Insurance, Health, and the Utilization of Medical Services*

by

Jonathan Meer, Stanford University
Harvey S. Rosen, Princeton University

CEPS Working Paper No. 85

October 2002

*We are grateful to Jeffrey Kling, Douglas Miller and Adriana Lleras-Muney for useful suggestions and to Princeton's Center for Economic Policy Studies for financial support.

ABSTRACT

Most data sets indicate a positive correlation between having health insurance and good health, and between having health insurance and utilizing health care services. Yet the direction of causality is not at all clear. If we observe a positive correlation between the utilization of health care services and insurance status, we do not know if this is because people who anticipate poor health buy more insurance (or take jobs with generous medical coverage), or because insurance lowers the cost of health care, increasing the quantity demanded. The direction of causation between health status and insurance is similarly unclear.

While a few attempts have been made to implement an instrumental variables (IV) strategy to deal with endogeneity, the instruments chosen have not been entirely convincing. In this paper we revisit the IV estimation of the reduced form relationships between insurance and both health care utilization and health status taking advantage of what we argue is a good instrument – the individual's self-employment status. We find that IV estimates of the impact of insurance on the utilization of a variety of health care services are greater than estimates that ignore endogeneity. However, instrumental variables leads to a small and statistically insignificant estimate of the impact of insurance on health status.

The validity of this exercise depends on the extent to which self-employment status is a suitable instrument. To make this case, we analyze panel data on transitions from wage-earning into self-employment and show that individuals who select into self-employment do not differ systematically from those who remain wage-earners with respect to either the utilization of health care or health status. That is, there appear to be no underlying differences that might lead to self-employment *per se* affecting health status or health services utilization.

key words: health, insurance, self-employment

Jonathan Meer
Department of Economics
Stanford University
Stanford, CA 94305
jmeer@stanford.edu

Harvey S. Rosen
Department of Economics
Princeton University
Princeton, NJ 08544
hsr@princeton.edu

You'd better have some medical insurance or you're gonna die. That's right. Everybody says, oh, you've got to eat right, exercise. No, you don't. You need some coverage.

--Chris Rock

1. INTRODUCTION

The Robert Wood Johnson Foundation (RWJF) sponsors a web site called Covering the Uninsured (*CoveringTheUninsured.org)*. The top of the home page delivers this message: "When you're uninsured, life turns out differently." In the center of the page is a picture of a downcast little girl. The text to the left of the girl states, "Her mom gets cancer. They find the tumor early. Her mom is OK." To the right it says, "Her mom gets cancer. She's diagnosed too late. Her mom is gone." The point is clear: Health insurance increases an individual's utilization of health care services which, in turn, leads to better health care outcomes. The RWJF web site certainly reflects the conventional wisdom in policy debates about health insurance.

However, there is not much in the way of convincing statistical measurement of the effect of insurance on utilization and health status. True, in most data sets there is a positive correlation between being insured and good health, and between being insured and utilizing health care services. Yet the direction of causality is not at all clear. As Gruber (2000) points out, "insurance coverage itself may be a function of health status, leading to endogeneity bias in estimates of the effects of insurance on health and the utilization of medical care." Thus, if we observe a positive correlation between the utilization of health care services and insurance status, we do not know if this is because people who anticipate poor health buy more insurance (or take jobs with generous medical coverage), or because insurance lowers the cost of health care, increasing the quantity demanded. The direction of causation between health status and insurance is similarly

unclear.

As Levy and Meltzer (2001) note in their review of the literature in this area, most studies of the links between insurance and health outcomes entirely ignore the possibility of endogeneity. They further observe that, while a few attempts have been made to implement an instrumental variables (IV) strategy to deal with endogeneity, the instruments chosen have not been entirely convincing. In this paper, we revisit the IV estimation of the reduced form relationships between insurance and both health care utilization and health status taking advantage of what we argue is a good instrument-- the individual's self-employment status.

Section 2 examines the previous literature on the subject. Section 3 describes the construction of the data set, which is drawn from the 1996 through 1998 waves of the Medical Expenditure Panel Survey (MEPS). Section 4 discusses econometric issues and presents the results. We find that IV estimates of the impact of insurance on the utilization of a variety of health care services are greater than estimates that ignore endogeneity. However, instrumental variables leads to a small and statistically insignificant estimate of the impact of insurance on health status.

Of course, the validity of these results depends on the extent to which self-employment status is a suitable instrument. Section 5 presents evidence that this is the case. Specifically, we analyze panel data on transitions from wage-earning into self-employment and show that individuals who select into self-employment do not differ systematically from those who remain wage-earners with respect to the utilization of health care and health status. That is, there appear to be no underlying differences that might lead to self-employment *per se* affecting health status or health services utilization. Sec-

tion 6 discusses the possibility that, regardless of health outcomes, lack of insurance presents serious financial difficulties. Again instrumenting using self-employment status, we find that this is generally not the case. Section 7 concludes with a summary and suggestions for future research.

2. PREVIOUS LITERATURE

A number of papers have examined the relationships between insurance status and health services utilization, and between health insurance and health status. Most have been observational studies, which analyze outcome differences between insured and uninsured populations. These papers generally show that having insurance increases the utilization of health care services (Brown *et al.* (1998)). However, Levy and Meltzer (2001) note that observational studies "are hopelessly confounded by both observable and unobservable difference between patients who do and do not have health insurance." For the usual reasons, this can lead to inconsistent estimates of the impact of insurance coverage on health care utilization.

The direction of the bias is not clear *a priori*. Anticipation of relatively high utilization of medical services might lead an individual to seek insurance, which would tend to impart an upward bias to the estimated impact of insurance on utilization. On the other hand, insurers may be able to identify people who will be intensive users of medical services and either decline to offer insurance or charge such a high price that they do not purchase it. An anecdote along these lines appeared recently in the *Wall Street Journal* (April 9, 2002). The story concerned a woman who had been paying $417 per month for health insurance, but whose rate increased to $1,881 per month after she was diagnosed with breast cancer. To the extent the tendencies present in the story are typical, the esti-

mated impact of insurance is biased downwards.

While the literature consistently shows a positive relationship between health services utilization and insurance, the empirical link between insurance and health status is less clear. Lurie *et al.*'s (1986) examination of people who lost Medi-Cal benefits finds that self-reported health status declined for people who lost coverage but did not for those who still had it. Similarly, Currie and Gruber (1996) show that expanding children's health insurance coverage improves child health as measured by mortality rates. On the other hand, Haas *et al.* (1993a, b) find that expanding coverage to pregnant women does not affect health outcomes for them or for their children. Similarly, Kaestner *et al.* (2001) find little support for the notion that the Medicaid expansion improved health. Gruber's (2000) assessment of the state of the literature is accurate: "the extent to which medical care has a positive effect on health is not clear."

The RAND Health Insurance Experiment (Newhouse (1993)) is the only study that randomly assigned people to different insurance regimes and observed the effects. An important finding is that those for whom health care was free utilized about 40 percent more health services than those who had some cost sharing, but this resulted in "little or no measurable effect on health status for the average adult" (p. 243). That is, lower utilization did not translate into worse health. However, "health among the sick poor - approximately the most disadvantaged 6 percent of the population - was adversely affected" (p. 339). This suggests that income may play a role in the causal network between insurance and health, an issue to which we will return below.

3. DATA

The data for this paper are drawn from the 1996, 1997 and 1998 waves of the

Medical Expenditure Panel Survey (MEPS), which has individual-level information on insurance coverage, utilization of health services, health status and self-employment status. The three waves consist of two overlapping two-year panels. We exclude persons younger than 18 and older than 62 in 1996. Those under 18 are unlikely to have a strong connection to the labor market, and those over 62 are facing retirement and have different health care options. We exclude the unemployed, as their situations are likely to be quite different from both wage earners and the self-employed. This leaves us with a sample of 37,331 observations, comprising 23,851 individuals, of whom 9.27 percent are self-employed. There are more observations than individuals because most people were followed for two years.

Each year, respondents were asked about their insurance status, utilization of health care, and a variety of other questions. Some questions, particularly those on health, were asked several times a year. Certain utilization questions, asked yearly, were not asked in 1997, and some questions deal with tests and procedures that are appropriate for only one gender or the other. Therefore, some models are not estimated with the entire sample. The health status question was asked of each individual three times ("rounds") per year (twice in 1996), giving us 92,972 observations comprising 22,514 individuals. Fewer individuals are present in this sample due to missing values.

The MEPS provides information on the utilization of a variety of health care services. They include visits to providers (such as physicians and dentists) and preventative care (such as blood pressure checks, flu shots, physicals, and breast exams). Respondents are asked at the end of the year if, for example, they visited a doctor or had a mammogram in that year. The insurance question is similarly asked on an annual basis. Subjects

are also asked whether the insurance is public or private, and whether it is provided by the employer or self-provided. If the individual has any kind of policy, we characterize him or her as being insured. Further, we construe an individual as being insured if the source of insurance is the spouse.

The MEPS uses a 1 to 5 scale for self-rated health, with 1 being excellent, 2 as very good, 3 as good, 4 as fair, and 5 as poor. Consistent with earlier literature, we use this information to create a dichotomous self-reported health variable which takes a value of one if the individual is in good, very good, or excellent health, and zero otherwise.[1] Clearly, no single number can capture every aspect of an individual's health. Further, "healthy" can mean different things to different people. Nevertheless, it is well-documented that self-reported measures of health have excellent explanatory power in predicting mortality rates. As Idler and Benyamini (1997) note in their comprehensive survey of the literature on self-reported health measures, "over two dozen studies have been published in the U.S. and international literature that test the association between simple, global health assessments and mortality in the samples used: Most find a significant, independent association that persists when numerous health status indicators and other relevant covariates are included."[2]

A description of the variables used in this analysis and the associated summary statistics can be found in Table 1. There is substantial variation in the utilization rates for various medical services. About 65 percent of the sample visited the office of a health care provider during the year; 43 percent had a physical exam; 72 percent had their blood

[1] Clearly, there is some arbitrariness to this procedure. In some experiments, we characterized the individual as being healthy only if she reported being in excellent or very good health, and found that this had no substantive effect on our results.

[2] Additional confirmation of this finding is reported in Hurd and McGarry (1997).

pressure checked; but only 18 percent had a flu shot. 92 percent of respondents consider themselves to be in good health, and 82 percent of the sample is insured for the entire year in any given year.

4. ANALYSIS

4.1 Preliminary Issues

We wish to estimate how a variety of medical services utilization and health status measures depend on insurance status and other covariates. In our models of the utilization of various health care services, the dependent variable, Y, takes a value of one if the individual used the service in question during that year, and zero otherwise. In models that focus on health status, Y is one if the individual is healthy, and zero otherwise. The independent variables in our basic model include an indicator variable for insurance status (Ins), region, family size, age, age-squared, gender, race, and education. We only include covariates that are very likely to be exogenous. Age clearly affects both health and utilization (Lakdawlka and Philipson (1998)), and education has been linked with both health and ability to pay (Taubman and Rosen (1982)). Race (Kass, Weinick, and Monheit (1996)), region (Skinner and Wennberg (1998)), family size (Taubman and Rosen (1982)), and gender (Verbrugge (1985)) have all been shown to have important effects on the variables of interest. Year effects (for utilization) and round effects (for health) are also included. We use the conventional probit model:

$$(1) \qquad \Pr(Y = 1) = \Phi[\alpha_0 + \alpha_1(Ins) + \alpha_2(X) + \varepsilon]$$

where X is the vector of exogenous covariates, ε is the error term, and $\Phi[\bullet]$ is the cumulative normal distribution. In computing confidence intervals for our parameter estimates, we wish to account for possible within-individual correlation of the errors and right hand

side variables. To do this, we perform a clustered procedure, with all years of an individual serving as the cluster.

To begin, we estimate the model without any correction for the endogeneity of insurance. The results, presented in Table 2, show that insurance has a positive and significant effect on most measures of utilization. (Only the coefficients and standard errors for the insurance variable are presented. The estimated coefficients on the other covariates are available on request.) For example, having insurance increases the probability of visiting an office-based care provider by 24.9 percentage points. Insurance also increases the likelihood of having a physical exam, mammogram, and a variety of other preventive procedures by as much as 25 percentage points. One cannot reject the hypothesis that it has no effect on visiting a chiropractor or using alternate care. This is unsurprising: insurance plans rarely cover these services, so it stands to reason that it should have no effect on their utilization. Insurance also has a statistically significant effect on hospital usage: 3.3 percentage points for overnight stays and 0.37 percentage points for outpatient visits. These figures are substantial given the small baseline proportions of hospital visits. (See Table 1.)

As for health, insurance coverage increases the probability of good health by 1.8 percentage points and one can reject the hypothesis that the impact is zero at conventional levels. Therefore, our preliminary results, which do not correct for endogeneity, are consistent with the prevalent view that ascribes both higher utilization rates and beneficial health results to insurance coverage.

However, as noted above, there is good reason to believe that insurance status is endogenous to both utilization and health care decisions. To deal with this in an IV

framework requires an instrument that is well correlated with the endogenous variable but not with the error term in the second stage (Bound *et al.* (1993)). The individual's self-employment status seems to fit the criteria for a good instrument. There is good reason to believe that the individual's self-employment status is correlated with whether or not he or she has insurance. In particular, several previous studies have shown that there is a strong negative correlation between self-employment and medical insurance (Health Insurance Association of American (2000); Perry and Rosen (2001b)).

In this context, though, one should note that the self-employed are not a homogeneous group with respect to the institutional environments in which they function. They operate in different organizational forms – sole proprietorships, partnerships, and corporations – and the probability of being insured could vary with organizational form. In particular, those who are incorporated might be more likely to have insurance for two reasons. First, their expenditures for health insurance are fully deductible; for members of partnerships and sole proprietors, they are not. Second, to the extent that corporate enterprises have more employees, the owners can purchase insurance at advantageous group rates.[3] Indeed, in our data, self-employed individuals who are organized as corporations have an insurance rate of 89.5 percent, as compared with 67.2 percent for sole proprietors and 71.3 percent for partners. Hence, organizational form potentially provides useful information, so instead of characterizing self-employment status by a single dichotomous variable, we use three, one for each of the organizational forms. In our sample 2.6 percent of the individuals have incorporated businesses, 5.7 percent are sole

[3] See Thomasson [2000] on the advantages of group coverage.

proprietors, and 0.97 percent are in partnerships, for a total of 9.27 percent self-employed.[4]

The other key issue in assessing the adequacy of self-employment status as an instrument is whether it exerts an independent effect on either health care utilization or health status. While there is no obvious reason that this should be the case, a possible problem is that there might be unobservable differences between wage earners and the self-employed that affect utilization and health. Perhaps people who are too unhealthy to hold jobs as wage-earners opt for self-employment. Alternatively, some sort of 'animal spirit' may drive both health and the propensity to be self-employed. Previous research using a variety of data sets has addressed this issue, and shown that neither health status nor utilization of health care services is a good predictor of whether or not a wage-earner will make a transition to self-employment (Holtz-Eakin, Rosen, and Penrod (1996) and Perry and Rosen (2001a)). We update these results in Section 6 and confirm that they hold in our data – there is no selection on the basis of health-related variables. While these findings cannot definitely exclude the possibility of unobservable heterogeneity, they certainly provide no support for the notion that people who select into self-employment are systematically different with respect to health-related attributes.

4.2 Instrumental Variables Estimates

As usual, we implement the two-stage estimation procedure by first estimating a reduced form regression of insurance status on all the covariates of the model. Impor-

[4] An interesting question is whether the differences in insurance rates by organizational form translate into differences in the utilization of health care services and health status. To investigate this question, we estimate a series of reduced form regressions, essentially substituting the self-employment variables for insurance status in Equation (1). The results, available upon request, show that incorporated individuals generally do not differ greatly from wage earners in their utilization of medical services, whereas sole proprietors and partners tend to utilize medical services less than wage earners. This is more or less what one

tantly, in this first-stage regression, one can strongly reject the joint hypothesis that the three self-employment variables have zero coefficients. Indeed, the associated F-statistic, 219.8, more than satisfies the usual criteria for a good fit in the first stage (Cawley (2000)). The next step is to estimate the probit equation (1) using the fitted values of the insurance variable from the first stage.[5] Rivers and Vuong (1988) and others have discussed the issues that arise in obtaining consistent standard errors within this framework. The most straightforward solution is to compute bootstrapped standard errors.[6] The bootstrap is based on random sampling; if the data are heteroscedastic, then each sample will have a different distribution, resulting in inconsistent point estimates and standard errors (Efron (1979)). To investigate whether this is a problem in our data, we estimate the bootstrapped model several dozen times to see if the results change substantially. They do not, and we conclude that the data are unlikely to be sufficiently heteroscedastic to render the bootstrap algorithm unreliable.

Table 3 reports the marginal effects for the bootstrapped two-stage probit model.[7] Comparing these results to their counterparts in Table 2, we see that, in general, instrumenting for insurance in the utilization equations increases the magnitude of its effect. For example, the effect of insurance on visiting a doctor increases from 24.9 percentage points to 35.2 percentage points. The change between the two specifications is even more pronounced for the preventative care measures. The insurance effect on the prob-

---

would expect given the pattern of insurance coverage across the organizational forms. Incorporated individuals have slightly higher self-reported general health than sole proprietors, partners, and wage earners.

[5] A Davidson-MacKinnon overidentification test indicates that we cannot reject the null hypothesis of exogeneity of the instruments ($p = 0.173$).

[6] As noted above, we perform a clustered procedure to allow for the possibility of within-individual correlation of the errors. To implement this in a bootstrapping context, we first create a list of individuals. For each bootstrap iteration, we then draw a set of individuals from this list, and use all years of data from the selected individuals to construct that iteration's dataset.

ability of have a cholesterol check, for example, increases by 13 percentage points. It appears, then, that ignoring endogeneity leads to underestimates of the impact of insurance coverage on utilization. As noted earlier, this is consistent with a scenario in which individuals who are likely to be intensive users of health services find themselves discouraged from obtaining insurance for any of a variety of reasons.

Turning to health status, the instrumental variables estimate of the insurance effect in Table 3 is smaller than its uncorrected counterpart in Table 2 – the point estimate falls from 1.8 percentage points (s.e. = 0.36) to -0.28 percentage points (s.e. = 2.52). In short, the finding of a positive and statistically significant effect of insurance on health status in the model that ignores endogeneity is not robust.

4.3 Alternative Specifications

We have shown that when one takes into account the endogeneity of insurance status, the estimate of the insurance effect on the utilization of a variety of health care services increases, while the effect on health status becomes very small and insignificantly different from zero. We estimate a number of variations on our basic model in order to assess the robustness of these results.

4.3.1 Income

A positive correlation between health status and income is well-documented, but the direction of causation is controversial. (See, for example, Feinstein (1993), Smith and Kington (1997), and Smith (1999)). In light of this controversy, we choose not to include income as a right-hand side variable in our basic models. However, given the widespread belief that low income leads to less or lower quality medical care, it seems

[7] Two-stage probit coefficients were calculated using a Stata module developed by Jonah Gelbach of the

worthwhile to re-estimate the model including family income as a covariate. The instrumental variables results are shown in Table 4. None of our substantive findings changes.

As noted in Section 2 above, an intriguing finding of the Rand experiment was that changes in the extent of insurance coverage had some impact on the health status of low-income people, but not for the rest of the population. In our context, this suggests that one might want to estimate a model that includes an interaction between insurance status and income. When we augment our basic models from Table 2 with both income and its interaction with insurance status, we find that the interaction term in the health status equation is both small and insignificant. The two-stage probit estimate of the interaction term is -0.0105 with a standard error of 0.0204 for the regression with office based provider visits on the left hand side. The negative point estimate is consistent with the notion that insurance plays a more important role in determining health status for low-income individuals, although one should note that it is imprecisely estimated. In any case, including the interaction term does not affect our substantive results with respect to insurance and the utilization of various health services.

4.3.2 Marital Status

It is easy to imagine that marital status affects health-related outcomes. Married people may have more stable home environments, better diets, and so on. However, it is equally easy to imagine that the direction of causality runs the other way. For example, healthy people may be more likely to find mates than unhealthy people (Goldman (2001)). Because of the possibility of joint determination, we leave marital status out of our basic model. But marital status has been included in other studies (e.g., Gruber and

---

University of Maryland.

Madrian (2002)), so it seems instructive to estimate a set of models that include a dichotomous marriage indicator. The results, reported in Table 5, are neither qualitatively nor quantitatively very different from those of the basic model.

4.3.3 Econometric Specification

A possible problem with our results is that they are a consequence of the particular assumptions underlying the two-stage probit model. In the probit model, the two-stage procedure generates consistent estimates only if the error terms in both the first- and second-stage equations are jointly normally distributed, and both equations are correctly specified. In a linear probability model, the conditions for consistency are less stringent – the right hand variables in the first-stage equation have to be uncorrelated with the error term in the second-stage equation, but consistent estimates may be obtained even if some variables that belong in the first-stage equation are omitted. Therefore, despite the well-known limitations of the linear probability model, it seems worthwhile to use it to check our estimates.

The results, reported in Table 6, are very similar to those obtained using the probit. For example, in Table 3 the insurance effect on the probability of visiting an office based provider is 35 percentage points; in Table 6 it is 31 percentage points. Thus, we feel confident that our results are not an artifact of the assumptions behind the two-stage probit model.

5. TRANSITIONS INTO SELF-EMPLOYMENT

As suggested above, there might be some unobserved variable that drives both health status and self-employment decisions. One can imagine, for example, that relatively unhealthy people are unable to hold jobs and hence enter self-employment. Alter-

natively, perhaps healthy, energetic people are particularly well-suited for running their own businesses. To the extent that either type of selection takes place, self-employment status is not a suitable instrument.

Examining the determinants of transitions from wage earning into self-employment can be instructive in ascertaining whether some underlying variable drives both health and self-employment status. Consider a group of wage-earners during a given time period. If the probability that an individual in this group transits to self-employment in the subsequent period is independent of his or her health status at the outset, then one can feel some confidence that selection into self-employment on the basis of health is not driving our results. On the other hand, if healthier or less healthy individuals are more likely to make transitions into self-employment, the interpretation of our findings becomes problematic. Using data from the Survey of Program Participation and the Panel Study of Income Dynamics, Holtz-Eakin, Penrod, and Rosen (1996) found no evidence for this latter hypothesis. Similarly, Perry and Rosen's (2001a) analysis of the 1996 and 1997 MEPS data suggested that the absence of health differences between wage earners and the self-employed did not appear to be due to the fact that people with relatively good health tend to select into self-employment.

In this section, we update these analyses and examine whether the results hold for the sample of individuals used in this study. Specifically, the overlapping two-year structure of the MEPS allows us to construct a pooled data set for two sets of transitions, from 1996 to 1997 and 1997 to 1998. There are 19,744 individuals, each with two years of information. Approximately 1 percent of wage earners leave their jobs in each year to become self-employed.

We model the probability that an individual who is a wage-earner in year t-1 makes a transition to self-employment in period t. The sample consists of individuals who are wage-earners in year t-1. The dependent variable is an indicator variable for whether the individual is self-employed (in any organizational form) in year t.[8] The right-hand side includes the covariates used in our canonical model dated year t-1, a time effect to take into account any systematic differences between the two transition periods, and some measure of the individual's health status and/or utilization of health services in year t-1.

Row (1) of Table 7 shows the results when self-reported health status is included on the right hand side; row (2) when an indicator variable for an office visit to a health-care provider is included, and row (3) when both are included. None of the health-related variables is significant in any of the specifications, and all are minuscule in magnitude. For example, the point estimate in row (1) suggests that being in good health makes one only 0.06 percentage points more likely to make a transition into self-employment, and row (2) indicates that having paid a visit to a health-care provider makes one only 0.02 percentage points more likely to make such a transition. We find similar results when we use other measures of health care utilization. In short, Table 7 confirms previous findings that neither health status nor health services utilization are predictors of whether wage-earners will become self-employed. While this does not prove that self-employment status is a good instrument in this context, it is nevertheless comforting to note that it is unlikely that some variable is driving both health and the self-employment decision.

[8] We also examined transitions from wage earning into particular organizational forms. The results were not affected.

6. EXPENDITURES

We have shown that once simultaneity is taken into account, there is no statistically discernible difference in health status between the insured and the uninsured. Our focus on this issue reflects perhaps the dominant issue in the public policy debate over the uninsured – the relationship between health insurance and health. However, the discussion over health care sometimes loses sight of the key function of insurance, namely, to spread consumption over different states of the world. Even if the uninsured are about as healthy as the insured, we cannot necessarily be sanguine about their situation if paying for health care causes serious reductions in their standard of living.

The MEPS includes information about out-of-pocket expenditures on health care. We wish to estimate how having insurance affects these expenditures, taking into account the endogeneity of insurance; as before, we use a set of self-employment variables as instruments. The most natural way to examine how health care expenditures affect the "standard of living" is to measure their magnitude relative to consumption expenditures. However, the MEPS does not include consumption information. Hence, we use income, which is recorded in the survey. We exclude families whose incomes are below $5,000 from this analysis, as we suspect that income is transitorily very low or mismeasured for such families.

To begin, we use ordinary least squares to estimate a regression with out-of-pocket medical expenditures on the left-hand side, and insurance and the same exogenous covariates as in Table 2 on the right-hand side. The coefficient on the insurance variable indicates that it reduces out-of-pocket medical expenditures by a mere $17.34 with a standard error of $12.16. When the model is estimated using two-stage least squares, the

effect increases substantially. Having insurance lowers out-of-pocket medical expenditures by $502.15 (s.e. = $172.52). It may be more instructive, however, to examine expenditures as a proportion of income. When we estimate an ordinary least squares regression of the proportion of income on the same covariates as above, we find that insurance reduces medical expenditures by 0.41 percent of income (s.e. = 0.065 percent). The two-stage least squares estimate is 0.61 percent of income with a standard error of 0.45 percent.

These calculations suggest that out-of-pocket medical expenditures are, on average, a manageable burden for the uninsured.[9] Nevertheless, they may be excessive for some of the uninsured. To investigate this possibility, we create an indicator that equals one if out-of-pocket expenditures as a proportion of income are greater than 20 percent and zero otherwise. Conventional probit estimation indicates that having insurance reduces the probability of reaching this threshold by 0.53 percentage points (s.e. = 0.13 percentage points). Instrumenting in a two-stage probit framework, we find that insurance reduces this probability by 2.3 percentage points, but with a standard error of 3.5 percentage points. One cannot reject the hypothesis that the effect is equal to zero. On the other hand, given that only 0.65 percent of the sample experiences out-of-pocket expenditures greater than 20 percent of their incomes, the large standard error encompasses some potentially important changes in the probability of a serious diminution in a household's standard of living. While this exercise provides no strong evidence that such a

[9] An interesting question is whether this finding holds for individuals with particularly severe health problems. The MEPS indicates whether or not the individual has a "priority condition," which is one of a number of serious illnesses such as cancer and heart disease. When we interact a dichotomous variable for the presence of a priority condition with the insurance variable, the IV estimate is -0.40 percentage points. That is, for a person with a priority condition, not having insurance increases out-of-pocket expenditures relative to income by 0.40 percentage points. However, one cannot reject the hypothesis that the interac-

phenomenon would occur, it is not inconsistent with a risk-aversion motive for purchasing health insurance.

Taken together, the results in this section do not provide support for the notion that lack of insurance greatly increases out-of-pocket expenditures relative to income. This is not too surprising given our findings from Table 3, which indicate that the uninsured are less likely than the insured to consume a variety of health care services. The expenditure results are perhaps more interesting against the backdrop of our previous finding that the uninsured have about the same health status as the insured. It seems that the uninsured are able to maintain their health without having to bear onerous out-of-pocket medical expenditures.

7. CONCLUSION

Using the Medical Expenditure Panel Survey, we have examined the effect of health insurance on the utilization of health care services and health status. We began by replicating previous studies which show that insured individuals are more likely to utilize a variety of health care services and to be in better health than those who do not, *ceteris paribus*. However, as several previous investigators have pointed out, insurance status is likely to be an endogenous variable in this context, so that attaching a causal interpretation to this statistical relationship is problematic. An instrumental variables strategy is a natural way to address this problem, and we argue that self-employment status is a suitable instrument.

We find that the instrumental variables estimates of the impact of insurance on utilization of a variety of health care services are larger than their non-instrumented counterparts. At the same time, the impact of insurance on health is both diminished and ren-

---

tion term is zero. The main effect of having a priority condition is 1.1 percentage points (s.d. = 0.78 percentage points).

dered statistically insignificant. This latter finding is consistent with a growing literature which argues that having health insurance does not necessarily improve health status. (See Joyce, Jaestner and Racine (1999), Ross and Mirowsky (2000), and Perry and Rosen (2001a), *inter alia*). The reason might be that access to health care is responsible for only a relatively small part of health, with more important determinants being genetics, environment, and health behaviors (Institute for the Future (2000, p. 23)). However, some caution is required here. Our findings pertain to short-term effects of insurance on health status. One can imagine that, after a number of years, the lower utilization of health care services associated with the lack of insurance could cumulatively have a negative impact on health. This observation is particularly cogent in light of our finding that insurance coverage has a substantial impact on the utilization of a variety of preventative care procedures. An important topic for future research is exploration of the long-term relationships among insurance, health care utilization, and health status.

REFERENCES

Bound, John, David A. Jaeger and Regina M. Baker. "The Cure Can Be Worse than the Disease: A Cautionary Tale Regarding Instrumental Variables." *Journal of the American Statistical Association*. Vol. 90, Issue 430, June 1995. 443-450.

Brown M., A. Bindman, and N. Lurie. "Monitoring the Consequences of Uninsurance: A Review of Methodologies." *Medical Care Research and Review*. Vol 55, No. 2. 1998. 177-210.

Cawley, John. "Body Weight and Women's Labor Market Outcomes." NBER Working Paper No. 7841. August 2000.

Currie, Janet and Jonathan Gruber. "Health Insurance Eligibility, Utilization of Medical Care and Child Health." *Quarterly Journal of Economics*. May 1996. 431-466.

Efron, Bradley. "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics*. Vol. 7, No. 1. January 1979. 1-26.

Feinstein, Jonathan. "The Relationship Between SES and Health: A Review of the Literature." *Milbank Quarterly*. Vol 71, No. 2 1993. 279-322

Gelbach, Jonah. "Probitiv: Stata module to perform instrumental variables probit." http://www.glue.umd.edu/~gelbach/ado/ .

Goldman, Noreen. "Social Inequalities in Health: Untangling the Underlying Mechanisms." Working Paper, Princeton University. February 2001.

Gruber, Jonathan. "Medicaid." NBER Working Paper 7829. August 2000.

Gruber, Jonathan and Brigitte Madrian. "Health Insurance, Labor Supply, and Job Mobility: A Critical Review of the Literature." NBER Working Paper 8817. February 2002.

Gruber, Jonathan and James M. Poterba. "Tax Incentives and the Decision to Purchase Health Insurance: Evidence from the Self-Employed." *Quarterly Journal of Economics*. August 1994. 701-733.

Haas, Jennifer, Steven Udvarhelyi and Arnold Epstein. "The Effect of Providing Health Coverage to Poor Uninsured Pregnant Women in Massachusetts." *Journal of the American Medical Association*. 1993a. 269: 87-91.

Haas, Jennifer, Steven Udvarhelyi and Arnold Epstein. "The Effect of Health Coverage for Uninsured Pregnant Women on Maternal Health and the Use of Caesarian Section." *Journal of the American Medical Association*. 1993b. 270: 61-64.

Harvey, P. Holly and Alan C. Monheit. "Sources of Health Insurance for the Self-Employed: Does Differential Taxation Make a Difference?" *Inquiry*. Fall 1993. 293-305.

Hausman, J.A. "Specification Tests in Econometrics." *Econometrica*. Vol. 46, No. 6 November 1978. 1251-1271.

Health Insurance Association of America, *Source Book of Health Insurance Data, 1999-2000,* Washington DC, 2000.

Holtz-Eakin, Douglas, John Penrod, and Harvey S. Rosen. "Health Insurance and the Supply of Entrepreneurs." *Journal of Public Economics*. October 1996. 209-235.

Hurd, Michael and K. McGarry, "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Survey," *Journal of Human Resources* 30, 1997. S268-292.

Idler, Ellen and Yael Benyamini. "Self-Related Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior*. March 1997. 21-37.

Institute for the Future. *Health and Health Care 2010 - The Forecast, The Challenge*. Jossey-Bass Publishers: San Francisco, 2000.

Joyce, Theodore, Robert Kaestner, and Andrew Racine. "Does Publicly Provided Health Insurance Improve the Health of Low-Income Children in the United States?" NBER Working Paper No. 6887. January 1999.

Kaestner, R., T. Joyce and A. Racine, "Medicaid Eligibility and the Incidence of Ambulatory Care Sensitive Hospitalizations for Children," *Social Science and Medicine* 52, 2001. 305-313.

Kass, B.L., R.M. Weinick, and A.C. Monheit. "Racial and Ethnic Differences in Health, 1996." Agency for Health Care Policy and Research. 1999. MEPS Chartbook No. 2. AHCPR Pub No. 99-0001.

Lakdawalla, Darius and Tomas Philipson. "The Rise in Old Age Longevity and the Market for Long-Term Care." NBER Working Paper 6547. May 1998.

Levy, Helen and David Meltzer. "What do We Really Know About Whether Health Insurance Affects Health?" Working Paper, University of Chicago. June 2001.

Lurie, N., *et al.* "Termination from Medi-Cal: Does it Affect Health?" *New England Journal of Medicine*. 16 August 1984. 480-484.

Newey, Whitney. "Efficient Instrumental Variables Estimation of Nonlinear Models." *Econometrica*. Vol 58, No 4 July 1990. 809-837.

Newhouse, Joseph P. *Free For All? Lessons from the RAND Health Insurance Experiment*. Harvard University Press: 1993.

Perry, Craig W. and Harvey S. Rosen. "The Self Employed are Less likely to Have Insurance than Wage Earners. So What?" NBER Working Paper 8316. June 2001a.

Perry, Craig W. and Harvey S. Rosen. "Insurance and the Utilization of Medical Services Among the Self Employed." NBER Working Paper 8490. August 2001b.

Rivers, Douglas and Quang H. Vuong. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics*. 39 (1988): 347-366.

Skinner, Jonathan, Elliot S. Fisher and John E. Wennberg. "The Efficiency of Medicare." NBER Working Paper 8395. July 2001.

Smith, James P. "Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status." *Journal of Economic Perspectives*. Vol. 13 No. 2 Spring 1999. 145-166.

Smith, James P. and Raynard S. Kington. "Race, Socioeconomic Status, and Health in Late Life." *Racial and Ethnic Differences in the Health of Older Americans*, Linda G. Martin and Beth J. Soldo, eds. National Academy Press: 1997.

Staiger, Douglas and James H. Stock. "Instrumental Variables Regression with Weak Instruments." *Econometrica*. Vol 65, No 3. May 1997. 556-586.

Taubman, Paul and Sherwin Rosen. "Healthiness, Education, and Marital Status." NBER Working Paper 611. August 1982.

Thomasson, Melissa A., "The Importance of Group Coverage: How Tax Policy Shaped U.S. Health Insurance," NBER Working Paper 8316, 2000.

Verbrugge, Lois M., "Gender and Health: An Update on Hypothesis and Evidence." *Journal of Health and Social Behavior*. Vol. 26, No. 3. September 1985. 156-182.

TABLES

Table 1
Summary Statistics

Variable	Description	Mean	Standard Deviation
Office-Based Provider	Respondent had an office-based provider visit in the last year.	0.6517	0.4764
Chiropractor	Respondent visited a chiropractor in the last year.	0.03662	0.1878
Prescription	Respondent received a prescription for drugs in the last year.	0.6023	0.4894
Alternate Care	Respondent used some form of alternate care (e.g. massage therapy or acupuncture) in the last year.	0.05425	0.2265
Night in Hospital	Respondent spent at least one night in a hospital in the last year.	0.05920	0.2360
Outpatient Hospital Stay	Respondent had an outpatient visit to a hospital in the last year.	0.004768	0.06889
Dentist	Respondent visited a dentist in the last year.	0.3728	0.4836
Optometrist	Respondent visited an optometrist in the last year.	0.04642	0.2104
Blood Pressure	Respondent had his or her blood pressure checked in the last year.	0.7227	0.4477
Cholesterol Check	Respondent had his or her cholesterol level checked in the last year.	0.3938	0.4886
Physical	Respondent had a physical in the last year.	0.4282	0.4948
Flu Shot	Respondent had a flu shot in the last year.	0.1792	0.3835
Prostate	Male respondent had a prostate exam in the last year.	0.2180	0.4129
Breast	Female respondent had a breast exam in the last year.	0.6354	0.4813
Mammogram	Female respondent had a mammogram in the last year.	0.5166	0.4998
Pap Smear	Female respondent had a pap smear in the last year.	0.6221	0.4849
Health	Self-reported health: 1 if healthy, 0 otherwise.	0.9193	0.2724
Midwest	Respondent lives in the Midwest.	0.2227	0.4161
South	Respondent lives in the south.	0.3535	0.4781
West	Respondent lives in the west.	0.2405	0.4274
Family Size	Respondent's annualized family size.	3.212	1.603

Male	Respondent is male.	0.5156	0.4998
Black	Respondent is black.	0.1318	0.3383
Other Race	Respondent is neither white nor black.	0.04682	0.2113
GED	Respondent has a GED.	0.04762	0.2129
High School Degree	Respondent has a high school degree.	0.5343	0.4988
College Degree	Respondent has a college degree.	0.1591	0.3657
Graduate Degree	Respondent has a graduate degree.	0.06982	0.2548
Other Degree	Respondent has some other degree.	0.03807	0.1914
Age	Respondent's age.	38.15	11.68
Log Family Income	Log of respondent's annual family income.	10.67	0.8343
Corporate	Respondent is self-employed and incorporated all year.	0.02604	0.1592
Sole Proprietor	Respondent is self-employed and a sole proprietor all year.	0.05692	0.2317
Partner	Respondent is self-employed and in a partnership all year.	0.009697	0.09799
Insurance Status	Respondent is insured all year.	0.8244	0.3805

These summary statistics are calculated from the MEPS for 1996 to 1998. Only individuals between the ages of 18 and 62 in 1996 are included.

Table 2
Probit Estimates of Health Services Utilization and Health Status (Marginal Effects)

Dependent Variable	Sample Size	Insurance Effect
Office-Based Provider	37583	*0.2491* *(0.007663)*
Chiropractor	37583	0.002425 (0.002743)
Prescription	27249	*0.2122* *(0.009059)*
Alternate Care	22497	-0.003548 (0.003989)
Night in Hospital	37583	*0.03286* *(0.002498)*
Outpatient Hospital Stay	37583	*0.003739* *(0.00002210)*
Dentist	37583	*0.2021* *(0.006509)*
Optometrist	37583	*0.02398* *(0.002326)*
Blood Pressure	22336	*0.2004* *(0.009038)*
Cholesterol Check	21977	*0.1746* *(0.008538)*
Physical	22269	*0.1795* *(0.008604)*
Flu Shot	21929	*0.07298* *(0.006265)*
Prostate Exam	10117	*0.1294* *(0.008630)*
Breast Exam	11073	*0.2327* *(0.01375)*
Mammogram	5096	*0.2575* *(0.02074)*
Pap Smear	11081	*0.2289* *(0.01367)*
Health	92972	*0.01773* *(0.003577)*

Figures show the effect of a discrete change from 0 to 1 in insurance coverage for the whole year. Each coefficient is from a probit equation in which the left-hand side variable is the variable in the associated row, and the other right-hand side variables include: region, family size, race, sex, education, age, age squared, and year effects. Standard errors are in parentheses. Those estimates significant at the 5 percent level are italicized.

Table 3
Two-Stage Probit Estimates of Health Services Utilization and Health Status (Marginal Effects)

Dependent Variable	**Sample Size**	**Insurance Estimate**
Office-Based Provider	37331	*0.3517* *(0.06453)*
Chiropractor	37331	*-0.1591* *(0.06011)*
Prescription	27077	*0.3541* *(0.07545)*
Alternate Care	22340	*-0.3198* *(0.08768)*
Night in Hospital	37331	*0.05369* *(0.01399)*
Outpatient Hospital Stay	37331	0.0066395 (0.01812)
Dentist	37331	*0.2076* *(0.04822)*
Optometrist	37331	*0.03790* *(0.01224)*
Blood Pressure	22177	*0.5194* *(0.07316)*
Cholesterol Check	21818	*0.3989* *(0.04097)*
Physical	22110	*0.3769* *(0.05094)*
Flu Shot	21776	*0.2270* *(0.02065)*
Prostate Exam	10062	*0.2301* *(0.03364)*
Breast Exam	10977	*0.4383* *(0.1307)*
Mammogram	5060	*0.3301* *(0.1373)*
Pap Smear	10985	*0.3787* *(0.1372)*
Health	92972	-0.002763 (0.02515)

Figures show the effect of a discrete change from 0 to 1 in insurance coverage for the whole year. Each coefficient is from a probit equation in which the left-hand side variable is the variable in the associated row, and the other right-hand side variables include: region, family size, race, sex, education, age, age squared, and year effects. The instruments for insurance are indicators for operating as a corporation, sole proprietor, or partner. The standard errors, shown in parentheses, are bootstrapped. Those estimates significant at the 5 percent level are italicized.

Table 4
Two-Stage Probit Estimates of Health Services Utilization and Health Status: With Income (Marginal Effects)

Dependent Variable	Sample Size	Insurance Estimate
Office-Based Provider	37203	*0.3549* *(0.0612)*
Chiropractor	37203	*-0.1581* *(0.06155)*
Prescription	26970	*0.3520* *(0.06835)*
Alternate Care	22226	*-0.3135* *(0.08069)*
Night in Hospital	37203	*0.05456* *(0.01380)*
Outpatient Hospital Stay	37203	0.006293 (0.003848)
Dentist	37203	*0.2143* *(0.04933)*
Optometrist	37203	*0.038631* *(0.01187)*
Blood Pressure	22066	*0.5065* *(0.06977)*
Cholesterol Check	21710	*0.3876* *(0.04028)*
Physical	22000	*0.3694* *(0.04871)*
Flu Shot	21666	*0.2236* *(0.02078)*
Prostate Exam	10017	*0.2215* *(0.03412)*
Breast Exam	10911	*0.44009* *(0.1205)*
Mammogram	5021	*0.33443* *(0.1252)*
Pap Smear	10920	*0.3889* *(0.1298)*
Health	92822	-0.006701 (0.02479)

Figures show the effect of a discrete change from 0 to 1 in insurance coverage for the whole year. The standard errors, which are in parentheses, are bootstrapped. Each coefficient is from a probit equation in which the left-hand side variable is the variable in the associated row, and the other right-hand side variables include: log of income, region, family size, race, sex, education, age, age squared, and year effects. The instruments for insurance are indicators for operating as a corporation, sole proprietor, or partner. The standard errors, which are in parentheses, are bootstrapped. Those estimates significant at the 5 percent level are italicized.

Table 5
Two-Stage Probit Estimates of Health Services Utilization and Health Status: With Marital Status
(Marginal Effects)

Dependent Variable	**Sample Size**	**Insurance Estimate**
Office-Based Provider	37325	*0.3618* *(0.06269)*
Chiropractor	37325	*-0.1532* *(0.03402)*
Prescription	27072	*0.3662* *(0.06913)*
Alternate Care	22336	*-0.3169* *(0.04628)*
Night in Hospital	37325	*0.05545* *(0.01324)*
Outpatient Hospital Stay	37325	0.006596 (0.003836)
Dentist	37325	*0.2141* *(0.05247)*
Optometrist	37325	*0.03809* *(0.01218)*
Blood Pressure	22174	*0.5226* *(0.06933)*
Cholesterol Check	21815	*0.3958* *(0.04044)*
Physical	22107	*0.3780* *(0.05244)*
Flu Shot	21773	*0.2260* *(0.02135)*
Prostate Exam	10059	*0.2286* *(0.03099)*
Breast Exam	10976	*0.4623* *(0.1134)*
Mammogram	5059	*0.3374* *(0.1330)*
Pap Smear	10984	*0.4116* *(0.1235)*
Health	92961	-0.0008350 *(*0.02538*)*

Figures are the effect of a discrete change from 0 to 1 in insurance coverage for the whole year. Each coefficient is from a probit equation in which the left-hand side variable is the variable in the associated row, and the other right-hand side variables include: marital status, region, family size, race, sex, education, age, age squared, and year effects. The instruments for insurance are indicators for operating as a corporation, sole proprietor, or partner. The standard errors, which are in parentheses, are bootstrapped. Those estimates significant at the 5 percent level are italicized.

Table 6
Two-Stage Least Squares Estimates of Health Services Utilization and Health Status (Marginal Effects)

Dependent Variable	Sample Size	Insurance Estimate
Office-Based Provider	37331	*0.3143* *(0.05819)*
Chiropractor	37331	*-0.1245* *(0.03488)*
Prescription	27077	*0.3249* *(0.07310)*
Alternate Care	22340	*-0.2214* *(0.04727)*
Night in Hospital	37331	*0.06742* *(0.02489)*
Outpatient Hospital Stay	37331	0.01106 *(*0.006441*)*
Dentist	37331	*0.2001* *(0.05932)*
Optometrist	37331	*0.05091* *(0.02329)*
Blood Pressure	22177	*0.4577* *(0.07549)*
Cholesterol Check	21818	*0.4827* *(0.07672)*
Physical	22110	*0.4140* *(0.07748)*
Flu Shot	21776	*0.3744* *(0.06089)*
Prostate Exam	10062	*0.3363* *(0.07480)*
Breast Exam	10977	*0.4308* *(0.1432)*
Mammogram	5060	*0.3175* *(0.1527)*
Pap Smear	10985	*0.3747* *(0.1456)*
Health	92972	-0.005460 *(*0.02924*)*

Figures are the effect of a discrete change from 0 to 1 in insurance coverage for the whole year. Each coefficient is from a linear probability model in which the left-hand side variable is the variable in the associated row, and the other right-hand side variables include: region, family size, race, sex, education, age, age squared, and year effects. The instruments for insurance are indicators for operating as a corporation, sole proprietor, or partner. Standard errors are in parentheses. Those estimates significant at the 5 percent level for the bootstrapped distribution are italicized.

Table 7
Probit Estimates of Transitions into Self-Employment (Marginal Effects)

Specification	Sample Size	Health Status	Office Based Provider Visit
1	12974	0.0005902 (0.002939)	-
2	12974	-	0.0001546 (0.001818)
3	12974	0.0006140 (0.002971)	0.0001851 (0.001841)

Each coefficient is from a probit equation in which the left-hand side variable is the probability of being self-employed in year t, and on the right-hand side are the following variables dated period t-1: region, family size, race, sex, education, age, and age squared. The sample consists only of those individuals who were wage-earners in period t-1. Hence, each coefficient shows the probability of making a transition from wage-earning to self-employment, *ceteris paribus*. Coefficients are the effect of a discrete change from 0 to 1 for the indicator variables for initial health status and for whether the individual visited a health services provider. Those coefficients significant at the 5% level are italicized.