PREVENTIVE HEALTH BEHAVIORS AMONG THE ELDERLY

by

Padmaja Ayyagari

Department of Economics

Duke University

Date: ______________________________

Approved:

____
Frank A. Sloan, Supervisor

____
Alessandro Tarozzi

____
Curtis R. Taylor

____
Ahmed Khwaja

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Economics
in the Graduate School of
Duke University

2008

ABSTRACT

PREVENTIVE HEALTH BEHAVIORS AMONG THE ELDERLY

by

Padmaja Ayyagari

Department of Economics

Duke University

Date: ______________________________

Approved:

____

Frank A. Sloan, Supervisor

____

Alessandro Tarozzi

____

Curtis R. Taylor

____

Ahmed Khwaja

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics in the Graduate School of Duke University

2008

Abstract

This dissertation consists of three essays that study preventive health behaviors among the elderly U.S. population.

The first essay studies the effect of Medicare coverage on demand for the influenza vaccine. I use a propensity score matching estimator to look at the effect of the 1993 Medicare part B coverage of the flu shot on demand. Using data from the Medicare Current Beneficiary survey, I find that the coverage increases demand by 12%. I also find that this effect varies by smoking status and by the presence chronic respiratory illnesses such as COPD, Asthma or Emphysema.

The second essay examines the effect of disease specific health shocks on risk perceptions and demand for the pneumonia vaccine. I find strong evidence of learning - individuals who experience a health shock are less likely to believe that they are not at risk of infection, conditional on prior beliefs. This change in beliefs is accompanied by a corresponding change in demand. Individuals who contract pneumonia or influenza are 60% more likely to vaccinate by the end of next year as compared to those who are not infected.

The third essay studies the relationship between education and health for a sample of elderly diabetics. We identify various mechanisms through which more education leads to improved health. We find that part of the strong positive correlation be-

tween educational attainment and health can be explained through differences in cognitive status, self-control and parental characteristics. However, some part of this relationship still remains unexplained.

Contents

List of Tables

List of Figures

Acknowledgements

During my time in graduate school, I benefited greatly from the support and encouragement of my family, mentors and friends. I would like to thank my advisor Frank A. Sloan for his guidance and for teaching me to ask important questions. I am grateful to Alessandro Tarozzi for excellent mentoring, long office hours and for his guidance. Thanks to Curtis Taylor for always being available and for his good advice, and to Ahmed Khwaja for his advice and suggestions.

I am grateful to participants of the applied Microeconomics lunch group and to all friends and colleagues for their suggestions and support. I owe many thanks to Martin Salm for some great discussions and for his humor.

My family has been a great source of support and inspiration to me. My special thanks goes to my mother, my brother Rajeev and sister-in-law Deepali. My late father has always been a tremendous source of inspiration to me. Thanks to him for his encouragement and for teaching me the value of hard work. This dissertation is dedicated to his memory.

Chapter 1

Introduction

This dissertation consists of three essays that focus on preventive health behaviors among the elderly and near-elderly U.S. population. Individuals face considerable uncertainty regarding their future health status and health expenditures. Preventive health behaviors such as vaccinations help reduce the probability of disease occurrence. Other behaviors such as exercise or dieting reduces the adverse outcomes of a disease. An individual's choice of health behaviors depends on several factors such as access to health care, price of health care, current health, perceptions about risk or effectiveness and so on. The first chapter of this dissertation studies the effect of a change in the price of the influenza vaccine on demand for it. The second chapter studies the role of risk perceptions in determining demand for the pneumonia vaccine and the third chapter focuses on the relationship between educational attainment and health behaviors among diabetics.

Medicare is a federal program of the U.S. government that provides health insurance coverage to persons over the age of 65 years or those who meet certain special criteria. Part B of Medicare, which is the optional outpatient insurance, started covering flu shots in 1993. This represents an exogenous change in the price of flu shots for the elderly population. Using data from the Medicare Current Beneficiary Survey (MCBS), I find that this change in coverage resulted in a modest increase in the demand for flu shots among persons aged 65 years or older.

According to the Centers for Disease Control and Prevention, influenza together with pneumonia is a leading immediate cause of death among the elderly in the U.S.

The cost of the vaccines and their administration, for both pneumonia and influenza, are fully covered by Medicare part B. Despite the high risk associated with the disease and the low cost of prevention, a large percentage of elderly persons report not being vaccinated. The second essay in this dissertation examines the role of risk perceptions in explaining rates of demand. Using data from the MCBS, I find that individuals who experience a health shock related to pneumonia or influenza in a particular year, are more likely to get vaccinated for pneumonia by the end of the subsequent year. I also find that individuals who experience a health shock are more likely to believe that they are at risk of getting infected, suggesting that risk perceptions plays a role in explaining demand for vaccines.

The third essay studies the role of educational attainment and cognitive status in explaining individual differences in health choices and subsequent health outcomes, for a sample of diabetics. Education might be correlated with health for a number of reasons. A direct causal relationship might exist because more educated persons might be more productive in investing in health or they might have access to better resources. On the other hand, it is possible that education and health are both correlated with other factors such as preferences which explain the positive association between education and health. Using data from the Health and Retirement Study, this chapter examines the role of educational attainment and cognitive status in individual health choices. This work is done jointly with Dr. Frank A. Sloan.

Chapter 2

Effect of Medicare Coverage on Flu Shot Demand

This essay studies the responsiveness of vaccine demand to a price change. Specifically, I focus on demand for the influenza vaccine among persons who are 65 years or older. According to the Centers for Disease Control and Prevention, each year between 5% to 20% of the U.S. population gets the flu. Influenza or complications arising from it are responsible for up to 36,000 deaths and more than 200,000 hospitalizations a year, on average (Fiore et al 2007). While all age groups are susceptible to the infection, rates of serious illness and death are highest among persons aged 65 years or older. On average, 90% of influenza-related deaths occur among this age group (Thompson et al 2003). Influenza vaccination is the most effective strategy to prevent influenza since it reduces the risk of influenza related hospitalizations and death among the elderly (Nordin et al 2001). Thus, there are large potential benefits from vaccination, especially for the elderly.

A feature of vaccines, in general, is that they exert a positive externality on the society. By getting vaccinated, not only does that person get immunity from the disease, but also anyone who comes in contact with him is no longer at risk of transmission. This externality is often not internalized and in fact, Brito et al (1991) show that the market equilibrium rate of vaccinations is not socially optimal. Thus, there is an incentive to subsidize vaccinations and increase demand. This is the motivation behind a natural experiment that I use to identify the price elasticity of demand. Medicare part B started covering flu shots for the elderly from May 1, 1993. Both the shot and its administration are fully covered and beneficiaries do not have to incur a co-payment for this benefit. Despite the full coverage, I find that about 39% of part B beneficiaries in 1993 report not receiving a flu shot.

Earlier studies (e.g, Philipson, 1996) have found that demand for vaccinations is responsive to the prevalence of the disease. A higher prevalence implies a higher risk of infection which in turn increases demand for preventive care. While several studies have focused on the prevalence elasticity of vaccine demand, not many discuss price elasticity. This study is, to my knowledge, the first to empirically estimate the responsiveness of vaccine demand to a change in price. The influenza virus frequently changes and the CDC recommends that flu shots be taken every flu season. This feature of the vaccine allows me to compare demand before and after coverage and hence identify the effect of the price change on demand. Flu shots are also relatively cheap and easily available. Hence, it is not clear, without an empirical investigation, that demand would be responsive to a price change.

I use data from the Medicare Current Beneficiary Survey (MCBS), which is a continuous survey of a representative sample of Medicare beneficiaries. The impact of the coverage is estimated using a propensity score matching estimator. The semi parametric nature of the estimator implies that we do not have to assume a linear form for the outcome equation. Most conventional regression methods rely on the linearity assumption. Also, matching allows me to compare similar individuals, which reduces selection bias. This also allows for the effects to vary across individuals.

2.1 Existing Literature

A large portion of the economic literature on infectious diseases and vaccines focuses on welfare effects. One of the important differences between the demand for vaccines and the demand for any other good is that vaccines generate a positive externality that is often not internalized by the people getting vaccinated. Conventional wisdom would suggest that vaccination be made compulsory to all in order to internalize this externality. However, Brito, Sheshinski and Intriligator (1991) show that compulsory vaccination is strictly dominated by the market equilibrium, for a very general class of models. The market equilibrium level of vaccinations is less than the socially optimal one, suggesting that the social optimum should be achieved by taxing those that do not vaccinate and subsidizing those that do. In contrast, P.J. Francis (1997) shows that this externality effect is present only when one considers a static model. He uses a dynamic model with a continuous time framework, where individuals can choose to vaccinate at any point in time and shows that in such a case the market equilibrium is the socially optimal one. The intuition behind this result is that once a threshold level of prevalence is reached, all individuals who prefer to vaccinate will do so instantaneously. So, what matters is not the proportion of the population that is vaccinated but the proportion that is infected. However, Francis makes the rather restrictive assumption that all individuals are homogenous.

An interesting aspect of vaccine demand is that it responds to prevalence of the disease in the population. If the prevalence is high then an individual is more likely to come in contact with another infected individual and contract the disease. So, when prevalence is high, people have a larger incentive to vaccinate. Several studies have focussed on the prevalence elasticity of vaccine demand and its effects. Geoffard and Philipson (1997) show that eradication in private markets cannot be achieved, since demand falls to zero as prevalence falls below a certain threshold level. Another interesting result is that price subsidies are less effective in the presence of prevalence elastic demand. This is because as price decreases and vaccinations increase, prevalence falls. This decreases demand and undermines the effect of the subsidy. Geoffard and Philipson suggest the use of a deficit-financed eradication program.

Philipson (1996) uses a proportional hazard model to test the effect of a measles epidemic time to first vaccination. He finds that the prevalence of measles in an individual's state of residence has a strong negative effect on his propensity to vaccinate.

More evidence of the prevalence elasticity of demand comes from a paper by Ahituv, Hotz and Philipson (1996), who show that condom-use increases with an increase in the prevalence of AIDS in ones state of residence.

Several studies focus on the cost-effectiveness of the influenza vaccine. Nichol (2001) focuses on healthy working adults between the ages of 18 and 64 years. She measures the direct costs of vaccination (vaccine and its administration) and indirect costs of work absenteeism to be vaccinated. Benefits are measured by costs averted for medical care, work absenteeism, reduced work effectiveness and death. Nichol finds that vaccinating leads to a mean cost saving of $13.66 per person vaccinated.

A study by Nordin et al (2001) uses administrative data on elderly (65 years and older) persons in three large health plans. They find that vaccinations significantly reduce hospitalizations for influenza and pneumonia and deaths for all causes. In a similar study, Hak et al (2002) find that influenza vaccination is associated with a reduction in hospitalization and death rates among elderly members of three large managed care organizations.

A big drawback of this literature is that it does not use a representative sample and often concentrates on a limited set of medical conditions such as respiratory disorders. By contrast, the MCBS consists of a nationally representative sample of Medicare beneficiaries and hence the results can be generalized. Moreover, I include several health conditions in my analysis.

While there is an extensive literature on the welfare effects of vaccination and relevant policies, very few studies focus on demand. One paper that does look at flu shot demand is by John Mullahy (1999). He studies the microeconomic determinants of flu shots, using the Health Promotion and Disease Prevention Supplement to the 1991 NHIS data. To estimate the causal effect of labor market variables on flu shot demand, Mullahy uses a two stage least squares methodology that instruments labor market variables with the state level unemployment rate and family structure. He finds that people with high time costs, as represented by wage rates and hours of work, are more likely to get immunized and he attributes this to their reluctance to lose working days because of sickness.

2.2 Data

I use data from the Medicare Current Beneficiary Survey (MCBS). This is a continuous survey of a nationally representative sample of Medicare beneficiaries. The sample is drawn from the Centers for Medicare and Medicaid Services (CMS) Medicare Enrollment file. The first stage of sampling was the selection of 107 geographic primary sampling units (PSUs). Within the PSUs, sampling was restricted to subareas corresponding to postal zip codes. Beneficiaries in these areas were selected by systematic random sampling within age strata. Sampling rates varied by age, in order to overrepresent the disabled (under 65 years of age) and the oldest-old (85 years of age or over) by a factor of about 1.5. The sample is replenished annually in the September-December round and the target size is of 12,000 persons with three years of cost and use information.

The MCBS consists of two files the Access to Care files and the Cost and Use files. The Access to Care files contain information on beneficiaries access to health care, satisfaction with care and source of care. The Cost and Use files contain complete expenditure and payment data on all health care services. They also include detailed information on demographics, supplemental insurance coverage and health characteristics.

I use data from the 1992 and 1993 Cost and Use files and exclude persons less than 65, who were eligible for Medicare because of disability or other reasons. Persons with missing values for the various covariates are also excluded. The final sample consists of 10,035 person-year observations, with 5,144 persons in 1992 and 4,891 persons in

1993.

The dependent variable is constructed from the answer to the question: Did you have a flu shot last winter? Since this is self-reported data it would include vaccinations that were obtained in health fairs or those for which claims were not filed. This is an advantage over administrative data, which is likely to suffer from under reporting. This is because if a person gets a flu shot in a health fair (or somewhere outside the facility) then that would not be observed in this kind of data.

Explanatory variables include several health characteristics. Self reported health ranges from 1 to 5, with 1 representing excellent health and 5 representing poor health. The survey also contains information on number of limitations in Activities of Daily Living (ADLs) and Instrumental Activities of Daily Living (IADLs).

Immunization propensity is likely to be correlated with certain chronic diseases. For example, persons with asthma are at higher risk of complications from influenza and so might be more likely to get vaccinated. Hence, I include dummy variables for several chronic diseases such as asthma, emphysema, diabetes and heart problems.

Supplemental insurance coverage also plays a role in demand for flu shots. I include dummy variables for private insurance, Medicaid, HMO enrollment and other insurance. I also construct a dummy variable for Medicare part B coverage. About 98% of the sample reports having part B coverage.

Prevalence of the disease in the population can be an important predictor of vaccine demand. If prevalence is high, an individual is more likely to come into contact with an infected person. The increased risk of infection induces greater demand for vaccines during a period of high prevalence. Unfortunately, information on the number of influenza cases was not available prior to 1997. Therefore, I use state mortality rates where the underlying cause of death was influenza or pneumonia, as a proxy for disease prevalence. These mortality rates were obtained from the National

Center for Health Statistics Compressed Mortality File based on ICD-9-CM codes for influenza and pneumonia. These are diagnostic codes based on the American Medical Associations International Classification of Diseases, 9th Revision, Clinical Modification.

2.3 Methods

2.3.1 Framework

To evaluate the effect of the Medicare coverage of flu shots, I use a propensity score matching estimator. This is a popular approach in evaluating labor market policies and has been widely applied. The quantity of interest is the average treatment effect on the treated (ATT):

$$ATT = E(Y_1 - Y_0|T = 1) \tag{2.1}$$

where, Y_1 represents the outcome with treatment and Y_0 represents the outcome without treatment. T is a dummy variable for treatment. The main problem in estimating (2.1) is that, for each person, only Y_1 or Y_0 is observed. If individuals are assigned to treatment randomly, then the counterfactual outcome may be estimated by the outcome for persons who do not receive the treatment. However, random assignment is not likely to be true in most cases. Thus, in order to use the outcomes of untreated persons to estimate the counterfactual outcomes of treated persons, it is important to account for selection into treatment. If one can control for such selection, through a vector of covariates, X, then the ATT (conditional on X) is represented by:

$$ATT(X) = E(Y_1 - Y_0|T = 1, X) \tag{2.2}$$

Under certain assumptions (discussed in more detail below), the ATT can be estimated as:

$$ATT = E\{E(Y_1|T = 1, X) - E(Y_0|T = 0, X)|T = 1\} \tag{2.3}$$

where the outer expectation is over the distribution of $(X|T = 1)$.

Rosenbaum and Rubin (1983) proved, further, that conditioning on X is equivalent to conditioning on the propensity score $P(X) = Prob(T = 1|X)$. Thus, the average treatment effect on the treated can be estimated as:

$$ATT = E\{E(Y_1|T = 1, P(X)) - E(Y_0|T = 0, P(X))|T = 1\} \tag{2.4}$$

where the outer expectation is over the distribution of $(P(X)|T = 1)$.

Conditioning on the one dimensional propensity score instead of the multi-dimensional vector X, reduces the curse of dimensionality and eases computation.

The following assumptions are required to derive equation (2.4):

1. Conditional Independence Assumption:

$$Y_1, Y_0 \perp T|X \tag{2.5}$$

   Using Rosenbaum and Rubins (1983) result on the propensity score, this can be written as:

$$Y_1, Y_0 \perp T|P(X) \tag{2.6}$$

   This assumption implies that selection into treatment is based solely on observable characteristics and that there are no unobservable characteristics that affect both treatment status and outcomes.

2. Balancing Property

$$T \perp X | P(X) \tag{2.7}$$

This property ensures that, given the propensity score, treatment is random. That is, treated and untreated persons with the same propensity score must have the same distribution of observable (and unobservable) characteristics.

3. Common Support

$$0 < P(X) < 1 \tag{2.8}$$

This assumption ensures that X does not perfectly predict treatment. That is, persons with the same characteristics have a positive probability of being both treated and untreated. Once the propensity score has been estimated, one can use matching on the propensity score to non-parametrically estimate the expectations in equation (3). A matching estimator matches each treated person with untreated persons based on their propensity scores and then computes the difference in outcomes. This difference is then averaged over all treated observations to obtain the estimate of the ATT. While there are various methods of matching on the propensity score, they all use a weighted average of the outcomes of the untreated observations to estimate the counterfactual outcome of treated observations. The various estimators differ only in the specific form of the weights used.

I use Kernel matching to estimate the ATT. This method uses a weighted average of all individuals in the control group to construct the counterfactual outcome for each treated individual. The weights used are inversely proportional to the difference between the propensity scores for the treated and untreated individuals. Thus, untreated persons who are very similar to the treated person get a higher weight while those who are very dissimilar get a low weight. The kernel matching estimator is given by:

$$ATT_K = \frac{1}{N^T}\sum_{i\in T}\left\{Y_i^T - \frac{\sum_{j\in C} Y_j^C G\left(\frac{p_j - p_i}{h_n}\right)}{\sum_{k\in C} G\left(\frac{p_k - p_i}{h_n}\right)}\right\} \tag{2.9}$$

where T is the set of treated observations and C is the set of untreated observations. $G(\cdot)$ is the kernel function and h_n is the bandwidth parameter. p is the estimated propensity score.

The kernel matching estimator uses all the observations in the control group whereas other matching estimators match on a single control observation. The advantage of using all control observations is lower variance since more information is used. This method also ensures that every treated observation has a matching control (group), which is often the problem with other matching estimators such as stratification or radius matching. In those methods, treated observations might be discarded since matching controls are not available. The drawback of the kernel method is that observations that are bad matches will also be used in the computation of the counterfactual. Imposing the common support condition can minimize bias from such bad matches.

2.3.2 Identification

In this paper, I analyze the coverage of flu shots by Medicare part B that began in 1993. The treatment variable is a binary indicator for whether or not an individual has part B coverage in 1993. I estimate the ATT for a samples of persons with part B coverage in 1992 and 1993. Thus, the control group is composed of persons with part B in 1992. Identification is based on the assumption that there are no unobserved characteristics that are correlated with both part B enrollment in 1993 and flu shot demand. In the full sample, 97.9% of the sample report having part B coverage in 1992 and 98.5% report part B coverage in 1993. Table 2.2 presents summary statistics for persons with part B coverage in 1992 and in 1993. As is evident, the samples are quite similar in terms of the various characteristics. The identifying assumption implies that, while there might be selection into part B coverage, there are no unobserved variables that make a person more likely to have part B coverage in 1993 (relative to 1992) and more (or less) likely to get a flu shot in 1993. This assumption would not be satisfied if individuals chose part B in 1993 because of the flu shot coverage. However, since the monthly premium for part B is much higher than the cost of a flu shot, this is unlikely. In 1993, part B also started covering therapeutic shoes. Though this might cause some selection into part B, it is unlikely to be correlated with flu shot demand and hence would not bias the estimation.

As an additional test, I also estimate the ATT for the full sample. The control group in this case includes persons without part B. Persons who purchase part B coverage are likely to be very different from persons without part B. Summary statistics for the full sample and for persons with and without part B coverage are presented in table 2.1. Persons without part B are, on average, younger, more likely to be married and richer. They are also less likely to have private insurance but are more likely to be enrolled in an HMO. In terms of health characteristics, persons without part B appear to be healthier than those without part B, though this difference is not large. I account for this selection by including these variables in the propensity score estimation. However, it is likely that persons with and without part B also differ in unobservable characteristics. Hence, a comparison between part B and non part B individuals in 1993 will yield a biased estimate of the ATT. Another drawback of using only persons without part B coverage as the control group is the small size of this group (less than 3% of the sample).

2.4 Results

2.4.1 Propensity Score

To estimate the propensity score I use a logit regression where the dependent variable is a binary indicator of treatment status. The propensity score is the predicted probability of being in the treatment group based on this regression. The regression uses sampling weights to account for the clustering and stratification structure of the MCBS data. The results from the propensity score regression are presented in table 2.3. The specification includes variables for age, race/ethnicity, education, marital status, income and household composition. I also include dummy variables for Medicaid, HMO coverage, private insurance and other insurance. Supplemental insurance coverage is likely to affect enrollment into part B. I control for health characteristics such as self-reported health, ADLs, IADLs and indicators for several chronic diseases, to account for the fact that health might be correlated with coverage choice. I also include the state mortality rate due to influenza and pneumonia. This serves as a proxy for the prevalence of the disease. The regression results suggest that there is some selection into treatment based on observable characteristics. Persons with a larger number of household members are more likely to have part B coverage in 1993. Income is positively correlated with part B coverage in 1993 while marital status is negatively correlated. The significance of marital status in the propensity score regression might reflect the availability of spousal insurance, which would reduce the demand for part B coverage. As expected, age is also an important predictor of part B coverage.

For the propensity score regression to be valid the balancing property must hold. The balancing property implies that for a given value of the propensity score, observations should have the same distribution of the covariates X independent of treatment status. This would imply that conditional on the propensity score, treated and control groups should on average be observationally identical. To test this, the sample is split into equally spaced intervals of the propensity score. Within each interval, I test whether the average propensity score of the treated and control groups are same and whether means of each characteristic are same between the treated and control groups. While this test is not a sufficient condition for the balancing property to hold, it is a necessary condition.

Figure 2.1 graphs the histogram of the propensity score distribution separately for the control and the treatment group. The graphs for both the groups exhibit a similar pattern. The mean propensity score for the treatment group is 0.492, while for the control group it is 0.480. The common support region is [0.356, 0.865]. I restrict the sample to observations that lie in the common support region since this is one of the conditions required for identification. This reduces bias from lingering selection on unobservables that is likely to have a larger effect for values that lie outside the common support region and improves the quality of the matches that are used to estimate the ATT.

2.4.2 Average Treatment Effect on the Treated

As mentioned above, I use Kernel matching to estimate the ATT. For the kernel function, I use a Gaussian function with a bandwidth of 0.06. The matching estimate of the effect of flu shot coverage by Medicare part B on demand is given in table 2.4. Bootstrapped standard errors and the numbers of treated and untreated observations are also presented. For the sample with part B coverage, I find a positive and significant effect of 7.3%. This represents an increase of approximately 12% over the baseline vaccination rate. Including persons without part B in the control does not change the estimated ATT much.

2.4.3 Heterogeneous Effects

The effect of the price change on vaccine demand is likely to vary across different types of individuals. For example, people who suffer from respiratory conditions such as asthma are at higher risk from influenza. Such persons might be more likely to get vaccination even without coverage and also might react less to a price change. To examine differences in the ATT across different groups of persons, I estimate it for different sub-samples. The results are presented in table 2.5.

As expected, persons without respiratory disorders are almost 2% more responsive to a price change. I also find that the average treatment effect is higher for non-smokers than smokers. However, if compared to the baseline, the percentage increase in vaccination rates is about 13.9% for smokers while it is 11.9% for non-smokers. The average treatment effect for persons who report being in fair or poor health is 5.8%. This implies a 10% increase in vaccination rates. For persons who report being in excellent to good health, the effect is much higher. I find an 8.3% effect, which implies an increase of 13.7% over the baseline. Viscusi and Evans (1990) find that the marginal utility of a given level of income is greater when healthy than when ill. If this is true, then persons in ill health will respond less to a price change whereas those in good health will respond more.

2.4.4 Robustness Checks

The specification presented above includes persons with other supplemental insurance that may cover flu shots. Persons with Medicaid or those enrolled in HMOs might already be covered for flu shots and hence the change in part B coverage would not affect their demand. Such supplemental coverage would not affect the estimation of the ATT as long as it remained the same between 1992 and 1993. If, however, flu shot coverage by supplemental insurance changed between 1992 and 1993, then the effect of that would be attributed to the Medicare part B coverage. To check this, I estimate the ATT for the sample of persons without HMO coverage and also for the sample with no additional insurance. These estimates are presented in table 2.5 and are 0.076 and 0.080, respectively. For the estimation using persons without any other supplemental insurance, the sample size is significantly smaller. This might be the reason for the larger estimate of the ATT for this sub-sample. However, the estimate is still positive and significant.

As a further check on the robustness of the estimate, I run the matching estimate for panel data. As mentioned above, the MCBS is a rotating panel survey. I use this feature of the data to restrict the sample to persons who are surveyed in 1992 and in 1993. Since the estimate now compares the same sample of persons before and after the treatment, this should control for unobserved heterogeneity in a better manner. The ATT for the panel sample is also very close to the ATT for the entire sample 0.076.

Coverage of flu shots by Medicare part B appears to have a moderate increase in demand for the vaccine. A study by Nichol et al (1994) finds that influenza vaccination results in average cost savings of $117 per person per year for the elderly. Assuming that the vaccine and its administration costs around $20 and given that the 1993 Medicare part B population consisted of 35 million persons, this implies a cost saving of $254.6 million to Medicare in 1993. The results also suggest that there is some heterogeneity across individuals. Persons with respiratory disorders and those who reported being in ill health were less elastic with respect to a price change.

Table 2.1: Summary Statistics

Variable	Full Sample	Std. Error	Part B	Std. Error	No Part B	Std. Error
Flu Shot	0.57	0.50	0.57	0.50	0.43	0.50
Age	75.28	7.17	75.34	7.16	72.29	7.08
# Hhd Members	2.08	1.13	2.07	1.12	2.39	1.51
# Hhd > 50 years	1.44	0.93	1.44	0.93	1.48	0.79
Married	0.60	0.49	0.60	0.49	0.65	0.48
Years of Education	10.91	3.84	10.91	3.83	10.85	4.40
Income	22.34	43.09	22.15	42.42	31.30	66.95
Male	0.58	0.49	0.58	0.49	0.68	0.47
White	0.88	0.32	0.89	0.32	0.76	0.43
Hispanic	0.05	0.22	0.05	0.22	0.15	0.35
Medicaid	0.10	0.30	0.10	0.30	0.08	0.28
Private Insurance	0.71	0.45	0.71	0.45	0.53	0.50
Other Insurance	0.07	0.25	0.07	0.25	0.05	0.21
HMO	0.13	0.34	0.13	0.34	0.19	0.40
Health (1-Ex, 5-Poor)	2.73	1.18	2.73	1.18	2.56	1.10
Current Smoker	0.22	0.42	0.22	0.42	0.26	0.44
Pneumonia Shot	0.23	0.42	0.23	0.42	0.20	0.40
Obese	0.13	0.34	0.13	0.34	0.09	0.29
IADL	0.35	1.02	0.35	1.02	0.27	0.89
ADL	0.02	0.25	0.02	0.25	0.01	0.15
High BP	0.51	0.50	0.51	0.50	0.46	0.50
MCI	0.17	0.38	0.17	0.38	0.12	0.32
Heart Problems	0.30	0.46	0.30	0.46	0.24	0.43
Diabetes	0.16	0.37	0.16	0.36	0.25	0.43
Alzheimer's	0.02	0.13	0.02	0.13	0.00	0.07
Emphysema, Asthma, COPD	0.17	0.38	0.17	0.38	0.15	0.35
N	10,035		9,823		212	

Table 2.2: Summary Statistics for Persons with Part B

Variable	1992	Std. Error	1993	Std. Error
Flu Shot	0.53	0.50	0.61	0.49
Age	74.90	7.33	75.81	6.95
# Hhd Members	2.01	1.00	2.14	1.23
# Hhd > 50 years	1.42	0.93	1.46	0.94
Married	0.61	0.49	0.60	0.49
Years of Education	10.91	3.81	10.92	3.84
Income	21.83	40.93	22.48	43.92
Male	0.58	0.49	0.58	0.49
White	0.89	0.32	0.88	0.32
Hispanic	0.05	0.22	0.05	0.22
Medicaid	0.10	0.30	0.10	0.30
Private Insurance	0.72	0.45	0.71	0.45
Other Insurance	0.07	0.25	0.07	0.26
HMO	0.13	0.33	0.14	0.34
Health (1-Ex, 5-Poor)	2.73	1.18	2.73	1.18
Current Smoker	0.23	0.42	0.21	0.41
Pneumonia Shot	0.23	0.42	0.23	0.42
Obese	0.13	0.33	0.13	0.34
IADL	0.34	0.99	0.37	1.05
ADL	0.02	0.24	0.03	0.26
High BP	0.50	0.50	0.53	0.50
MCI	0.17	0.37	0.18	0.39
Heart Problems	0.29	0.45	0.31	0.46
Diabetes	0.15	0.36	0.16	0.37
Alzheimer's	0.02	0.13	0.02	0.13
Emphysema, Asthma, COPD	0.17	0.37	0.18	0.38
N	5,015		4,808	

Table 2.3: Propensity Score Regression

Part B in 1993	Coefficient	Std. Error
Household Members	0.111**	0.024
Members > 50years	0.061	0.039
Medicaid	-0.007	0.087
Private Insurance	-0.028	0.065
Other Insurance	-0.003	0.085
HMO	0.062	0.074
High BP	0.068	0.045
MCI	0.058	0.060
Heart Problems	0.049	0.050
Diabetes	0.006	0.061
Alzheimer's	0.017	0.168
Emphysema, Asthma, COPD	0.070	0.058
Health Status	-0.016	0.021
Current Smoker	-0.050	0.052
Pneumonia Shot	-0.028	0.052
Obese	0.095	0.065
IADL	-0.025	0.025
ADL	0.039	0.097
Income	0.028**	0.010
Married	-0.170**	0.071
Education	-0.003	0.007
Male	-0.063	0.048
White	0.076	0.074
Hispanic	-0.029	0.101
Age 70-74	0.245**	0.059
Age 75-79	0.186**	0.062
Age 80-84	0.267**	0.067
Age 85-89	0.292**	0.087
Age 90+	0.289**	0.122
Constant	-0.572**	0.141
N	10,041	
Log Likelihood	-6907.12	

Table 2.4: Average Treatment Effect on the Treated

	ATT	% Change	Std. Error	t	N Treatment	N Control
Part B sample	0.073	12.05	0.011	6.563	4,808	5,012
Full sample	0.075	12.38	0.011	7.137	4,808	5,227
Heterogeneous Effects						
W/ Resp Disorders	0.061	8.84	0.024	2.545	847	869
W/o Resp Disorders	0.079	13.44	0.010	7.534	3,961	4,357
Smokers	0.066	13.69	0.022	3.035	1,017	1,221
Non Smokers	0.076	11.89	0.013	6.057	3,791	4,008
Fair/Poor Health	0.058	9.59	0.022	2.656	1,274	1,359
Excellent/Good Health	0.083	13.70	0.012	6.911	3,534	3,869

Table 2.5: Robustness Checks

	ATT	Std. Error	t	N Treatment	N Control
No HMO	0.076	0.014	5.529	4,149	4,556
No other insurance	0.080	0.036	2.240	423	487
Panel data	0.076	0.012	6.167	3,933	5,211

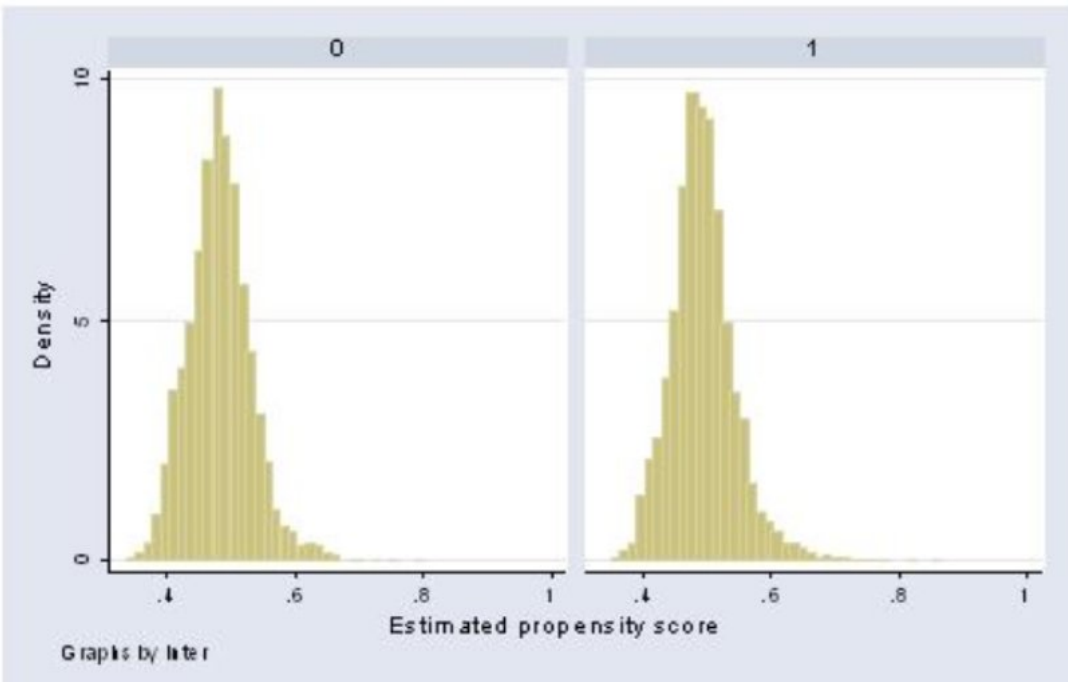

Figure 2.1: Histogram of Propensity Score for Control and Treatment Groups

Chapter 3

Do Health Shocks Affect Preventive Behavior?

This essay examines whether health shocks provide information to individuals about the risk of disease and whether this new information affects health behaviors. Specifically, I study the effect of getting pneumonia or influenza on beliefs about infection risk and on vaccine demand. Markets such as health care are characterized by imperfect information and consumer learning can play an important role in such markets. For example, individuals are uncertain about whether or not they will get pneumonia and their decision to vaccinate will depend on their beliefs about infection risk. Since the pneumonia vaccine is usually a one-time vaccine, the earlier a person vaccinates the longer he remains immune and so, the higher the benefit. Vaccination decisions exhibit considerable heterogeneity - individuals who are similar in observable characteristics choose to vaccinate at different times. A model with no learning would attribute this variation in vaccination time to unobserved heterogeneity across individuals. However, an alternative explanation could be differences in beliefs about the risk of infection. If beliefs change over time then this would explain why some individuals choose to vaccinate later.

Several earlier studies have examined consumer learning, mainly in the context of an experience good. For example, Erdem and Keane (1996) study learning about the attributes of different laundry detergent brands. They find that past usage experience as well as advertising exposure provides information about product characteristics and thus affects future demand. Ackerberg (2003) uses the same data to distinguish between "informative" and "prestige" effects of advertisements and finds that the primary effect is informative. One major difference between these studies and mine is that vaccination is not an experience good. In fact, the pneumonia vaccine is usually taken only once in a person's lifetime. As discussed below, in my model individuals learn about the risk of disease as long as they choose not to vaccinate. This approach is similar to that of Israel (2005) who studies learning about service quality using data from an automobile insurance company. Individuals learn about service quality from their claims experience as long as they choose not to switch insurers.

An individual's decision to vaccinate may be influenced by several factors such as availability of the vaccine or discomfort associated with the shot. Pneumonia is an illness that affects the lungs and respiratory system, caused by infection with bacteria and can impose significant health costs especially on the elderly or on persons with certain chronic diseases (e.g. diabetes mellitus, emphysema, chronic liver disease, congestive heart failure)[*]. Together with influenza, pneumonia is the sixth leading cause of death among the elderly in the United States (National Vital Statistics Reports (2006)). The Advisory Committee on Immunization Practices (ACIP) recommends that all persons aged 65 or over should get the pneumococcal polysaccharide vaccine (Morbidity and Mortality Weekly Report (1997)). The pneumonia vaccine is usually taken only once in a person's lifetime. However, the ACIP recommends that individuals who were vaccinated before age 65 should be revaccinated if it has been more than 5 years since the primary vaccination. Due to data restrictions, this study focuses on an individual's first decision to vaccinate. The cost of the shot and its administration is fully covered by Medicare part B. Despite the low (pecuniary) cost of the vaccine and the high risk of the disease, vaccination rates remain

[*]See Almirall et al. (1999), Fine et al. (1997).

low. According to statistics obtained from the 1999 National Health Interview Survey, only about 50% of adults aged 65 or more reported being vaccinated for pneumonia†. The infectious nature of the disease implies that there is considerable uncertainty about the probability of contracting it. Thus, subjective beliefs about the risk of getting pneumonia might play an important role in explaining observed vaccination rates. These subjective beliefs are likely to vary considerably across individuals and those who believe that they face a low risk of infection might be less likely to vaccinate. An earlier study by Kenkel (1990) finds that better information about certain disease symptoms increases the probability of medical care use, though the quantity of care conditional on use is not affected.

This study is closely related to a paper by Smith et al. (2001), in which they examine whether smokers update their longevity expectations in response to health shocks. They find that when smokers experience a smoking related health shock they revise their longevity expectations downwards much more dramatically than do former smokers or non smokers. However, smokers do not react as strongly to general health shocks whereas former and never smokers react comparably to both types of health shocks. Khwaja et al. (2006) extend this study to find a similar pattern in the decision to quit smoking - individuals quit in response to personal shocks. However they find that spousal shocks have no effect suggesting that only highly personalized anti-smoking messages are likely to be effective.

I use data from the Medicare Current Beneficiary Survey (MCBS) that surveys a nationally representative sample of Medicare beneficiaries. I restrict the analysis to persons who are 65 years of age or older. The MCBS data is a rotating

† http://www.cdc.gov/vaccines/stats-surv/downloads/general-99.pdf

panel and this panel structure allows me to identify the effect of new information on beliefs and behavior. I use answers to a survey question about why a person did not get a pneumonia shot, to measure perceptions about the risk of infection. I assume that new information about disease risk arrives in the form of exogenous health shocks. The MCBS survey data are linked to administrative Medicare claims which can be used to identify the health shocks. The estimation procedure allows for unobserved heterogeneity and also conditions on a large number of observable characteristics. To account for the fact that the pneumonia vaccine is taken only once in a lifetime, I use a hazard model to estimate the effect of health shocks on vaccination rates.

I find substantial heterogeneity across individuals - those who choose to vaccinate tend to be older, more educated and wealthier than those who do not. Self reported health and the presence of chronic conditions such as emphysema and asthma are also predictors of vaccination status. I do find strong evidence of learning - conditional on prior perceptions, individuals who experience health shocks are less likely to report that they think they are not at risk of infection. This change in risk perceptions is accompanied by a corresponding change in vaccine demand. Those who experience a health shock are much more likely to vaccinate in the next period, conditional on not having vaccinating till date. This suggests that personal experiences can significantly affect behavior. I also find that individuals with more outpatient care are more likely to vaccinate suggesting that time costs might be an important factor. Individuals might be more likely to combine vaccinations with visits for other care. Consistent with previous literature, I find that vaccine demand is higher the higher the prevalence of the disease. However, the effect of disease prevalence is much smaller than the effect of a personal health shock. This is consistent with the theory that individuals react more strongly to a personal risk than they do to the average risk of a disease.

3.1 Conceptual Framework

I consider a simple model of belief updating and vaccination choice. At the beginning of each time period, an individual is either susceptible (not vaccinated) or immune through previous vaccination. Given subjective beliefs about the risk of infection, each susceptible person decides whether or not to vaccinate that period. If the person decides not to vaccinate then he remains susceptible to infection and may contract the disease. This health shock provides new information about the risk of infection and the individual updates his beliefs in response to it. In the next period, the person again chooses whether or not to vaccinate based on the updated beliefs. Once a person has vaccinated he is immune for life and does not make the choice again.

3.1.1 Updating

Following previous literature, I assume that subjective beliefs follow a Bayesian updating process[‡]. Individuals start with prior beliefs about the risk of infection. Experiencing a health shock provides information about disease risk. For example, someone who is hospitalized due to pneumonia in the current period might be more likely to believe that they will be infected in the next period as well. The posterior probability of infection risk will then be a function of the prior probability and the risk equivalent of new information.

The exact form of the posterior probability depends on the distributions of the

[‡]The Bayesian framework has been used extensively to study learning. See Viscusi (1985), Viscusi and Evans (1998), Evans and Viscusi (1991), Smith et al. (1990), Viscusi and O'Connor (1984).

prior probabilities and the new information. Crawford and Shum (2005) assume normal distributions for both the prior and the shock, in their study of learning about effectiveness of various prescription drugs. Israel (2005) assumes a beta prior and a bernoulli distribution for the shock in his study on consumer departures from automobile insurance after an accident claim. The beta distribution is quite flexible and so is a popular choice for the prior distribution (see Viscusi (1979), Viscusi and O'Connor (1984), Gayer et al. (2000)).

Following this literature, I assume that individuals' prior beliefs follow a beta distribution with parameters a and b. The expectation of this distribution can be represented as $\pi_0 = a/(a+b)$. Individuals then update this prior after observing health shocks which are assumed to be draws from a Bernoulli distribution. The posterior distribution will then be a beta distribution with expectation(DeGroot (1970)):

$$\pi_1 = \omega\pi_0 + (1-\omega)S \tag{3.1}$$

where S is a binary indicator for the observed health shock. The term $0 < \omega < 1$ represents the precision of the prior and is defined as $\omega = (a+b)/(a+b+1)$.

Note that the general form of equation (3.1) is consistent with a variety of distributional assumptions. The exact definition of the weights however depends on the choice of distributions. For example, with a normal prior and normal shocks, the posterior distribution would also be normal with expectation equal to a weighted average of the prior expectation and health shock. The weights in that case would be proportional to the inverse of the each distribution's variance (DeGroot(1970)). In the empirical analysis described below, I estimate the effect of health shocks on a proxy measure of individual subjective probabilities. In this section I make the assumption of a beta prior and a bernoulli shock for the sake of concreteness, however the estimation procedure does not rely on these distributional assumptions.

To allow for differences across individuals, I assume that other factors such as chronic health conditions, vaccination status or disease prevalence affect risk perceptions. Previous research has found that demand for vaccines is prevalence elastic (Philipson (1996), Li et al. (2004)). For infectious diseases such as pneumonia, the higher the prevalence of the disease the higher is the probability of contracting it. So, individuals who live in areas where there is high prevalence of the disease would be more likely to believe that they are at risk of infection and also more likely to vaccinate.

Health status affects the probability of falling ill because people in bad health are more susceptible to disease. Also, certain chronic conditions such as diabetes, cardiac or pulmonary disease are complicated by pneumonia (Almirall et al. (1999), Fine et al. (1997)). Persons with such conditions are strongly advised to get vaccinated (Recommendations of the Advisory Committee on Immunization Practices, MMWR (1997)). Such individuals might have a higher subjective probability of infection than healthy persons. Another factor that affects how individuals process new information is the level of education. Kahn (1998) finds that education has a positive impact on diabetic health investments.

Vaccination status also affects the updating process. A person who did not vaccinate and contracted pneumonia is likely to update his subjective probability of illness upwards. On the other hand, if a person does not fall ill, he might not change his perceptions or might even revise them downwards. If a person has received a pneumonia shot in the past then assuming that the vaccine is fully effective, the probability of infection is zero. The vaccine recommended for adults is the pneumococcal polysaccharide vaccine and it protects against 23 types of pneumococcal bacteria. However, it is still possible to contract pneumonia from other sources and so the vaccine is not 100% effective. In the empirical analysis, I also study the effect of shocks on individual beliefs about the effectiveness of the vaccine.

To capture the effect of various factors on beliefs, I allow beliefs to depend on individual time varying characteristics and a fixed effect. Assuming a linear form, the regression counterpart of 3.1 can be written as:

$$P_{it} = \alpha_0 + \alpha_1 P_{it-1} + \alpha_2 S_{it-1} + \alpha_3 X_{it} + \eta_i \tag{3.2}$$

where P_{it} is individual i's subjective probability of infection in period t, S_{it-1} is an indicator for the health shock in period $t-1$, X_{it} includes observed factors such as prevalence, education and presence of chronic conditions and η_i is a unobserved individual fixed effect. The health shock is lagged to ensure that it occurs before the updated beliefs are reported. The main parameter of interest is α_2 which captures the effect of health shocks on perceptions, conditional on prior perceptions, observable and unobservable characteristics. A positive effect implies that shocks increase risk probabilities.

3.1.2 Vaccination Choice

In this section, I provide a theoretical framework that outlines the various factors that determine individual immunization decisions. The demand for vaccines is derived from the demand for good health, since vaccination increases the probability of good health by providing immunity from infection. Beliefs about the risk of infection affect vaccine demand through their effect on expected health conditional on not vaccinating. Other factors that affect the decision to vaccinate include the cost of vaccination and individual characteristics.

Following Viscusi and Evans (1990), I categorize health into the good health state (not infected), denoted by H^0, and the bad health state (infected), denoted by H^1, and I assume that $H^0 > H^1$. Utility is higher in the good health state: $U(H^0, X) > U(H^1, X)$, where X denotes other variables that affect utility. Since the pneumonia vaccine is usually taken only once in a person's lifetime, I assume that if a person vaccinates he is permanently immune from the disease, that is, he enjoys H^0 for life. Individuals do not know whether or not they will get pneumonia before they choose to vaccinate and so, their decision is based on expected health in that period.

Since one can only identify differences in utility, I assume that the net benefit from vaccinating is given by:

$$\Delta U_{it} = u(E[\Delta H|\Theta_{it}], C_{it}, X_{it}, \eta_i, \varepsilon_{it})$$

Where, Θ_{it} denotes the information available to individual i at time t, C_{it} denotes the cost of vaccinating and X_{it} denotes other individual characteristics that might affect the decision to vaccinate. Unobserved individual specific preferences are captured by the term η_i and ε_{it} represents the error term. $E[\Delta H|\Theta_{it}]$ is the expected health from vaccinating less the expected health from not vaccinating, conditional on information available in that period. This term is assumed to be strictly positive to capture the fact that vaccination reduces the probability of infection.

The cost of vaccination involves the monetary cost of the vaccine, time cost involved in getting the shot and the discomfort from the shot. Since the study sample consists of Medicare beneficiaries, for whom the pneumonia shot and its administration is covered, the monetary cost of getting the shot does not play an important role in the decision to vaccinate. On the other hand, non-pecuniary costs are likely to have important deterrent effects. In fact, Yoo and Frick (2005) find that individuals who have to visit the doctor for some other medical condition will be more likely to combine influenza vaccination with the visit and save time. But, if a person has to go to a doctor solely to get a flu shot then they might be less likely to do so. On the other hand, Mullahy (1997) finds a positive relationship between labor supply and demand for flu shots suggesting that persons with a high time value are more likely to vaccinate to avoid loss of work days due to illness. However, this effect is not likely to be strong for persons aged 65 or older since most of them will be retirees.

Individuals will choose to vaccinate if the net utility from doing so in that period is positive. Given that the pneumonia vaccine is taken only once in a lifetime, the vaccination decision can be thought of as an optimal stopping problem. That is, individuals optimally choose the time at which to vaccinate. Therefore, I analyze the hazard rate of vaccination which is the probability of vaccinating conditional on not having done so till date. Let $T > 0$ be a random variable with density $f(t)$ that denotes the time to vaccination. The hazard rate of vaccination, defined as $h(t) = f(t)/(1 - F(t))$, will be a function of net utility and is given by:

$$h_{it} = h\left(u(E[\Delta H|\Theta_{it}], C_{it}, X_{it}, \eta_i)\right) \tag{3.3}$$

The exact form of $h(\cdot)$ is specified in section 4.2. Here, I provide some intuition on the mechanisms through which pneumonia or influenza shocks affect the hazard rate of vaccination.

The expected change in health depends on individual beliefs about the probability of infection. In the case where the vaccine is completely effective, $E[\Delta H|I_{it}] = H^0 - (P_{it}H^1 + (1-P_{it})H^0)$. Therefore, an increase in P_{it} decreases the expected health from not vaccinating (since $H^0 > H^1$ and $0 < P_{it} < 1$) which in turn increases the net benefit from vaccination. Thus, if experiencing a health shock increases individual perceptions about the risk of infection, then health shocks should have a positive effect on the hazard rate of vaccination. In other words, individuals who revise their beliefs about the risk of getting pneumonia upwards in response to a health shock, will be more likely to vaccinate.

The analysis so far only accounts for uncertainty regarding the risk of infection. However, it is likely that individuals are uncertain about other factors such as the disutility of getting pneumonia. Consider the case where a health shock provides information about the disutility of pneumonia and not about the risk of getting pneumonia. Suppose that individuals who experience a health shock revise their beliefs about the disutility of pneumonia such that health in the infected state changes from H^1 to H^2, where $H^2 < H^1$. This would increase the net utility from vaccination and so would increase the hazard rate. On the other hand, if individuals who experience a health shock revise their beliefs about health in the infected state upwards ($H^2 > H^1$), then this would have a negative effect on the hazard rate of vaccination.

Thus, health shocks might increase the hazard rate through their effect on beliefs about both infection risk and the disutility of pneumonia. Hence, it is not possible to separately identify these two effects without additional restrictions on the model. I make the additional assumption that while a health shock provides information about both the risk and the disutility of pneumonia, the severity of the shock only affects perceptions about the disutility of pneumonia.

This assumption allows me to separately identify the effect of learning about disutility from the effect of learning about risk.

Of course, any factor that affects the expected change in health from vaccinating will affect the hazard rate of vaccination. For example, an individual who believes that the vaccine is not very effective in providing immunity from the disease will have a lower value of $E[\Delta H|\Theta_{it}]$ and so a lower hazard rate of vaccination than someone who believes that the vaccine is effective. Health shocks by themselves do not provide any information about the effectiveness of the vaccine. However, it is possible that increased doctor-patient interaction as a result of the health shock affects beliefs about other aspects such as the effectiveness of the vaccine. I examine this empirically using self reported data on perceptions.

3.2 Data

3.2.1 Sample

I use data from the Medicare Current Beneficiary Surveys (MCBS), which is a continuous survey of a nationally representative sample of Medicare Beneficiaries. MCBS data consists of two sets of files - the Access to Care and the Cost and Use files. The Access to Care (ATC) files sample the "always enrolled" Medicare population, which consists of those enrolled in one or both parts of Medicare on January 1 of that year and who remain enrolled through the end of December. The Cost and Use (CNU) files sample the "ever enrolled" population, which includes persons who enrolled in the program at any time during the year. The MCBS was initially designed to be a longitudinal survey with no predetermined limit on participation by the sample persons. However, due to high attrition, in 1993 it was decided to phase out the earlier panels after no more than six years and to limit the future panels to no more than four years of interview. This created an unbalanced data set with upto six years of data on individuals who were sampled early on and only four years of data on those included later.

To identify the effect of health shocks on vaccination I use data from the 1992-2003 Cost and Use files. In 1996, the Access to Care surveys introduced a question on why individuals chose not to vaccinate. The answers to this question help identify risk perceptions. This question was skipped in 2000. For the regression on risk perceptions I use data from the 1996-1999 Access to Care files.

I restrict the sample to persons who were living in the community at the time of the interview. That is, I exclude persons living in facilities such as nursing homes, retirement homes or other assisted living facilities at the time of the interview. This is because individuals living in a facility might all be vaccinated as part of the facility's care and so vaccination would not be their choice. I also exclude persons who were less than 65 years of age. These individuals are eligible for Medicare for reasons such as disability or end stage renal disease and are likely to be very different from the 65+ population. Since the analysis involves lagged variables, I further restrict the sample to persons who are interviewed in at least two consecutive years. The sample construction details are given in tables 3.1 and 3.2. Summary statistics for the community sample aged 65+ are presented in table 3.3. In the regressions described below, I exclude individuals who have missing data on any of the independent variables. The exact number of persons excluded depends on the specification. I report the sample size for each regression in the tables.

3.2.2 Perceptions and Vaccinations

Vaccination choice is identified from the answer to a survey question: "Have you ever had a shot for pneumonia?". An advantage of self reported data over administrative data is that would include vaccinations for which claims were not filed, such as those obtained for free at health fairs.

In order to identify the causal effect of health shocks on preventive behavior, I restrict the analysis to vaccinations that occur after the health shock. Since the survey data does not include information on the date of vaccination, I regress vaccination choice on lagged health shocks. In this analysis, a time period is one year, so that pneumonia shots in one year are regressed on health shocks in the previous year and other characteristics. It is possible that a person falls ill early in the year and then decides to get a pneumonia shot later on in the same year. To the extent that health shocks have a positive effect on updating, my estimate will be a lower bound of the effect of health shocks.

From 1996, the survey included an additional question that was asked to those who answered no to the question on pneumonia shots: "Why didn't you ever have a shot for pneumonia?". Respondents were allowed to give any answer which were then coded by the interviewer into all the categories that applied. One of the categories was "pneumonia not serious/ would not get pneumonia anyway / not at risk". I measure risk perceptions by a binary variable that takes the value one if an individual reports not vaccinating because they thought they were not at risk or that pneumonia is not serious. Other categories are reported in table 3.4 along with the percentage of respondents with answers in each category.

3.2.3 Health Shocks

An advantage of the MCBS is that the survey data are linked to administrative data on Medicare claims. I use the claims data to identify health shocks. I construct a binary health shock variable that takes the value one if the person has at least one Medicare claim with a diagnosis of influenza or pneumonia during the year, and is zero otherwise. Influenza or pneumonia related claims are identified using diagnosis codes (480-487) based on the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). I also constructed a binary variable for health shocks related to any respiratory illness using ICD-9-CM codes 460 through 519.

Both influenza and pneumonia are highly infectious diseases and there is a lot of variation in the prevalence of each disease over time and across regions. The infectious nature of the disease allows me to model it as an exogenous shock that individuals face. This approach is similar to that taken by Israel (2005) who uses automobile insurance claims where the consumer is not at fault as exogenous shocks that provide information. He identifies learning behavior about service quality by studying consumers' likelihood of switching to another insurer after a shock. Smith et al. (2001) use self-reports of serious health conditions such as heart attacks to measure health shocks.

3.2.4 Health Measures

The MCBS data also includes self reported chronic conditions such as diabetes, cardiovascular disease, emphysema and so on. These conditions increase the risk of pneumonia and so it is necessary to control for them. Since there are a large number of such self reported health conditions, I use principal component analysis to reduce the number of variables to three factors. The factor loadings (after varimax rotation), shown in table 3.5, help to interpret these factors. The first factor loads heavily on cardiovascular diseases such as hardening of arteries, myocardial infarction, coronary heart disease and other heart conditions. The second factor loads heavily on stroke, alzheimer's disease and psychiatric disorders whereas the third factor loads positively on conditions such as arthritis, rheumatoid arthritis and obesity.

I also include other self-reported measures of overall health such as a binary indicator for fair or poor health, the number of limitations in activities of daily living (ADL), in instrumented activities of daily living (IADL) and in Nagi disability measures. ADLs include ability to bathe, dress, eat, get in and out of a bed/chair, walk and use the toilet without difficulty. IADLs include ability to use the telephone, do housework, prepare meals, shop and manage money without difficulty. The Nagi Measure includes difficulty stooping/crouching/kneeling, lifting about 10 pounds, walking 2-3 blocks and writing or handling small objects (Nagi (1976)). For each of these variables, I consider a person to have a limitation in a particular activity if they report any difficulty in performing the activity, or that they require help to perform the activity or if do not perform the activity due to health reasons.

3.2.5 Prevalence

Since I do not have data on the actual prevalence of the disease for all the years, I use a proxy measure - age-adjusted state-specific mortality rates where influenza or pneumonia was the underlying cause of death. These rates were obtained from the Compressed Mortality File of the National Center for Health Statistics, Centers for Disease Control and Prevention. I used ICD-9-CM codes

480-487 to identify deaths caused by influenza or pneumonia for the years 1992 through 1998. Deaths during 1999 through 2003 were identified using ICD-10 codes J10-J18. This measure of prevalence has been used as a proxy for prevalence by previous studies (see Li et al. (2004)). Since the mortality files do not provide data for Puerto Rico, I include an indicator variable for persons living in Puerto Rico who have missing data on prevalence.

3.3 Estimation

3.3.1 Updating

Since the data does not include information on perceived probabilities of infection, I approximate the probability of infection by the "no risk" variable. This assumes that individuals who report not vaccinating because they believe that they are not at risk of getting pneumonia, have a lower perceived probability than those who do not give "no risk" as a reason for not vaccinating.

The empirical specification of the risk updating model described above is:

$$NoRisk_{it} = \alpha_0 + \alpha_1 NoRisk_{it-1} + \alpha_2 S_{it-1} + \alpha_3 X_{it} + \eta_i + \delta_t + \varepsilon_{it} \qquad (3.4)$$

The δ_ts are time fixed effects that capture aggregate shocks in each year; X_{it} includes demographics, insurance coverage and other observable individual characteristics.

In the above specification, the lagged dependent variable is endogenous. This is because it depends on the individual fixed effect and so will be correlated with the current error term ($\eta_i + \varepsilon_{it}$) through its dependence on η_i. I control for this endogeneity using methods suggested by Arellano and Bover (1995), Blundell and Bond (1998) and Windmeijer (2005). Assuming that $E[\eta_i] = E[\varepsilon_{it}] = E[\eta_i \varepsilon_{it}] = 0$ and that ε_{it} is not serially correlated, we can use the following moment conditions to estimate the model:

$$\begin{aligned} E[\Delta\varepsilon_{it}] &= 0 \\ E[NoRisk_{it-2}\Delta\varepsilon_{it}] &= 0 \\ E[S_{it-1}\Delta\varepsilon_{it}] &= 0 \\ E[x_{it}\Delta\varepsilon_{it}] &= 0 \qquad (3.5) \end{aligned}$$

This estimation procedure involves first differencing equation 3.4 which eliminates the fixed effect. The differenced lag term is still endogenous and is instrumented by a two period lag of the dependent variable to account for this endogeneity. The two period lag of the dependent variable ($NoRisk_{it-2}$) will be correlated with the differenced lag term ($NoRisk_{it-1} - NoRisk_{it-2}$) but should be orthogonal to the error term ($\varepsilon_{it} - \varepsilon_{it-1}$). Orthogonality of the instruments to the error term depends on the assumption that ε_{it} is not serially correlated. Lagged health shocks as well as other independent variables such as demographic, insurance and health status variables are assumed to be strictly exogenous. Strictly exogenous variables are uncorrelated with ε_{it} but might still be correlated with the fixed effect η_i.

Assuming that $E[x_{it}\eta_i]$ and $E[NoRisk_{it}\eta_i]$ are time-invariant generates additional moment conditions:

$$\begin{aligned} E[(\eta_i + \varepsilon_{it})\Delta NoRisk_{it-1}] &= 0 \\ E[(\eta_i + \varepsilon_{it})\Delta x_{it-1}] &= 0 \qquad (3.6) \end{aligned}$$

These additional moment conditions help identify the effect of time invariant variables such as gender and race that would otherwise drop out in (3.5). For data series that have near unit root properties, lagged levels form weak instruments for first differences causing finite sample biases. I use the combined moment conditions, (3.5) and (3.6), to estimate the model ("system GMM") since this method is shown to have much smaller finite sample bias and greater precision when the data series is highly persistent[§].

Since the "no risk" variable is observed only for those persons who report not being vaccinated, the above model is estimated for a subset of the entire sample. While persons who choose to vaccinate might differ significantly from those who do not, this will not bias the results as long as the probability of selection for each person is time invariant. To see this, denote the sample selection bias by $\lambda_{it} = E[\varepsilon_{it}|v_{it} = 0] \neq 0$, where v_{it} is an indicator of vaccination status. Assuming $\lambda_{it} = \lambda_{it'} = \lambda_i$ implies that (3.4) can be re-written as:

$$NoRisk_{it} = \alpha_0 + \alpha_1 NoRisk_{it-1} + \alpha_2 S_{it-1} + \alpha_3 X_{it} + \tilde{\eta}_i + \delta_t + \tilde{\varepsilon}_{it}$$

where, $\tilde{\eta}_i = \eta_i + \lambda_i$ and $E[\tilde{\varepsilon}_{it}|v_{it} = 0] = 0$. This is essentially the same model as before.

3.3.2 Vaccination

To analyze the effect of a health shock on vaccination rates I assume that the hazard function in equation 3.3 takes a proportional form. The dependent variable in this regression is $\tilde{t} = (t_v - 65)$, where t_v is the age at vaccination. I use $(t_v - 65)$ since the population of interest in this analysis is the 65 years or older population. The variable $\tilde{t}$ captures the duration since a person is eligible for Medicare coverage of vaccination and is recommended to get the shot. This strategy is similar to Philipson's (1996) study of measles vaccination. Philipson

[§]See Blundell and Bond (1998) and Bond (2002).

used a proportional hazard model to analyze the relationship between prevalence of measles and age in months at which first measles vaccination occurs. I use both the Cox proportional hazard model and a parametric proportional hazard model assuming a Weibull distribution.

The empirical specification for the proportional hazard model is given by:

$$h_i(\tilde{t}) = h_0(\tilde{t}) exp(\delta_1 S_{it-1} + \delta_2 X_{it}) \tag{3.7}$$

where δ_1 captures the effect of a health shock on the hazard rate of vaccination. The Cox model does not parameterize the baseline hazard $h_0(\tilde{t})$ and instead assumes that the shape of the hazard is the same for everyone. The baseline hazard is not estimated in this model. The Weibull model, on the other hand, assumes that the baseline hazard is of the form: $h_0(\tilde{t}) = p\tilde{t}^{p-1} exp(\delta_0)$. The Weibull distribution is consistent with a variety of shapes of the hazard function which depends on the estimated p. One restriction that the Weibull model places is that the hazard rate is monotonic. Preliminary analysis suggested that this was true for this particular sample of persons. S_{it-1} denotes the lagged health shock and X_{it} is a vector of all other explanatory variables that affect vaccine demand.

The above specification does not allow for unobserved heterogeneity across individuals. However, individuals might differ in unobserved attitudes towards preventive health care, which would affect their baseline probability of getting a vaccination. For example, elderly persons who live with young children might know that they are at higher risk of contracting pneumonia and so might be more likely to get a shot. To account for this, I estimate a second specification that incorporates an unobserved observation specific effect (η_{it}):

$$h_i(\tilde{t}) = h_0(t)exp(\delta_1 S_{it-1} + \delta_2 X_{it})\eta_{it} \tag{3.8}$$

This model assumes that η_{it} follows a Gamma distribution with mean one (for identifiability) and variance θ. The variance θ is estimated from the data. It is also assumed that η_{it} is not correlated with the other explanatory variables S_{it-1} and X_{it}.

Individuals who remain unvaccinated during the survey period are treated as censored. The likelihood contribution of an uncensored person i is the density function $f(\tilde{t}) = dF(\tilde{t})/d\tilde{t}$. For censored observations, the likelihood contribution is the survivor function $S(\tilde{t}) = 1 - F(\tilde{t})$, which is the probability that the duration exceeds $\tilde{t}$. Note that these functions are conditional on the independent variables and parameters described above which I omit to simplify notation. The parameters of the model corresponding to equation 3.7 are estimated by maximizing the following likelihood function:

$$L = \prod_i f(\tilde{t}_i)^{d_i} S(\tilde{t}_i)^{(1-d_i)}$$

where, d_i is an indicator for censoring status.

The likelihood function for the model incorporating unobserved heterogeneity (equation 3.8) is given by:

$$L = \prod_i \int f(\tilde{t}_i|\eta)^{d_i} S(\tilde{t}_i|\eta)^{(1-d_i)} g(\eta) d\eta$$

3.4 Results

3.4.1 Perceptions

The results for the risk perception regression 3.4 are presented in table 3.6. I find a significant negative effect of pneumonia or influenza related health shocks on the dependent variable. Individuals who experience a shock in the previous year have a 2.5% lower probability of reporting that they did not vaccinate because they thought they were not at risk as compared to those who did not experience a shock. Since only about 5% of persons believe they are not at risk, this represents a significant change in perceptions. The health shocks in these regressions are restricted to those that occurred after the prior year's interview to ensure that lagged perceptions represent beliefs before the individual experiences a shock. The results suggest that individuals update the probability of infection upwards after a health shock. The coefficient on the lagged risk variable is positive and significant suggesting that there is persistence in perceptions. This is consistent with the prediction of the Bayesian model of learning. In OLS regressions with the same specification, I find a slightly larger effect of the lagged dependent variable. This suggests that part of the persistence in risk perceptions is due to individual fixed effects.

Overall measures of health such as fair/poor health and number of ADLs have weak effect on perceptions. The number of IADLs and the Nagi disability measure do not affect perceptions. However, the presence of respiratory illness such as emphysema, COPD or asthma does affect risk perceptions. Persons with such diseases are considered to be at higher risk from pneumonia. Surprisingly, having less than a high school education decreases the likelihood that an individual thinks that they are not at risk of disease. In column 2, I restrict the sample to persons who did not use a proxy during the survey interview. This is to ensure that the reported risk perceptions represent those of the sample persons and not of the proxy respondents. The results are essentially unchanged.

The validity of the instruments used in the GMM estimation depends on the assumption that the error terms ε_{it} are not serially correlated. With only 3 years of data I cannot test for second order serial correlation, however, I do test for the validity of the instruments using the Hansen-Sargan test of overidentifying restrictions[¶]. A p-value of 0.999 in the first regression and that of 0.989 in the second regression suggests that the null hypothesis of valid instruments cannot be rejected. A second concern is that lagged levels are weak instruments for first differences if the data is highly persistent. This problem is mitigated by the use of the "system GMM" which is shown to have smaller finite sample bias (Bond (2002)). The coefficient on the lagged "no risk" variable is estimated to be 0.086 in the first regression and 0.082 in the second. These estimates are significantly different from one implying that the data series is not a near unit root process and so weakness of the instruments is not likely to be a problem.

As mentioned above, other perceptions such as those about the effectiveness of the vaccine might also play a role in determining vaccine demand. If health shocks provide information mainly through increased doctor-patient interactions then one would expect them to affect other perceptions regarding the vaccine. To examine this I estimate the effect of health shocks on : (a) a report that the individual did not vaccinate because the shot would not prevent pneumonia (b) a report that the individual did not vaccinate because the shot could cause pneumonia and (c) a report that the individual did not vaccinate because the doctor did not recommend it. The results are presented in table 3.7. Other

[¶]First order serial correlation is expected because of first differencing.

variables included in these regressions are the same as those reported in table 3.6. The coefficients on these covariates are very similar to those in table 3.6 and so I do not report them[‖]. As can be seen, I do not find a significant effect of health shocks on any of these measures of perceptions. This suggests that health shocks provide information only about the risk of infection and not about other perceptions, which implies that interactions with the doctor do not act as the source of information. While the coefficient on the health shock is not significant in any of these regressions, it is relatively large for the regression on doctor recommendations. Thus, it is difficult to completely rule out doctor interactions since the lack of significance might be due to insufficient variation in the data. The lagged perception variable in each column of table 3.7 corresponds to the lag of the dependent variable in that column. Surprisingly, the coefficient on the lagged term is not significant in the case of perceptions about the effectiveness of pneumonia or for perceptions that the vaccine can cause pneumonia. In a simple linear regression, however, I do find a significant effect of the lagged term. This suggests that persistence in these perceptions is completely due to individual fixed effects.

The above regressions estimate a very simple linear equation. However, it is possible that the effect of prior perceptions changes with age or with the presence of diseases such as emphysema, asthma or COPD. The effect of health shocks might also vary with individual characteristics. Unfortunately, lack of sufficient variation in the dependent variable does not allow me to estimate a more flexible model with multiple interactions. Thus, I only present results from a model with no interactions and leave the more flexible model to future research.

[‖]Available on request.

3.4.2 Vaccination

A non-parametric examination of the effect of a health shock on vaccination is shown in figure 1. This graph plots the Nelson-Aalen cumulative hazard estimate for persons who experienced a flu or pneumonia shock and for those that didn't. The Nelson-Aalen estimator is defined as $\hat{H}(t) = \sum_{j|t_j \leq t} \frac{v_j}{r_j}$, where d_j is the number of persons who vaccinated at time t_j and r_j is the number of persons who are still susceptible (unvaccinated and uncensored) at time t_j. The figure shows that at each age, the cumulative hazard is higher for persons who receive a shock than it is for those who do not. Figure 2 shows that Kaplan-Meier failure rate for persons with and without health shocks. This estimator is defined as $\hat{F}(t) = 1 - \sum_{j|t_j \leq t} \frac{r_j - v_j}{r_j}$ and represents the empirical cumulative density function. The figure shows that persons who experienced a health shock are more likely to vaccinate than those that did not experience a shock.

Table 3.8 presents the estimates from different specifications of the hazard model. The first column presents the hazard ratio estimates from a Cox proportional hazard regression. Experiencing a flu or pneumonia shock increases the hazard of vaccination by 45%. In other words, at any age a person who has experienced a health shock is 45% more likely to vaccinate in the next year than someone who has not experienced a shock, conditional on not having vaccinated previously and on other characteristics. The hazard of vaccination also increases with an increase in the number of outpatient events that the person has in that time period. This suggests that the time cost plays an important role in determining immunization decisions. On the other hand, a hospitalization decreases the probability of vaccinating in that year. Those who are married are 14% more likely to get vaccinated. This is probably because marriage increases the disutility of ill health. Persons with less than a high school education are about 9% less likely to vaccinate. Self reported health, number of ADLs and IADLs do not have a significant effect on the hazard of vaccination. However, the presence of respiratory illnesses, hardening of arteries and high blood pressure increases the vaccination hazard.

Supplemental health insurance also increases the hazard rate of vaccination. Persons with other public insurance face a hazard that is 23% higher, while those in an HMO face a hazard that is 26% higher. Prevalence has a small effect of 1% on the probability of vaccination. This suggests that while individuals do respond to the average risk of infection, personal experiences matter a lot more.

The second column reports the hazard ratios for a Weibull proportional hazard model. The results are quite similar to the Cox model estimates. Number of IADLs now has a significant negative effect on the probability of vaccination. The final column presents the results of a Weibull regression model that accounts for unobserved heterogeneity. I assume that the heterogeneity term η_{it} follows a gamma distribution with variance θ. The effect of a health shock is much larger once we control for heterogeneity. Experiencing a pneumonia or influenza related shock in the previous year increases the hazard of vaccination in the current year by 59%. The presence of emphysema, COPD or asthma also has a larger effect.

The effect of disease prevalence is now significant at the 1% level. I also find that the effect of the indicator for missing prevalence is very large (70%) and significant at the 1% level. These are persons living in Puerto Rico for whom data on prevalence was not available. To test the sensitivity of my results to these person, I estimate the same model excluding persons with missing data on prevalence. I find that the results are essentially unchanged. These results are not presented here but are available on request.

The shape parameter of the Weibull distribution is estimated to be greater than one. This implies that the baseline hazard rate is increasing over time. That is, conditional on all other factors, as individuals get older they are more likely to vaccinate. The effect of age on vaccination rates is an empirical issue. As a person gets older, he has shorter life expectancy and so can enjoy the benefits of vaccination (immunity) for a shorter time period. This would have a negative effect on demand. On the other hand, as one gets older the health cost of getting pneumonia is higher, thus increasing the demand for vaccination. The estimate of 1.439 implies that the latter effect dominates.

As mentioned earlier, one concern in interpreting the effect of health shocks on vaccine demand is that it might represent learning about the disutility or cost of getting pneumonia and not risk of infection. However if that is the case, one would expect a more severe health shock to have a larger effect on the vaccine demand. I test for this by including a binary variable for health shocks that resulted in a hospitalization. The results are shown in first column of table 3.9. I find an insignificant effect suggesting that the severity of the shock provides no additional information to the individual. I take this as support of my hypothesis that individuals learn about the risk of infection from pneumonia or influenza shocks. One concern with this interpretation is that the estimated hazard ratio for the shocks that resulted in hospitalizations is quite large and this might simply imply that the sample does not have sufficient power to help identify a significant effect. However, I have a large sample size of 17,706 persons with 439 hospitalizations and so I do not expect low power to be driving this result. In the second column of table 3.9, I test the robustness of my results to a different measure of health. I include the health factors shown in table 3.5 in the regression. The results do not change substantially.

Overall, the results suggest that individuals do respond to health shocks by updating their risk perceptions. These results are consistent with a Bayesian learning model. Health shocks also increase the hazard rate of vaccination, suggesting that the increase in risk perceptions affects an individual's vaccination choice. The results are robust to model specification. The severity of the shock, however, does not appear to have any additional effect on the probability of vaccinating. Individuals with a larger number of outpatient visits in that year are also more likely to get vaccinated, suggesting that the time cost of vaccination might play a role in decision-making.

Table 3.1: Sample Construction for Perception Estimation (Access to Care)

Year	Total	Community	Age $\geq$ 65	2 Years' Data
1996	17,794	16,518	14,060	-
1997	18,330	17,078	14,479	8,448
1998	20,889	19,651	16,789	8,971
1999	17,936	16,670	14,120	9,063
Total	74,949	69,917	59,448	26,482

Table 3.2: Sample Construction for Vaccine Estimation (Cost and Use)

Year	Community	Age $\geq$ 65	Reg Sample (2 Yrs)
1992	12,110	9,899	-
1993	11,359	9,446	7,722
1994	11,749	9,757	8,201
1995	11,096	9,198	5,727
1996	10,869	9,059	5,533
1997	11,500	9,647	5,433
1998	12,048	10,118	5,985
1999	12,148	10,255	6,263
2000	12,000	10,036	5,988
2001	11,893	9,924	5,903
2002	11,750	9,846	5,920
2003	11,614	9,682	5,756
Total	140,136	116,867	68,431

Table 3.3: Summary Statistics (Means)

Variable	ATC Sample	CNU Sample
Vaccinated	0.56	0.51
Flu Shock	0.03	0.07
Outpatient		3.58
Inpatient		0.38
Age	76.40	77.12
Male	0.42	0.42
Hispanic	0.07	0.06
White	0.88	0.87
< High School	0.37	0.39
Married	0.52	0.51
Fair/Poor Health	0.23	0.25
# of ADLs	0.67	0.77
# of IADLs	0.86	0.99
Nagi Disability	0.85	0.96
Current Smoker	0.11	0.11
Former Smoker	0.47	0.48
Emphysema	0.13	0.15
Artery	0.11	0.14
High BP	0.54	0.58
MI	0.15	0.16
Angina / CHD	0.14	0.16
Other Heart	0.27	0.27
Stroke	0.11	0.12
Medicaid	0.09	0.13
Public	0.05	0.07
Private	0.61	0.47
HMO	0.22	0.21
Prevalence	30.45	28.49
Prev Missing	0.01	0.02

Table 3.4: Reasons for not vaccinating (Percentages)†

Reason	1996	1997	1998	1999	Total
Pneumonia not serious / would not get pneumonia anyway / not at risk	**4.7**	**4.9**	**4.8**	**5.5**	**4.9**
Didn't know it was needed	62.1	64.3	59.3	55.5	60.6
Shot could cause pneumonia	3.3	4.1	4.0	6.0	4.2
Shot could have side effects or cause disease	2.5	3.4	3.8	4.4	3.5
Didn't think it would prevent pneumonia/ could get pneumonia anyway	4.8	5.0	5.3	5.0	5.0
Doctor did not recommend shot	14.0	12.5	12.4	12.2	12.8
Doctor recommended against getting shot / allergic to shot / medical reasons	1.0	1.0	1.3	1.2	1.1
Don't like shots or needles / concerns about soreness or rash / local reactions	2.9	2.6	2.6	3.7	2.9
Inconvenient to get shot/ unable to get to location	0.7	0.8	1.0	0.9	0.9
Didn't think about it / forgot /missed it	10.8	7.3	8.8	8.0	8.8
Cost of shot / not worth the money	0.3	0.2	0.3	0.2	0.2
Other	2.5	1.9	1.6	4.4	2.5
N	7,122	6,501	6,709	5,070	25,402

† Percentages in each column do not sum to a 100 because a respondent's answer could be coded into multiple categories.

Table 3.5: Principal Components Analysis- Factor Loadings

Variable	Factor 1	Factor 2	Factor 3
Obese	-0.0270	-0.1880	**0.4100**
Hardening of Arteries	**0.3985**	0.0689	0.0221
High Blood Pressure	0.1743	-0.0652	0.3339
Myocardial Infarction	**0.5097**	-0.0281	-0.0627
Coronary Heart Disease	**0.5100**	-0.0280	-0.0627
Other Heart Conditions	**0.3693**	0.0174	0.0763
Stroke	0.2392	**0.3482**	-0.0701
Skin Cancer	0.0913	0.0513	-0.0443
Cancer	0.0351	0.0317	0.0713
Diabetes	0.1511	-0.0875	0.2604
Rheumatoid Arthritis	-0.0008	0.0538	**0.4022**
Arthritis	-0.0285	0.0097	**0.5283**
Mental Retardation	-0.0608	0.2152	0.0092
Alzheimer's Disease	-0.0376	**0.4595**	-0.0507
Psychiatric Disorder	-0.0673	**0.4037**	0.1088
Osteoporosis	-0.1259	0.2256	0.3660
Broken Hip	-0.0884	0.3213	0.0760
Parkinson's Disease	-0.0542	0.3292	-0.0012
Emphysema, Asthma or COPD	0.0776	0.0645	0.1737
Paralysis	0.1454	**0.3621**	-0.0714

Table 3.6: Perceptions - Effect of Health Shocks on "No Risk" of Pneumonia

	(1)	Std. Error	(2)	Std. Error
Lag Pneu / Flu Shock	-0.025***	(0.007)	-0.027***	(0.008)
Lagged Risk Perception	0.086**	(0.039)	0.082***	(0.041)
Age 70-74	-0.011	(0.008)	-0.014	(0.009)
age 75-79	-0.012	(0.009)	-0.011	(0.009)
age 80-84	-0.013	(0.009)	-0.013	(0.009)
age 85-89	-0.007	(0.011)	-0.010	(0.012)
Age 90+	0.008	(0.015)	-0.019	(0.019)
Male	-0.002	(0.006)	-0.001	(0.007)
Hispanic	-0.010	(0.010)	-0.012	(0.011)
White	0.007	(0.006)	0.003	(0.007)
< High School	-0.014**	(0.006)	-0.015**	(0.006)
Income > $25K	0.002	(0.007)	0.005	(0.007)
Married	0.002	(0.006)	0.000	(0.006)
Fair/Poor Health	-0.013*	(0.007)	-0.013**	(0.007)
# of ADLs	0.006*	(0.004)	0.006	(0.004)
# of IADLs	-0.002	(0.003)	-0.003	(0.004)
Nagi Disability	-0.002	(0.003)	-0.002	(0.003)
Current Smoker	0.014	(0.011)	0.014	(0.011)
Former Smoker	-0.010*	(0.006)	-0.012	(0.007)
Emphysema	-0.016**	(0.008)	-0.017*	(0.009)
Artery	-0.010	(0.008)	-0.010	(0.009)
High BP	-0.006	(0.006)	-0.006	(0.006)
MI	-0.009	(0.008)	-0.011	(0.008)
Angina / CHD	-0.011	(0.007)	-0.012	(0.008)
Other Heart	-0.003	(0.007)	-0.003	(0.007)
Stroke	-0.008	(0.007)	-0.004	(0.008)
Medicaid	0.006	(0.009)	0.011	(0.011)
Public	-0.013	(0.011)	-0.016	(0.011)
Private	-0.013*	(0.007)	-0.017**	(0.008)
HMO	-0.023***	(0.008)	-0.024***	(0.008)
Prevalence	-0.001	(0.001)	-0.001	(0.001)
Prev Missing	-0.043**	(0.022)	-0.042*	(0.024)
Hansen-Sargan Test (p-value)	0.999		0.989	
Year Fixed Effects	Yes		Yes	
Region Fixed Effects	Yes		Yes	
N	5,331		4,839	

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 3.7: Other Perceptions - Some Robustness Checks

	Prevent	Std. Error	Cause	Std. Error	Dr. Reco	Std. Error
Lag Pneu / Flu Shock	-0.003	(0.049)	-0.009	(0.050)	-0.033	(0.047)
Lagged Perception	0.052	(0.039)	0.012	(0.060)	0.147***	(0.041)
Hansen-Sargan Test (p-value)	0.317		0.843		0.990	
Year Effects	Yes		Yes		Yes	
Region Effects	Yes		Yes		Yes	
Other Covariates	Yes		Yes		Yes	
N	5,331		5,331		5,331	

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 3.8: Hazard Models - Effect of Pneumonia or Flu Shock on Vaccination

	Cox	Std. Error	Weibull	Std. Error	Gamma Frailty	Std. Error
Lag Pneu / Flu Shock	1.451***	(0.089)	1.397***	(0.082)	1.591***	(0.138)
Outpatient	1.011***	(0.001)	1.010***	(0.001)	1.020***	(0.003)
Inpatient	0.957**	(0.020)	0.947***	(0.017)	0.928***	(0.024)
Male	1.010	(0.032)	0.991	(0.034)	1.005	(0.045)
White	1.167***	(0.059)	1.209***	(0.055)	1.287***	(0.079)
Hispanic	0.936	(0.070)	0.920	(0.066)	0.904	(0.084)
Married	1.138***	(0.033)	1.214***	(0.039)	1.260***	(0.055)
< High School	0.914**	(0.033)	0.871***	(0.028)	0.848***	(0.037)
Fair/Poor Health	1.027	(0.047)	1.044	(0.042)	1.046	(0.058)
# of ADLs	0.997	(0.020)	0.988	(0.018)	0.984	(0.026)
# of IADLs	0.976	(0.017)	0.955***	(0.015)	0.951**	(0.023)
Nagi Disability	1.006	(0.017)	0.998	(0.018)	1.014	(0.025)
Current Smoker	0.906*	(0.050)	0.962	(0.050)	0.955	(0.062)
Former Smoker	1.014	(0.037)	1.045	(0.035)	1.050	(0.047)
Obese	0.994	(0.035)	1.081**	(0.042)	1.083	(0.055)
Emphysema	1.297***	(0.059)	1.342***	(0.058)	1.487***	(0.092)
Artery	1.134***	(0.054)	1.144***	(0.052)	1.192***	(0.076)
High BP	1.095***	(0.033)	1.098***	(0.034)	1.111***	(0.045)
Heart Attack	0.972	(0.046)	0.966	(0.045)	0.952	(0.061)
Angina/ CHD	1.070	(0.049)	1.091*	(0.050)	1.122*	(0.072)
Other Heart	1.088**	(0.039)	1.089**	(0.038)	1.106**	(0.053)
Stroke	1.043	(0.054)	1.063	(0.050)	1.118	(0.077)
Cancer	1.003	(0.038)	0.998	(0.038)	0.989	(0.052)
Diabetes	1.074*	(0.043)	1.093**	(0.044)	1.096*	(0.059)
Medicaid	1.000	(0.058)	1.024	(0.057)	1.049	(0.079)
Public	1.225***	(0.073)	1.197***	(0.066)	1.289***	(0.104)
Private	1.093**	(0.046)	1.081*	(0.043)	1.089	(0.058)
Private Missing	0.636	(0.273)	0.829	(0.385)	0.693	(0.434)
HMO	1.275***	(0.059)	1.325***	(0.053)	1.458***	(0.080)
Prevalence	1.010**	(0.005)	1.008***	(0.003)	1.012***	(0.004)
Prev Missing	1.841	(0.877)	1.514***	(0.219)	1.701***	(0.326)
Year Effects	Yes		Yes		Yes	
p	-		1.161***	(0.020)	1.439***	(0.054)
θ	-		-		0.163	(0.031)
Log Likelihood	-10527.308		-1782.8174		-1765.4702	
N	17,706		17,706		17,706	

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 3.9: Vaccine Demand - Some Robustness Checks

	(1)	Std. Error	(2)	Std. Error
Lag Pneu / Flu Shock	1.505***	(0.145)	1.537***	(0.149)
Lag Shock × Hospitalized	1.259	(0.252)	1.330	(0.277)
Outpatient	1.020***	(0.003)	1.021***	(0.003)
Inpatient	0.926***	(0.024)	0.931***	(0.024)
Male	1.004	(0.045)	1.037	(0.047)
White	1.287***	(0.079)	1.301***	(0.080)
Hispanic	0.905	(0.084)	0.882	(0.082)
Married	1.261***	(0.056)	1.259***	(0.055)
< High School	0.848***	(0.037)	0.858***	(0.038)
Fair/Poor Health	1.045	(0.058)	1.045	(0.058)
# of ADLs	0.983	(0.027)	0.981	(0.026)
# of IADLs	0.950**	(0.023)	0.947**	(0.023)
Nagi Disability	1.014	(0.025)	0.999	(0.025)
Current Smoker	0.955	(0.063)	1.000	(0.065)
Former Smoker	1.050	(0.047)	1.051	(0.047)
Obese	1.085	(0.055)	-	
Emphysema	1.485***	(0.092)	-	
Artery	1.190***	(0.076)	-	
High BP	1.111**	(0.045)	-	
Heart Attack	0.952	(0.061)	-	
Angina/ CHD	1.123*	(0.072)	-	
Other Heart	1.106**	(0.053)	-	
Stroke	1.119	(0.077)	-	
Cancer	0.990	(0.052)	-	
Diabetes	1.094*	(0.060)	-	
Health Factor 1	-		1.073***	(0.017)
Health Factor 2	-		1.037*	(0.019)
Health Factor 3	-		1.133***	(0.020)
Medicaid	1.049	(0.080)	1.024	(0.078)
Public	1.290***	(0.105)	1.319***	(0.109)
Private	1.089	(0.058)	1.092*	(0.058)
HMO	1.458***	(0.081)	1.418***	(0.080)
Prevalence	1.011***	(0.004)	1.010**	(0.004)
Prev Missing	1.707***	(0.328)	1.982*	(0.744)
Year Fixed Effects	Yes		Yes	
Region Fixed Effects	No		Yes	
p	1.442	(0.055)	1.439	(0.054)
θ	0.165	(0.031)	0.165	(0.030)
Log Likelihood	-1764.9287		-1748.7837	
N	17,706		17,684	

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

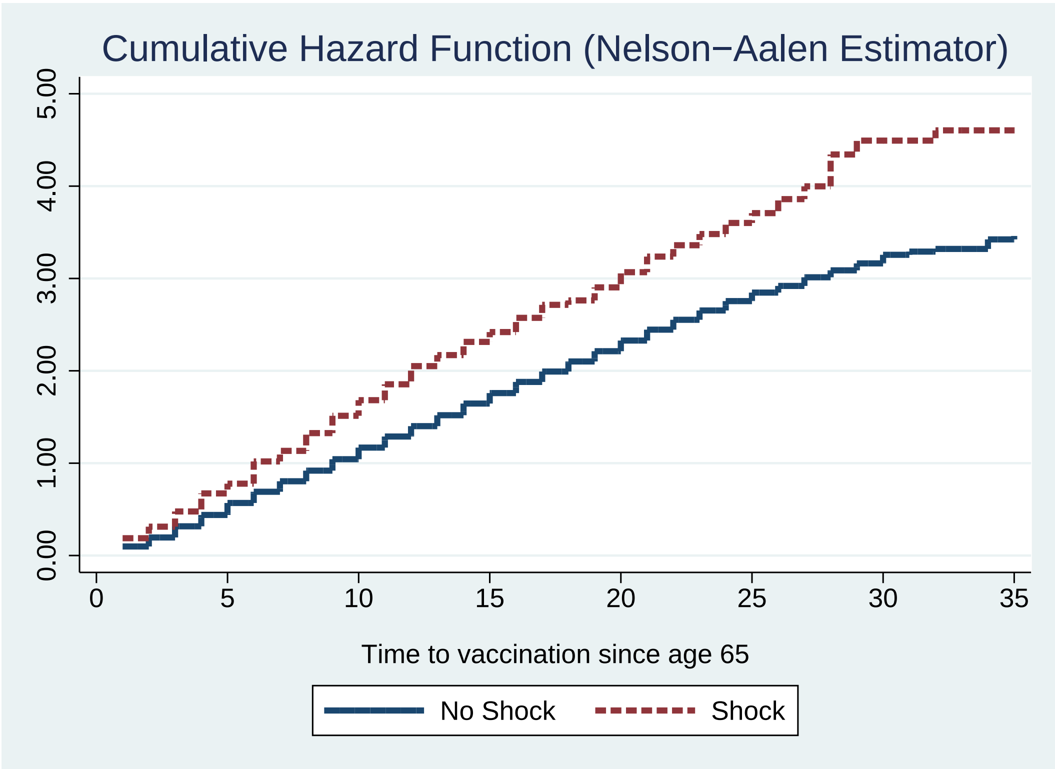

Figure 3.1: Cumulative hazard of vaccination for persons with and without shocks

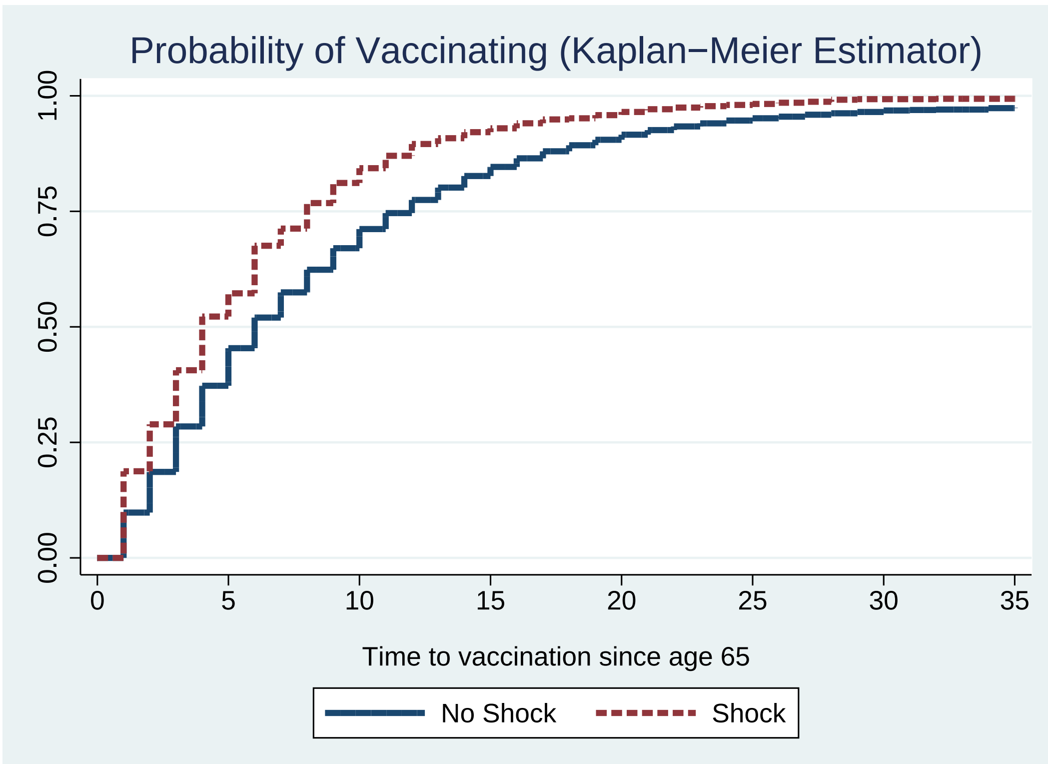

Figure 3.2: Probability of vaccination for persons with and without shocks

Chapter 4

Education and Health: A Peek Inside the Black Box

The notion that there is a relationship between years of schooling and health has a long history in economics (Schultz 1962). Conceptually, the correlation between schooling and health operates through (1) the effect of schooling on wealth which in turn affects health* and (2) through more educated persons being more productive in investing in health (Grossman 1972, Becker 2007).

Productivity in health production may be increased for a combination of reasons. First, educated persons may be more productive in producing good health because they have been knowledge about the health production function. According to this view, holding factors correlated with education such as the price of time, less educated persons would make the same choices as their more educated counterparts if they had better knowledge. It is not that more educated and less educated persons have different preferences, rather it is simply that the educated know more. Another aspect of health knowledge is beliefs about the benefits of treatment or harmful side effects. If more educated persons are more likely to believe in the benefits of treatment then they would be more likely to adhere to recommendations. This implies that disseminating health messages should result in people making better health decisions. More knowl-

*More educated persons secure better jobs and earn higher income. In the U.S. health insurance is frequently provided through employment; more educated persons are more likely to obtain more complete health insurance through their employment (Olson 2002), which is turn reduces the price of medical care at the point of service.

edgeable people are more X-efficient (Leibenstein 1966) because they are better in understanding what the doctor is telling them, they are better at reading the instructions and warning labels on prescription drug packages, they know the signs and symptoms of disease complications and adverse side effects of treatment. So they obtain more health output from a given investment in health inputs. More knowledgeable persons may also make better decisions about input allocations as well, i.e. may be more allocatively efficient. In practice, in this context, there is a fine line between X- and allocative efficiency[†].

Second, preferences which in turn affect choices people make about health care and health behaviors, may be influenced by individuals' educational attainment. Such preferences include time and risk preferences and the value that people attach to being in good health versus other goods (Fuchs 1982; Becker 1996; Becker and Mulligan 1997; Barsky et al. 1997).

Third, education may improve an individual's ability to execute choices. In this sense, the educational process may improve individuals' self control and impulsivity (Goldman and Smith 2002; Laibson 1997; Laibson 2001; Loewenstein and O'Donoghue 2005), the ability to plan, and the ability to carry out plans (Ameriks, Caplin et al. 2003; Khwaja et al. 2007), and through mechanisms not fully understood, reduce the rate of cognitive decline with age (Sloan and Wang 2005).

Fourth, education may neither affect preferences nor self-control nor cognitive functioning, but rather be correlated with some variable influencing both educational attainment and health. A variant of this view is that experiences at birth and during childhood also may affect the amount of education received

[†]For a general discussion of education and health production and the two types of efficiency, see e.g., Grossman (2000).

during youth and health outcomes in later life (e.g., Currie and Hyson, 1999; Costa 2003)[‡]. If so, public policies designed to boost educational attainment of the population would not improve population health, except perhaps to the extent that maternal education improves birth outcomes which in turn improve health in later life.

The polar views are the first-people are ill-informed and the second and part of the fourth. According to the latter view, people are rational and forward-looking. Choices differ because preferences differ. Somewhere in between are perspectives from behavioral economics: At best people are imperfectly rational and differ in the extent to which anticipated future consequences of current actions guide current decisions.

Although there is little conceptual disagreement about the factors that may underlie the observed positive relationship between schooling and health, with some important exceptions, empirical evidence on the underlying mechanisms is lacking. Kenkel (1991) provides empirical evidence that education improves the allocative efficiency of health production. He shows that at least part of the positive correlation between health behaviors (smoking, drinking and exercise) and schooling is partly explained by differences in health knowledge. Education may affect use of new health care technologies, presumably because more educated individuals are among the first to learn about their existence. Lleras-Muney and Lichtenberg (forthcoming) find that among individuals who repeatedly purchase drugs for a given condition, more educated persons are more likely use drugs more recently approved by the U.S. Food and Drug Administration.

[‡]There is a related body of economic research on the relationship between socioeconomic status, of which educational attainment is part, and health. See e.g., Adams et al. (2003).

Certain health behaviors, such as seatbelt use, smoking, heavy alcohol consumption, lack of exercise, and poor diet have negative marginal effects on health and are positively correlated with educational attainment (Fuchs 1982, Leigh 1990; Kenkel 1994; Sloan, Smith, and Taylor 2003; Arcidiacono, Sieg and Sloan 2007).

Educational attainment is correlated with cognitive skills which in turn allows individuals to produce health more efficiently (Auld and Sidhu 2005). Better cognitive ability permits a better understanding of the complex health care regimens commonly recommended for diabetics. Both education and health might be affected by genetic factors or childhood and family background. For example, individuals with good genes might be more educated since in childhood they would have missed fewer school days due to sickness. Or, individuals from disadvantaged backgrounds (e.g. low maternal education) might be less educated and unhealthier because of low parental investments on both dimensions (Thomas et al. 1991, (Almond 2006). Since past health has a significant effect on current health we would observe a correlation between current health and education[§]. To learn more about these mechanisms, this study focuses on near elderly and elderly persons who have been diagnosed with diabetes mellitus, using longitudinal data from the Health and Retirement Study (HRS). We use data from a supplement to the HRS, the Diabetes Study (HRS-DS), administered to sample persons who had previously reported having been told by a physician that they had diabetes conducted by mail during late 2003 and early 2004. The HRS-DS obtained detailed information on general preferences and beliefs and preferences and health knowledge as well as those pertaining to diabetes, and on self-care practices and on utilization of personal health

[§]The HRS includes information on parental education and parental age at death. We use these as measures of family background and genetic factors, respectively.

services. Especially when combined with the data on educational attainment, cognitive status, health, and individual attributes on the same persons from the HRS which was administered in every even-numbered year from 1992-2006, we have an unusually rich amount of information on actual consumer decisions and factors underlying these decisions.

In this study, we focus on three main research questions. One, is the effect of education on longevity and health status. Second, we study the relationship between educational attainment and various factors such as knowledge, preferences, health behaviors and social interactions. Finally, we examine whether the observed positive correlation between education and health can be explained, at least partly, by differences in the various factors that we consider.

4.1 Background on Diabetes

Diabetes is a chronic disease that affects approximately 16 million persons in the U.S.[¶] . It was the sixth leading cause of death in U.S. in 2005[‖]. Focusing on diabetes offers several advantages for analysis of the relationship between years of schooling and health. First, and most importantly, diabetes frequently results not only in earlier death but in eye, kidney, heart, and peripheral vascular disease, and stroke. There is considerable empirical evidence that by taking certain precautions, death and disease can be at least postponed. Yet the care regimens are complex and burdensome to those with the disease making health investments costly. The regimens include diet and exercise, regular testing, and awareness of when to seek medical health. To the extent that more educated

[¶]Source: http://www.cdc.gov/diabetes/statistics/prev/national/figpersons.htm, accessed 18 April 2008

[‖]Source: National Vital Statistics Reports, Vol. 56 No. 10

persons are more efficient in health production in general, they are very likely to be more efficient in health production if they have diabetes in particular.

A negative relationship between educational attainment and diabetes prevalence as well as in the fraction of persons with undiagnosed diabetes has been documented (Smith 2007). Also, there is evidence that educational attainment relates to the course of the disease. Goldman and Smith (2002) report that among persons with self-reported diabetes, self-reported health deteriorated more over an eight year period among persons with lower than higher educational attainment. They find that less educated persons were less good in diabetes self-management, although several parameter estimates on educational attainment variables are not quite statistically significant at conventional levels. Less educated persons in a well-known randomized clinical trial benefit more than more education persons from the treatment protocols, presumably because they would be less likely to manage their disease effectively without the guidance of treatment protocols that accompany a randomized clinical trial.

4.2 Data

The HRS is a panel study of persons in the U.S. in birth cohorts 1931 through 1941 and their spouses if married who could be of any age. The HRS oversamples blacks, Hispanics and residents of Florida. Participants were interviewed every two years from 1992-2006. In 1998, a cohort of persons born during 1942-47 was added to HRS. Assets and Health Dynamics among the Oldest Old (AHEAD), which collected information on persons born in 1913-23, was initially conducted in 1993. Sample persons were re-interviewed in 1995. In 1998, the HRS and AHEAD samples were combined. Since that year, HRS and AHEAD sample interviews have received identical interviews which are conducted biennially.

The HRS-DS was conducted by mail with a sample of person reporting diabetes in the 2002 wave. HRS-DS was administered in 2 stages. First, a self-administered questionnaire about diabetes care, self-management, and health care utilization was obtained from respondents. Second, a mail-in kit with a finger-stick dried blood spot sample to measure levels of clinical biomarker of glucose control, glycosylated hemoglobin (HbA1c) was obtained from respondents. HbA1c is the most widely used summary measure of diabetes control although other measures are also relevant, including blood pressure and cholesterol. Of the 3,194 respondents reporting a diagnosis of diabetes in 2002, 2,341 persons were eligible to participate in the HRS-DS after excluding individuals for death or random assignment to another study. Of the eligible cases, 1,901 returned a questionnaire and 1,233 had valid laboratory data.

We construct 2 analysis samples. The first ("baseline sample") consists of all respondents to the 1992 wave of the HRS who were 50 to 65 years old in that year and reported that a physician had told them that they had diabetes (N=1,067). The second ("diabetes survey sample") includes all respondents to the HRS-DS (N=1,901). Summary statistics for the 2 samples are presented in Table 1. On average, persons in the baseline sample were much younger than those in the diabetes survey sample (56 and 69, respectively, on average). For this reason, a much higher share of individuals had some form of health insurance coverage in the diabetes survey than in the baseline sample.

The key explanatory variables in this analysis are number of years of schooling and cognitive status. In addition, we control for age, gender, black and Hispanic race/ethnicity, household income, and lack of private or public health insurance.

The cognitive score is a sum of 4 separate measures of cognitive status. These are measures of immediate and delayed word recall, a working memory measure based on serial 7s subtraction and a modified version of the Telephone Interview for Cognitive Status or TICS score (Brandt, Spencer et al. 1988). The word recall test required respondents to remember as many words as possible from a list of 10 words that were read to them by the interviewer. Respondents were asked to recall the words immediately and 5 minutes after the list was read to them. The serial 7 subtraction test required respondents to subtract 7 from 100 five times. They were assigned a value of one for each correct subtraction. The TICS score was a count of correct responses to various questions that measured knowledge, language and orientation. Respondents were asked the month, day, year and day of the week of the interview. They were asked to name the following: the thing used to cut paper, the prickly plant that grows in the desert, the current U.S. president and the current U.S. vice president. Respondents were also asked to count backwards starting from 20 to 10. Respondents were assigned a value of 2 for counting correctly the first time, 1 for counting correctly on the second try and 0 for an incorrect count both times. The cognitive score for the HRS sample is constructed in a different manner. It is the sum of 3 measures–the immediate and delayed memory recall measures and a measure of language. Respondents were asked to characterize relationships between words, i.e., an orange and a banana, a table and a chair, an eye and a ear, an egg and a seed, air and water, a fly and a tree and praise and punishment. They were assigned a value of 2 for a fully correct answer, 1 for a partially correct answer and 0 for an incorrect answer.

Next we describe the dependent variables, which are grouped by major categories.

4.2.1 Dependent Variables

Knowledge and Beliefs

Knowledge We construct a general measure of diabetic-specific knowledge. The measure is a count of the number of "don't knows" and "uncertains." Individuals are assigned a value of 1 for the item if they answered "don't know" or "uncertain" to questions concerning any of the following: age at which the person was told by a doctor that s/he had diabetes; type of diabetes (type 1 or type 2); value of the person's last HbA1c test-a measure of the person's blood glucose level and the most widely used single indicator of diabetes control; whether or not the person ever had diabetes education other than during routine doctor or nurse visits; the number of days in the last month the person had symptoms of low or high blood sugar; when the person's last blood pressure reading was performed; what the person's blood pressure reading was; and when the last cholesterol test was performed.

Understanding DM Care Processes The second knowledge measure gauges the individual's understanding of diabetes case processes. Respondents were asked to rate on a 4-point scale how well s/he understood each of the following areas of diabetes care: what each of his or her prescribed medications do; how to choose the food s/he should eat, how to read nutrition labels on food; and how to exercise. A rating of 1 represented "I don't understand it at all" and a rating of 4 was for "I understand completely." We create binary variables set to 1 if the respondent replied either "I understand pretty well," or "I understand completely,, and is 0 otherwise. The dependent variable is the sum of these values.

Beliefs Regarding Effectiveness of Diabetes Treatment The HRS-DS asked how much respondents agreed or disagreed with a series of statements to the effect that following a prescribed diabetes treatment plan closely makes a big difference in preventing various DM complications: eye problems; kidney problems; foot problems; hardening of arteries; heart disease; and stroke. The responses were elicited on a 5-point Likert scale with 1 representing "strongly disagree" and 5 representing "strongly agree." We create binary variables set to 1 if the person responded "agree" or "strongly agree," with the statement and is 0 otherwise. Our measure is the sum of these binary variables. Beliefs Regarding Importance of Following DM Treatment Plan. The HRS-DS asked how much the respondent agreed or disagreed with these statements: it is important for the person to follow the meal plan carefully, take medicine as recommended, take care of the person's feet, get enough physical activity, test blood sugar as recommended, go to the doctor or nurse for all appointments, and keep weight under control. Again, we convert the 5-point Likert response to binary variables taking the value 1 if the person responded "agree" or "strongly agree," and 0 otherwise. The final dependent variable is the sum of these values.

Preferences

Time Preference Respondents to the HRS-DS were asked whether they agreed or disagreed with the following statement: "I live life one day at a time and don't think much about the future." Responses were elicited on a 5-point scale with 1 representing "strongly disagree" and 5 representing "strongly agree." We create a binary variable with 1 for "agree" or "strongly agree" and 0 representing other responses. Thus, persons with a value of 1 to this question are relatively impatient. In contrast to the two preference measures described below, the HRS-DS question is not domain-specific (i.e., health, financial, other domain).

Financial Planning Horizon The HRS asks the following question regarding their financial horizon: "In deciding how much of their (family) income to spend or save, people are likely to think about different financial planning periods. In planning your (family's) savings and spending, which of the time periods listed in the booklet is most important to you [and your (husband/wife/partner)]?" We create a binary variable to represent a short financial horizon, which is 1 if the person's horizon was a year or less.

Risk Preference To measure risk aversion, the HRS asks respondents to choose among 4 different gambles:

"Suppose you are the only income earner in the family, and you have a good job You are given the opportunity to take a new and equally good job, with a 50-50 chance that it will double your income and a 50-50 chance that it will reduce your income by a third. Would you take the new job?" If the answer was "no", the interviewer asked: "Suppose the chances were a 50-50 chance that it would double your income and a 50-50 chance that it would cut your income by 20 percent. Would you still take the new job?" If the answer to the first question was "yes", the interviewer asked: "Suppose the chances were a 50-50 chance that it would double your income and a 50-50 chance that it would cut your income by half. Would you still take the new job?"

Following (Barsky, Juster et al. 1997), we construct a measure of risk tolerance that took the value 0.15 if the sample person rejected both the one-third and one-fifth gambles, 0.28 if the person rejected the one-third gamble but accepted the one-fifth gamble, 0.35 if the person accepted the one-third gamble but rejected the one-half gamble and 0.57 if the person accepted both the one-third and one-half gambles. If a value for risk preferences was not available for the respondent, we set the risk tolerance variable to 0 and defined a separate explanatory variable for missing equal to 1. Barsky et al. (1997) show that this measure of risk preference is related to a number of behaviors. (Khwaja, Sloan et al. 2006) find that smokers are more risk tolerant on average.

Measures of Self-Control

Overall Self-Control The HRS-DS asked how much the respondent agreed or disagreed with the following statement "I have little control over things that happen to me". Responses were elicited on a 5-point scale with 1 representing "strongly disagree" and 5 representing "strongly agree." We create a binary variable that took the value one if the person responded "agree" or "strongly agree" and zero otherwise.

Self-Confidence Respondents were also asked how much they agreed or disagreed with the following statement "I can do just about anything I set out to do." Responses were also to be made on a 5-point scale from which we construct a binary variable with one representing "agree" or "strongly agree" and zero representing other responses.

Self-care activities in Last 7 Days The HRS-DS asked on how many days of the last seven days did respondents: take their recommended insulin or diabetes pills; take all of their recommended doses of insulin or number of diabetes pills; follow a healthful eating plan; eat five or more servings of fruits and vegetables; eat high fat foods such as red meat or full-fat dairy products; eat two or more servings of snack or dessert foods such as chips, cookies, cake or pie; and test their blood sugar as often as their doctor has recommended. We create binary variables for each self-care activity taking the value of 1 if the individual responded that for s/he did the task for 5+ for at least 5 days per week, and is 0 otherwise. The measure we use is the sum of these binary variables. Since some of two activities, eating red meat and snacking-eating deserts contribute to poor health, these items enter with a negative sign. To make the index nonnegative for each respondent, we add 2 to the sum.

Difficulty in following the Doctor's Recommendations HRS-DS respondents were asked over the last 6 months, how difficult it has been to do the following tasks exactly as suggested by their doctors: take diabetes medications (pills and/or insulin); exercising regularly; follow recommended eating plan; check blood sugar; check feet for wounds or sores; take medication for blood pressure; and see their doctors or other providers. Answers are on a 5-point scale with 1 representing "So difficult that I couldn't do it all" and 5 representing "Not difficult: I got it exactly right." Our measure is a count of affirmative responses.

Sources of Information and Social Networks

Social Support The HRS-DS asked how much the respondent agreed or disagreed with the statement that s/he can count on his/her family or friends to help and support him a lot with following his meal plan, taking his medicine, taking care of his feet, getting enough physical activity, testing his sugar, going to the doctor or nurse, keeping his weight under control and handling his feeling about diabetes. We create binary variables with the value 1 if the person responded "agree" or "strongly agree," which is set to 0 otherwise and sum the values to obtain a score for social support.

Patient Input during Physician Visit HRS-DS respondents were asked how often, over the past 6 months, in discussions with their physicians, they were asked for ideas about making their treatment plans, given choices about treatment to think about, asked to talk about goals in caring for their diabetes, helped to set specific goals to improve their eating or exercise, sure that their doctor or nurse thought about their values and traditions when they recommended treatment, helped to make a treatment plan that they could do in their daily life and helped to set a goal with their doctor or nurse. Responses were made on a 5-point scale with 1 representing "never" and 5 for "very often." We created a binary variable for each answer, with 1 representing a response of "often" or "very often" and zero representing "never", "rarely" or "sometimes." The variable is a count of these responses.

Content of Discussion with Doctor Individuals were asked if their doctor or nurse discussed the following aspects of care in the past six months: when and how to take insulin or diabetes pills; when and how to check blood sugar; how to time meals; what to eat; how to check and care for their feet; how to increase physical activity; how to make changes in medications; how to deal with emotional demands of diabetes; and where to find community resources to help with diabetes. We construct a count of topics that were discussed with the respondent's physician during the past 6 months.

Sources of Information About Diabetes and Diabetes Care HRS-DS respondents were asked how often they got diabetes information from the following sources: television, internet, newspapers or magazines, books and friends. They were also asked how often they bring up with their doctor any information they heard or saw that might affect their treatment. Responses were elicited on a 5 point scale, with 1 representing "never" and 5 representing " very often.. We create binary variables with the value 1 if the person answered "sometimes," "often," or "very often, and 0 otherwise. Each of these sources of care is analyzed separately.

4.3 Statistical Analysis

To estimate the effects of education attainment on longevity and health, we use the baseline sample. We use a log-logistic accelerated failure time (AFT) model to estimate the effect of educational attainment on longevity and health. The model incorporates a gamma frailty term that accounts for unobserved heterogeneity at the individual level. The log-logistic model is a flexible model that permits non-monotonic hazard functions and preliminary analysis suggested that the hazard rate was not monotonically increasing or decreasing for this sample. Time to failure is defined alternatively as the time from 1992 to a major diabetes-related health shock, or to death, whichever occurred first. Individuals who do not report any DM related health condition and who survive till the end of the survey period (2006) are right censored. Diabetes-related health shocks include the first report of blindness, stroke, heart condition (heart attack, coronary heart disease, angina, congestive heart failure or other heart problems) or onset of fair or poor health. In addition to educational attainment, measured as years of schooling completed, we control for age, black, Hispanic, marital status, household income, lack of health insurance, cognitive score, parental longevity measured as age at death or current age if alive.

The hazard model excludes persons who have already experienced a major health shock or those who died by the end of 1992. To analyze the effect of this selection, we ran a logit regression where the dependent variable took the value one if the person was excluded from the hazard regression, and zero otherwise. We also present the results for a regression model in which failure time is defined as time from 1992 to death or to 2006 if the person survives. We use logit analysis for binary variables and Poisson regression for count variables.

4.4 Results

4.4.1 Effect of Educational Attainment on Survival and Health

Educational attainment has no effect on survival but persons with lower cognition at baseline (1992) experienced higher rates of survival through 2006 (Table 4.2, col. 1). Even after excluding cognitive status from the analysis (not shown), the parameter estimate for educational attainment was not statistically significant at conventional levels. Blacks and Hispanics were less likely to survive through 2006. As anticipated, older persons and men tend to die earlier.

The second column in table 4.2 shows that baseline health status varies considerably across indivudals. More educated persons as well as those with a higher cognitive score were much less likely to report a DM related health shock or fair or poor health. Older persons and higher income individuals were also less likely to have experienced such DM-related shocks by the baseline HRS interview.

Eliminating persons who had reported having a major diabetes related health condition at baseline (60.0% of the sample used in the analysis of survival) for our analysis of time to death or to onset of a health condition whichever occurred first (cols. 3-6), educational attainment decreases onset of health condition or death in all regressions. The parameter estimate on educational attainment is rather insensitive to changes in specification. Higher cognitive status at baseline improves health and survival. Neither age nor household income have effects on health and survival. However, persons who lacked health insurance coverage at baseline experienced worse health/survival. Longevity of parents has no effect on health/survival.

4.4.2 Knowledge

Among the sample persons who were interviewed in the 2003 supplemental survey, those with higher educational attainment were more likely to be diagnosed with the disease later conditional on current age (Table 4.3). This result is fairly robust to the inclusion of other explanatory variables. It is evident that although all sample persons eventually got the disease, being more highly educated helped to postpone disease onset.

Overall, the results in Tables 4.2 and 4.3 are consistent with the findings of previous literature reporting a positive relationship between health and schooling for persons diagnosed with diabetes (Smith 2002) as well as findings from studies of more general populations which find this type of correlation. However, we find a stronger relationship between health and educational attainment than for survival and education.

As reported by others for other HRS samples (e.g., Sloan and Wang), there is a positive relationship between cognition at baseline and educational attainment in the HRS (Table 4.4, col. 1). Black race and Hispanic ethnicity are associated with lower cognition, holding other factors constant. Older persons with DM have a higher cognitive score than their younger counterparts.

More educated individuals score higher on each of the 4 components of the HRS measure of cognitive status (cols. 2-5). The higher cognitive status on age masks different relationship among the components of the total score. Older persons do better on the knowledge component (TICS), but experience worse memory with age.

With or without inclusion of cognition defined here for 2002, more educated persons give fewer "don't know" responses to questions about their level of diabetes control and about whether or not they have received specific tests which are essential for good diabetes care (col. 1). Inclusion of cognition reduces the effect of educational attainment on the number of don't knows, implying that part of the result for education is accounted for by cognitive status at the time the HRS-DS was conducted. By contrast, cognition accounts for much of the relationship between education and understanding of basic DM care processes, the parameter estimate for which is statistically significance without inclusion of cognition as an explanatory variable (col. 2). Educational attainment has no effect on either of the belief variables with out without inclusion of cognition (cols. 3 and 4). Parameter estimates on cognition also lack statistical significance at conventional levels. These results suggest that while beliefs regarding treatment do not vary systematically across different education groups, more educated persons do have better knowledge about DM related care.

4.4.3 Preferences

Using the HRS-DS measure of time preference, more educated persons are more patient than others (Table 4.6, cols. 1 and 2). The parameter estimate on educational attainment does not materially change when the cognitive score and other covariates are included. Likewise, more educated persons tend to have a longer planning horizon and are more risk tolerant. On risk tolerance, it is noteworthy that educational attainment seems to be a proxy for household income when the latter variable is excluded as an explanatory variable. Not surprisingly, more affluent individuals are less risk averse. Also, the HRS measure of risk tolerance measures risk preferences from risk neutrality to risk aversion. As the questions were constructed, no respondent can be a risk lover. Whether or not education causes differences in preferences cannot be determined from these data. A conservative interpretation is that these preferences are correlated with educational attainment.

4.4.4 Self-Control and Self Management

The parameter estimates on educational attainment are negative in the analysis of self-control (Table 4.7, col.1) and self-confidence (col.2), implying that individuals who have completed more schooling believe that they have more control over their life outcomes than do others. However, the result for self-confidence is sensitive to inclusion of the explanatory variable for cognitive status. The parameter estimate is only statistically significant when cognition is included in the analysis. There is no relationship between educational attainment and cognition and either of the two measures of diabetes self-management (cols. 3 and 4).

When asked reasons for missing an important form of diabetes self-management-testing one's own blood sugar level- more educated persons were more likely to say "I forgot" (not shown) and slightly less likely to state that they "cannot do this alone." The reason "I forgot" suggests the importance of higher opportunity costs of tests and more highly educated persons' activity level rather than being a measure of self-control, even though the parameter estimate for educational attainment in the "no time" regression is not statistically significant at conventional levels.

4.4.5 Sources of Information and Social Networks

More highly educated persons are more likely to obtain information about diabetes from the internet, newspapers and magazines, and from their physicians and less likely to receive information from friends, than are those with lower educational attainment (Table 4.8). More educated persons were also more likely to discuss any new information they had obtained with their doctor (col. 6).

More educated persons are less likely to be able to rely on family or friends to help support them (Table 4.9, col. 1). The result in Table 4.8 that more educated persons are less likely to rely on friends for diabetes information is consistent with this finding. Without inclusion of cognition, educational attainment does not have statistically significant effect on the probability that the physician asked the person for information about treatment plans during visits (col. 2). However, when cognition is included, education has a statistically negative coefficient, implying that more educated persons are less likely to be asked for their input in planning the course of treatment. Health care providers are also less likely to discuss various aspects of care with more educated persons.

Thus, overall, while information sources and social networks differ according to the person's educational attainment, there is no clear evidence that the educated have more active interaction with their physicians. There are clear differences in information sources, but this may be attributable to greater reliance on written sources of information among the more highly educated in general.

4.4.6 Health Choices

More educated persons were more likely to miss blood sugar testing (Table 4.10, col. 2), but were more likely to have tried to lose weight (col. 4), holding other factors including a binary variable identifying obese persons, and to engage in exercise (col. 5). With cognition included, differences in the probability of being a current smoker are not statistically significant at conventional levels (col. 6), but dropping cognitive, the parameter estimate for smoking becomes negative and statistically significant at conventional levels. Cognition is positively related to having an annual HbA1c test, trying to lose weight and exercising. Cognition is negatively related to the probability of being a current smoker. However, when we examined relationships between components of cognition and the probability of being a smoker (not shown), it was the memory components, not the knowledge component or the subtraction test that mattered. There is some empirical evidence that smoking impairs memory (Richards et al, 2003). Thus, cognition is likely to be endogenous to being a current smoker. Thus, higher levels of education and cognition appear to be associated with some healthy behaviors though not with all the ones we examine.

4.4.7 The Contribution of Educational Attainment to Health

In table 4.11, we analyze the contribution of various factors to overall self-reported health. Taken alone, an additional year of schooling reduces the probability of being in fair or poor health by 0.042 (Table 4.11). Controlling for variables such as age, gender, race, income and insurance status reduces the coefficient on education to -0.33. This implies that part of the overall effect of education on health can be explained through differences in income, insurance and other demographic variables. Including other variables further reduces the coefficient on education. When all of the factors listed in Table 4.11 are included in a regression for the probability of being in fair or poor health, the coefficient on educational attainment changes to -0.011 (not shown). Much of the reduction appears to be attributable to inclusion of measures of self-control as explanatory variables. When these measures are included as explanatory variables in addition to educational attainment and demographics, the marginal effect of an additional year of schooling drops to -0.020. In particular, self-confidence has a large effect on fair or poor health.

Consistent with previous literature, we find a positive correlation between educational attainment and health. However, in this study, the link appears to be stronger between self-reported health and years of schooling than it is between survival and schooling. Educational attainment affects some but not all the health behaviors we evaluated. There is a strong positive relationship between educational attainment and cognitive status of these near elderly and elderly persons who have been diagnosed with diabetes. In some cases, educational attainment appears to serve as a proxy for cognition. More educated persons tend to think they have more self-control and self-confidence about the conduct of their lives. Surprisingly, we do not find a relationship between diabetes self-management and educational attainment. We also find that interaction with health care providers does not appear to benefit more educated persons disproportionately. Our results suggest that part of the positive correlation between education and health can be explained through differences in some of the factors we have considered.

Table 4.1: Summary Statistics

Panel A: Explanatory Variables	HRS Sample	Std. Error	HRS-DS Sample	Std. Error
Educational Attainment	11.03	3.50	11.52	3.54
Cognitive Score	16.81	6.55	20.41	4.92
Age	56.69	3.81	69.80	8.96
Male	0.50		0.47	
Black	0.29		0.19	
Hispanic	0.13		0.11	
Married	0.71		0.63	
Household Income ('0000$)	3.70	3.49	4.04	5.35
No health insurance	0.20		0.06	
N	1,067		1,901	
Panel B: Dependent Variables	HRS-DS Sample	Std. Error		
Frequency of Don't Know or Uncertain	2.22	1.92		
Understanding DM Care Processes	3.35	1.09		
Beliefs - Effectiveness of Treatment	4.98	1.96		
Beliefs - Importance of Following Treatment Plan	7.21	1.55		
Time Preference	0.41			
Short Financial Planning Horizon	0.35			
Risk Tolerance	0.23	0.13		
Overall Self-Control	0.31			
Self-Confidence	0.55			
Self-care Activities in Last 7 Days	5.40	1.38		
Difficulty in Following Dr's Recommendations	5.63	1.36		
Social Support	5.89	2.86		
Patient Input During Physician Visit	2.11	2.47		
Doctor Discussion	5.20	2.94		
HbA1c+Uranalysis+Eye Exam	2.33	0.82		
Miss Blood Sugar Testing	0.38			
Take Aspirin Daily	0.57			
Tried to Lose Weight	0.56			
Exercise	0.58			
Smoking	0.10			
Fair/ Poor Health in 2003	0.45			
N	1,901			

Cognitive score in the HRS sample is constructed differently from that in the HRS-DS sample (see text)
Household income is normalized to 2003 US Dollars

Table 4.2: Log-logistic Accelerated Failure Time Model with Gamma Frailty

	Survival Only	Shock at Baseline	Survival/ Major Health Shock Onset			
	(1)	(2)	(3)	(4)	(5)	(6)
Educational attainment	-0.007	-0.071***	0.055***	0.053***	0.039**	0.041*
	(0.013)	(0.025)	(0.017)	(0.019)	(0.020)	(0.021)
Standardized cognitive score	0.243***	-0.390***			0.197**	0.192**
	(0.063)	(0.115)			(0.081)	(0.084)
Age	-0.022**	-0.042**		0.001	0.006	0.006
	(0.010)	(0.018)		(0.014)	(0.014)	(0.014)
Male	-0.345***	0.308**		0.092	0.135	0.165
	(0.078)	(0.144)		(0.108)	(0.108)	(0.110)
Black	0.200**	0.184		-0.021	0.087	0.151
	(0.090)	(0.165)		(0.126)	(0.132)	(0.136)
Hispanic	0.344***	0.121		0.025	0.100	0.167
	(0.127)	(0.237)		(0.206)	(0.204)	(0.212)
Married	0.324***	-0.112		0.142	0.144	0.144
	(0.088)	(0.174)		(0.124)	(0.123)	(0.125)
Household income	0.051***	-0.135***		-0.004	-0.008	-0.011
	(0.015)	(0.032)		(0.014)	(0.014)	(0.017)
No health insurance	0.134	-0.138		-0.324**	-0.304**	-0.275**
	(0.094)	(0.173)		(0.132)	(0.131)	(0.135)
Mother's longevity						0.006
						(0.004)
Father's longevity						0.004
						(0.004)
Constant	3.726***	4.090***	1.116***	1.056	0.836	0.150
	(0.593)	(1.085)	(0.221)	(0.860)	(0.856)	(0.960)
N	1,051	1,054	418	414	414	400

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level
Standard errors in parentheses
Regressions in columns (1), (2) , (5) and (6) also include covariate for cognitive score missing

Table 4.3: Relationship between educational attainment and age of diagnosis

Education	N	Controls
0.076	1,502	None
(0.109)		
0.357***	1,502	Age
(0.092)		
0.361***	1,501	Age, Male, Black, Hispanic, Married
(0.106)		
0.328***	1,499	Age, Male, Black, Hispanic, Married, Income, No health insurance
(0.107)		

Robust Standard errors in parentheses
*** Significant at 1% level

Table 4.4: Cognitive Status for HRS-DS Sample

	Cognitive Score (1)	Immediate Memory (2)	Delayed Memory (3)	Serial 7 Subtraction (4)	TICS (5)
Educational Attainment	0.435***	0.168***	0.173***	0.047***	0.067***
	(0.042)	(0.013)	(0.015)	(0.008)	(0.022)
Age	0.188***	-0.051***	-0.057***	-0.005*	0.262***
	(0.013)	(0.005)	(0.005)	(0.003)	(0.008)
Male	-0.765***	-0.483***	-0.526***	0.169***	0.105
	(0.242)	(0.079)	(0.092)	(0.053)	(0.137)
Black	-1.978***	-0.484***	-0.582***	-0.393***	-0.535***
	(0.312)	(0.097)	(0.116)	(0.070)	(0.190)
Hispanic	-1.075**	-0.111	-0.201	-0.218**	-0.226
	(0.433)	(0.134)	(0.160)	(0.095)	(0.242)
Married	0.009	0.003	0.030	-0.088	0.290**
	(0.255)	(0.084)	(0.098)	(0.057)	(0.139)
N	1,457	1,651	1,621	1,896	1,699

Robust Standard errors in parentheses
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level
Marginal Effects from Poisson Regression

Table 4.5: Knowledge and Beliefs

	Frequency of Don't Know or Uncertain	Understanding DM Care Processes	Beliefs	
			Effectiveness of Treatment	Importance of Following Treatment Plan
	(1)	(2)	(3)	(4)
Educational attainment	-0.079***	0.029	0.002	-0.021
	(0.017)	(0.027)	(0.019)	(0.015)
Cognitive score	-0.041***	0.048***	0.015	0.006
	(0.010)	(0.016)	(0.012)	(0.010)
Age	0.021***	-0.016	-0.022***	0.000
	(0.006)	(0.010)	(0.007)	(0.006)
Male	-0.023	-0.389***	-0.257**	-0.146*
	(0.094)	(0.151)	(0.107)	(0.087)
Black	0.256**	0.447**	-0.021	0.357***
	(0.121)	(0.198)	(0.143)	(0.104)
Hispanic	-0.023	-0.692***	-0.077	-0.023
	(0.166)	(0.268)	(0.212)	(0.181)
Married	-0.048	0.218	-0.076	0.102
	(0.100)	(0.158)	(0.116)	(0.093)
Household income	-0.040**	0.000	0.011*	0.006
	(0.019)	(0.008)	(0.006)	(0.004)
No health insurance	0.078	-0.255	-0.013	-0.242
	(0.205)	(0.322)	(0.225)	(0.222)
N	1,457	1,255	1,346	1,240
Without cognitive score				
	Frequency of Don't Know or Uncertain	Understanding DM Care Processes	Beliefs	
			Effectiveness of Treatment	Importance of Following Treatment Plan
	(1)	(2)	(3)	(4)
Educational attainment	-0.110***	0.094***	0.026	0.006
	(0.013)	(0.023)	(0.017)	(0.013)
N	1,894	1,623	1,723	1,617

Robust Standard errors in parentheses
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 4.6: Time Preference, Planning Horizon, and Risk Tolerance

	Time preference		Financial Planning Horizon			Risk Tolerance		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Educational attainment	-0.016***	-0.017***	-0.024***	-0.022***	-0.013*	0.005***	0.003*	0.003*
	(0.005)	(0.004)	(0.006)	(0.007)	(0.008)	(0.002)	(0.002)	(0.002)
Cognitive score	-0.008**			0.003	0.004			-0.003**
	(0.003)			(0.006)	(0.006)			(0.002)
Age	0.008***	0.007***			0.002			
	(0.002)	(0.002)			(0.005)			
Male	-0.023	0.008						
	(0.029)	(0.026)						
Black	-0.013	-0.007						
	(0.037)	(0.032)						
Hispanic	0.060	0.043						
	(0.053)	(0.045)						
Married	-0.061**	-0.064**						
	(0.031)	(0.028)						
Household income	-0.017***	-0.018***			-0.019***		0.002**	0.003**
	(0.005)	(0.005)			(0.007)		(0.001)	(0.001)
No health insurance	0.038	0.100						
	(0.058)	(0.054)						
N	1,408	1,814	591	527	527	569	569	512

Robust Standard errors in parentheses

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 4.7: Self-Control and Self-Management

	Overall Self-Control	Self-Confidence	Self-care Activities in Last 7 days	Difficulty in following Dr's Recommendations
Educational attainment	-0.014***	-0.016***	0.017	-0.029
	(0.005)	(0.005)	(0.017)	(0.019)
Cognitive score	-0.007**	0.012***	0.012	0.012
	(0.003)	(0.003)	(0.010)	(0.011)
Age	0.005***	-0.003*	0.022***	0.030***
	(0.002)	(0.002)	(0.006)	(0.007)
Male	-0.067**	-0.003	-0.451***	0.113
	(0.027)	(0.029)	(0.096)	(0.097)
Black	0.042	0.112***	-0.010	0.165
	(0.036)	(0.036)	(0.125)	(0.117)
Hispanic	0.063	0.050	0.015	-0.026
	(0.052)	(0.050)	(0.199)	(0.200)
Married	-0.015	-0.002	0.190*	0.154
	(0.028)	(0.031)	(0.101)	(0.101)
Household income	-0.004	0.017***	-0.001	-0.008
	(0.004)	(0.005)	(0.006)	(0.012)
No health insurance	0.043	0.104*	-0.102	-0.198
	(0.058)	(0.056)	(0.217)	(0.246)
N	1,363	1,408	942	816
	Without cognitive score			
Educational attainment	-0.015***	0.004	0.033**	0.016
	(0.004)	(0.004)	(0.014)	(0.017)
N	1,765	1,815	1,216	1,060

Robust Standard errors in parentheses
*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 4.8: Source of Information

	Television	Internet	Newspapers/ Magazines	Books	Friends	Discuss with doctor
Educational attainment	-0.006	0.024***	0.018***	0.009	-0.014**	0.010**
	(0.006)	(0.004)	(0.005)	(0.006)	(0.006)	(0.005)
Cognitive score	0.002	0.006**	0.006	0.002	0.0001	0.002
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Age	-0.001	-0.011***	-0.003	-0.003	-0.007***	-0.006***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Male	-0.017	-0.030	-0.079	-0.099***	-0.036	-0.029
	(0.031)	(0.023)	(0.030)	(0.031)	(0.031)	(0.028)
Black	0.296***	0.034	0.118	0.212***	0.151***	0.034
	(0.034)	(0.034)	(0.036)	(0.036)	(0.041)	(0.037)
Hispanic	0.264***	0.040	0.089	0.088*	0.140**	-0.025
	(0.048)	(0.058)	(0.050)	(0.053)	(0.058)	(0.052)
Married	-0.029	0.008	-0.062	-0.006	-0.090***	-0.058*
	(0.034)	(0.024)	(0.031)	(0.033)	(0.034)	(0.030)
Household income	-0.007*	0.002	-0.002	0.002	0.002	0.002
	(0.004)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
No health insurance	0.037	-0.023	-0.091	-0.005	-0.093	-0.018
	(0.068)	(0.039)	(0.066)	(0.066)	(0.064)	(0.060)
N	1,241	1,035	1,247	1,210	1,194	1,430

Robust Standard errors in parentheses

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 4.9: Social Networks and Patient-Physician Interactions

	Social Support	Patient Input during Physician Visit	Doctor Discussed
Educational attainment	-0.122***	-0.062***	-0.100***
	(0.027)	(0.024)	(0.028)
Age	0.041***	-0.024***	-0.023**
	(0.011)	(0.009)	(0.011)
Male	0.399**	0.161	0.113
	(0.177)	(0.141)	(0.173)
Black	0.769***	0.824***	0.908***
	(0.235)	(0.216)	(0.232)
Hispanic	0.608**	0.583**	0.718**
	(0.272)	(0.293)	(0.319)
Married	0.770***	-0.007	-0.271
	(0.193)	(0.149)	(0.180)
Household income	-0.032*	0.012**	0.008
	(0.019)	(0.006)	(0.010)
No health insurance	-0.136	-0.722***	-0.513
	(0.388)	(0.237)	(0.363)
Cognitive score	-0.009	0.014	0.001
	(0.018)	(0.016)	(0.019)
N	1,223	1,252	1,253
Without cognitive score			
	Social support	Patient Input during Physician Visit	Doctor discussed
Educational attainment	-0.103***	-0.027	-0.097***
	(0.021)	(0.019)	(0.022)
N	1,590	1,609	1,616

Robust Standard errors in parentheses

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 4.10: Health Choices

	HbA1c, Ur, Eye (1)	Miss blood sugar testing (2)	Take aspirin daily (3)	Tried to lose weight (4)	Exercise (5)	Smoking (6)
Educational attainment	0.006	0.011**	-0.005	0.009*	0.015***	-0.003
	(0.008)	(0.005)	(0.004)	(0.005)	(0.005)	(0.002)
Cognitive score	0.013***	-0.002	-0.003	0.013***	0.006*	-0.003**
	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Age	0.005*	-0.010***	0.003*	-0.012***	-0.002	-0.006***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Male	-0.095**	0.076**	-0.015	-0.116***	0.045	0.036**
	(0.047)	(0.030)	(0.022)	(0.030)	(0.028)	(0.014)
Black	0.083	-0.065*	-0.047	-0.076*	0.046	-0.005
	(0.063)	(0.038)	(0.035)	(0.041)	(0.038)	(0.016)
Hispanic	0.133	0.047	-0.134**	0.036	0.121**	-0.027
	(0.091)	(0.057)	(0.058)	(0.054)	(0.048)	(0.019)
Married	0.080	-0.047	-0.020	0.002	0.014	-0.048***
	(0.050)	(0.032)	(0.023)	(0.033)	(0.030)	(0.017)
Household income	0.006**	0.000	0.003	0.005**	0.001	-0.002
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
No health insurance	-0.207**	0.073	0.025	0.043	0.031	0.024
	(0.108)	(0.068)	(0.040)	(0.062)	(0.060)	(0.027)
Obese				0.307***	-0.133***	
				(0.026)	(0.028)	
N	1,262	1,202	864	1,410	1,355	1,457
Without cognitive score						
Educational attainment	0.018***	0.007	-0.005*	0.016***	0.019***	-0.004**
	(0.007)	(0.004)	(0.003)	(0.004)	(0.004)	(0.002)
N	1,618	1,545	1,120	1,813	1,740	1,894

Robust Standard errors in parentheses

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Table 4.11: Fair or Poor Health in 2003

Coefficient on Educational Attainment	Std. Error	N	Controls Included
-0.042***	(0.003)	1,855	Education Only
-0.033***	(0.004)	1,852	Education, Demographics
-0.029***	(0.005)	1,428	Education, Demographics, Cognitive Score
-0.028***	(0.004)	1,435	Education, Demographics, Knowledge
-0.032***	(0.004)	1,778	Education, Demographics, Preferences
-0.020***	(0.006)	783	Education, Demographics, Self-control
-0.034***	(0.004)	1,360	Education, Demographics, Social Networks
-0.030***	(0.004)	1,284	Education, Demographics, Health Choices
-0.027***	(0.004)	1,424	Education, Demographics, Parental Characteristics

Robust Standard errors in parentheses

*** Significant at 1% level ** Significant at 5% level * Significant at 10% level

Demographic variables are age, male, black, Hispanic, married, household income and no health insurance

Chapter 5

Conclusion

This dissertation studies various determinants of preventive health behaviors among the the elderly U.S. population.

The first study finds that coverage of flu shots by Medicare part B increased demand by 12.4%. Applying the findings of earlier studies on cost savings per person, a 12.4% increase in vaccination rates translates to a cost saving of $254.6 million to Medicare in 1993. I also find that the responsiveness of demand to a price change varied across different groups, suggesting that the cost savings might vary across groups. A consideration in interpreting these results is that I am not able to separate a pure price effect from the total effect of the policy change. It is possible that demand for flu shots increased because the introduction of the Medicare coverage caused individuals to become more aware of the flu shots and of the harmful effects of influenza. To the extent that one is interested in the overall impact of a policy change, these results are interesting.

The estimated effect might be sensitive to method used. One advantage of the approach used in this study is that it is a semi-parametric estimation method. Hence, I do not need linearity or distributional assumptions to identify the effect as would be required in the case of a differences-in-differences or other such approaches. Moreover, matching on the propensity score improves the quality of the estimate by reducing any selection bias. Matching also allows for the effect of the treatment to vary across individuals. This is an improvement over conventional regression models, which make the implicit assumption that the treatment effect is the uniform across all individuals in the treatment and control groups.

A limitation of this study is that the estimates from the part B sample might be driven by a time trend. The analysis using the full sample attempts to account for this. However, to the extent that individuals with only Medicare part A form an imperfect control group for persons with Medicare part B, these estimates might suffer from a bias.

The second chapter of this dissertation studies the effect of health shocks on perceptions about risk and their effect on vaccine demand. I use self reported data to identify individual perceptions about the risk of getting pneumonia and use the panel structure of the data to identify learning about disease risk. I find evidence that individuals do respond to health shocks by updating their risk perceptions. Consistent with a Bayesian learning model, I also find significant persistence in risk perceptions. While this implies that personal health shocks act as a source of information, the exact mechanism is not clear. For example, it is possible that when an individual contracts pneumonia his doctor recommends immunization and that this is the source of information and not the health shock per se. If health shocks merely serve to increase interactions with the doctor who is the source of information then one would expect them to affect other perceptions such as beliefs about the effectiveness of the vaccine and not just the risk of disease. However, I find that health shocks do not affect perceptions about the effectiveness of the vaccine nor do they affect the misconception that the vaccine might cause the disease. Doctor recommendations about vaccination also do not vary significantly across persons with and without health shocks. This implies that the shocks themselves act as a source of information to individuals.

The updating of risk perceptions is also accompanied by a corresponding change in vaccine demand. Individuals who experience a health shock are much more likely to vaccinate than those who do not. These results are robust to model specification and accounting for individual-specific unobserved heterogeneity increases the effect. However, once the shock is controlled for, the severity of the shock, measured by whether or not it resulted in a hospitalization, does not affect the probability of vaccinating. If shocks provide information about the disutility of getting pneumonia then one would expect a more severe shock to have a larger effect on perceptions about the disutility than a less severe shock and so affect vaccine demand. Thus, individuals seem to learn about the risk of getting the disease but not about the disutility of the disease or the effectiveness of the vaccine.

The hazard rate of vaccination also increases with the number of outpatient visits in that year suggesting that the time cost of vaccination affects choice. One concern with interpreting the effect of the health shocks on vaccination rates is that this might reflect physician induced demand. That is, individuals who are diagnosed with pneumonia might be urged by their doctors to vaccinate and this might explain the higher rates of vaccination. However, since vaccinations work by stimulating the body's natural immune system, it is unlikely that doctors would vaccinate at the time that the person is suffering from pneumonia since his immune system is weakened at this time. Most persons probably likely to visit the doctor at a later date in order to get vaccinated. This decision would depend on their risk perceptions. However, doctor's recommendations are also likely to play a role as with most health care decisions. This analysis does not attempt to separate these two effects.

Consistent with previous literature, I find that demand increases as the prevalence of the disease increases. However, this effect is substantially smaller than the effect of a health shock suggesting that individuals react much more strongly to a personal risk of the disease than to the average population risk. It is possible that while individuals have information about the general risk of getting pneumonia they are very uncertain about their own idiosyncratic risk of contracting the disease and a health shock provides information about this idiosyncratic aspect of risk.

The estimated coefficients change in specifications that allow for unobserved heterogeneity when compared to specifications that do not allow for it. This suggests that unobserved time invariant effects, such as individual attitudes towards preventive care, play a role in determining both beliefs and demand. Even after accounting for such effects beliefs as well as demand for preventive care are responsive to new information. This has implications for policies aimed at increasing demand. If heterogeneity in demand across individuals was completely determined by unobserved fixed effects, then an appropriate policy might be a price subsidy. Such a policy would increase demand by decreasing the cost of vaccination. On the other hand, if differences in beliefs play a role in determining demand and if these beliefs are sensitive to new information then this provides support for policies such as information campaigns. Of course, evaluating the exact impact of each type of policy requires further research.

The third chapter of this dissertation examines the role of educational attainment in explaining variation in health. Consistent with previous literature, we find a strong positive relationship between educational attainment and health. However, in this study, the link is stronger between self-reported health and years of schooling than it is between survival and schooling. We find that educational attainment is related to some health behaviors but not all of the health behaviors we evaluated. We also find that there is a positive relationship between educational attainment and cognitive status and that in some cases, cognitive status trumps educational attainment suggesting that without cognitive status included as a covariate, educational attainment is merely a proxy for cognition. But often cognitive status does not trump educational attainment, e.g., in analysis of social networks as they pertain to health care. Although much of the variation in health is attributable to differences in cognitive status, this only explains half of the relationship between cognition and health. More educated persons tend to think they have more self-control and self-confidence about the conduct of their lives. We cannot say, however, whether or not education is causal. This study finds several surprising results and/or results for which empirical evidence has previously been unavailable. For example, the notion that more educated persons are more knowledgeable about DM care does not receive empirical support. Knowledge seems to be widespread. This result is consistent with other research findings on smoking persons (Khwaja et al. 2008). Persons with diabetes seem to know what to do to maintain their health. An interesting result is that sources of health information are quite different by educational attainment, holding other factors constant. More educated individuals are more likely to obtain their health information from reading. Less educated persons obtain their information more often from oral communication. There tends to be more, not less discussion, between physicians and patients who have lower educational attainment, at least when cognitive status is held constant. One does not gain the impression from these findings that interactions between physicians and patients are far better for more educated persons, holding such factors as lack of health insurance constant. While the results on overall self-control imply that those with more schooling have greater control over their lives, we do not obtain similar results for diabetes self-management. Thus, we do not see that greater self-control directly translates into care behaviors that would lead to improved DM outcomes. One surprising result, is that more educated individuals were more, not less likely, to forget to take their medications. This study suggests that simple warnings such as "not following your doctor's advice will lead to serious complications" are not likely to be productive. Persons diagnosed with DM know this. Higher educational attainment may help individuals know where to look when they need disease-specific information relevant to their self-care. Or perhaps it helps them recognize signs and symptoms of complication onset that may lead to more timely and hence action with higher marginal product. Unfortunately, the HRS-DS did not delve into these latter issues.

Although we find some relationships between individual preferences and educational attainment, this study provides no new insights on causal pathways. Whether or not preferences influence educational attainment or educational attainment molds preferences or both, these processes occurred at a stage in the life cycle much earlier than is captured by the HRS and HRS-DM. This is a fundamental question, but one requiring data on younger individuals. Our contribution in this study has been to show that a link between self-control and personal health exists. Even if people are forward-looking in their health decisions, they may often lack the wherewithal to carry their best intentions out. To the extent that education affects longevity, this provides an additional period over which returns to education accrue to the individual. This implies that individuals with higher life expectancy will have a higher incentive to invest in education (Echevarria 2004, Castello-Climent and Domenech 2008). While this source of reverse causality is ruled out for our sample of elderly persons, current life expectancy is likely to be correlated with life expectancy at birth which would have affected incentive to invest in human capital, including investments in health and education. While in neoclassical models of household decision making, individual decisions are guided by maximization of utility subject to constraints, with the well-being of persons outside the household having no effect on decision making or on wellbeing, there is some evidence that individuals are affected by their standing relative to others (see e.g., Marmot et al (1978), Marmot et al. (1991) and Eibner and Evans (2001)). For one, more educated persons may have better access than others to high quality medical care through their professional networks. However, if lower educated individuals have networks with others who also have less education and hence have a lower opportunity cost of time, access to care giving from family and friends may be better for such persons as our findings imply. Economists have long recognized that humans and human decision making are interdependent (see e.g., Duesenberry 1949), and the recent literature on peer effects recognizes this as well. Nevertheless, measurement of this phenomenon remains challenging. Finally, to really document returns to schooling, it is indeed important to gauge its effects on efficiency in household production. Although a link between education and health undoubtedly exists, the underpinnings of this relationship are more complex than economists have generally recognized. Particularly in view of these complexities, conducting case studies such as we have done for diabetes can yield useful insights. Monitoring decision making for a variety of chronic diseases would be a highly worthwhile undertaking as is empirical analysis of individuals much earlier in the life cycle to undercover relationships with educational attainment and variables we have assumed here to be exogenous.

This dissertation provides some insights into the determinants of preventive health behaviors among the elderly and near-elderly. However, individual decisions regarding health related choices is a complex issue and there remain several avenues for future research.

Bibliography

[AB91] M. Arellano and S. Bond. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58:277–297, 1991.

[AB95] M. Arellano and O. Bover. Another look at the instrumental variables estimation of error-components models. *Journal of Econometrics*, 68:29–51, 1995.

[Aba05] A. Abadie. Semiparametric difference-in-differences estimators. *Review of Economic Studies*, 72(1):1–19, 2005.

[ABBG99] J. Almirall, I. Bolibar, X. Balanzo, and C.A. Gonzalez. Risk factors for community-acquired pneumonia in adults: A population-based case-control study. *European Respiratory Journal*, 13:349–355, 1999.

[Ack03] D.A. Ackerberg. Advertising, learning, and consumer choice in experience good markets: A structural empirical examination. *International Economic Review*, 44:1007–1040, 2003.

[ACL03] J. Ameriks, A. Caplin, and J. Leahy. Wealth accumulation and the propensity to plan. *Quarterly Journal of Economics*, 118(3):1007–1047, 2003.

[AH82] T.W. Anderson and C. Hsiao. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18:47–82, 1982.

[AH01] M. Arellano and B. Honore. *Handbook of Econometrics*, volume 5, chapter 53. 2001.

[AHP96] A. Ahituv, J. Hotz, and T. Philipson. The responsiveness of the demand for condoms to the local prevalence of aids. *Journal of Human Resources*, 31(4):869–897, 1996.

[AI06] S. Athey and G. Imbens. Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2)(2):431–498, 2006.

[Alm06] D. Almond. Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 u.s. population. *Journal of Political Economy*, 114(4):672–712, 2006.

[Arr63] K. Arrow. Uncertainty and the welfare economics of medical care. *American Economic Review*, 53:941–973, 1963.

[AS05] M.C. Auld and N. Sidhu. Schooling, cognitive ability and health. *Health Economics*, 14(10):1019–1034, 2005.

[ASS07] P. Arcidiacono, H. Sieg, and F. Sloan. Living rationally under the volcano? an empirical analysis of heavy drinking and smoking. *International Economic Review*, 48:37–65, 2007.

[BB98] R. Blundell and S. Bond. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87:111–143, 1998.

[BDPer] A. Bryson, R. Dorsett, and S. Purdon. The use of propensity score matching in the evaluation of active labour market policies. *Department for Work and Pensions*, 4, 2002 Working Paper.

[Bec96] G. Becker. Accounting for tastes. *Harvard University Press*, 1996.

[Bec07] G. Becker. Health as human capital: synthesis and extensions. *Oxford Economic Papers*, 59(3):379–410, 2007.

[BI02] S. Becker and A. Ichino. Estimation of average treatment effects based on propensity score. *The Stata Journal*, 2002.

[BJKS97] R. Barsky, F. Juster, M. Kimball, and M. Shapiro. Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement study. *Quarterly Journal of Economics*, 112(2):537–579, 1997.

[BM97] G. Becker and C. Mulligan. The endogenous determination of time preference. *Quarterly Journal of Economics*, 112(3):729–758, 1997.

[BO77] J. Bachman and P. O'Malley. Self-esteem in young men: A longitudinal analysis of the impact of educational and occupational attainment. *Journal of Personality and Social Psychology*, 35(6):365–380, 1977.

[Bon02] S. Bond. Dynamic panel data models: A guide to micro data methods and practice. *Centre for Microdata Methods and Practice, Institute for Fiscal Studies, working paper no. CWP09/02*, 2002. Available at http://ideas.repec.org/p/ifs/cemmap/09-02.html.

[BS04] D. Black and J. Smith. How robust is the evidence on the effects of college quality? evidence from matching. *Journal of Econometrics*, 121(1-2):99–124, 2004.

[BSea88] J. Brandt, M. Spencer, and et al. The telephone interview for cognitive status. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, 1(2):111–117, 1988.

[BSI91] D. Brito, E. Sheshinski, and M. Intriligator. Externalities and compulsory vaccinations. *Journal of Public Economics*, 45:69–90, 1991.

[CCD08] A. Castello-Climent and R. Domenech. Human capital inequality, life expectancy and economic growth. *Economic Journal*, 118(528):653–677, 2008.

[CDLMng] D. Cutler, A. Deaton, and A. Lleras-Muney. The determinants of mortality. *Journal of Economic Perspectives*, Forthcoming.

[CE03] A. Caplin and K. Eliaz. Aids policy and psychology: a mechanism-design approach. *RAND Journal of Economics*, 34(4):631–646, 2003.

[CHL04] M. Cropper, M. Haile, and J. Lampietti. The demand for a malaria vaccine: evidence from ethiopia. *Journal of Development Economics*, 75(1):303–318, 2004.

[CLY95] W.S. Chern, E.T. Loehman, and S.T. Yen. Information, health risk beliefs, and the demand for fats and oils. *Review of Economics and Statistics*, 77(3):555–564, 1995.

[Cos03] D. Costa. Understanding mid-life and older age mortality declines: Evidence from union army veterans. *Journal of Econometrics*, 112(1):175–192, 2003.

[CS05] G.S. Crawford and M. Shum. Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173, 2005.

[DeG70] M. DeGroot. *Optimal Statistical Decisions*. New York: McGraw-Hill, 1970.

[Ech] C. Echevarria. Life expectancy, schooling time, retirement and growth. *Economic Inquiry*, 42(4):602–617.

[EE01] C.E. Eibner and W. Evans. Relative deprivation, poor health habits and mortality. *Journal of Human Resources*, 50(3):591–620, 2001.

[EK96] T. Erdem and M. Keane. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15:1–20, 1996.

[EV91] W. Evans and W.K. Viscusi. Estimation of state dependent utility functions using survey data. *Review of Economics and Statistics*, 73(1):94–104, 1991.

[EW89] Z. Eckstein and K. Wolpin. The specification and estimation of dynamic stochastic discrete choice models: a survey. *Journal of Human Resources*, 24(4):562–598, 1989.

[FAY+97] M.J. Fine, T.E. Auble, D.M. Yealy, B.H. Hanusa, L.A. Weissfeld, D.E. Singer, C.M. Coley, T.J. Marrie, and W.N. Kappor. A prediction rule to identify low-risk patients with community-acquired pneumonia. *New England Journal of Medicine*, 336(4):243–250, 1997.

[FBQ93] B. Fischhoff, A. Bostrom, and M.J. Quadrel. Risk perception and communication. *Annual Review of Public Health*, 14:183–203, 1993.

[Fin04] A. Finkelstein. Static and dynamic effects of health policy: Evidence from the vaccine industry. *The Quarterly Journal of Economics*, 119(2):527–564, 2004.

[Fra97] P. Francis. Dynamic epidemiology and the market for vaccinations. *Journal of Public Economics*, 63(3):383–406, 1997.

[FSHea07] A. Fiore, D. Shay, P. Haber, and et al. Prevention and control of influenza: Recommendations of the advisory committee on immunization practices 2007. *Mortality and Morbidity Weekly*, 56(RR06), 2007.

[Fuc82] V. Fuchs. Time preference and health: an exploratory study. *Economic Aspects of Health*, 1982.

[GHW00] T. Gayer, J.T. Hamilton, and Viscusi W.K. Private values of risk tradeoffs at superfund sites: Housing market evidence on learning about risk. *Review of Economics and Statistics*, 82(3):439–451, 2000.

[Gil98] D. Gilleskie. A dynamic stochastic model of medical care use and work absence. *Econometrica*, 66(1):1–46, 1998.

[GK97] M. Grossman and R. Kaestner. Effects of education on health. *The Social Benefits of Education*, 1997.

[GP95] P. Geoffard and T. Philipson. The empirical content of canonical models of infectious diseases: the proportional hazard specification. *Biometrika*, 82(1):101–111, 1995.

[GP97] P. Geoffard and T. Philipson. Disease eradication: Private versus public vaccination. *The American Economic Review*, 87(1):222–230, 1997.

[Gro00] M. Grossman. The human capital model. *Handbook of Health Economics*, 2000.

[GS02] D. Goldman and J. Smith. Can patient self-management help explain the ses health gradient. *Proceedings of the National Academy of Sciences of the United States of America*, 99(16):10929–10934, 2002.

[GS06] P. Gertler and T. Simcoe. Disease management. *Working paper*, 2006.

[HIST98] J. Heckman, H. Ichimura, J. Smith, and P. Todd. Characterizing selection bias using experimental data. *Econometrica*, 66(5):1017–1098, 1998.

[HIT98] J. Heckman, H. Ichimura, and P. Todd. Matching as an econometric evaluation estimator. *Review of Economic Studies*, 65(2):261–294, 1998.

[HNL04] J. Heckman and S. Navarro-Lozano. Using matching, instrumental variables, and control functions to estimate economic choice models. *The Review of Economics and Statistics*, 86(1):30–57, 2004.

[HNW+05] E. Hak, J. Nordin, F. Wei, J. Mullooly, S. Poblete, R. Strikas, and K. Nichol. Influence of high-risk medical conditions on the effectiveness of influenza vaccinations among elderly members of 3 large managed-care organization. *Clinical Infectious Diseases*, 35(4):370–377, 2005.

[HV05] J. Heckman and E. Vytlacil. Structural equations, treatment effects, and econometric policy evaluation. *Econometrica*, 73(3):669–738, 2005.

[Imb04] G. Imbens. Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics*, 86(1):4–29, 2004.

[Isr05] M. Israel. Services as experience goods: An empirical examination of consumer learning in automobile insurance. *American Economic Review*, 95(5):1444–1463, 2005.

[Kah98] M. Kahn. Education's role in explaining diabetic health investment differentials. *Economics of Education Review*, 17(3):257–266, 1998.

[Ken] D. Kenkel. Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99(2):287–305.

[Ken00] D. Kenkel. Prevention. *Handbook of Health Economics*, 2000.

[Khw02] A. Khwaja. Health insurance, habits and health outcomes: moral hazard in a dynamic stochastic model of investment in health. *PhD. Dissertation draft*, 2002.

[KLM06] D. Kenkel, D. Lillard, and A. Mathios. The roles of high school completion and ged receipt in smoking and obesity. *Journal of Labor Economics*, 24(3):635–660, 2006.

[KSC06] A. Khwaja, F.A. Sloan, and S. Chung. Learning about individual risk and the decision to smoke. *International Journal of Industrial Organisation*, 24(4):683–699, 2006.

[KSS06] A. Khwaja, F. Sloan, and M. Salm. Evidence on preferences and subjective beliefs of risk takers: The case of smokers. *International Journal of Industrial Organization*, 24(4):667–682, 2006.

[KSSW06] A. Khwaja, D. Silverman, F. Sloan, and Y. Wang. Smoking, wealth accumulation and the propensity to plan. *Economic Letters*, 94(1):96–103, 2006.

[KSWng] A. Khwaja, F. Sloan, and Y. Wang. Do smokers value their health and longevity less? *Journal of Law and Economics*, 2008 forthcoming.

[Kyr01] E. Kyriazidou. Estimation of dynamic panel data sample selection models. *Review of Economic Studies*, 68(3):543–572, 2001.

[Lai97] D. Laibson. Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–477, 1997.

[Lai01] D. Laibson. A cue-theory of consumption. *The Quarterly Journal of Economics*, 116(1):81–120, 2001.

[Lei66] H. Leibenstein. Allocative v. x-efficiency. *American Economic Review*, 56(3):392–415, 1966.

[Lei90] P. Leigh. Schooling and seat-belt use. *Southern Economic Journal*, 57(1):195–207, 1990.

[Leo07] K. Leonard. Idle chatter or learning? evidence of social learning about doctor quality from rural tanzania. *Working Paper*, 2007.

[LM04] A. Lleras-Muney. The relationship between education and adult mortality in the united states. *Review of Economic Studies*, 72(1):189–221, 2004.

[LMFM03] J.A List, D.L. Millimet, P.G. Fredriksson, and W.W. McHone. Effects of environmental regulations on manufacturing plant births: evidence from a propensity score matching estimator. *Review of Economics and Statistics*, 85(4):944–952, 2003.

[LMLng] A. Lleras-Muney and F. Lichtenberg. The effect of education on medical technology adoption: Are the more educated more likely to use new drugs? *special issue of the Annales d'Economie et Statistique in memory of Zvi Griliches*, Forthcoming.

[LND04] Y-C. Li, E.C. Norton, and W.H. Dow. Influenza and pneumococcal vaccination demand responses to changes in infectious disease mortality. *Health Services Research*, 39(4):905–925, 2004.

[LO05] G. Loewenstein and T. O'Donoghue. Animal spirits: Affective and deliberative processes in economic behavior. *Working Paper*, 2005.

[Loc07] L. Lochner. Individual perceptions of the criminal justice system. *American Economic Review*, 97(1):444–460, 2007.

[McG03] T. McGuire. Setting prices for new vaccines (in advance). *International Journal of Health Care Finance and Economics*, 3(3):207–224, 2003.

[MRSH78] M. Marmot, G. Rose, M. Shipley, and P. Hamilton. Employment grade and coronary heart disease in british civil servants. *Journal of Epidemiology and Community Health*, 32(4):244–249, 1978.

[MSSea91] Marmot M., G. Smith, S. Stansfeld, and et al. Health inequalities among british civil servants: the whitehall ii study. *Lancet*, 337(8754):1387–1393, 1991.

[Mul99] J. Mullahy. It'll only hurt a second? microeconomic determinants of who gets flu shots. *Health Economics*, 8:9–24, 1999.

[Nag76] S.Z. Nagi. An epidemiology of disability among adults in the u.s. *Millbank Memorial Fund Quarterly: Health and Society*, 54(4):439–467, 1976.

[Nic01] K. Nichol. Cost benefit analysis of a strategy to vaccinate healthy working adults against influenza. *Annals of Internal Medicine*, 161(5):749–759, 2001.

[NMP+01] J. Nordin, J. Mullooly, S. Poblete, R. Strikas, R. Petrucci, F. Wei, B. Rush, B. Safirstein, D. Wheeler, and K. Nichol. Influenza vaccine effectiveness in preventing hospitalizations and deaths in persons 65 years and older in minnesota, new york and oregon: data from 3 health plans. *Journal of Infectious Diseases*, 184(6):665–670, 2001.

[NMWVS94] K. Nichol, K. Margolis, J. Wuorenma, and T. Von Sternberg. The efficacy and cost effectiveness of vaccination against influenza among elderly persons living in the community. *New England Journal of Medicine*, 331(12):778–784, 1994.

[Ols02] C. Olson. Do workers accept lower wages in exchange for health benefits? *Journal of Labor Economics*, 20(S2):S91–S114, 2002.

[Phi96] T. Philipson. Private vaccination and public health: An empirical examination for u.s. measles. *Journal of Human Resources*, 31(3):611–630, 1996.

[RJTW03] M. Richards, M. Jarvis, N. Thompson, and M. Wadsworth. Cigarette smoking and cognitive decline in midlife: Evidence from a prospective birth cohort study. *American Journal of Public Health*, 93(6), 2003.

[Roo06] D. Roodman. How to do xtabond2: An introduction to "difference" and "system" gmm in stata. *Center for Global Development*, Working Paper Number 103, 2006.

[RR83] P. Rosenbaum and D. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

[Sch62] T. Schultz. Reflections on investment in man. *Journal of Political Economy*, 70(5), 1962.

[Smi05a] J. Smith. Consequences and predictors of new health events. *Analyses in the Economics of Aging*, 2005.

[Smi05b] J. P. Smith. The impact of ses on health over the life course. *Journal of Human Resources*, 42(4):739–764, 2005.

[Tab06] Table7. National vital statistics reports. *Centers for Disease Control and Prevention*, 54(19), 28 June 2006.

[TSH91] D. Thomas, J. Strauss, and M. Henriques. How does mother's education affect child height? *Journal of Human Resources*, 26(2):183–211, 1991.

[TSWea03] W. Thompson, D. Shay, E. Weintraub, and et al. Mortality associated with influenza and respiratory syncytial virus in the united states. *Journal of the American Medical Association*, 289, 2003.

[VE90] W.K. Viscusi and W.N. Evans. Utility functions that depend on health status: Estimates and economic implications. *American Economic Review*, 80(3):353–374, 1990.

[VE99] W. K. Viscusi and W. Evans. Estimation of revealed probabilities and utility functions for product safety decisions. *Review of Economics and Statistics*, 80(1):28–33, 1999.

[Vis85] W. K. Viscusi. A bayesian perspective on biases in risk perceptions. *Economic Letters*, 17(1):59–62, 1985.

[Vis90] W. K. Viscusi. Do smokers underestimate risks? *Journal of Political Economy*, 98:1253–1269, 1990.

[VO84] W.K. Viscusi and C. O'Connor. Adaptive responses to chemical labeling: Are workers bayesian decision makers? *American Economic Review*, 74(5):942–956, 1984.

[Win05] F. Windmeijer. A finite sample correction for the variance of linear efficient two-step gmm estimators. *Journal of Econometrics*, 126:25–51, 2005.

Biography

Padmaja Ayyagari was born on March 4, 1978 in Visakhapatnam, India. She received her B.Sc. in Economics from Calcutta University in 2000. After her bachelors degree, she obtained a masters degree in Quantitative Economics from the Indian Statistical Institute in 2002. She earned a second masters degree from Duke University in 2005. She has received various fellowship awards. These include a Government of India fellowship and annual contigency grant from the Indian Statistical Institute from 2000 to 2002, the Frank T. DeVyver Endowment fellowship from the Department of Economics at Duke University in 2002-2003, a summer dissertation research fellowship from the Department of Economics at Duke University in 2005, a graduate school conference travel fellowship from Duke University in 2006 and the Philip Jackson Baugh fellowship in aging and human development from Duke University in 2006-2007.

She has co-authored two papers during her tenure at Duke University. These are "Effects of Diagnosed Dementia on Costs to Medicare and Medicaid Program Payments", with Martin Salm and Frank A. Sloan, which in published in Inquiry. A second paper is "Rates of Glaucoma Medication Utilization Among Persons with Primary Open Angle Glaucoma, 1992-2002", with J.D. Stein, F.A. Sloan and P.P.Lee, which is published in Ophthalmology.