METU Studies in Development, 45 (August), 2018, 209-238

The effect of medical care on health capital

Şevket Alper Koç
Kocaeli University, Department of Economics, İzmit, Kocaeli,
sevketkoc@hotmail.com

Çağatay Koç
Cornerstone Research: Economic and Financial Consulting and Testimony, Washington D.C., USA
ckoc@cornerstone.com

Abstract

This paper analyzes the effect of medical care on the stock of health capital by estimating the health investment production function. An ordered probit model for the stock of health with instrumental variables is estimated using the Two-Stage Residual Inclusion method. We argue that risk tolerance and the opportunity cost of time are suitable instruments for the change in medical care consumption. In contrast to majority of the empirical work, which does not uncover that medical care has a positive effect on the stock of health, the results suggest that physician visits significantly *increases* the probability of excellent health (or *decreases* the probability of poor health), in accordance with Grossman's (1972a) demand for health capital model.

Key words: Medical care, probit model, risk tolerance, opportunity cost of time.

JEL classification: G22, I11.

1. Introduction

Grossman's (1972a) demand for health capital model suggests that medical care consumption is positively associated with the stock of health capital; nevertheless, majority of the empirical work does not uncover this conclusion [Grossman (2000)]. Estimating the effect of medical care on health stock is challenging due to the possible endogeneity in medical care consumption. There are differences in the distribution of health across individuals and some of these health characteristics may be known to individuals but not to the researcher [Rosenzweig and Schultz (1983)]. These unobservables may affect both the stock of health and medical care utilization, so that the observed relationship between health stock and medical care would be spurious. In fact, Grossman (1972b) demonstrates that the estimate of the effect of medical care on the stock of health would be *biased downward* if the endogeneity in medical care consumption is left uncontrolled. The health capital model suggests that the rate of health depreciation in the initial period is unobserved and that medical care utilization is positively correlated with it. But since an increase in the initial rate of depreciation lowers the amount of health stock, the estimated coefficient of the effect of medical care on the stock of health would be biased downward. Consider, for instance, individuals with lower genetic endowments or poor nutrition during early childhood. Due to such characteristics, these individuals may demand medical services and still be endowed with low levels of health capital. If these characteristics are unobserved, the estimated effect of medical care on the stock of health would likely be understated.

There is a substantial amount of prior literature that analyzes the association between medical care consumption and the stock of health. One strand of the literature estimates the conditional demand function for physician visits to assess this association [e.g., Wagstaff (1986), Erbsland *et al*. (1995)] and suggests that health status and medical care utilization are negatively related. The latent variable health, however, is treated as exogenous in these papers.[1] Another strand of the literature estimates the effect of medical care on the stock of health by estimating the health investment production function [e.g., Grossman (1972b), Kemna (1987), Strattman (1999), Hu and Wolfe (2002)]. Kemna (1987) treats medical care as exogenous and finds that medical care utilization and health status are negatively associated. Grossman (1972b), using size-adjusted family income as an instrument, shows that the sign of the association between medical care and the stock of health can be correctly reversed if medical care is treated as endogenous. Strattman (1999) and Hu and Wolfe (2002) use health insurance indicators as instruments and find that physician visits has a positive effect on measures of health stock.[2] The instrumental variables results are dependent on the assumption that the instruments can be excluded from the main regression and there may be some problems with these instruments regarding this issue. Family income serves as a proxy for missing inputs in the health production function; therefore, it may not be a valid instrument.[3] Health insurance variables may also not provide exogenous variation in medical care as they are very likely to be related to health status.[4] In this paper we propose

[1] See Grossman (2000) for a detailed discussion of this strand of the literature.

[2] Hu and Wolfe (2002) also use the opportunity cost of time as a further instrument.

[3] Due to this issue, Grossman (1972b) asks the readers to interpret the results with "extreme caution."

[4] See Bhattacharya *et al*. (2003), Deb *et al*. (2006) for evidence to the endogeneity of health insurance in the demand for medical care.

to deal with this endogeneity by using *risk tolerance* and the *opportunity cost of time* as instrumental variables for variations in medical care consumption. Both of these variables are correlated with the change in an individual's medical care utilization, but are plausibly unrelated to changes in his stock of health capital.

Employing risk tolerance as an instrument for the change in medical care consumption is motivated by the presence of uncertainty in the health care industry. It arises because the consumer is uncertain as to what his level of health stock is due to randomness in the exogenous determinants of health [Arrow (1963)]. The current literature analyzing the role of uncertainty in the demand for medical care suggests that individuals increase their medical care consumption because of the potential of having a shock in their health stock and such behavior is increasing in the degree of risk aversion [Dardanoni and Wagstaff (1990) and Picone *et al.* (1998)]. The degree of risk aversion, therefore, would serve as a suitable instrument since it would only indirectly affect the stock of health through investments in medical care. Unfortunately, the Medical Expenditure Panel Survey, the data base used in this paper, does not provide a measure for it. Cutler, Finkelstein and McGarry (2008), however, find that more risk averse individuals exhibit lower tolerance for risk. Thus, this finding, coupled with the findings of the literature analyzing the effects of uncertainty in medical care demand, suggest that an individual who has less tolerance for risk is likely to consume more medical care. We represent risk tolerance by a variable that indicates the *use of seat belts* [Cutler *et al*. (2008)] and use it as an instrument to identify the effect of medical care on the stock of health. The second instrument used is the opportunity cost of time, which affects access to medical care [Ruhm (2000, 2005), Hu and Wolfe (2002), Chou *et al*. (2004)]. To represent this factor, we create a variable that indicates whether the individual's employer *paid sick leave*. If an individual has paid sick leave, he is expected to use more medical care due to decreased time price of medical care, since sick leave insures the time costs of medical care.[5]

A continuous variable for the stock of health is not observed in the data. Instead, we observe an ordered response for health stock. To simultaneously deal with the ordinal ranking for health stock and the endogeneity in medical care consumption, we estimate an ordered probit model with instrumental variables using the Two-Stage Residual Inclusion method, which allows for consistent estimation of nonlinear models in the presence of endogeneity. The endogeneity-uncorrected ordered probit model suggests that physician visits decreases the probability of excellent health (or increases the probability of poor health). The instrumental variables (IV) ordered probit model, on the other hand, suggests that

[5] In the absence of individual-specific medical care prices, paid sick leave indicates a measure of price or opportunity cost to the patient.

physician visits increases the probability of excellent health (or decreases the probability of poor health). The IV results indicate that for an average individual an additional physician visit increases the predicted probability of being in excellent health by 5.3 percent, while it decreases the predicted probability of being in poor health by 10 percent.

We proceed as follows. Section 2 delineates the empirical model. Section 3 provides a description of the data containing the working definition of the stock of health capital. Section 4 starts with a discussion on the specification test results dealing with the relevance of instruments. Next, this section presents results regarding the effect of physician visits on the stock of health. The section then presents a series of tests dealing with the validity of instruments and assessing the stability of main estimates. Section 5 contains robustness analyses assessing the effect of excluding health insurance variables in explaining the variation in the first-stage physician visits regression. Section 6 concludes.

2. The data

We analyze a sample of adults between the ages of 18 and 64, drawn from the Household Component of the 2002 Medical Expenditure Panel Survey (MEPS) and its Medical Conditions file. MEPS is co-sponsored by the Agency for Health Care Policy and Research, and National Center for Health Statistics. It is a nationally representative survey of the U.S. population that provides data on demographics, health status, medical conditions, medical care utilization, health insurance coverage, income, and employment. Family income and family size variables are constructed before the deletion of individuals younger than 18 and older than 64.[6] Observations containing veterans and individuals who are covered by Tricare insurance are removed from the data set since their medical care demand and access to medical care distinctly differs from the general population. Observations that are being designated as non-key and out-of-scope are also removed from the data set.[7] This leaves a sample of 16,583 individuals.

---

[6] In this paper, a family is defined as a *health insurance eligibility unit*, which includes adults plus those family members who would typically be eligible for coverage under the adults' private health insurance family plans. Health insurance eligibility units include adults, their spouses and their unmarried natural/adoptive children age 18 and under. Children under 24 who are full time students are also included.

[7] An individual is considered as *inscope* during a round of interview if he is a member of the U.S. civilian, non-institutionalized population during that round. An individual is *key* if he is linked to the set of National Health Interview Survey sampled households designated for inclusion in MEPS. Only individuals who are inscope, key and responded for the full period in which they are inscope are assigned positive personal weights by MEPS.

The dependent variable is the stock of health capital. Five categories of explanatory variables are included in the estimation of the health production function: environmental factors, education, factors affecting health depreciation, the presence of medical conditions, and medical care consumption. The definitions of all regressors along with the definition of the dependent variable are reported in Table 1.

The stock of health capital is measured using the self-reported health index. This measure is based on the answer to the following question: "In general, would you say your health is excellent, very good, good, fair or poor?" The answer to this question is coded on a 1-5 scale, with 1 being excellent, 2 as very good, 3 as good, 4 as fair, and 5 as poor, which provides an ordinal ranking of the perceived stock of health.[8] Self-reported health index is a very good indicator of an individual's overall stock of health. As Johnson and Wolinsky (1993) argue, as diseases are detected, there is a natural progression from body to mind. The effect of a disease, through physical disability, moves into activity limitations and ultimately leads to the relative perception of health and illness. Self-reported health index may also reflect the transitional status of acute medical conditions unrelated to the more stable influences of disease and disability [Johnson and Wolinsky (1993)]. Furthermore, it has been shown to be an important predictor of subsequent mortality [Idler and Benyamini (1997)] and of subsequent medical services utilization [Connely *et al*. (1989), Wolinsky and Johnson (1991), van Doorslaer *et al*. (2000, 2004)].

Following Buckley *et al*. (2004), environmental factors are controlled for by using a dummy variable that indicates whether the individual lives in an urban area and by four regional dummy variables that indicate the census region the individual resides in. A dummy variable is included to control for the education of the individual that takes on the value 1 if the individual is at least a high school graduate. Age, gender, race, marital status, employment status, family income and family size are included as variables affecting the rate of health depreciation. Gender is represented by a dummy variable that indicates whether the individual is male and race is represented by a dummy variable that takes on the value 1 for white individuals.

Medical conditions are identified by disease, disability, acute conditions and obesity. Disease is identified by the existence of chronic or life-threatening medical conditions.[9] We recognize the presence of disability based on functional limitation

---

[8] We use this ordinal ranking instead of a dichotomization approach (such as 1 if fair or poor health), since not all health variation contained in the self-reported health index is used in the dichotomous variable.

[9] Disease indicates whether the individual has one of the following priority conditions: long-term life threatening conditions such as cancer, diabetes, emphysema, high cholesterol, HIV/AIDS, hypertension, ischemic heart disease, and stroke; chronic manageable conditions such as arthritis, asthma, gall bladder

status and use a variable that indicates whether the individual has a functional limitation at any time during the year.[10] Disease and disability are two important dimensions of health and are expected to affect the individual's overall stock of health [Johnson and Wolinsky (1993)]. We also include a dummy variable that identifies the existence of acute conditions of any sort and minor chronic conditions (ACUTE). Such conditions, unrelated to the more stable influences of disease and disability, may affect an individual's stock of health in the short run [Johnson and Wolinsky (1993)].[11] Obesity is included as a further medical condition measure, since it is considered a risk factor for several diseases. It is identified using the body mass index (defined as the ratio of weight in kilograms to height squared in meters) and is a dummy variable that takes the value one if the body mass index is greater than or equal to 30 (OBESE) [Ruhm (2005)].

The number of office based visits to a medical physician for the full year defines the medical care utilization variable. These include only consultations with a medical doctor and exclude non-physician visits such as chiropractors, nurse and nurse practitioners, optometrists, physician's assistants, psychologists, and physical or occupational therapists.

3. Empirical model

Grossman's (1972a) health capital model suggests that consumers derive utility from a stock of health capital which depreciates over time. The stock can be augmented by a household production of gross health investment using time and medical care as inputs. Grossman (1972b) shows that the gross health investment production function can be estimated by replacing gross health investment with the stock of health and derives that it is a function of medical care consumption, education, factors affecting rate of depreciation (e.g., age, environmental factors,

disease, stomach ulcers, and back problem of any kind; and Alzheimer's disease or other dementias, depression and anxiety disorders.

[10]This is a combined measure that indicates whether the individual has one or more of the following limitations:

Needs help with instrumental activities of daily living such as doing laundry or taking medications.

Needs help with activities of daily living such as bathing or dressing.

Limitations in work, housework or school.

Cognitive limitations such as confusion or memory loss

Sensory limitations such as visual or hearing impairments.

[11] Using a clinical classification code based on disease classification codes (ICD-9-CM condition codes), MEPS classifies 260 mutually exclusive medical conditions. Some of these conditions, identified by the disease variable (see footnote 16 for a list of these conditions), are called *priority* conditions, which are designated as such due to their prevalence, expense, or relevance to policy. Acute indicates whether the individual has any one of the 260 conditions but the priority conditions.

etc.) and the depreciation rate in the initial period. The health capital model asserts that medical care consumption is positively associated with the stock of health. Higher levels of education may increase the efficiency with which individuals are able to produce health by increasing health knowledge and by improving the ability to process health information [Grossman's (1972a) *productive efficiency* hypothesis]. Environmental factors may not only directly impact the stock of health but may also have an indirect effect by affecting the productivity of medical care.[12] Following Hu and Wolfe (2002), we assume that the stock of health also depends on a set of medical conditions, such as chronic and acute conditions, disabilities, etc. These conditions are important dimensions of health and are expected to affect the individual's overall stock of health capital [Johnson and Wolinsky (1993)].

To represent the health production function, we assume a linear relationship between the continuous latent health variable H^* and the independent variables given by

$$H^* = \alpha + \beta_1 \log(DOC) + \beta_2 COLLEGE + \beta_3 X_1 + \beta_4 X_2 + \varepsilon \tag{1}$$

where X_1 is a vector of variables including environmental factors and other controls related to health depreciation, X_2 is a set of variables representing medical conditions, COLLEGE indicates whether the individual is at least a high school graduate, and ε is an error term which is assumed to be normally distributed, independent and homoscedastic. Following Kemna (1987) and Lu and McGuire (2002) we use the natural logarithm of the number of physician visits (DOC) to reduce skewness in this variable and allow for diminishing marginal productivity. In fact, Grossman's model suggests that the stock of health is related to medical care consumption through a health production function which is increasing and concave in medical care. See, for example, Dardanoni and Wagstaff (1987), Selden (1993) and Chang (1996).

The latent outcome H^* is not observed in our data set. Instead, we observe an ordered self-reported response with five categories (1=excellent health, 2=very good health, 3=good health, 4=fair health, 5=poor health).[13] In other words, we observe an indicator of the category in which the latent outcome falls. This can be expressed as

---

[12] For example, exposure to air-pollutants may exacerbate lung and cardiovascular diseases and thus medical care for these diseases may be less effective than otherwise it would.

[13] Grossman (1972b) also uses such an ordered self-reported health index as a proxy for the stock of health capital. The next section provides a detailed explanation for why the self-reported health index is a good proxy for an individual's stock of health.

$$H = j \quad \text{if} \quad \xi_{j-1} < H^* < \xi_j, \quad j = 1,...,5$$

with $\xi_0 = -\infty$, $\xi_j \le \xi_{j+1}$, and $\xi_5 = \infty$. An ordered probit specification is used to model the self-reported health index. Thus, using the probabilities of observing the particular self-reported health index category for each individual, the parameters in (1) can be estimated employing the Maximum Likelihood method.[14]

The above ordered probit model may not yield unbiased estimates of the effect of physician visits on health capital. As discussed above, there might be differences in the distribution of health across individuals that may affect both health stock and medical care consumption, so that the estimated effect of medical care on the stock of health would be biased downward. To deal with this problem, we estimate the ordered probit model with instrumental variables using the Two-Stage Residual Inclusion (2SRI) method, which allows for consistent estimation of nonlinear models in the presence of endogeneity [Terza *et al*. (2008)]. In the first stage of the 2SRI method, we estimate the determinants of the endogenous variable $\log(DOC)$ using the linear regression model

$$\log(DOC) = \theta + \gamma_1 COLLEGE + \gamma_2 X_1 + \gamma_3 X_2 + \gamma_4 Z + \tau \tag{2}$$

where Z is a vector of exogenous variables that influence physician visits but do not directly affect the individual's stock of health and τ is a stochastic error term. we next obtain the predicted residuals $\hat{\tau}$ (RESDOC) from the linear regression model in (2). In the second stage of the 2SRI estimation method, we use $\hat{\tau}$ as an additional regressor to the original vector of covariates *including* the endogenous variable $\log(DOC)$ in (1) and estimate the ordered probit model using the Maximum Likelihood method. $\hat{\tau}$ is a consistent estimate of the unobservables that affect both the stock of health and physician visits. Thus, by including it among the set of covariates the source of the endogeneity problem is directly modeled. Since we use a predicted regressor in this second stage, we estimate the standard errors using the Bootstrap method.

As explained in some detail above, we use SICKPAY, a variable that indicates whether the individual's employer paid sick leave, and SEATBELT, a variable that indicates the use of seat belts, as instruments that influence physician visits but do not directly affect health stock. SICKPAY represents the opportunity cost of time

---

[14] Note that such an ordered probit model only identifies $(\xi_j - \alpha)$. A common normalization is to set $\alpha = 0$ and estimate the remaining coefficients and all the cut points.

and affects access to medical care.[15] SEATBELT represents heterogeneity in risk tolerance and is based on the answer to the following statement: "wear seat belt when drives or rides in a car." The response to this statement is coded on a 1-5 scale with 1 as being always, 2 as nearly always, 3 as sometimes, 4 as seldom, and 5 as never. Descriptive statistics for these instruments appear in Table 1.

Note that private and public insurance indicators are not used as instruments to identify the effect of physician visits on health stock. Since health insurance is likely to be endogenous to the demand for physician visits, including it in the set of instruments would not provide variation in physician visits unrelated to health status. However, since health insurance is an important variable relevant to an individual's ability to pay for medical care, excluding it from the set of covariates in (2) might create an omitted variable bias that may be carried over to the estimation of (1). To analyze the consequences of this exclusion, two robustness analyses are presented later in the paper. First, to analyze the effect of excluding private health insurance from the set of covariates in (2), we conduct the entire empirical analysis using a variable that is a good predictor of private insurance coverage but may not be directly related to health status as an instrument in addition to seat belt use and paid sick days. Second, to analyze the effect of excluding public health insurance, we eliminate individuals who have public insurance from the sample and re-estimate all econometric models. This exercise is motivated by the fact that the Medical Expenditure Panel Survey does not provide the state in which the individual resides and therefore established instruments for public insurance – state-specific variables that influence the ease with which individuals can obtain public insurance – cannot be constructed.

[15] Any factor representing the price of medical care or access to medical care may be correlated with other missing health inputs (e.g., housing, recreation, excessive alcohol use, etc.) and therefore SICKPAY may violate validity. Following Grossman (1972b), family income is used as a proxy for missing inputs.

Table 1
Variable Definitions and Descriptive Statistics

Variable	Description of Variable	Mean	Std. Dev.
Demographics			
MALE	1 if male	0.42	0.49
AGE	Number of years old	38.87	12.57
AGE2	Age squared divided by 1,000	1.67	1.01
WHITE	1 if white	0.79	0.41
MARRIED	1 if married	0.56	0.50
COLLEGE	1 if at least high school graduate	0.42	0.49
NOREAST	1 if resides in the Northeast	0.16	0.36
MIDWEST	1 if resides in the Midwest	0.20	0.40
SOUTH	1 if resides in the South	0.38	0.49
WEST	1 if resides in the West	0.26	0.44
URBAN	1 if the individual lives in an urban area	0.79	0.41
EMPLOYED	1 if employed	0.69	0.46
INCOME	Family income divided by 1,000	46.75	44.32
SIZE	Family size	2.54	1.49
Stock of Health Capital and Medical Conditions			
HSTOCK	Self-reported health index	2.33	1.08
DISEASE	1 if at least one priority or long-term life threatening condition	0.44	0.50
DISABILITY	1 if at least one functional limitation	0.21	0.41
ACUTE	1 if at least one acute or minor chronic condition	0.74	0.44
OBESE	1 if the body mass index is greater than or equal to 30	0.27	0.45
Instruments			
SEATBELT	Index of risk tolerance	1.42	0.92
SICKPAY	1 if paid sick leave	0.37	0.48
Health Insurance			
PRIVATE	1 if privately insured	0.68	0.47
PUBLIC	1 if publicly insured	0.13	0.33
Physician Services Use			
DOC	Number of physician office visits	3.13	5.57

This table presents descriptive statistics for the sample that includes adults between the ages of 18 and 64, drawn from the Household Component of the 2002 Medical Expenditure Panel Survey (MEPS) and its Medical Conditions file. The sample includes 16,583 observations.

4. Results

4.1. Instrument relevance: The reduced form regression for physician visits

The relevance of instruments requires that SICKPAY and SEATBELT must be correlated with physician visits after conditioning on other variables affecting physician visits. If the instruments are weakly correlated with the endogenous explanatory variable, then the IV estimates are biased in the same direction as the endogeneity-uncorrected estimates. The magnitude of this bias depends on the *R*-squared between the instruments and the endogenous variable in a *linear* model: as this multiple correlation increases, the bias of the IV estimator decreases. The finite sample bias of the linear IV estimator can also be expressed in terms of the *F*-statistic on the instruments [Bound *et al*. (1995)]. Indeed, Stock *et al*. (2002) suggest that the first-stage *F*-statistic on the instruments can be used to ascertain whether these instruments are weak. It is therefore common practice to report the test statistic on the joint significance of the instruments in the first stage of *nonlinear* IV models [e.g., Deb and Trivedi (2006)].

The results of the reduced form regression for physician visits appear in Table 2. The majority of the explanatory variables are highly significant and their parameter estimates are consistent with those obtained in previous studies. In contrast to female adults, adult males demand significantly fewer physician visits. The age coefficient is negative and the coefficient of age squared is positive. It appears that as age increases the demand for physician visits first decreases but later it increases. Employed individuals demand fewer physician visits. Individuals who are married and with higher incomes demand more physician visits. Medical condition variables have a large positive effect on the demand for physician visits and they also have very high significance levels.

The results of the reduced form regression for physician visits also reveal that the identifying variables are highly significant and have the predicted signs. Physician visits increase with the availability of paid sick days. As the likelihood of seat belt use increases (i.e., as the actual value of the SEATBELT variable decreases, which suggests a decrease in risk tolerance), the demand for physician visits increases. As mentioned above, the partial *R*-squared and the joint significance test on the identifying variables are useful as guides to the quality of IV estimates. The partial *R*-squared of instruments is 0.0043 and the *F*-statistic with two degrees of freedom is 34.51. According to Stock *et al*. (2002), if the number of instruments is 2, the 5% critical value to reject the null hypothesis that instruments are weak is 11.59. Thus, the instruments have useful prediction and hence are relevant.

Table 2
The Reduced Form Regression for Physician Visits

Variable	Coefficient	*t*-stat
CONSTANT	-7.63***	-17.95
MIDWEST	-0.58***	-4.50
SOUTH	-0.68***	-5.85
WEST	-1.05***	-8.41
URBAN	-0.18*	-1.82
AGE	-0.05**	-2.14
AGE2	0.90***	3.46
MALE	-1.57***	-18.59
WHITE	0.04	0.42
EMPLOYED	-0.52***	-5.01
MARRIED	0.49***	4.72
COLLEGE	0.33***	3.73
INCOME	0.009***	8.85
SIZE	-0.06*	-1.77
DISEASE	2.84***	31.33
DISABILITY	1.05***	9.97
ACUTE	5.31***	54.19
OBESE	0.28***	3.15
SEATBELT	-0.16***	-3.54
SICKPAY	0.73***	7.43

This table presents the coefficient estimates and *t*-statistics estimated using the linear regression model specified in (2) for the demand for physician visits. The results of this regression are especially used for assessing the relevance of the instruments SEATBELT and SICKPAY. The partial R-squared of the instruments is 0.0043 and the F-statistic testing the joint significance of these instruments (with two degrees of freedom) is 34.51. *t*-stats are calculated using robust standard errors. *** indicates statistical significance at .01 or better. ** indicates statistical significance at .05. * indicates statistical significance at .10.

4.2. Estimating the effect of medical care on the stock of health capital

We estimate the health production function using an ordered probit model with and without taking into account the endogeneity of physician visits. The estimation results for both models are reported in Table 3. The signs of the estimated coefficients in the ordered probit model are informative only for the rightmost and leftmost choices [Greene (2003)]. In this paper's application, the sign of a variable's estimated coefficient indicates the direction this variable changes the probability that the individual would report "Excellent Health" (leftmost choice) or

"Poor Health" (rightmost choice). If the coefficient estimate is positive (negative), then the probability of excellent health decreases (increases), whereas the probability of poor health increases (decreases). For the remaining three middle choices, the effect of a variable on the probabilities of these choices are ambiguous and marginal effects should be calculated in order to ascertain the qualitative effects of explanatory variables.

Table 3
Health production function regressions

	Ordered Probit Model		***IV Ordered Probit Model***	
Variable	**Coefficient**	***t*-stat**	**Coefficient**	***t*-stat**
DOC	0.01***	6.12	-0.12***	-4.13
RESDOC			0.13***	4.46
MIDWEST	0.02	0.55	-0.07*	-1.82
SOUTH	0.08***	2.94	-0.01	-0.40
WEST	0.13***	4.62	-0.01	-0.30
URBAN	-0.03	-1.36	-0.04*	-1.71
AGE	0.05***	10.05	0.04***	8.38
AGE2	-0.48***	-8.27	-0.39***	-5.85
MALE	0.02	0.96	-0.20***	-3.70
WHITE	-0.12***	-5.91	-0.13***	-5.18
EMPLOYED	-0.23***	-11.64	-0.26***	-10.29
MARRIED	0.03	1.18	0.09***	3.16
COLLEGE	-0.27***	-14.60	-0.22***	-8.71
INCOME	-0.004***	-17.77	-0.003***	-7.00
SIZE	0.01*	1.94	0.004	0.40
DISEASE	0.42***	22.13	0.81***	9.05
DISABILITY	0.66***	28.53	0.80***	19.97
ACUTE	0.13***	5.72	0.84***	5.18
OBESE	0.26***	13.70	0.30***	12.51
Cut Points	**Coefficient**	**Std. Error**	**Coefficient**	**Std. Error**
CUT-1	0.24	0.091	1.34	0.272
CUT-2	1.23	0.092	2.34	0.272
CUT-3	2.28	0.093	3.39	0.274
CUT-4	3.19	0.095	4.30	0.272
Pseudo *R*-Squared	0.1033		0.1039	

This table presents the estimation results for the health production function specified in (1). The endogeneity-uncorrected ordered probit model is estimated using the Maximum Likelihood method. The endogeneity-corrected ordered probit model is estimated employing the Two-Stage Residual Inclusion method, using SEATBELT and SICKPAY as instruments. *t*-stats for the IV ordered probit model are calculated using the standard errors estimated by the Bootstrap method. *** indicates statistical significance at 01 or better. * indicates statistical significance at 10.

The IV ordered probit model provides an explicit test for the *exogeneity* of physician visits. The *t*-test for the coefficient of the estimated residual is the test for the null hypothesis that physician visits is exogenous. If this hypothesis is rejected, then the appropriate specification is the IV ordered probit model. We find evidence to reject exogeneity of physician visits in the health production function regression equation, since the estimated residual (RESDOC) is highly significant in the IV ordered probit model. Furthermore, this endogeneity matters. Comparison of the standard ordered probit model formulation results with those emanating from the IV ordered probit model shows that the estimated sign of the effect of physician visits on the stock of health contradicts the theory if the endogeneity is ignored. Lastly, the IV ordered probit model indicates that the estimated residual decreases the probability of excellent health (or it increases the probability of poor health). This result is consistent with the endogeneity story outlined above, as it suggests that the unobservables are negatively associated with the stock of health.

Medical condition variables are highly statistically significant in the IV ordered probit model. Having a functional limitation, a chronic condition, an acute condition, or being obese significantly decreases the probability of excellent health (or increases the probability of poor health). As to the health depreciation variables, being male, employed and white significantly increases the probability of excellent health (or decreases the probability of poor health). Being married, on the other hand, significantly decreases the probability of excellent health (or increases the probability of poor health). Marginal effects suggest that being married significantly increases the probability of good health. Age has a significant effect on the stock of health. As age increases, the probability of excellent health first decreases and later it continues to decrease at an increasing rate. Both education and income significantly increases the probability of excellent health.

The standard ordered probit model suggests that physician visits significantly *decreases* the probability of excellent health (or significantly *increases* the probability of poor health). On the other hand, the coefficient of the IV ordered probit model suggests that physician visits significantly *increases* the probability of excellent health (or significantly *decreases* the probability of poor health), in accordance with Grossman's (1972a) demand for health capital model. Thus, if one does not take into account the endogeneity of physician visits, one may severely underestimate its effect on the stock of health, so that one would be lead to believe that its effect on health stock is contrary to what the theory predicts.

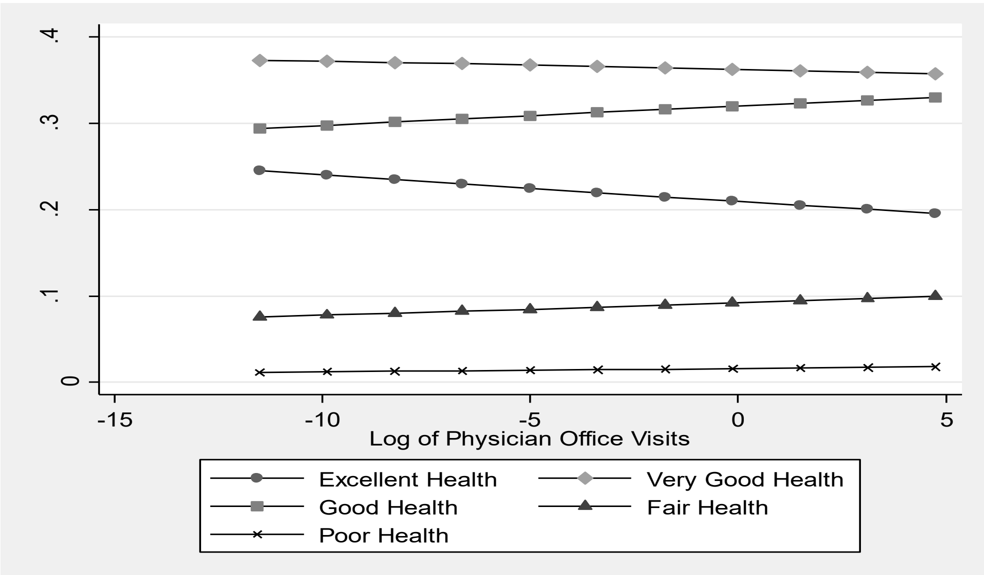
The predicted probabilities for each health category as a function of physician visits are graphed in Figure 1. The standard ordered probit model suggests that the predicted probabilities of being in excellent health or very good health decrease as the number of physician visits increases. This model also suggests that the predicted probabilities of being in fair health or poor health increase with the number of physician visits. These findings are contrary to what Grossman's (1972a) demand for health capital model suggests. The results of the IV ordered probit model, on the other hand, are consistent with Grossman's (1972a) model. The IV model suggests that the predicted probability of being in excellent health increases as the number of physician visits increases, while the predicted probabilities of being in fair health or poor health decrease with the number of physician visits. The predicted probability of being in very good health increases with physician visits until about the mean value of this variable and then it starts to decrease with further increases in the number of physician visits. The predicted probability of being in good health increases in the number of physician visits for low values of this variable and decreases in it for most of the values in its range.

The quantitative effects of a change in physician visits on the stock of health appear in Table 4. The standard ordered probit model suggests that the marginal effects of *logarithm of physician visits* are -0.003, -0.0009, 0.002, 0.001 and 0.0003 on health status outcomes of excellent, very good, good, fair and poor, respectively. In order to evaluate how big these marginal effects are, we calculate the percent effect of an *additional physician visit* on the predicted probability of being in each health category.[16] For example, an additional physician visit decreases the predicted probability of being in excellent health by 0.4 percent, whereas it increases the predicted probabilities of being in fair and poor health by 0.5 and 0.8 percent, respectively. The quantitative effects confirm that the effect of physician visits on health stock is severely biased downward if the endogeneity in physician visits is not taken into account. In fact, the marginal physician visit seems to shift individuals from excellent and very good outcomes to good, fair and poor health outcomes.

[16] The percent effect of an additional physician visit is given by $\left[\text{marginal effect of } \log(DOC) * 1/\overline{DOC} * 1/\overline{\hat{H}_i}\right]$, where $\overline{DOC}$ is the mean of physician visits and $\overline{\hat{H}_i}$ is the mean of the predicted probability for the health status category i.

Figure 1
Predicted Probabilities of Health Status as a Function of Physician Visits

A. Ordered Probit Model



B. Ordered Probit Model

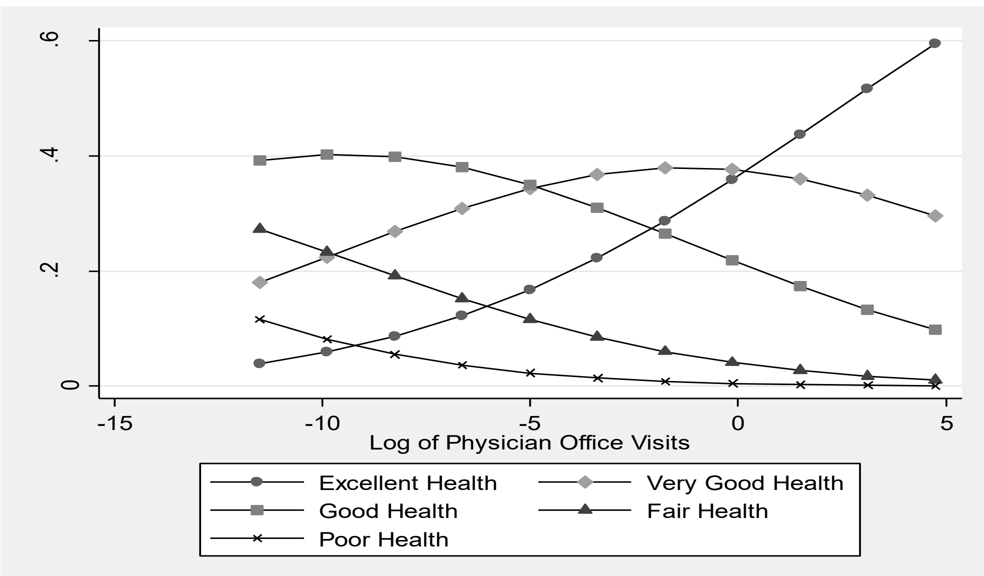


Table 4

Effects of a Change in Physician Visits on the Stock of Health Capital

A. Ordered Probit Model

	Pr[Excellent]	Pr[Very Good]	Pr[Good]	Pr[Fair]	Pr[Poor]
Marginal Effect of log(DOC)	-0.0030***	-0.0009***	0.0022***	0.0014***	0.0003***
t-stat	-6.11	-5.98	6.08	6.07	5.92
Percent Effect of DOC	-0.4	-0.08	0.2	0.5	0.8

B. Ordered Probit Model

	Pr[Excellent]	Pr[Very Good]	Pr[Good]	Pr[Fair]	Pr[Poor]
Marginal Effect of log(DOC)	0.0365***	0.0115***	-0.0262***	-0.0173***	-0.0045***
t-stat	4.15	4.11	-4.16	-4.22	-4.09
Percent Effect of DOC	5.3	1.0	-2.7	-6.4	-10.0

In each panel the first row presents the marginal effect of logarithm of physician visits on the predicted probability of being in each health status category. The third row presents the percent effect of an additional physician visit on the predicted probability of being in each health status category. *t*-stats for the IV Ordered Probit model are calculated using the standard errors estimated by the Bootstrap method. *** indicates statistical significance at 01 or better.

The IV ordered probit model indicates that the marginal effects of logarithm of physician visits are 0.037, 0.012, -0.026, -0.017 and -0.005 on health status outcomes of excellent, very good, good, fair and poor, respectively. An additional physician visit increases the predicted probability of being in excellent health by 5.3 percent, increases the predicted probability of being in very good health by 1 percent, decreases the predicted probability of being in good health by 2.7 percent, decreases the predicted probability of being in fair health by 6.4 percent, and decreases the predicted probability of being in poor health by 10 percent. The signs of the effects are consistent with Grossman's (1972a) model and their sizes are much larger than those indicated by the standard ordered probit model. Lastly, marginal effects suggest that the marginal physician visit shifts individuals from good, fair and poor outcomes to very good and excellent health outcomes. The most significant decrease in the percentage of individuals occurs in the category of poor health.

4.3. Validity of instruments

In order for the instruments to identify the effect of physician visits on the stock of health, it must be the case that these variables are validly excluded from the health stock regression. Standard tests of overidentifying restrictions exist for linear models and nonlinear models estimated by the Generalized Method of Moments. However, for the type of nonlinear model estimated in this paper these test are less straightforward. Therefore, to assess the validity of our instruments we conduct *four* separate analyses. We first ignore the discrete and ordered nature of the health stock variable, run a linear two-stage least squares regression and obtain Hansen's *J*-statistic to test for the overidentfying restrictions. A rejection of this test casts doubt on the validity of identifying variables. Hansen's *J*-statistic (with one degree of freedom) is 1.17 with a *p*-value of 0.28. This is a tolerably small *J*-statistic. Thus, the hypothesis of correct specification is not rejected, which suggests that the model is reasonably well specified and the overidentifying restrictions have not been violated.

The second analysis to assess the validity of instruments provides the reduced form regression of health stock on exogenous covariates in the model *and* the instruments to check the intuition behind the identification story by analyzing the estimated coefficients of the instruments from this regression. If the reduced form estimate of the coefficient of an instrumental variable is not significantly different from zero, or has a sign that is not compatible with the instrument's intuition, then the presumption should be that the effect of interest is either absent or the instrument is too weak to detect it [Angrist and Krueger (2001), Meer *et al*. (2003)]. The results of the reduced form regression for health stock appear in Table 5. The probability of excellent health significantly increases (or the probability of poor health significantly decreases) with the availability of sick pay. This result is in accord with the identification story. Since the availability of sick pay suggests decreased time price of medical care, an individual who has paid sick leave is expected to use more medical services; thus, the individual's stock of health is expected to be higher. As the likelihood of seat belt use increases (i.e., as the actual value of the seat belt variable decreases, which suggests a decrease in risk tolerance), the probability of excellent health significantly increases (or the probability of poor health significantly decreases). This result is also in accord with the identification story. An individual who has a lower risk tolerance is expected to use more medical care and hence his stock of health is expected to be higher.

Table 5
The Reduced Form Regression for the Stock of Health Capital

Variable	Coefficient	*t*-stat
SEATBELT	0.03***	3.40
SICKPAY	-0.08***	-3.59
MIDWEST	0.007	0.23
SOUTH	0.07***	2.78
WEST	0.12***	4.46
URBAN	-0.02	-0.89
AGE	0.05***	10.37
AGE2	-0.49***	-8.49
MALE	-0.01	-0.56
WHITE	-0.13***	-6.13
EMPLOYED	-0.20***	-9.04
MARRIED	0.03	1.48
COLLEGE	-0.26***	-13.64
INCOME	-0.004***	-16.68
SIZE	0.01	1.60
DISEASE	0.45***	24.40
DISABILITY	0.67***	28.77
ACUTE	0.18***	9.06
OBESE	0.26***	13.67
Cut Points	**Coefficient**	**Std. Error**
CUT-1	0.42	0.092
CUT-2	1.41	0.093
CUT-3	2.46	0.094
CUT-4	3.37	0.096
Pseudo *R*-Squared	0.1031	

This table presents the reduced form regression of health stock on exogenous covariates in the model *and* the instruments to check the intuition behind the identification story. The regression model is an ordered probit model and is estimated by the Maximum Likelihood Method. *** indicates statistical significance at .01 or better.

One common approach to test the validity of instruments when standard overidentification tests are not available is to estimate an exactly-identified model and then include the remaining instruments among the explanatory variables in the second-stage regression. Under the null hypothesis of correct specification, the addition of the remaining potential instruments should have little effect on the explanatory power of the second-stage regression. Thus, if one fails to reject the hypothesis that these remaining potential instruments all have zero coefficients in the second-stage regression, the validity of these variables as instruments would be supported. This is the third analysis that we carry out to assess the validity of instruments.

The choice of the instrument to exclude does not influence the test result [Bollen *et al*. (1995)]; therefore, we only present the results of the IV ordered probit regression that is exactly identified by SICKPAY (Table 6). This regression includes SEATBELT in its vector of covariates.[17] The test result suggests that the validity of instruments has not been violated. We fail to reject the null hypothesis that seat belt use has zero coefficient. Moreover, the estimate of the effect of physician visits on the stock of health is very similar to the original estimate. The marginal effects of a change in physician visits on the predicted probabilities of health categories also appear in Table 6. The results are very similar to those reported in Table 4: small changes in physician visits have a positive impact on health stock and the marginal physician visit makes an increase in the percentage of individuals in excellent and very good health categories at the expense of a decrease in the percentage of individuals in the remaining categories, with the largest decrease occurring in the category of poor health.

[17] The results of the exactly-identified IV ordered probit regression by SEATBELT, which includes SICKPAY in its vector of covariates, are available upon request.

Table 6
The IV Ordered Probit Regression Exactly Identified by SICKPAY, which Includes SEATBELT in its Set of Covariates

Variable	Coefficient	*t*-stat
DOC	-0.10***	-3.11
RESDOC	0.11***	3.40
SEATBELT	0.02	1.26
MIDWEST	-0.05	-1.48
SOUTH	0.001	0.03
WEST	0.01	0.25
URBAN	-0.04	1.49
AGE	0.04***	8.82
AGE2	-0.40***	-6.12
MALE	-0.17***	-2.97
WHITE	-0.13***	-5.40
EMPLOYED	-0.26***	-10.38
MARRIED	0.08***	2.84
COLLEGE	-0.22***	-9.00
INCOME	-0.003***	-7.25
SIZE	0.005	0.58
DISEASE	0.75***	7.66
DISABILITY	0.78***	18.33
ACUTE	0.74***	4.07
OBESE	0.29***	11.99
Cut Points	**Coefficient**	**Std. Error**
CUT-1	1.21	0.287
CUT-2	2.20	0.287
CUT-3	3.26	0.288
CUT-4	4.17	0.287
Pseudo *R*-Squared	0.1040	

Marginal Effects of a Change in Physician Visits on the Stock of Health Capital:

	Pr[Excellent]	Pr[Very Good]	Pr[Good]	Pr[Fair]	Pr[Poor]
Marginal Effect of log(DOC)	0.0307***	0.0097***	-0.0221***	-0.0145***	-0.0037***
t-stat	3.13	3.13	-3.11	-3.09	-3.08
Percent Effect of DOC	4.5	0.8	-2.3	-5.4	-8.4

This table presents the results of the IV ordered probit regression exactly identified by SICKPAY. This regression includes SEATBELT in its vector of covariates. Under the null hypothesis of correct specification, the addition of SEATBELT should have little impact on the explanatory power of this regression. The regression model is estimated employing the Two-Stage Residual Inclusion method. *t*-stats are calculated using the standard errors estimated by the Bootstrap method. *** indicates statistical significance at .01 or better.

The final analysis that checks the validity of instruments compares the effect of physician visits on health stock from two sets of IV ordered probit models using SICKPAY and SEATBELT as instruments one at a time. In other words, we re-estimate the health production function regression using the smaller set of instruments, excluding each of two instruments from the full set one at a time. The intuition behind this analysis is as follows: the basis of the standard overidentification test is that if two instruments are valid, then they both yield consistent estimates of the effect of physician visits on the stock of health and thus the difference between the estimates should be small. If not, then at least one of the instruments is not valid. As a consequence, if the effect of physician visits on health stock using different instruments provides the same interpretation of the data, then the credibility of the instruments is enhanced [Murray (2006)].

The effect of physician visits on the stock of health for the regression specifications when each instrument is used one at a time appears in Table 7. The effect estimates are quite close to the original estimate reported in Table 3. The marginal effects of a change in physician visits on the predicted probabilities of health categories are similar to the original ones reported in Table 4. Lastly, each instrument is highly significant (F-stats are 56.52 and 13.85 for SICKPAY and SEATBELT, respectively) and has the predicted sign in their respective first stage regressions and hence is relevant.

Note that all of the overidentification tests performed above rest on the assumption that at least one of the instruments is valid. Consequently, these tests are in suspect if both of our instruments share a common characteristic, since if one of them is invalid, it casts doubt on the other one. If, however, these instruments are grounded on different rationales, the overidentification tests performed would provide more comfort about their validity. Availability of sick pay suggests decreased time price of medical care, whereas wearing seat belts suggests a decrease in risk tolerance. Thus, the instruments are grounded on different rationales and the overidentification test results may provide some comfort about the validity of these identifying variables.

Table 7
IV Ordered Probit Models Exactly Identified by Using Each Instrument One at a Time

A. *IV Ordered Probit Model Exactly Identified by SICKPAY*
First Stage:
Coefficient of SICKPAY: 0.74*** (*t*-stat: 7.52, *F*-stat: 56.52)
Second Stage:
Coefficient of DOC: -0.11*** (*t*-stat: -3.18)
Coefficient of RESDOC: 0.12*** (*t*-stat: 3.47)
Effects of a Change in Physician Visits on the Stock of Health Capital:

	Pr[Excellent]	**Pr[Very Good]**	**Pr[Good]**	**Pr[Fair]**	**Pr[Poor]**
Marginal Effect of log(DOC)	0.0311***	0.0098***	-0.0223***	-0.0147***	-0.0038***
t-stat	3.21	3.16	-3.19	-3.20	-3.17
Percent Effect of DOC	4.5	0.9	-2.3	-5.4	-8.5

B. *IV Ordered Probit Model Exactly Identified by SEATBELT*
First Stage:
Coefficient of SEATBELT: -0.16*** (*t*-stat: -3.72, *F*-stat: 13.85)
Second Stage:
Coefficient of DOC: -0.20* (*t*-stat: -1.90)
Coefficient of RESDOC: 0.21** (*t*-stat: 2.00)
Effects of a Change in Physician Visits on the Stock of Health Capital:

	Pr[Excellent]	**Pr[Very Good]**	**Pr[Good]**	**Pr[Fair]**	**Pr[Poor]**
Marginal Effect of log(DOC)	0.0593*	0.0187*	-0.0425*	-0.0281*	-0.0073*
t-stat	1.90	1.93	-1.89	-1.90	-1.92
Percent Effect of DOC	8.6	1.6	-4.3	-10.3	-16.3

This table presents the results of two sets of IV ordered probit model estimates of the effect of physician visits on the stock of health using SICKPAY and SEATBELT as instruments one at a time. *** indicates statistical significance at .01 or better. ** indicates statistical significance at .05. * indicates statistical significance at .10.

t-stats for the second-stage estimates and for marginal effects are calculated using the standard errors estimated by the Bootstrap method.

5. Robustness analyses

This section provides robustness analyses regarding the exclusion of private and public insurance indicators as explanatory variables in the first-stage physician visits demand regression. Whether the individual has health insurance is an important explanatory variable relevant to his ability to pay for medical care.[18] Thus, excluding it from the set of covariates in the first-stage regression might create an omitted variable bias that may be carried over to the estimation of the health production function.

Since private insurance is likely to be endogenous to the demand for physician visits, including it in the set of instruments (and hence in the set of covariates in the first-stage regression) would not provide exogenous variation in physician visits (i.e., variation unrelated to health status), if this endogeneity problem is left uncontrolled. As a consequence, private health insurance would not be a valid instrument for the change in medical care consumption.[19] One way to get around this problem is to use a variable that is a good predictor of private insurance coverage but may not be directly related to health status. One such variable that has been used in the recent literature is the size of the company where the individual works (FIRMSIZE).[20] Greater risk pooling and economies of scale in the purchase and administration lowers the price per worker for purchasing fringes. Thus, large firms are more likely to offer health insurance benefits to their employees.

[18] Since health insurance contracts take the form of a price reduction at the time the insured purchases medical care, an insured individual would respond to the price decrease by purchasing more medical care than he would have purchased at the market price, ceteris paribus. This effect of health insurance on medical care demand is known as the *ex post moral hazard effect*.

[19] In fact, unreported regression results (which are available upon request) suggest that neither private nor public insurance is a valid instrument for the change in medical care consumption. Hansen's *J*-statistic (with one degree of freedom) from the linear model is 119.21! The subset of instruments analysis suggests that public insurance is negatively associated with the unobservables in the health stock regression. This result is not surprising since there may be unhealthy individuals without health insurance "spending down" their resources to qualify for Medicaid. The subset of instruments analysis indicates that private insurance is positively associated with the unobservables in the health stock regression, same as the findings of Deb *et al*. (2006).

[20] See, for example, Bhattacharya (2003), Deb and Trivedi (2006), and Deb *et al*. (2006). The mean and standard deviation of this variable in our sample are 95.64 and 163.08, respectively.

Table 8
Robustness Analysis – 1: The Exclusion of Private Insurance from the Set of Covariates in the First-Stage Physician Visits Regression

First Stage:
Coefficient of SICKPAY: 0.65*** (*t*-stat: 6.35)
Coefficient of SEATBELT: -0.15*** (*t*-stat: -3.51)
Coefficient of FIRMSIZE: 0.0008*** (*t*-stat: 2.81)
F-stat (three degrees of freedom): 25.41
Second Stage:
Coefficient of DOC: -0.11*** (*t*-stat: -4.07)
Coefficient of RESDOC: 0.12*** (*t*-stat: 4.43)
Hansen's *J*-statistic (two degrees of freedom): 4.46 (*p*-value: 0.11)
C-statistic (one degree of freedom): 3.29 (*p*-value: 0.07)
Subset of Instruments Analysis:
Coefficient of FIRMSIZE: 0.0001 (*t*-stat: 1.63)
Coefficient of DOC: -0.14*** (*t*-stat: -3.93)
Coefficient of RESDOC: 0.15*** (*t*-stat: 4.20)
Effects of a Change in Physician Visits on the Stock of Health Capital:

	Pr[Excellent]	**Pr[Very Good]**	**Pr[Good]**	**Pr[Fair]**	**Pr[Poor]**
Marginal Effect of log(DOC)	0.0322***	0.0101***	-0.0231***	-0.0152***	-0.0039***
t-stat	4.13	4.04	-4.05	-4.11	-4.33
Percent Effect of DOC	4.7	0.9	-2.4	-5.6	-8.8

This table presents a summary of the entire empirical analysis using FIRMSIZE, SEATBELT and SICKPAY as instruments. *** indicates statistical significance at .01 or better. *t*-stats for the second-stage estimates, subset of instruments analysis regressions and marginal effects are calculated using the standard errors estimated by the Bootstrap method.

We carry out the entire empirical analysis using FIRMSIZE, SICKPAY and SEATBELT as instruments. The results appear in Table 8. All three instruments are highly significant and have the predicted signs in the first-stage physician visits regression. The estimate of the effect of physician visits on the stock of health is very similar to the original estimate. The quantitative effects of a change in physician visits on the predicted probabilities of health categories are also very similar to the original effect estimates. Thus, using firm size as an additional instrument to capture changes in private health insurance in the first-stage regression does not change the conclusions of the paper. This result, however, should be interpreted with some caution, since strong conclusions cannot be made regarding the validity of firm size. Hansen's *J*-statistic (with two degrees of freedom) from the linear model is 4.46 (with a p-value of 0.11). Given that Hansen's J-statistic for the original (linear) model is 1.17, this suggests that the validity of firm size is in suspect. WeFmy therefore test the validity of firm size using the C-statistic, which tests the exogeneity of a subset of instruments.[21] The C-statistic (with one degree of freedom) is 3.29 (with a p-value of 0.07), which indicates that firm size may not be a valid instrument at the 10 percent significance level. The subset of instruments analysis where firm size is included among the explanatory variables in the health production function regression suggests that as the size of the firm the individual works for increases, the probability of being in excellent health decreases (with a t-stat of 1.63). This exercise also suggests that firm size may not be a valid instrument at the 10 percent significance level.

The second robustness analysis is similar and deals with the exclusion of public insurance indicator in the first-stage regression. Public insurance is also likely to be endogenous to the demand for physician visits; thus, including it in the set of instruments would not provide exogenous variation in physician visits. While one may reasonably claim that Medicare insurance is exogenous (because only the elderly and disabled are eligible), it is possible that Medicaid coverage is not. The reason is that since basic Medicaid eligibility is via poverty thresholds, there may be unhealthy individuals without (adequate) health insurance "spending down" their resources in order to qualify for Medicaid. Established instruments that affect the probability of obtaining public insurance but are unrelated to health status are state-policy variables that influence the ease with which individuals can obtain public insurance, such as state income eligibility threshold and state income threshold for the medically-needy program [Bhattacharya *et al*. (2003)]. Unfortunately, MEPS does not provide the state in which the individual resides and hence these variables cannot be constructed. Therefore, to explore the sensitivity of our results to the exclusion of public insurance as an explanatory variable from the first-stage regression, individuals who have public insurance are eliminated from the sample and the econometric models are re-estimated. The results of this robustness analysis are reported in Table 9. All results - the effect of physician visits on the stock of health, the marginal effect of physician visits on the predicted probabilities of health categories, and the relevance and validity of instruments - are very similar to those of the original ones.

---

[21] C-statistic is defined as the difference of the J-statistic of the regressions with the full set of instruments and with the smaller set of instruments. Under the null hypothesis that both the smaller set of instruments and the suspect instruments are valid, the C-statistic is distributed as chi-squared in the number of suspect instruments tested.

Table 9
Robustness Analysis – 2: The Exclusion of Public Insurance from the Set of Covariates in the First- Stage Physician Visits Regression

First Stage:
Coefficient of SICKPAY: 0.76*** (*t*-stat: 7.49)
Coefficient of SEATBELT: -0.17*** (*t*-stat: -3.53)
F-stat (two degrees of freedom): 34.95
Second Stage:
Coefficient of DOC: -0.12*** (*t*-stat: -3.80)
Coefficient of RESDOC: 0.13*** (*t*-stat: 4.12)
Hansen's *J*-statistic (one degree of freedom): 1.01 (*p*-value: 0.31)
Subset of Instruments Analysis:
Coefficient of SEATBELT: 0.02 (*t*-stat: 1.05)
Coefficient of DOC: -0.10*** (*t*-stat: -2.78)
Coefficient of RESDOC: 0.11*** (*t*-stat: 3.06)
Effects of a Change in Physician Visits on the Stock of Health Capital:

	Pr[Excellent]	Pr[Very Good]	Pr[Good]	Pr[Fair]	Pr[Poor]
Marginal Effect of log(DOC)	0.0373***	0.0073***	-0.0270***	-0.0145***	-0.0030***
t-stat	3.85	3.65	-3.80	-3.82	-3.75
Percent Effect of DOC	5.4	0.7	-3.4	-7.4	-11.4

This table presents a summary of the entire empirical analysis excluding publicly insured from the sample and using SEATBELT and SICKPAY as instruments. Number of observations is 14,477. *** indicates statistical significance at .01 or better. *t*-stats for the second-stage estimates, subset of instruments analysis regressions and marginal effects are calculated using the standard errors estimated by the Bootstrap method.

6. Summary and conclusions

This paper analyzes the effect of medical care on the stock of health capital by estimating the health investment production function. To simultaneously deal with the ordinal ranking for the stock of health and the endogeneity in physician visits, an ordered probit model for health stock with instrumental variables is estimated using the Two-Stage Residual Inclusion method. The effect of physician visits on the stock of health is identified by changes in risk tolerance and the opportunity cost of time. The variation in risk tolerance is achieved by using a variable that indicates the use of seat belts. The opportunity cost of time, which affects access to medical care, is represented by a variable that indicates whether the individual's employer paid sick leave.

The standard ordered probit model suggests that physician visits significantly decreases the probability of excellent health, whereas the IV ordered probit model indicates that physician visits significantly increases this probability, in accordance with Grossman's (1972a) demand for health capital model and in contrast to majority of the empirical work which does not uncover that medical care has a positive effect on the stock of health. Thus, if one does not take into account the endogeneity of physician visits, one may severely underestimate its effect on health stock, so that one would be lead to believe that its effect on the stock of health is contrary to what the theory predicts. The results indicate that for an average individual an additional physician visit increases the predicted probability of being in excellent health by 5.3 percent, while it decreases the predicted probabilities of being in fair health and poor health by 6.4 and 10 percent, respectively. The marginal physician visit seems to shift individuals from good, fair and poor outcomes to very good and excellent health outcomes.

References

ARROW, J.K. (1963), "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 53, 941-973.

ANGRIST, J.D. and KRUEGER, A.B. (2001), "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, 15, 69-85.

BHATTACHARYA, J., GOLDMAN, D. and SOOD, N. (2003), "The Link between Public and Private Insurance and HIV-related Mortality," *Journal of Health Economics*, 22, 1105-1122.

BOLLEN, K.E., GUILKEY, D. K. and MROZ, T.A. (1995), "Binary Outcomes and Endogenous Explanatory Variables: Tests and Solutions with an Application to the Demand for Contraceptive Use in Tunisia," *Demography*, 32, 111-131.

BOUND, J., JAEGER, D.A. and BAKER, R.M. (1995), "Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90, 443-450.

BUCKLEY, N.J., DENTON, F.T., ROBB, A.L. and SPENCER, B.G. (2004), "The Transition from Good to Poor Health: An Econometric Study of the Older Population," *Journal of Health Economics*, 23, 1013-1034.

CHANG, F.R. (1996), "Uncertainty and Investment in Health," *Journal of Health Economics*, 15, 369-376.

CHOU, S., GROSSMAN, M. and SAFFER, H. (2004), "An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System," *Journal of Health Economics*, 23, 565-587.

CONNELY, J., PHILBRICK, J., SMITH, R., KAISER, D. and WYMER, A. (1989), "Health Perceptions of Primary Care Patients and the Influence on Health Care Utilization," *Medical Care*, 27, S99-S109.

CUTLER, D. M., FINKELSTEIN, A. and MCGARRY, K. (2008), "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance," *American Economic Review*, 98, 157-162.

DARDANONI, V. and WAGSTAFF, A. (1987), "Uncertainty and the Demand for Medical Care," *Journal of Health Economics*, 6, 283-290.

DARDANONI, V. and WAGSTAFF, A. (1990), "Uncertainty, Inequalities in Health and the Demand for Health," *Journal of Health Economics*, 9, 23-38.

DEB, P. and TRIVEDI, P.K. (2006), "Specification and Simulated Likelihood Estimation of a Non-normal Treatment-outcome Model with Selection: Application to Health Care Utilization," *Econometrics Journal*, 9, 307-331.

DEB, P., MUNKIN, M.K. and TRIVEDI, P.K. (2006), "Private Insurance, Selection, and Health Care Use: A Bayesian Analysis of a Roy-Type Model," *Journal of Business and Economic Statistics*, 24, 403-415.

ERBSLAND, M., RIED, W. and ULRICH, V. (1995), "Health, Health Care and the Environment: Econometric Evidence from German Micro Data," *Health Economics*, 4, 169-182.

GRENEE, W. H. (2003), "Econometric Analysis," 5th ed., Prentice Hall, New Jersey.

GROSSMAN, M. (1972a), "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80, 223-255.

GROSSMAN, M. (1972b), "The Demand for Health: A Theoretical and Empirical Investigation," Columbia University Press for the National Bureau of Economic Research, New York.

GROSSMAN, M. (2000), "The Human Capital Model," in: J. A. Culyer and J. P. Newhouse, eds., *Handbook of Health Economics*, Elseiver, Amsterdam, Chapter 7.

HU, Y.H. and WOLFE, B. (2002), "Health Inequality between Black and White Women," IRP Discussion Paper 1251-02.

IDLER, E.L. and BENYAMINI, Y. (1997), "Self-rated Health and Mortality: A Review of Twenty-seven Community Studies," *Journal of Health and Social Behavior*, 38, 21-37.

JOHNSON, R.J. and WOLINSKY, F.D. (1993), "The Structure of Health Status among Older Adults: Disease, Disability, Functional Limitation, and Perceived Health," *Journal of Health and Social Behavior*, 34, 105-121.

KEMNA, H. (1987), "Working Conditions and the Relationship between Schooling and Health." *Journal of Health Economics*, 6, 189-210.

LU, M. and MCGUIRE, T.G. (2002), "The Productivity of Outpatient Treatment for Substance Abuse," *Journal of Human Resources*, 37, 309-335.

MEER, J., MILLER, D., L. and ROSEN, H.S. (2003), "Exploring the Health-Wealth Nexus," *Journal of Health Economics*," 22, 713-730.

MURRAY, M. P. (2006), "Avoiding Invalid Instruments and Coping with Weak Instruments," *Journal of Economic Perspectives*, 20, 111-132.

PICONE, G., URIBE, M. and WILSON, R. M. (1998), "The Effect of Uncertainty on the Demand for Medical Care, Health Capital and Wealth," *Journal of Health Economics*, 17, 171-185.

ROSENZWEIG, M.R. and SCHULTZ, T.P. (1983), "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and their Effects on Birth Weight," *Journal of Political Economy*, 91, 723-746.

RUHM, C. J. (2000), "Are Recessions Good for Your Health?" *Quarterly Journal of Economics*, 115, 617-650.

RUHM, C. J. (2005), "Healthy Living in Hard Times," *Journal of Health Economics*, 24, 341-363.

SELDEN, T. M. (1993), "Uncertainty and Health Care Spending by the Poor: The Health Capital Model Revisited," *Journal of Health Economics*, 12, 109-115.

STOCK, J.H., WRIGHT, J.H. and YOGO, M. (2002), "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20, 518-529.

STRATTMAN, T. (1999), "What Do Medical Services Buy? Effects of Doctor Visits on Work Day Loss," *Eastern Economic Journal*, 25, 1-16.

TERZA, J.V., BASU, A. and RATHOUZ, P.J. (2008), "Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling," *Journal of Health Economics*, 27, 531-543.

VAN DOORSLAER, E. and WAGSTAFF, A. *et al*. (2000), "Equity in the Delivery of Health Care in Europe and the U.S.," *Journal of Health Economics*, 19, 553-583.

VAN DOORSLAER, E., JONES, A.M. and KOOLMAN, X. (2004), "Explaining Income-related Inequalities in Doctor Utilization," *Health Economics*, 13, 629-647.

WAGSTAFF, A. (1986), "The Demand for Health: Some New Empirical Evidence," *Journal of Health Economics*, 5, 195-233.

WOLINSKY, F. D. and JOHNSON, R. J. (1991), "The Use of Health Services among Older Adults," *Journal of Gerontology: Social Issues*, 46, S345-S357.

Özet

Sağlık sermayesi üzerine tıbbi bakımın etkisi

Bu makalede, sağlık yatırımı üretim fonksiyonunun tahmin edilmesiyle tıbbi bakımın sağlık sermayesi stoğu üzerindeki etkisi incelenmiştir. Enstrümantal değişkenlerle sağlık stoğu için düzenli bir probit modeli, İki Aşamalı Residual Inclusion yöntemi kullanılarak tahmin edilir. Risk toleransı ve zamanın fırsat maliyetinin, tıbbi bakım tüketimindeki değişim için uygun araçlar olduğunu tartışıyoruz. Tıbbi bakımın sağlık stoğu üzerinde olumlu bir etkiye sahip olmadığını ortaya çıkaran ampirik çalışmaların çoğunluğunun aksine, sonuçlar, doktor ziyaretlerinin mükemmel sağlık olasılığını önemli ölçüde artırdığını (ya da sağlıksızlığın olasılığını azalttığını) göstermektedir ki bu da Grossman'ın (1972a) sağlık sermaye modeli talebiyle uyumludur.

Anahtar kelimeler: Tıbbi tedavi, probit modeli, risk toleransı, zamanın fırsat maliyeti.

JEL kodları: G22, I11.