The Internet and Healthcare Use:

How Access to Information Affects the Demand for Medical Services

Nathan Grace

Abstract

The examination of imperfect information and its relationship to health care decisions only truly began after Arrow (1963). Since then, economists have been especially interested in health care markets, specifically because they suffer substantially more from a lack of consumer knowledge than most others. The primary difficulty in examining this phenomenon is determining a useful measure of information. This paper approaches the issue of consumer knowledge from a new angle. The integration of the internet in society has granted the public much greater access to information than was previously possible. Therefore is this investigation, an individual's use of the internet, specifically searches for health care information, is used as a proxy for the individuals overall information. The individual's exposure to information via the web is then compared to that individuals medical care use as a result of that information. The results suggest that there is a meaningful relationship between online searches for health information and healthcare use.

Introduction

A survey in the year 2000 estimated that 70% of those who have searched the web for health information say that information has influenced their decision concerning the treatment of an illness or condition (Fox & Rainie, 2000). The same study determined that 92% of those seeking healthcare information online say they found something useful during their search (Fox & Rainie, 2000).In the years since then, our ability to access the internet has grown. This increased access to electronic information has enormous potential to alleviate many of the uncertainties we face in our daily lives. Because the internet has become so widely used as a source of healthcare information it is important to understand the impact of the internet on the way we behave. The internet may serve to bridge some of the informational gap from the consumers' perspective. Using information available online, the consumer ought to be better equipped to understand their own state of health and how to improve upon it. Understanding how the internet influences our actions will allow us to better understand the way information impacts our health decisions. Through this understanding, we can develop better means of information distribution, reducing uncertainty in the healthcare market, and therefore improving market function.

The purpose of this study is to analyze the effect of internet use on the consumer's demand for healthcare. The internet may serve to bridge some of the informational gap from the consumers' perspective. Using information available online, the consumer ought to be better equipped to understand their own state of health and how to improve upon it. Two conflicting theories concerning the effect of health information exist. One theory assets that a more informed consumer is likely to demand medical care more than a less informed person, as the less informed are more likely to underestimate the marginal benefit of healthcare services (Kenkel, 1990).However, the possibility also exists that easy access to health care information may preclude the need for a doctor visit. In such a scenario, online information would serve as a substitute for the information purchased from physicians, thus reducing demand for healthcare services. Because conflicting theories exist, this study will attempt to ascertain whether or not the use of the internet will result in a greater demand for healthcare. As of this time, no study has analyzed the effect of internet use on the demand for healthcare. The effects of information and uncertainty of healthcare demand have been studied (Arrow, 1963), (Kenkel, 1990), (Goldner, 2006), (Pohlmeier & Ulrich, 1995), and this study will attempt to build on this knowledge. For the remainder of this paper, both healthcare and medical care are defined as a patient's decision to visit a physician.

Review of Literature

Kenneth Arrow (1963) lays a framework for study concerning the way uncertainty affects the healthcare industry. The overarching theme of Arrow's work is that imperfections in information result in market failure (Arrow, 1963). Arrow's work laid the foundation for virtually all future discussions of information and healthcare. Arrow's contribution to the medical information discussion was so profound that his work has been summarized and built upon by others, including Deborah Haas-Wilson who updated Arrow's theoretical work in 2001 (Haas-Wilson, 2001).

Arrow understood that uncertainty was particularly prevalent in healthcare markets, and that this uncertainty makes it especially difficult for patients to make appropriate health care decisions (Haas-Wilson, 2001).Because patients are uncertain about the effectiveness of different treatments they have difficulty evaluating the utility of medical services (Arrow, 1963), (Haas-Wilson, 2001). In such situations informatiuon becomes a marketable good itself (Arrow, 1963). However the obscure nature of medical information presents a substantial difficulty for patients and their ability to acquire and process information (Arrow, 1963). Information is indeed part of what a patient purchases from a physician, however patients are often unaware of the value of this information because they cannot evaluate the usefullness of information they do not yet posess (Arrow, 1963). This situation makes physician information relatively non-marketable, as it is difficult for patients to determine what information they require, let alone who is best fitted to supply it (Arrow, 1963). Several economists have noted that as one's health deteriorates they have increasing incentive to procure health information (Goldner, 2006), (Kenkel, 1990). However, those who are ill may have a reduced capacity to acquire and interpret the very information that might help them (Haas-Wilson, 2001). Additionally, many treatments are most effective if applied in the early stages of illness (Haas-Wilson, 2001). Often, patients cannot evaluate the utility of a particular treatment until after they have purchased it, if ever at all (Haas-Wilson, 2001). This uncertainty surrounding treatment utility and the non-marketability of the information physicians supply, leads to market failure (Arrow, 1963), (Haas-Wilson, 2001). Patients are unable to determine what services they require, nor do the understand the utility of those services. Under these conditions it is theorized that patients will underutilize healthcare given their relative inability to understand the benefits of treatement versus the cost (Arrow, 1963), (Goldner, 2006).

Deborah Haas-Wilson (2001) provides a synopsis of Arrow's theoretical work and builds upon it, describing the changes that have occurred since Arrow and how they apply to modern situations (Haas-Wilson, 2001). Wilson highlights the changes in our ability to acquire, measure and assimilate information. Consistent with Arrow, Wilson argues that information of any variety is difficult to measure. There are two commonly used proxies for information; the number of physicians in a given area, and the frequency of referral (Haas-Wilson, 2001). As the number of physicians in an area increases, the likelihood that an individual will be familiar with someone who has visited a particular physician diminishes (Haas-Wilson, 2001). As ‘word-of-mouth’ is one of the few indications of physician quality an individual has access to, this measure is defendable but limited (Haas-Wilson, 2001). The referral rate a physician receives represents the level of proficiency his peers and previous patients consider him to have (Haas-Wilson, 2001). This measure is also limited as it is based on individual perceptions of quality, rather than tangible criteria.

Wilson asserts that patients have access to three categories of information; information concerning the source of an illness, the effectiveness of treatments, and which physician provides the best information (Haas-Wilson, 2001). All of these categories suffer the same informational deficiencies discussed by Arrow, but each one allows us to consider the different aspects of a patient’s healthcare decision. Wilson contends that the decision concerning which physician a patient ought to acquire information from is the most important decision a patient can make (Haas-Wilson, 2001). Also of interest is Wilson’s mention of the internet, which has exploded as a source of information, with 26,000 health related web sites active in 2001 (Haas-Wilson, 2001). The growth of the internet should allow patients to better explore methods of treatment, as well as information regarding quality of care. Therefore, it is possible that the expansion of the internet will make patients better equipped in all three categories of information discussed by Wilson.

Grossman’s article on the demand for health is one of the quintessential works on the topic of healthcare. Grossman presents a theoretical and empirical analysis of the demand for health. A central theme in Grossman's work is that healthcare is produced by individuals with inputs of market goods and their own time (Grossman, 1972). Therefore, any demand estimated for healthcare is really derived demand. This view of market production is preferable for analysis of healthcare because it recognizes that demand in healthcare markets is truly the demand for health, rather than the demand for healthcare. Grossman also points out that health is demanded not only because it is part of the individuals preference, but because sickness represents a disutility (Grossman, 1972). Put simply, it is preferable to be healthy rather than sick. Also, illness incurs a cost in time required to get well. This time would be better spent working for a wage or consuming pursuing leisure activities (Grossman, 1972).

Grossman finds that education has a positive effect on the demand for healthcare, as education lowers the cost of additions to health stock (Grossman, 1972). The result that age is has a negative effect on health stock is also featured and is congruent with the findings of Kenkel and Goldner (Grossman, 1972), (Kenkel, 1990), (Goldner, 2006). In Grossmans analysis the results indicate that income is actually negatively correlated with good health. Grossman suggests that this phenomenon may be associated to the greater ability to purchase goods which reduce health stock, and simultaneously increase demand for medical care (Grossman, 1972). For example, the wealthy can afford more convenient and luxurious, but less healthy, eating habits.

Don Kenkel (1990) directed his attention to a specific manifestation of uncertainty, imperfect information. Imperfect information is similar to asymmetric information, in that both represent a lack of sufficient information. The difference is that in asymmetric information one party has an informational advantage over another. With imperfect information neither party is fully informed about the other. For example, a patient is more familiar with their symptoms than a doctor; therefore if the doctor must rely on the patient to describe symptoms then the doctor suffers from imperfect information. The patient may overlook details or symptoms that are important because they do not recognize their significance due to a lack of medical knowledge. Neither party has complete information. Kenkel argues that there are two possible theories that describe the effect of information on healthcare. Option one is that information increases an individual's ability to recognize the marginal benefit of treatment thus increasing the likelihood that they will use healthcare (Kenkel, 1990). The second option is that information makes the individual better able to make decisions concerning the intensity of care, allowing the individual to be less dependent on the physician and more active in decisions concerning intensity of care. In this second option, the patient is less susceptible to physician induced demand as is less likely to use medical care as a result of increased information.

Kenkel (1990) uses a modified Grossman model to estimate an individual's information from a survey data set measuring the difference between patient's perceptions of symptoms and the conventional medical interpretation of those symptoms. This proxy is then used to estimate the probability of physician visits dependant on information and environmental factors. While this method does provide reliable results, the difficulty is finding an instrument for information. In his estimation Kenkel argues that education is a suitable instrument to represent an individual's overall information (Kenkel, 1990). In his conclusions Kenkel suggests that greater information does increase the likelihood that an individual will consume more medical care because the marginal benefits of care are more apparent (Kenkel, 1990). However, Kenkel also says that there is not sufficient evidence to suggest that information has an impact on an individual's intensity decision (Kenkel, 1990).

Melinda Goldner (2006) focused her study on the effects of health status on the demand for healthcare information, specifically information found on the internet. Primarily, she finds that those with lower health status are the most likely to seek healthcare information (Goldner, 2006). The significance of this cannot be understated. Those with the highest incentive to seek healthcare information are also those whom should be most influenced by the information they find. Golner's finding that poorer health status is associated with greater use of the internet for information has great implications for physicians. Physicians must educate their patients about the uneven quality of information available on the internet, specifically which sources are trustworthy and which are not (Goldner, 2006). As a result of her findings, Goldner recommends that future research should concentrate on the effect of healthcare information on the demand for healthcare itself. It is this task which this study seeks to accomplish.

Authors Pohlmeier and Ulrich (1995) have added to previous studies concerning the demand for healthcare, using data from the German experience. The authors suggest that the decision to use health care is part of a two-stage process. In the first stage, the patient must decide whether or not to use health care at all, the contact decision. In the second stage, it is the doctor, not the patient, who determines the frequency or intensity of treatment, the intensity decision (Pohlmeier & Ulrich, 1995). The authors assert that an investigation of healthcare demand that does not account for both the contact and intensity effect is incomplete and will produce incorrect results (Pohlmeier & Ulrich, 1995).

In their estimation Pohlmeier and Ulrich have used a two part approach to estimating healthcare demand (Pohlmeier & Ulrich, 1995). In their estimation, the authors submit that the decision to use healthcare consists of two separate components, the decision to use healthcare at all (contact decision), and the decision about how often to use healthcare (intensity decision).

This is an important distinction. Simply measuring the amount of information an individual has against the likelihood of a visit or the frequency of visits, is incomplete. It is important to know both whether and how often an individual visits a physician dependant on their information. There are certain instances where a patient requires a follow up visit or a series of treatments resulting in several visits. There is also the possibility that the patient checks for information before regularly scheduled exams in order to be better informed during the exam. If only one facet of the patient's healthcare use is examined, crucial effects may be overlooked.

Theoretical Framework

In this paper, the goal is to estimate the probability that an individual will visit a physician conditional on that individual's search for healthcare information on the internet. We know from the works of Wilson and Arrow that uncertainty in healthcare markets creates a number of unfavorable market conditions resulting in market failure (Arrow, 1963), (Haas-Wilson, 2001). From Kenkel's work we expect that an individual is more likely to seek healthcare if they have greater information because they understand the marginal benefit of treatment. Further, it is clear from Pohlmeire and Ulrich's study that healthcare decisions must be treated as a two-stage process, encompassing both the contact and intensity decisions.

The following theoretical model is derived from the works of previous authors:

$$\mathbf{DocVis} = \beta_0 + \beta_1 \mathit{Search} + \beta_2 \mathit{Hlthstat} + \beta_3 \mathit{Educ} + \beta_4 \mathit{Age} + \beta_7 \mathit{Exp} + \mu$$

The model shows that an individual's decision to see a physician will be a function of that individuals search for health information on the internet, health status, education, age and an array of socio-economic factors, plus the error term. The dependant variable DocVis measures whether the individual consulted a physician based on an affirmative response to the search variable. The parameter labeled Search is a binary variable which indicates whether or not an individual searched for healthcare information online. The parameter Hlthstat represents an individual's perception of their own health status as being poor, fair, good, or excellent. The Educ parameter measures an individual's highest completed level of education. Responses were separated into five categories: less than a high school degree, a high school degree, an associate's degree or vocational training, a four year college degree, or greater than a four year college degree. The Age variable is a pure measure of an individual's age in years ranging from 18 to 98. Finally, the Exp parameter represents a vector of socio-economic factors including race, marital status, ethnicity, income, student status, children and gender. A full list of these variables and their expected signs is included below in the variable legend.[1]

As the research question concerns a patient's decision to search for health information online, whether or not an individual searched for health information online will be an important factor in this study. Additionally, in order to test the effect of the internet on healthcare use, the contributions of previous authors must be accounted for wherever possible.

In his paper, Kenkel stressed the importance of education when considering an individual's health care decisions. According to his work, education is a measure of our ability to locate and process information (Kenkel, 1990). This finding is concurrent with that of Grossman, who found that the marginal cost of additions to health stock are lower for those with higher levels of education (Grossman, 1972).

---

[1] Please refer to the table below entitled 'Variable Legend' for a complete list of the variables and their expected signs.

Goldner demonstrated in her analysis that those with the poorest health status are those who have the most to gain from health information and have the greatest incentive to seek out that information (Goldner, 2006). A variable accounting for an individual's health must be considered in this model, as health status is a determinant of the incentive to search for healthcare information.

Multiple authors have determined that age negatively affects one's health stock and therefore is an important determinant of our healthcare decisions. Grossman maintains that increasing age reduces health and thus increases healthcare use (Grossman, 1972). This view is supported by both Kenkel and Goldner as well (Kenkel, 1990), (Goldner, 2006). Age is shown to have an effect on an individual's demand for healthcare use and therefore their demand for physician visits. Thus a measure of age will be included in this model as well.

As evidenced by Pohlmeier and Ulrich, an analysis of healthcare demand that does not consider the two-stage decision making process is not the most appropriate (Pohlmeier & Ulrich, 1995). Unfortunately, the data does not contain any tangible measure of a patient's intensity decision. Suitable data exists concerning the contact decision only, so this study will be constrained to the analysis of the contact decision only.

Empirical Model

The majority of data used in this estimation originate from a telephone survey of 2931 individuals conducted in 2006 by the Pew Research Center as part of the Internet and American Life project. The data are organized by state as designated by the FIPS[2] system used by the American Government. No measure of cost is currently available through the Pew data set. For

[2] Federal Information Processing Standard; A unique series of numbered codes used by the U.S. Census Bureau and Emergency Alert System to identify states or territories within the United States.

this reason, the estimation used here will attempt to account for ability to pay by including a dummy variable for insurance coverage. While not perfect, this approach makes sense. Cost in the health care market does not behave as it does in other markets. In the healthcare market the consumer is largely separated from the cost of medical care by insurance. Therefore, in lieu of an accurate measure of cost, insurance coverage will represent an individual's ability to pay.

The data used for this study contains sufficient observations to carry out an analysis similar to Kenkel's. Kenkel uses a two-stage least squares regression, using an OLS regression for the first stage to create an instrument for information using education as a proxy for total information. For the second stage Kenkel uses a maximum likelihood probit estimation to find the probability that an individual visits a physician given the information available. However, the same statistical methods will not be useful in this estimation, as the focus is specifically information obtained from the internet, not an individual's total information. As this is a very specific definition of information, using education as a proxy for information would be inappropriate, as it is far too broad. Because it is not correct use education as a proxy for information in this estimation, and no other suitable instrument exists in the data set, a two-stage least squares estimation will not be used here. The probit estimation, on the other hand, is still possible and will be used to estimate demand, conditional the search for health information online.

Following Kenkel's work, this estimation will use maximum likelihood probit to estimate the probability that an individual visits a physician conditional on that individual's access to online information. The probit procedure is the most appropriate estimation for this study given time and data constraints. There is insufficient data to carry out a two-stage least squares analysis and insufficient time to perform the Heckit procedure. Following the probit procedure, marginal effects analysis will be used to simplify reporting of the results. Equation (1), estimates this portability where "V" is the likelihood of a physician visit, "I" is the individual's use of online health information, and "Y" is a set of environmental factors.

(1)

In equation (1) "Y" represents a set of observable characteristics including age, educational attainment and health status. Of these, age and health status are included to account for differences of incentive to seek heath information (Kenkel, 1990). As a person ages, or as their health status deteriorates, there is an increasing incentive to seek medical care (Goldner, 2006). Education is included as a proxy for any health related schooling an individual may have, as well as a measure of an individual's ability to locate access and understand health information through other sources such as literature and news or public service messages (Kenkel, 1990). All other variables are included to explore any socio-economic differences that exist.

The data set obtained from the Pew Research Center describes several variables of interest with multiple choice questions. In order to isolate the effects of each response, the answers to each question have been converted into binary dummy variable and then grouped into simplified categories. A breakdown of these variables is presented in the variable legend below, along with the expected sign for each variable.

For this study, the hypothesis that the variable for search will have a positive sign is based on Kenkel's results, which indicated that greater information leads to a better understanding of the marginal benefit of healthcare[3]. Concerning health rate, it is expected that

[3] Please refer to the table below entitled 'Variable Legend' for a complete list of the variables and their expected signs.

as status worsens an individual will have more incentive to seek health care, while those with better status will have less incentive to do so (Goldner, 2006). The presence of children represents a parental responsibility. Parents are expected to have considerably less time to make and keep medical appointments for all but the most serious cases, and will therefore be more apt to forgo physician visits than non-parents. Concerning marital status, those in a partnership will have access to greater resources, both monetary and informational, making physician visits more likely than those who lack such resources. It is anticipated that lower levels of education will have a negative effect on the likelihood of a physician visit as the lesser educated have a diminished ability to find and interpret health information, while the more educated are better equipped to do so. It is assumed that greater education leads to greater healthcare use as the more educated better understand the utility of treatment (Kenkel, 1990).

It is not clear what effect being a student may have on the likelihood of a physician visit. On one hand students have a far greater exposure to information that do non-students, however the student devotes a significant amount of time and resources to academic study which may reduce their ability to search for health information or to see a physician. The variables for Hispanic ethnicity, gender and race are not expected to have any clear effect at this point. These variables are included for the sake of socio-economic interest. Those with greater income would be expected to have better access to healthcare as they are better able to purchase it. Thus it is expected that lower levels of income will have a negative impact on physician visits, while higher incomes are expected to facilitate healthcare use. Finally, the parameter for age is expected to have a positive sign, as an individual's age increases their health status will tend to decrease, resulting in greater incentive to seek medical care.

Variable Legend			
Variable	**Values**	**Dummy Variable Label**	**Expected Sign**
Physician Visits	1 = Visit	DocVis	
Internet Search	1 = Searched online for Health information	Search	+
Health Rate	1 = Excellent Health	ExcelHlth	-
	2 = Good Health	GoodHlth	Reference Group
	3 = Fair Health	FairHlth	+
	4 = Poor Health	PoorHlth	+
Kids	1= Kids	kids	-
Marital Status	1 = married/living with partner	WPartner	Reference Group
	2 = no longer married	Separated	-
	3 = never married	NeverMar	-
Education	1 = Less than HS education	LessHSed	-
	2 = HS education	HSed	Reference Group
	3= Associates Degree or Vocational training	HSedPlus	+
	4=4yr degree	IVyrDegree	+
	5= > 4yr degree	IVyrDegreePlus	+
Student	1 = Student	Student	-/+
Hispanic	1 = Hispanic	hispc	X
Race	1 = white	White	X
	2 = black	Black	X
	3 = non -- White	NonWhite	X
Income	1= < 40,000	LoIncome	-
	2 = 40,001 – 75,000	MdIncome	Reference Group
	3= >75,000	HiIncome	+
Age	The individual's age	AGE	+
Gender	1= male	Male	X

Results

The data from table C clearly suggests that online searches for healthcare information have a strong, positive and significant effect on an individual's healthcare use. The variable Search is highly significant with a parameter estimate of 0.1098. This result suggests that an individual who has searched for health information on the internet is nearly 11% more likely to visit a physician based on that search. The implications of this effect become very clear when compared to the results for the education variables. Traditionally, education has been regarded as an adequate measure of an individual's information. In Kenkel's analysis education is used as a proxy for an individual's information when determining healthcare demand (Kenkel, 1990). A great deal of emphasis has been placed on education, however the results of this estimation suggest that an individual with a four year college degree is only about 4.5% more likely to visit a physician than one who has only a high school education. The results imply then, that the effect of online information searches is roughly twice that of education. The implications of this finding are substantial; an individual's ability to locate information online may have a greater impact on their decision making than their education.

In this analysis the variable for health status is insignificant. This is somewhat counter intuitive and appears to contradict Goldner's work (Goldner, 2006). The case can be made however, that health status does not matter in this particular case. Those individuals who are already sick will likely seek medical attention whether or not they have searched for health information online. This is the most likely explanation, given the findings of this study as well as Goldner's work.

Finally, the variable for age returned significant, but the effect was small and negative. This is contrary to expectations, one would expect a person to have greater health problems as age increases, thus giving greater incentive to find information and visit a physician. The result for age is small and negative, however the mean age in table A is fifty two. The data set is skewed toward those of older age, who is not as familiar with the internet as their younger counterparts and are therefore less likely to use the internet as a source of information. As this analysis estimates demand conditional on internet searches, it makes sense that age would have a negative sign, because of the older mean age in the sample.

We can see from table B that none of the socio-economic variables returned a significant result. While socio-economic differences would have been interesting to analyze, and may have provided interest for future research, there are no clear distinctions to be made based on an individual's race, gender, marital status or otherwise. This is not entirely surprising given the pervasive nature of the internet in current American society; nearly everyone has some means of accessing the internet.

Variable Means					
Variable	**Observations**	**Mean**	**Std. Error**	**Min**	**Max**
DocVis	2925	0.0920	0.2890	0.0	1.0
Search	2800	0.2068	0.4051	0.0	1.0
ExcelHlth	2928	0.3221	0.4673	0.0	1.0
GoodHlth	2928	0.4788	0.4996	0.0	1.0
FairHlth	2928	0.1414	0.3485	0.0	1.0
PoorHlth	2928	0.0471	0.2120	0.0	1.0
LessHSed	2928	0.0905	0.2870	0.0	1.0
HSed	2928	0.3125	0.4636	0.0	1.0
HSedPlus	2928	0.2596	0.4385	0.0	1.0
IVyrDegree	2928	0.2087	0.4064	0.0	1.0
IVyrDegreePlus	2928	0.1199	0.3249	0.0	1.0
Insur	2915	0.1029	0.3039	0.0	1.0
kids	2914	0.2718	0.4450	0.0	1.0
WPartner	2928	0.5956	0.4909	0.0	1.0
Separated	2928	0.2415	0.4280	0.0	1.0
NeverMar	2928	0.1496	0.3567	0.0	1.0
Student	2922	0.0530	0.2242	0.0	1.0
hispc	2903	0.0513	0.2207	0.0	1.0
White	2931	0.8335	0.3726	0.0	1.0
Black	2931	0.0948	0.2931	0.0	1.0
NonWhite	2931	0.0464	0.2104	0.0	1.0
HiIncome	2931	0.2112	0.4082	0.0	1.0
MdIncome	2931	0.3016	0.4590	0.0	1.0
LoIncome	2931	0.2521	0.4343	0.0	1.0
AGE	2859	52.0063	18.2289	18.0	97.0
Male	2931	0.4190	0.4935	0.0	1.0

Table A

Analysis of Maximum Likelihood Parameter Estimates				
Parameter	**Estimate**	**Standard Error**	**Pr > ChiSq**	**Parameter**
Intercept	0.9034	0.1905	22.48	<.0001
Search	-0.5409	0.0771	49.24	<.0001
ExcelHlth	0.0437	0.0797	0.30	0.5835
FairHlth	-0.0735	0.1111	0.44	0.5081
PoorHlth	0.1818	0.2071	0.77	0.3800
LessHSed	-0.1704	0.1535	1.23	0.2671
HSedPlus	-0.3153	0.0980	10.35	0.0013
IVyrDegree	-0.3377	0.1038	10.59	0.0011
IVyrDegreePlus	-0.1994	0.1318	2.29	0.1305
Insur	-0.0687	0.1162	0.35	0.5545
kids	-0.0480	0.0879	0.30	0.5848
Separated	0.1375	0.1031	1.78	0.1822
NeverMar	0.2005	0.1190	2.84	0.0921
Student	-0.1248	0.1429	0.76	0.3823
hispc	-0.0522	0.1534	0.12	0.7337
Black	-0.1762	0.1171	2.26	0.1325
NonWhite	-0.1572	0.1604	0.96	0.3270
HiIncome	-0.0206	0.0901	0.05	0.8196
LoIncome	0.0339	0.0958	0.13	0.7236
AGE	0.0151	0.0029	26.96	<.0001
Male	0.0478	0.0729	0.43	0.5124

Table B

Marginal Effects: Parameter Estimates				
Parameter	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**
Intercept	0.1534	0.0297	5.16	<.0001
Search	0.1098	0.0140	7.82	<.0001
ExcelHlth	-0.0113	0.0128	-0.88	0.3790
FairHlth	0.0071	0.0168	0.42	0.6726
PoorHlth	-0.0212	0.0265	-0.80	0.4238
LessHSed	0.0287	0.0209	1.37	0.1699
HSedPlus	0.0450	0.0147	3.05	0.0023
IVyrDegree	0.0460	0.0164	2.80	0.0051
IVyrDegreePlus	0.0191	0.0202	0.94	0.3451
Insur	0.0046	0.0192	0.24	0.8087
kids	0.0121	0.0150	0.80	0.4224
Separated	-0.0155	0.0146	-1.07	0.2863
NeverMar	-0.0242	0.0191	-1.27	0.2051
Student	0.0303	0.0255	1.19	0.2347
hispc	0.0141	0.0255	0.55	0.5817
Black	0.0288	0.0193	1.49	0.1351
NonWhite	0.0322	0.0276	1.17	0.2432
HiIncome	0.0004	0.0153	0.02	0.9802
LoIncome	-0.0001	0.0141	-0.01	0.9924
AGE	-0.0020	0.0004	-4.58	<.0001
Male	-0.0081	0.0114	-0.71	0.4785
_Sigma	0.2856	0.0039	73.52	<.0001

Table C

Conclusions

The data clearly suggests a strong, positive relationship between online searches for health information and healthcare use. This conclusion is even more dramatic when compared to education as a measure of individual information. The effect of online healthcare searches is more than twice as great as that of education. The implications of this are immense, based on the results it would appear that education alone does not accurately measure an individual's information. Future research concerning uncertainty and information should consider the implications of multiple informational sources when attempting to measure information.

It is plain to see how the advent of the internet has helped to alleviate some of the asymmetric information problems brought up by the likes of Arrow and Wilson (Arrow, 1963), (Haas-Wilson, 2001). Some questions remain unanswered however. The most prevalent of these is the two part decision making framework developed by Pohlmeier and Ulrich (Pohlmeier & Ulrich, 1995). It would be interesting to see what affect internet information has on the intensity decision mentioned by both Pohlmeier and Kenkel (Pohlmeier & Ulrich, 1995), (Kenkel, 1990).

Limitations

There are several limitations to the current estimation. The foremost being that no adequate measure of consumer cost was available in the data. In typical demand estimations cost is undeniably important. It is possible that cost would have a minimal effect on the model as patients are largely separated from the cost of medical services by insurance. This is not to say that cost does not belong in the model. It is suggested here that the effect of a cost variable on the results would be minimal; however the omission of the cost variable is not in any way preferable. Even if the assertions as to the impact of cost on the model are correct, the variable should be included to substantiate this claim. As no such variable is included in the model, the lack of a suitable cost variable remains a limitation of this research. Future research should account for cost using data external to that which is used here. Possible measures of cost might include average premiums in the individual's state of residence, or average co-pay rates.

The data contains variables that measure if an individual did or did not search online for healthcare information, concerning a single instance. In addition, the data accounts for whether a person did or did not visit a physician as a result of that specific instance of searching. Another limitation of this analysis is that the data does not measure the physician visits of those who did not search for healthcare information. Given this, the data contains a sample selection bias, neglecting the physician visits of those who did not search for health information prior to visiting a physician. If the physician visitation rate of those who did not search for health information online is similar to the rate of those who did search, then the results of this analysis would be quite different. As no data exists, no positive statements to that effect can be made at this time. Future research should attempt ascertain if any difference exists in physician visitation rates between those who search online for health information and those who do not. Specifically, the Heckit procedure should be employed to control for sample selection bias (Heckman, 1979).

Finally, Pohlmeire and Ulrich have provided strong evidence that health care decisions are part of a two-stage process (Pohlmeier & Ulrich, 1995). The current data set does not contain the necessary observations to control for intensity. Therefore, the current estimation is not complete according to this two-stage model. Future research should attempt to control for the intensity decision, which is omitted here due to a lack of relevant data.

Bibliography

Arrow, K. J. (1963). *Uncertainty and the Welfare Economics of Medical Care.* Retrieved from The American Economic Review, Vol. 53, No. 5 (Dec., 1963), pp. 941-973: http://www.jstor.org/stable/1812044

Fox, S., & Rainie, L. (2000). *The Online Health Care Revolution:How the Web Helps Americans Take Better Care of Themselves.* Retrieved from The Pew Internet & American Life Project, (Nov 26, 2000): http://www.pewinternet.org/Reports/2000/The-Online-Health-Care-Revolution.aspx

Goldner, M. (2006). *Using the Internet and Email for Health Purposes: The Impact of Health Status.* Retrieved from Social Science Quarterly, Vol. 87, Issue 3, pp. 690 - 710: http://www3.interscience.wiley.com/cgi-bin/fulltext/118550945/PDFSTART

Grossman, M. (1972). *The Demand for Health: A Theoretical and Empirical Investigation.* Retrieved from NBER Books, National Bureau of Economic Research, Inc, number gros72-1, June.: http://www.nber.org/books/gros72-1

Haas-Wilson, D. (2001). *Arrow and the Information Market Failure in Health Care: The Changing Content and Sources of Health Care Information.* Retrieved from Journal of Health Politics, Policy and Law, Vol. 26, No. 5, October 2001.

Heckman, J. J. (1979, January). *Sample Selection Bias as a Specification Error.* Retrieved from Econometrica, Vol. 47, No. , pg.153-162.

Kenkel, D. (1990). *Consumer Health Information and the Demand for Medical Care.* Retrieved from The Review of Economics and Statistics, Vol. 72, No. 4 (Nov., 1990), pp. 587-595: http://www.jstor.org/stable/2109598

Pohlmeier, W., & Ulrich, V. (1995). *An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care.* Retrieved from The Journal of Human Resources, Vol. 30, No. 2 (Spring, 1995), pp. 339-361: http://www.jstor.org/stable/146123

Appendix A: SAS Output Tables

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
DocVis		2925	0.0919658	0.2890271	0	1.0000000
Search		2800	0.2067857	0.4050728	0	1.0000000
ExcelHlth		2928	0.3220628	0.4673467	0	1.0000000
GoodHlth		2928	0.4788251	0.4996368	0	1.0000000
FairHlth		2928	0.1413934	0.3484865	0	1.0000000
PoorHlth		2928	0.0471311	0.2119555	0	1.0000000
LessHSed		2928	0.0905055	0.2869536	0	1.0000000
HSed		2928	0.3125000	0.4635916	0	1.0000000
HSedPlus		2928	0.2595628	0.4384696	0	1.0000000
IVyrDegree		2928	0.2086749	0.4064309	0	1.0000000
IVyrDegreePlus		2928	0.1198770	0.3248732	0	1.0000000
Insur		2915	0.1029160	0.3039012	0	1.0000000
kids		2914	0.2717914	0.4449593	0	1.0000000
WPartner		2928	0.5956284	0.4908538	0	1.0000000
Separated		2928	0.2414617	0.4280427	0	1.0000000
NeverMar		2928	0.1495902	0.3567302	0	1.0000000
Student		2922	0.0530459	0.2241633	0	1.0000000
hispc		2903	0.0513262	0.2207003	0	1.0000000
White		2931	0.8335039	0.3725889	0	1.0000000
Black		2931	0.0948482	0.2930551	0	1.0000000
NonWhite		2931	0.0464005	0.2103869	0	1.0000000
HiIncome		2931	0.2111907	0.4082230	0	1.0000000
MdIncome		2931	0.3016035	0.4590324	0	1.0000000
LoIncome		2931	0.2521324	0.4343109	0	1.0000000
AGE	AGE	2859	52.006295	18.228891	18.000000	97.000000
Male		2931	9	2	0	0
			0.4189696	0.4934746	0	1.0000000

The SAS System

The Probit Procedure

Model Information	
Data Set	WORK.TOTAL
Dependent Variable	DocVis
Number of Observations	2703
Name of Distribution	Normal
Log Likelihood	-773.5554487

Number of Observations Read	2931
Number of Observations Used	2703
Missing Values	228

Class Level Information		
Name	**Levels**	**Values**
DocVis	2	0 1

Response Profile		
Ordered Value	**DocVis**	**Total Frequency**
1	0	2444
2	1	259

The SAS System

The Probit Procedure

PROC PROBIT is modeling the probabilities of levels of DocVis having LOWER Ordered Values in the response profile table.?

Type III Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Search	1	49.2392	<.0001
ExcelHlth	1	0.3006	0.5835
FairHlth	1	0.4380	0.5081
PoorHlth	1	0.7708	0.3800
LessHSed	1	1.2314	0.2671
HSedPlus	1	10.3538	0.0013
IVyrDegree	1	10.5868	0.0011
IVyrDegreePlus	1	2.2862	0.1305
Insur	1	0.3492	0.5545
kids	1	0.2985	0.5848
Separated	1	1.7799	0.1822
NeverMar	1	2.8375	0.0921
Student	1	0.7632	0.3823
hispc	1	0.1158	0.7337
Black	1	2.2628	0.1325
NonWhite	1	0.9608	0.3270
HiIncome	1	0.0520	0.8196
LoIncome	1	0.1251	0.7236
AGE	1	26.9649	<.0001
Male	1	0.4291	0.5124

The SAS System

The Probit Procedure

Analysis of Maximum Likelihood Parameter Estimates							
Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	0.9034	0.1905	0.5299	1.2768	22.48	<.0001
Search	1	-0.5409	0.0771	-0.6920	-0.3898	49.24	<.0001
ExcelHlth	1	0.0437	0.0797	-0.1125	0.1999	0.30	0.5835
FairHlth	1	-0.0735	0.1111	-0.2914	0.1443	0.44	0.5081
PoorHlth	1	0.1818	0.2071	-0.2240	0.5876	0.77	0.3800
LessHSed	1	-0.1704	0.1535	-0.4713	0.1306	1.23	0.2671
HSedPlus	1	-0.3153	0.0980	-0.5073	-0.1232	10.35	0.0013
IVyrDegree	1	-0.3377	0.1038	-0.5412	-0.1343	10.59	0.0011
IVyrDegreePlus	1	-0.1994	0.1318	-0.4578	0.0591	2.29	0.1305
Insur	1	-0.0687	0.1162	-0.2965	0.1591	0.35	0.5545
kids	1	-0.0480	0.0879	-0.2202	0.1242	0.30	0.5848
Separated	1	0.1375	0.1031	-0.0645	0.3396	1.78	0.1822
NeverMar	1	0.2005	0.1190	-0.0328	0.4337	2.84	0.0921
Student	1	-0.1248	0.1429	-0.4049	0.1552	0.76	0.3823
hispc	1	-0.0522	0.1534	-0.3530	0.2485	0.12	0.7337
Black	1	-0.1762	0.1171	-0.4058	0.0534	2.26	0.1325
NonWhite	1	-0.1572	0.1604	-0.4715	0.1571	0.96	0.3270
HiIncome	1	-0.0206	0.0901	-0.1972	0.1561	0.05	0.8196
LoIncome	1	0.0339	0.0958	-0.1538	0.2216	0.13	0.7236
AGE	1	0.0151	0.0029	0.0094	0.0208	26.96	<.0001
Male	1	0.0478	0.0729	-0.0951	0.1907	0.43	0.5124

The SAS System

The Probit Procedure

Summary Statistics of Continuous Responses							
Variable	**Mean**	**Standard Error**	**Type**	**Lower Bound**	**Upper Bound**	**N Obs Lower Bound**	**N Obs Upper Bound**
DocVis	0.095819	0.294398	Regular				

Model Fit Summary	
Number of Endogenous Variables	1
Endogenous Variable	DocVis
Number of Observations	2703
Missing Values	228
Log Likelihood	-447.99584
Maximum Absolute Gradient	8.2842E-11
Number of Iterations	0
Optimization Method	Newton-Raphson
AIC	939.99167
Schwarz Criterion	1070

Parameter Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**
Intercept	1	0.153433	0.029727	5.16	<.0001
Search	1	0.109807	0.014044	7.82	<.0001
ExcelHlth	1	-0.011269	0.012810	-0.88	0.3790
FairHlth	1	0.007103	0.016811	0.42	0.6726
PoorHlth	1	-0.021225	0.026538	-0.80	0.4238
LessHSed	1	0.028697	0.020906	1.37	0.1699
HSedPlus	1	0.044963	0.014720	3.05	0.0023
IVyrDegree	1	0.046002	0.016435	2.80	0.0051
IVyrDegreePlus	1	0.019071	0.020197	0.94	0.3451
Insur	1	0.004641	0.019169	0.24	0.8087
kids	1	0.012069	0.015044	0.80	0.4224
Separated	1	-0.015544	0.014577	-1.07	0.2863
NeverMar	1	-0.024218	0.019110	-1.27	0.2051
Student	1	0.030266	0.025468	1.19	0.2347
hispc	1	0.014072	0.025540	0.55	0.5817
Black	1	0.028788	0.019266	1.49	0.1351
NonWhite	1	0.032169	0.027564	1.17	0.2432
HiIncome	1	0.000378	0.015268	0.02	0.9802
LoIncome	1	-0.000134	0.014104	-0.01	0.9924
AGE	1	-0.001961	0.000428	-4.58	<.0001
Male	1	-0.008062	0.011375	-0.71	0.4785
_Sigma	1	0.285590	0.003884	73.52	<.0001