0.011** | 0.017** | 0.007** | 0.002* |
| R-squared | 0.18 | 0.24 | 0.21 | 0.16 | 0.14 | 0.09 | 0.1 | 0.06 | 0.01 |

| | Expected Expenditures Conditioning on Age and Sex | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Individual | | | Small Group | | | Large Group | | |
| | Low | Medium | High | Low | Medium | High | Low | Medium | High |
| Deciles 3-4 | 0.002 | 0.106 | 0.054 | 0.004 | 0.012 | 0.047 | 0.007 | 0.022 | 0 |
| | [0.03] | [1.44] | [1.09] | [0.08] | [0.45] | [2.35]* | [0.20] | [1.45] | [0.04] |
| Deciles 5-6 | 0.091 | 0.161 | 0.005 | 0.136 | 0.083 | 0.084 | 0.088 | 0.048 | 0.009 |
| | [1.81]+ | [2.39]* | [0.10] | [3.54]** | [3.47]** | [4.49]** | [2.99]** | [3.44]** | [1.53] |
| Deciles 7-8 | 0.048 | 0.28 | 0.063 | 0.09 | 0.076 | 0.088 | 0.065 | 0.045 | 0.008 |
| | [0.97] | [3.81]** | [1.09] | [2.26]* | [2.88]** | [4.77]** | [2.18]* | [3.15]** | [1.17] |
| Decile 9 | 0.219 | 0.328 | 0.179 | 0.225 | 0.187 | 0.078 | 0.131 | 0.084 | 0.009 |
| | [3.52]** | [4.61]** | [3.44]** | [4.37]** | [4.77]** | [2.90]** | [2.28]* | [4.09]** | [1.04] |
| Decile 10 | 0.216 | 0.322 | 0.104 | 0.24 | 0.038 | 0.055 | -0.002 | 0.073 | 0.011 |
| | [3.07]** | [4.41]** | [1.87]+ | [3.97]** | [0.79] | [2.26]* | [0.04] | [3.70]** | [1.60] |
| Constant | 0.048 | -0.205 | 0.144 | 0.127 | 0.299 | 0.624 | 0.414 | 0.69 | 0.908 |
| | [0.50] | [1.75]+ | [1.31] | [1.10] | [4.14]** | [8.45]** | [5.74]** | [15.73]** | [43.27]** |
| Observations | 1155 | 804 | 1044 | 2057 | 2547 | 2413 | 1981 | 4208 | 5892 |
| D34=D56=D78=D9=D10 | [4.94]** | [5.00]** | [4.12]** | [5.41]** | [6.42]** | [2.43]* | [2.19]+ | [2.76]+ | [1.14] |
| D34=D56=D78=D9=D10=0 | [4.44]** | [7.68]** | [3.96]** | [7.84]** | [6.63]** | [6.05]** | [3.02]* | [5.21]** | [1.3] |
| Slope>0 | 0.025** | 0.039** | 0.013* | 0.024** | 0.014** | 0.008** | 0.010* | 0.008** | 0.002* |
| R-squared | 0.19 | 0.24 | 0.21 | 0.16 | 0.14 | 0.08 | 0.09 | 0.06 | 0.01 |

| | Expenditures Attributable to Year 1 Conditions | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Individual | | | Small Group | | | Large Group | | |
| | Low | Medium | High | Low | Medium | High | Low | Medium | High |
| Deciles 3-4 | 0.126 | 0.086 | 0.017 | 0.066 | 0.067 | 0.007 | 0.051 | 0.011 | -0.002 |
| | [2.47]* | [1.69]+ | [0.40] | [1.72]+ | [2.24]* | [0.42] | [1.35] | [0.80] | [0.24] |
| Deciles 5-6 | 0.076 | 0.096 | 0.023 | 0.129 | 0.074 | -0.008 | 0.091 | 0 | 0.007 |
| | [1.34] | [1.36] | [0.46] | [2.71]** | [2.24]* | [0.36] | [2.24]* | [0.02] | [0.93] |
| Deciles 7-8 | 0.091 | 0.02 | 0.091 | 0.086 | 0.082 | 0.037 | 0.085 | 0.026 | 0.005 |
| | [1.49] | [0.26] | [2.02]* | [1.84]+ | [2.27]* | [1.70]+ | [2.25]* | [1.68]+ | [0.77] |
| Decile 9 | 0.129 | 0.181 | 0.069 | 0.092 | 0.12 | 0.06 | 0.118 | 0.01 | 0.008 |
| | [1.98]* | [2.56]* | [1.46] | [1.61] | [3.63]** | [3.37]** | [2.75]** | [0.52] | [1.11] |
| Decile 10 | 0.107 | 0.172 | 0.075 | 0.177 | 0.129 | 0.017 | 0.118 | 0.025 | 0.011 |
| | [2.21]* | [2.90]** | [1.35] | [3.76]** | [3.79]** | [0.79] | [2.83]** | [1.78]+ | [1.83]+ |
| Constant | -0.047 | -0.162 | 0.177 | 0.022 | 0.234 | 0.689 | 0.414 | 0.681 | 0.902 |
| | [0.37] | [1.24] | [1.47] | [0.17] | [2.89]** | [8.72]** | [4.44]** | [15.04]** | [35.73]** |
| Observations | 1155 | 804 | 1044 | 2057 | 2547 | 2413 | 1981 | 4208 | 5892 |
| D34=D56=D78=D9=D10 | [0.28] | [1.73] | [1.12] | [2.08]+ | [1.63] | [4.21]* | [1.08] | [1.2] | [1.05] |
| D34=D56=D78=D9=D10=0 | [1.74] | [2.92]+ | [1.38] | [3.23]* | [3.7]* | [3.85]* | [2.39]* | [1.26] | [1.11] |
| Slope>0 | 0.011* | 0.016* | 0.011* | 0.014* | 0.013** | 0.005* | 0.013** | 0.003+ | 0.001+ |
| R-squared | 0.22 | 0.26 | 0.23 | 0.18 | 0.15 | 0.09 | 0.11 | 0.06 | 0.02 |

Absolute value of t statistics in brackets
+ significant at 10%; * significant at 5%; ** significant at 1%
Note: Models include controls for race, education, marital status, income, income^2, panel year and (when appropriate) market

**Table 8: MNL results by market**
Sample: Continuously insured or uninsured
Marginal effects shown
Health Risk measured as expected expenditures conditional on age, sex, and conditions
Dependent variable is type of coverage (group or individual) relative to uninsured all year

ALL

| | Pr(Group Coverage) | | Pr(Individual Coverage) | |
|---|---|---|---|---|
| | dy/dx | P>\|z\| | dy/dx | P>\|z\| |
| Deciles 3-4 | 0.026 | 0.000 | -0.001 | 0.735 |
| Deciles 5-6 | 0.043 | 0.000 | 0.001 | 0.737 |
| Deciles 7-8 | 0.055 | 0.000 | -0.001 | 0.847 |
| Decile 9 | 0.066 | 0.000 | -0.003 | 0.315 |
| Decile 10 | 0.073 | 0.000 | -0.005 | 0.050 |

By Market

Individual Market

| | Pr(Group Coverage) | | Pr(Individual Coverage) | |
|---|---|---|---|---|
| | dy/dx | P>\|z\| | dy/dx | P>\|z\| |
| Deciles 3-4 | 0.064 | 0.250 | 0.020 | 0.631 |
| Deciles 5-6 | 0.117 | 0.030 | 0.062 | 0.172 |
| Deciles 7-8 | 0.118 | 0.034 | 0.093 | 0.067 |
| Decile 9 | 0.208 | 0.000 | 0.041 | 0.349 |
| Decile 10 | 0.199 | 0.000 | 0.052 | 0.275 |

Small Group Market

| | Pr(Group Coverage) | | Pr(Individual Coverage) | |
|---|---|---|---|---|
| | dy/dx | P>\|z\| | dy/dx | P>\|z\| |
| Deciles 3-4 | 0.039 | 0.005 | -0.004 | 0.550 |
| Deciles 5-6 | 0.066 | 0.000 | 0.001 | 0.910 |
| Deciles 7-8 | 0.089 | 0.000 | -0.012 | 0.080 |
| Decile 9 | 0.102 | 0.000 | -0.001 | 0.913 |
| Decile 10 | 0.113 | 0.000 | -0.016 | 0.025 |

Large Group Market

| | Pr(Group Coverage) | | Pr(Individual Coverage) | |
|---|---|---|---|---|
| | dy/dx | P>\|z\| | dy/dx | P>\|z\| |
| Deciles 3-4 | 0.015 | 0.002 | -0.001 | 0.721 |
| Deciles 5-6 | 0.021 | 0.000 | -0.002 | 0.387 |
| Deciles 7-8 | 0.028 | 0.000 | -0.002 | 0.419 |
| Decile 9 | 0.028 | 0.000 | -0.002 | 0.167 |
| Decile 10 | 0.037 | 0.000 | -0.005 | 0.004 |

Note: Models include controls for race, education, marital status, income, income^2, and panel year