NBER WORKING PAPER SERIES

CONSUMER DEMAND FOR HEALTH INFORMATION ON THE INTERNET

M. Kate Bundorf
Laurence Baker
Sara Singer
Todd Wagner

Working Paper 10386
http://www.nber.org/papers/w10386

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2004

Consumer Demand for Health Information on the Internet
M. Kate Bundorf, Laurence Baker, Sara Singer, and Todd Wagner
NBER Working Paper No. 10386
March 2004
JEL No. I1, D83

ABSTRACT

The challenges consumers face in acquiring and using information are a defining feature of health care markets. In this paper, we examine demand for health information on the Internet. We find that individuals in poor health are more likely than those in better health to use the Internet to search for health information and to communicate with others about health and health care. We also find that individuals facing a higher price to obtain information from health care professionals are more likely to turn to the Internet for health information. Our findings indicate that demand for consumer health information depends on the expected benefits of information and the price of information substitutes.

M. Kate Bundorf
Stanford University School of Medicine
HRP Redwood Building
Room 257
Stanford, CA 94305-5405
and NBER
bundorf@stanford.edu

Laurence Baker
Stanford University School of Medicine
HRP Redwood Building
Room T-110
Stanford, CA 94305-5405
and NBER
laurence.baker@stanford.edu

Sara Singer
Center for Health Policy and
Center for Primary Care and Outcomes Research
Stanford University
singer@healthpolicy.stanford.edu

Todd Wagner
VA Palo Alto Health Care System
twagner@stanford.edu

I. Introduction

The challenges consumers face in acquiring and using health information is a defining feature of health care markets. In his seminal article, Arrow (1963) identified two characteristics of health information that affect the functioning of health care markets. First, because the production and transmission of health information is costly, differences in the value of health information among consumers results in an unequal distribution of information or knowledge about health and medical care. Second, difficulties consumers have in valuing information, "if [the consumer] knew enough to value the information, he would know the information itself (pg. 946)", lead to problems of the marketability of consumer health information.

A number of subsequent studies have examined the implications of these features of consumer demand for health information. Arrow's observations regarding the existence of unequally informed consumers underlie studies of the relationship between information or knowledge and demand for inputs into the production of health. Either measuring consumer knowledge about health and health care directly or using education as a proxy for knowledge, these studies have found that more informed individuals use more health care than those who are less informed (Kenkel 1990; Kenkel 1994; Hsieh and Lin 1997), suggesting that less informed individuals potentially underestimating the productivity of medical care (Kenkel 1990). In a study of the relationship between technological change and health disparities, Goldman and Lakdawalla (2001) demonstrate that innovations that reduce the amount of time necessary to improve health may reduce health disparities while innovations that require greater time investments to improve health may increase health disparities. Their results are based on the assumption that the strong, positive empirical relationship between education and heath represents greater efficiency among those with higher levels of education in producing health. Finally, a study of the expansion of coverage for preventive services in the U.S. Medicare

program demonstrates that consumer knowledge about health insurance coverage for preventive services is associated with greater demand for these services (Parente, Salkever et al. 2003).

The information asymmetry between physicians and patients identified by Arrow has led to a large and unresolved literature examining the extent to which physicians "induce demand" on the part of patients, exploiting their information advantage to make patients consume more services than they otherwise would (McGuire 2000). Kenkel (1990) links variation in information or knowledge across consumers to information asymmetries by proposing that more informed consumers are less likely to experience demand inducement, although the empirical findings of his study do not support this hypothesis.

Finally, based on the work of Stigler (1961), studies of the relationship between consumer search for price information and health care prices have generally demonstrated that increased search by more informed consumers lowers the level of and reduces the variation in health care prices (Pauly and Satterthwaite 1981; Haas-Wilson 1990; Sorensen 2000).

Although the literature demonstrates the importance of information seeking behavior for the functioning of health care markets, surprisingly few studies have examined Arrow's underlying proposition regarding differences among consumers in their demand for health information. Marquis (1985) tests the hypothesis that consumers facing higher cost sharing are more likely to search for lower cost providers, but finds minimal supporting empirical evidence. In their analysis of a community-wide informational intervention, Wagner and colleagues (2001) find that the distribution of free health information increased demand for self-care resources and reduced demand for information from health professionals. Consistent with the proposition that those that benefit most from health information are more likely to use it, they find that poor health status and the presence of children in the household were positively associated with the use of self care resources. They also find that those facing higher costs of accessing information from physicians, as measured by travel time to usual source of care, were also more likely to use self-

care resources, although they do not find evidence that the price of physician services as measured by insurance status has a positive effect on the use of self-care resources.

The emerging role of the Internet as a source of health information provides new impetus to study the determinants of demand for health information. One of the primary uses of the Internet by consumers is to obtain information about health and health care. In a 2001 survey, 40% of Internet users indicated that they had used the Internet within the last year to obtain information about health and medical care (Baker, Wagner et al. 2003). In addition, a number of studies from health care as well as other industries suggest that the availability of Internet shopping has reduced consumer prices (Brown and Goolsbee 2000; Zettelmeyer, Morton et al. 2001; Pauly, Herring et al. 2002), although exceptions exist (Clay, Krishan et al. 2001). In general, these types of studies suggest that the availability of health information on the Internet may have major implications for consumer behavior and market efficiency.

In this paper, we examine differences among consumers in their demand for health information, focusing on information from the Internet. We find that, consistent with the underlying economic theory of demand for information, those who potentially benefit the most from health information and those who face higher costs of obtaining information from other sources are more likely to turn to the Internet as a source of health information.

II. Theoretical Framework

Demand for Health Information

Our theoretical framework is based on a human capital model of demand for health (Grossman 1972). Demand for health information is derived from demand for health, and individuals gain utility either directly from being healthy or indirectly from the greater productivity in both labor and leisure associated with better health. Although many possible inputs for the production of health exist, including healthy behaviors, medical care from health care professionals, and self-care products, consumers are often poorly informed about the

available alternatives. In addition, uncertainty exists in the effectiveness of any particular treatment or input for a given individual. The lack of knowledge on the part of consumers about the potential alternatives in the production of health and their likely effectiveness creates demand for health information. In summary, consumers seek information to improve their knowledge about the production of health. They use this knowledge, in turn, to choose among different potential inputs into the production of health.

Demand for health information is affected by the expected costs and benefits to consumers of acquiring that information. In particular, individuals in poor health or those who are likely to experience poor health are more likely to benefit from knowledge about the production of health than those in better health. Thus, we would expect those in poor health to have greater demand for health information. More highly educated individuals are also likely to have greater demand for health information. This is because schooling may enhance allocative efficiency in the production of health, increasing demand for the inputs into the production of health. This is particularly likely to be true for information regarding self-care (Goldman and Lakdawalla 2001). In addition, higher levels of education may be associated with lower costs of acquiring health information, particularly when acquiring information requires identifying, accessing and processing information from multiple sources.

Sources of Health Information

Although health information is available from a variety of sources, traditionally physicians and health care professionals have been the primary source for consumers. Physicians have superior information about the production of health due to both their extensive training and their experience in treating patients. Because health production is highly customized based on the characteristics of individual patients, both sources of information create a gap between physicians and patients in their ability to obtain, understand, and make recommendations based on health information. This asymmetry in information between physicians and consumers creates a market for health information from physicians.

Physicians, however, also provide health care services for which they are more likely to be reimbursed than for the information itself. This has created concern over the extent to which physicians act as objective agents for their patients in choosing inputs into the production of health given their financial incentives and superior knowledge. However, asymmetries in information between physicians and patients in the opposite direction may also exist. Patients often have greater information than physicians about their individual characteristics, such as their symptoms, medical history, likely adherence to treatment, and preferences among alternatives that may affect the optimal choice of inputs into the production of health (Haas-Wilson 2001).

Consumers may also obtain health information through non-physician channels. For example, consumer medical guides, the popular press, and recommendations from family and friends all serve as alternative sources of health information. Other sources include professional journals or publications by health plans or not for profit organizations. In 2001, 38% of Americans looked for or obtained information about a personal health concern from a source other than their doctor (Tu and Hargraves 2003).

Non-physician information may serve as either a substitute for or a complement to health information from a physician. Consumers may forego contacts with physicians in response to the availability of information from alternative sources depending on the value they place on these alternative sources. This is more likely to be the case for consumers for whom the opportunity costs of obtaining information from a physician are high either due to a lack of health insurance coverage for physician services or greater time costs associated with physician visits. Consumers facing these costs may turn to consumer health information as a substitute for health information provided by a physician.

Non-physician information may also complement that provided by a physician. For example, a physician may recommend a particular type of diet or treatment and a patient may seek out other patients on the same regime for more detailed advice on how to incorporate these instructions into their daily living. Consumers may also search for information to evaluate the

quality of either the information or the services they receive from a physician, effectively attempting to address information asymmetry in favor of the physician. Consumer health information may also address information asymmetries in the other direction. For example, patients learning more about a particular condition may identify alternative treatments that better match their underlying preferences than those recommended by a physician. The ability of patients to gain information from alternative channels, including second opinions from other physicians, limits the extent to which physician behavior can deviate from that desired by the patient. Finally, consumers, particularly those facing high out-of-pocket costs for health care services, may search for information on the prices of health care providers.

The emergence of the Internet as a channel to distribute health information has placed greater importance on understanding consumer demand for health information from non-physician sources. Electronic dissemination may increase the supply of consumer health information both by dramatically reducing the costs of distributing information and by addressing public good problems that reduce the supply of information through website sponsorships, advertising, or requiring consumer payments for information, particularly when it is customized. Although the number of websites providing consumer health information is difficult to quantify due to both definitional issues and difficulty tracking sites in the face of turnover and consolidation, high estimates suggest that thousands of on-line health-related sites exist (Danzon and Furukawa 2001).

The extent of demand for Internet-based health information, its effect on demand for information from other sources, and its ultimate impact on the productivity of health care, however, are less clear. One possibility is that the availability of Internet-based information may dramatically reduce consumer search costs, increasing demand for consumer health information. If consumers are able to locate valuable information on the Internet, this potentially has the effect of reducing the information asymmetries that characterize the relationship between physician and patients and, thereby, reducing the value and the overall demand for physician information.

Valuable health information from the Internet may also complement demand for physician services. Not only will consumers be better able to evaluate the quality of physician services, they will also be able to be more effective participants in managing their care, ultimately increasing the productivity of medical care. Alternatively, if patients read low quality information and delay seeking needed care or bring the bad information to physicians, then this may actually reduce the production of health and/or productivity of medical care.

Consumer demand for health information from the Internet, however, may be limited. Electronic information sources are unlikely to be of much assistance to patients in acute settings. In addition, individuals seeking health information on the Internet may, in reality, face substantial search costs given the large number of potential sources of information available and the difficulty they face in evaluating the quality of the information available from any single source. Consistent with these potential difficulties in accessing Internet-based health information, although 38% of consumers in a national survey conducted in 2002 looked for or obtained health information from a source other than their doctor, only 16% identified the Internet as a source of information (Tu and Hargraves 2003). Consumers were more likely to report seeking information from books or magazines or friends or relatives.

In this paper, we examine consumer demand for health information from non-physician sources, focusing on the Internet. The discussion above suggests that characteristics of individuals that affect both demand for health information overall and demand for information from the non-physician sources will affect demand for information from the Internet. In particular, those currently in poor health or those whose higher expectations of future poor health are more likely to benefit from health information and thus have greater demand for health information than those in better health. Education is expected to be positively associated with both demand for any health information and demand for health information from non-physician sources due to both the greater allocative efficiency of more highly educated individuals in producing health, the lower cost to these individuals of accessing information without the

assistance of a health professional, and their greater ability to self manage their care. Finally, the price of substitutes for consumer health information is also a determinant of demand for health information on the Internet. Individuals facing higher costs, either out of pocket or opportunity costs, in accessing physicians and health care providers are more likely to search for health information on the Internet.

III. Empirical Methods

a. Data

The data source for this analysis was a survey of consumer use of the Internet for health conducted during December 2001 and January 2002. Respondents were selected from an Internet-enabled panel of over 60,000 households created and maintained by a survey research firm. The panel was created by randomly selecting individuals to be offered free Internet access via Web-TV in exchange for answering occasional surveys administered through the Web-TV. At the time of our survey, the panel recruitment response rate was 50.6% and the household connection rate was 70.2%.[1] Our survey was sent to 12,878 panel members, including a random sample of adults over age 21 and over-samples of individuals over age 50 and veterans. The survey completion rate was 69.4%. In our analyses, we use post-stratification weights, which correct the distribution of respondents to match the known distribution of the U.S. population on age, sex, race, education, region, metropolitan residence, and veteran status, to account for the over-sampling and non-response. Comparisons between the study sample and other nationally representative data sources along a number of dimensions of demographic, socioeconomic, and health status indicate that the survey panel is representative of the U.S. population along these dimensions (Baker, Bundorf et al. 2003).

[1] The household connection rate is the percentage of recruited households for which an adult completed the initial demographic profile survey.

Our study sample is unique in that all members have at least a minimum level of Internet access through the WebTV, although some respondents also have access through a computer they obtained independent of panel membership. This has two important implications for our study. First, the cost of Internet access in terms of out-of-pocket payments for basic Internet service is identical among survey respondents. Thus, access to health information on the Internet among the study sample is not affected by the monetary cost of Internet access. Second, our results are not restricted to the population currently with Internet access. Because our sample is comprised of a random sample of individuals offered free Internet access, our findings apply to a broader population than current Internet users. Of course, some of those randomly selected to receive Internet access declined the offer, so that our sample is representative of the population of individuals responding to free Internet access rather than the general population.

In this study, we restrict our analysis to adults 21-64 years of age, excluding older adults because health insurance status is a key study variable and individuals over 64 in the U.S. are nearly universally covered by the Medicare program. The characteristics of our study sample are presented in Table 1.

b. Empirical Models

Our objective in the empirical analysis is to test models of consumer demand for health information on the Internet, and we test two hypotheses. First, individuals who obtain greater benefits from health information in general are more likely to turn to the Internet for health information. Second, individuals facing higher costs in accessing health information from health care providers in traditional settings are more likely to use the Internet for health information. For the first hypothesis, we use a measure of health status as a proxy for demand for health information. We propose that those with chronic conditions are more likely to benefit from health information. For the second hypothesis, we use two proxies for the costs of accessing information from providers in traditional settings, including health insurance status and distance from an individual's usual source of care. Individuals without health insurance coverage face

higher out-of-pocket payments for both physician visits and physician services. Travel time to the individual's regular source of care serves as a measure of the opportunity cost of visits to providers in traditional settings (Acton 1975; Cauley 1987; Currie and Reagan 2003).

The probability of a consumer searching for health information on the Internet, however, is the product of two components:

Pr(Search Internet for health information)= Pr(Search for consumer health information)*
Pr(Use the Internet|Search for consumer health information) (1)

A subset of the population will search for any consumer health information and a subset of those searching will use the Internet in their search. Demand for health information on the Internet will be affected by individual characteristics that affect both probabilities. This is important to note in the context of this analysis since we measure and student the frequency of using the Internet for health information (the left hand side of equation 1), but do not have data that allows us to measure separately the two components on the right hand side. Specifically, we do not observe the frequency of searching for health information from other sources. As a result, our estimates should not be construed as estimates of the frequency of searching for any information. In addition, we are not able to disentangle the effects of individual characteristics on the probability of searching for any consumer health information and the probability of using the Internet conditional on searching.

The basic empirical model for our analysis is:

$$Y_i = \alpha + \beta_1 H_i + \beta_2 P_i^S + \delta X_i + \varepsilon_i \quad (2)$$

where

Y_i is an indicator of use of the Internet for consumer health information,

H_i is a binary indicator of health needs,

P_i^s is the price of substitutes for consumer health information

X_i is a vector of control variables that affect individual demand for health information, and

ε_i is the error term.

One potential issue in estimating this model is that the error term may contain information about unobserved preferences for health and health care that are correlated with both P_i^s, health insurance status in particular, and demand for health information on the Internet. For example, if P_i^s is an indicator that the survey respondent has no health insurance coverage, it will be negatively correlated with unobserved preferences for health and health information in general. This will create a downward bias in the estimates of the effects of the price of information substitutes on demand for health information on the Internet. To address this, we estimate the following alternative version of the empirical model:

$$Y_i = \alpha + \beta_1 H_i + \beta_2 P_i^S + \beta_3 H_i * P_i^S + \delta X_i + \varepsilon_i \quad (3)$$

By interacting the indicators of health and insurance status, we use a differences-in-differences approach to identifying the effect of price of the substitutes for health information, as measured by health insurance status, on demand for health information on the Internet. β_1, which measures the effect of no health insurance coverage on demand for health information on the Internet, captures the effect of unobserved preferences for health and medical care that affect both demand for health information and health insurance. The coefficient on the interaction term (β_3), in contrast, measures the incremental effect of the price of information substitutes among those expected to have greater demand for health information.

We perform additional analyses on the subset of the study sample indicating they used the Internet for consumer health information in the last year intended to provide additional tests for our main findings. In these models, we examine the types of applications for which consumers turn to the Internet and the self-reported effects on the utilization of physician

services. Although the empirical models used in these analyses are similar to Models 1 and 2, the models are estimated on the subsample of Internet users.

c. Study Variables

The dependent variables in our analysis are measures of use of the Internet and e-mail for health and health care. We asked survey respondents, "Within the past year, how often did you look on the Internet for information or advice about health or health care?" From this question, we constructed two measures of utilization of the Internet for health information based on the frequency with which they turned to the Internet. Individuals who "ever searched for information on the Internet" are those who indicated they had done so at least once within the last year. Individuals who "sometimes searched for information on the Internet" are those who indicated they had done so every few months or more frequently. We also asked individuals, "In the past year, about how often did you use email or the Internet to a) communicate with a doctor or other health care provider; b) communicate with a family member or friend about health or health care; or c) communicate with other people who have health conditions or concerns like mine. For each of these questions, we identify users as those indicating they did this at least once in the past year.

For those indicating they had looked on the Internet for information or advice about health or health care at least once in the past year, we followed up with questions about for whom they were looking for health information (myself, child, parent, someone else) during the most recent time they used the Internet and the type of information for which they were looking. For those looking for information for themselves, the mutually exclusive choices for individuals not indicating a self-reported chronic condition were: 1) Information about symptoms you were experiencing or a condition you thought you had; 2) Information about diseases you are at risk of getting or that you want to prevent; 3) Information about doctors or hospitals; 4) Information about healthy lifestyles, fitness, or nutrition; and 5) Other information. Individuals with at least one chronic condition were provided with these choices as well as information about {condition}, where condition was customized based on the health conditions the respondent had reported

earlier in the survey. In cases in which a respondent reported more than one chronic health condition, the condition to which this question referred was randomly chosen. Among those who indicated having a chronic condition, we collapsed the responses indicating search for information about symptoms or conditions the individuals thought they had and search for information about a particular condition into a single measure of searching for information about a health condition.

These questions about the type of information individuals were seeking provide additional evidence on whether health information from the Internet serves as a complement to or a substitute for information from a health care professional. In particular, evidence that uninsured individuals are more likely to search for the types of information that health care professionals generally provide, such as information about symptoms or a condition that an individual has, is consistent with information from the Internet serving as a substitute for information from a health professional for individuals facing higher costs to access health care providers. If those with insurance are more likely to search for particular types of information on the Internet, this provides evidence of complementary relationships between the sources of information.

The key independent variables in our analysis are measures of the costs and benefits of consumer health information on the Internet. These include a measure of health needs, which is an indicator of at least one of five self-reported chronic conditions including high blood pressure or hypertension; diabetes or high blood sugar; cancer; a heart attack, coronary heart disease, angina, heart failure, or other heart problems; and depression. For the first four conditions, survey respondents were asked whether a doctor or other health care provider had ever told them they had the condition. For depression, respondents were asked whether they ever had or a doctor told them they ever had the condition.

The study includes two measures of the price of information from providers in traditional settings. The first is health insurance status. We asked individuals about their current health insurance status using a hierarchical question format intended to assign individuals to a single

category. We also asked a follow-up question to those not indicating any source of coverage to verify that they were uninsured. The categories we use in our analysis are uninsured, any private (including employer-sponsored and individually purchased), and public coverage (which includes individuals indicating an unknown source of coverage as well as public programs). We focus on the comparison between the uninsured and the privately insured, given that the third group is potentially heterogeneous in the characteristics of their coverage. We also use travel time to the individual's usual source of care as a measure of the opportunity cost of obtaining information in traditional provider settings. The categories for this variable are less than 15 minutes, 15-29 minutes, and 30 minutes or more each way.

We include variables in our model intended to control for individual characteristics that affect demand for health information overall and demand for information on the Internet. These include education (<12 years, 12+ years), income (<35,000, 35-75,000, >$75,000), sex (binary indicator of male), and age (21-30, 31-40, 41-50, 51-64). We also include an indicator of whether the individual had used the Internet prior to obtaining access through the survey research firm to control for general familiarity with the Internet. This also controls for mode effects – some respondents may use the Web TV to access the internet while others may use computers.

d. Model Estimation

We estimate maximum likelihood logistic regression models for questions regarding the use of Internet and e-mail for health information and communication since the dependent variables are binary indicators of use. These models are estimated using the entire study sample. We present the coefficients as well as the marginal probabilities for key study variables. The marginal probabilities are the difference of the average predicted probability of the study sample holding the independent variable constant at the indicated value. For models of the type of use, conditional on any use in the last year and the last visit was for oneself, we use multinomial logit models since the dependent variables are a mutually exclusive set of choices. For these models, we present only the marginal effects. All results are weighted to be nationally representative.

IV. Results

a. Descriptive statistics for dependent variables

Although 34% of the study sample had used the Internet in the last year to search for health information, only 13% were frequent users of the Internet for this purpose (Table 2). Individuals with chronic conditions were more likely to use the Internet for health information and the uninsured were less likely to use it for this purpose. Fewer survey respondents used the Internet or e-mail for communication about health matters than searching for information and they were more likely to use it to communicate with family members or friends (23%) than with health care providers (6%) or with other people with similar health conditions or concerns (11%). The uninsured were less likely to use the Internet for each type of communication than the insured and those with chronic conditions were more likely than those without to use it for communication. Both individuals with and without chronic conditions who used the Internet for health information were most likely to use it to obtain information about symptoms they were experiencing and the uninsured were more likely than the average respondent to use it for this purpose.

b. Search for health information and communication with health providers

Those with chronic conditions were more likely to use the Internet to search for health information than those without chronic conditions, consistent with the hypothesis that those with greater underlying health needs have greater demand for health information (Table 3). The effects of the indicators for a chronic condition are statistically significant in both the model of ever searching for information in the past year and for searching more frequently. In the models with only the main effects of insurance status, we find that the uninsured were less likely to use the Internet as a source of health information, although the effects were not statistically significant. Including the interaction of health insurance status with the indicator of a chronic

condition demonstrates that the effect of insurance status depends on health needs. Although the uninsured without chronic conditions were less likely than the insured without chronic conditions to search for health information, among those with chronic conditions, the uninsured were more likely than the insured to turn to the Internet for health information. Because the uninsured face higher prices for accessing providers, these findings are consistent with the existence of positive cross-price elasticities with respect to the price of obtaining information in traditional provider settings. Although travel time to usual source of care is positively associated with searching for health information on the Internet, the effects are not statistically significant.

We also found that more highly educated individuals are more likely to turn to the Internet for health information than those with less education, consistent with education being correlated with greater marginal productivity from self-care (Table 3). Men are less likely than women to turn to the Internet for health information.

The results for communication about health and health care using the Internet or e-mail with family members or friends and people with similar types of health conditions are similar to those for search for health information (Table 4). For these purposes, individuals with chronic conditions were more likely to use the Internet or e-mail. In addition, the uninsured without chronic conditions were less likely than the insured without chronic conditions to use the Internet or e-mail for health care communication, while the uninsured with chronic conditions were more likely than the privately insured with chronic conditions to use the Internet for these types of communication. The interaction effect of the uninsured indicator and the chronic condition indicator, however, is only statistically significant in the case of communication with family members or friends. In the case of communication with health care providers, although the directions of the effects of health needs, insurance status, and their interaction are similar to those for the other dependent variables, they are not statistically significant.

The most striking finding with respect to health-related electronic communication is the effect of travel time, which had a positive and statistically significant effect on each of the forms

of Internet and e-mail communication we examined (Table 4). These results are consistent with they hypothesis that individuals who face higher opportunity costs of accessing health care providers in traditional settings are more likely to use the Internet and e-mail for communication about health care.

In Table 5, we demonstrate that the magnitudes of many of the effects we find are relatively large. Privately insured individuals with a chronic condition were 6 percentage points more likely to have ever looked for health information on the Internet and 4 percentage points more likely to be frequent users of the Internet for health information (Table 5a). The uninsured without chronic conditions were 8 percentage points less likely to have ever used the Internet in the last year to search for information and 4 percentage points less likely to be frequent users of the Internet for health information than the privately insured without chronic conditions. The magnitude of the interaction effect was 7 percentage points both for ever searching for information and for more frequent use of the Internet.

In the case of communication with health care providers, individuals with the greatest distance to their usual source of care (>30 minutes) were 3 percentage points more likely to report use of the internet or e-mail for this purpose and this represents a 67% increase (Table 5b). Individuals characterized by longer travel times to their usual source of care were also more likely to report using the Internet for communication with family members or friends (8 percentage points) as well as others with similar health conditions (6 percentage points).

c. The types of health information individuals access on the Internet

We find that insurance status affects the type of health information for which individuals search using the Internet and the effects are similar for individuals with and without chronic conditions. In particular, among both individuals with (Table 6) and without (Table 5) chronic conditions, the uninsured were more likely to look for information about diagnosing symptoms (or information about their condition in the case of those with an identified chronic condition) and

less likely to look for information about health care providers than to look for other types of information. Travel time to a usual source of care has statistically significant effect on the type of information sought for those with chronic conditions (Table 6), but not for those without chronic conditions (Table 5). In particular, individuals with longer travel times to their usual source of care were less likely to have sought information about symptoms they were experiencing or their chronic condition than other, unspecified types of information.

V. Conclusions

The Internet is an important source of information for individuals likely to seek health information. We find that individuals in poor health, defined as having at least one of five self-reported chronic conditions, are more likely than those without a chronic condition to use the Internet to search for health information and to communicate with others about health and health care. These results support the use of indicators of the presence of chronic conditions as proxies for the benefits of health information as well as our hypothesis that demand for health information on the Internet is higher among those likely to benefit most from health information in general. The alternative result was possible. Individuals with chronic conditions may already have had access to adequate information through alternative sources such as physicians and other health care providers. Our findings, in contrast, suggest that the Internet is an important source of health information even among those who may be highly informed about their health status through other sources.

Individuals facing a higher price to obtain information from health care professionals are more likely to turn to the Internet for health information. Our findings for insurance status provide evidence that a higher out-of-pocket price for physician services increases demand for health information from alternative sources. Among those with one of five common chronic conditions, the uninsured were more likely to turn to the Internet for health information than the insured. Among individuals without chronic conditions, in contrast, the uninsured were less

likely to seek health information on the Internet. We interpret this as evidence of a positive cross price elasticity of demand for health information on the Internet, with the uninsured without chronic conditions serving as controls for unobserved preferences for health and health care among the uninsured. We also found that the price as measured by the opportunity cost of time in visiting providers in traditional settings affects demand for health information on the Internet, primarily in the form of communication with others. Those reporting longer travel times to their usual source of care were more likely to report using the Internet or e-mail to communication with provider, family and friends, and other patients. This is consistent with past research that analyzed at self-care books, telephones advice nurses and computerized self-care information (Wagner, Hu and Hibbard, 2001).

Differences between the insured and the uninsured in the types of information they seek provide further support for the proposition that the price of information alternatives affects information seeking behavior. Among those using the Internet for health information, the uninsured, both with and without chronic conditions, were more likely to turn to the Internet to diagnose symptoms they were experiencing or to obtain information about their chronic condition. This indicates that the uninsured use the Internet to obtain information that those with insurance would generally seek from a physician, consistent with the existence of positive cross price elasticities for information from a physician and self-care information. Those facing high costs of information from providers in traditional settings are likely to turn to the Internet for important health care needs.

Our study had some limitations that could affect the results. First, we asked about five common chronic illnesses, but the non-chronically ill group could have another chronic illness that we did not identify. This suggests that our results may be conservative, as misclassification would attenuate the findings. In addition, our study is based on a sample of individuals randomly chosen to receive Internet access in exchange for periodically filling out surveys. While this design reduces the likelihood that our findings with respect to information seeking behavior are

driven by the price of Internet access, it does introduce the possibility that our sample is not nationally representative due to selection into the survey research panel. If this is the case, we anticipate that these individuals are more likely to be interested in obtaining information from the Internet. The implication is that our findings may overestimate the extent of Internet use for health information among the general population but underestimate the extent among those obtaining Internet access on their own.

The results of our study generally support economic models of information seeking behavior on the part of consumers. Consumers consider the costs and benefits of information when making decisions about information seeking. Demand for health information is higher among those more likely to benefit from it and the price of alternative sources of information, including both out-of-pocket costs and the opportunity costs of time, affects individual choice of information sources. Based on these findings, we would expect that a shift toward increased patient cost sharing at the point of service would result in greater demand among plan enrollees for self-care information. Our findings also imply that Internet-based resources may be most attractive to individuals facing high opportunity costs, such those in remote areas or high-income individuals.

Our findings also indicate that Internet is an important source of information for the chronically ill and the uninsured, suggesting that the Internet serves populations with serious health needs facing significant barriers to accessing health care in traditional settings. Although this creates an opportunity to provide resources to previously underserved populations, it also points to the potential importance of tools to evaluate and monitor the quality of health information on the Internet and to direct consumers to reliable sources of information.

Acknowledgement: The authors gratefully acknowledge funding from the Department of Veterans Affairs, Stanford University Office of Technology and Licensing, and the National Institute on Aging, Grant # AG17253. Bundorf was also supported by grant KO2-HS11668-01 from the Agency for Healthcare Research and Quality. A preliminary version of this paper was presented at the 2003 International Health Economics Conference, the 2003 International Society for Technology Assessment in Healthcare, the 2003 meeting of the American Association of Public Health.

References:

Acton, J. P., 1975. Nonmonetary factors in the demand for medical services-some empirical evidence. Journal of Political Economy 83(3): 595-614.

Arrow, K. J., 1963. Uncertainty of the Welfare Economics of Medical Care. American Economic Review 53(5): 941-73.

Baker, L., M. Bundorf, et al. 2003. Validity of the Survey of Health and Internet and Knowledge Network's Panel and Sampling. Stanford, CA, Stanford University: 1-13.

Baker, L., T. H. Wagner, et al., 2003. Use of the Internet and E-mail for Health Care Information: Results from a National Survey. JAMA 289(18): 2400-2406.

Brown, J. R. and A. Goolsbee, 2000. Does the Internet Make Markets More Competitive?, National Bureau of Economic Research: 1-29.

Cauley, S. D., 1987. The Time Price of Medical Care. The Review of Economics and Statistics 69(1): 59-66.

Clay, K., R. Krishan, et al., 2001. Prices and Price Dispersion on the Web: Evidence from the Online Book Industry, National Bureau of Economic Research: 1-42.

Currie, J. and P. B. Reagan, 2003. Distance to Hospital and Children's Use of Preventive Care: Is Being Closer Better and for Whom? Economic Inquiry 41(3): 378-391.

Danzon, P. M. and M. F. Furukawa, 2001. Health Care: Competition and Productivity. The Economic Payoff from the Internet Revolution. R. E. Litan. Washington D.C., Brookings Institution Press: 189-234.

Goldman, D. and D. Lakdawalla, 2001. Understanding Health Disparities across Education Groups, National Bureau of Economic Research: 1-57.

Grossman, M., 1972. On the Concept of Health Capital and the Demand for Health. Jounal of Political Economy 80(2): 223-243.

Haas-Wilson, D., 1990. Consumer Information and Providers' Reputations: An Empirical Test in the Market for Psychotherapy. Journal of Health Economics 9: 321-333.

Haas-Wilson, D., 2001. Arrow and the Information Market Failure in Health Care: The Changing Content and Sources of Health Care Information. Journal of Health Politics, Policy, and Law 26(5): 1031-1042.

Hsieh, C.-R. and S.-J. Lin, 1997. Health information and the demand for preventive care among the elderly in Taiwan. Journal of Human Resources 32(2): 308-333.

Kenkel, D. S., 1990. Consumer Health Information and Demand for Medical Care. The Review of Economics and Statistics 72(4): 587-595.

Kenkel, D. S., 1994. The demand for preventive medical care. Applied Economics 26: 313-325.

Marquis, M. S., 1985. Cost-Sharing and Provider Choice. Journal of Health Economics 4: 137-157.

McGuire, T. G., 2000. Physician Agency. Handbook of Health Economics. A. J. Culyer and J. P. Newhouse. Amsterdam, Elsevier Science, B.V. 1A: 461-528.

Parente, S. T., D. Salkever, et al., 2003. The role of consumer knowledge of insurance benefits in the demand for preventive health, National Bureau of Economic Research: 1-32.

Pauly, M. V., B. J. Herring, et al., 2002. Health Insurance on the Internet and the Economics of Search, National Bureau of Economic Research: 1-27.

Pauly, M. V. and M. A. Satterthwaite, 1981. The Pricing of Primary Care Physicians' Services: A Test of the Role of Consumer Information. Bell Journal of Economics 12(2): 488-506.

Sorensen, A., 2000. Equilibrium Price Dispersion in Retail Markets for Prescription Drugs. Journal of Political Economy 108(4).

Stigler, G., 1961. The Economics of Information. The Journal of Political Economy: 213-225.

Tu, H. T. and J. L. Hargraves ,2003. Seeking Health Care Information: Most Consumers still on the Sidelines, Center for Studying Health System Change: 1-4.

Wagner, T. H., T.-W. Hu, et al., 2001. The demand for consumer health information. Journal of Health Economics 20: 1059-1075.
Zettelmeyer, F., F. S. Morton, et al., 2001. Cowboys or Coward: Why are Internet Car Prices Lower?, National Bureau of Economic Research: 1-33.

Table 1: Sample Characteristics

N=6,289

Variable	Mean
Public or other insurance	0.155
Private Insurance	0.677
Uninsured	0.167
Prior Internet Access	0.580
Travel time < 15 minutes	0.500
Travel Time 15-29	0.354
Travel Time 30+	0.146
Rural	0.225
Chronic Condition	0.461
Education < 12 years	0.556
Education>=12 years	0.444
Income <35K	0.377
Income 35-75K	0.391
Income 75+K	0.135
Income Missing	0.096
Male	0.475
Age 21-30	0.201
Age 31-40	0.240
Age 41-50	0.285
Age 51-64	0.273

Table 2: Descriptive Statistics - Dependent Variables

	Full Sample		Chronics	Uninsured
	N	%	%	%
Use of the Internet for Health Information within the past year				
Ever Search for Information on the Internet	6289	0.340	0.379	0.273
Sometimes Search for Information on the Internet	6289	0.131	0.163	0.111
Use of the Internet or E-mail ever within the last year to:				
Communicate with a doctor or other health care provider	6262	0.056	0.060	0.046
Communicate with a family member or friend about health or health care	6261	0.229	0.274	0.195
Communicate with other people who have health conditions or concerns like mine	6253	0.105	0.133	0.105
Among those who ever searched within the last year and most recent search was for self				
Non-chronics				
Information about symptoms	560	0.389		0.597
Information about prevention	560	0.077		0.067
Information about doctors or hospitals	560	0.045		0.004
Information about healthy lifestyles, fitness, or nutrition	560	0.286		0.242
Other information	560	0.203		0.089
Chronics				
Information about {condition} or symptoms	776		0.469	0.651
Information about prevention	776		0.051	0.061
Information about doctors or hospitals	776		0.028	0.004
Information about healthy lifestyles, fitness, or nutrition	776		0.179	0.126
Other information	776		0.273	0.158

Table 3: Search for information on the internet

	Ever		Sometimes	
	Main Effects	Interaction	Main Effects	Interaction
Uninsured	-0.218	-0.399**	-0.04	-0.444
	[0.134]	[0.182]	[0.183]	[0.278]
Public	0.024	0.033	0.428***	0.442***
	[0.129]	[0.129]	[0.157]	[0.157]
Chronic Condition	0.342***	0.285***	0.454***	0.352***
	[0.086]	[0.090]	[0.118]	[0.126]
Uninsured*Chronic Condition		0.373		0.703**
		[0.255]		[0.352]
Travel Time 15-29	0.072	0.072	-0.054	-0.054
	[0.089]	[0.089]	[0.123]	[0.123]
Travel Time 30+	0.185	0.18	0.162	0.154
	[0.129]	[0.129]	[0.163]	[0.163]
Rural Indicator	-0.091	-0.093	-0.252	-0.258
	[0.115]	[0.115]	[0.160]	[0.161]
Previous Use of Internet	0.530***	0.528***	0.325**	0.322**
	[0.093]	[0.093]	[0.126]	[0.127]
Education 12+ years	0.571***	0.572***	0.450***	0.451***
	[0.087]	[0.087]	[0.119]	[0.120]
Income 35-75K	-0.244**	-0.246***	-0.404***	-0.409***
	[0.095]	[0.095]	[0.133]	[0.133]
Income 75+K	-0.045	-0.05	-0.205	-0.214
	[0.120]	[0.120]	[0.167]	[0.167]
Male	-0.664***	-0.661***	-0.502***	-0.497***
	[0.082]	[0.082]	[0.114]	[0.114]
age3140	0.045	0.039	0.079	0.069
	[0.135]	[0.136]	[0.185]	[0.185]
age4150	0.183	0.177	0.174	0.165
	[0.130]	[0.130]	[0.179]	[0.179]
age5164	0.113	0.115	0.187	0.192
	[0.125]	[0.125]	[0.168]	[0.168]
Constant	-1.116***	-1.085***	-2.297***	-2.236***
	[0.153]	[0.154]	[0.207]	[0.210]
Observations	6289	6289	6289	6289

Logistic Regression - coefficients shown

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Communication using internet or e-mail

	Doctor or other HCP		Family Member or Friend		People w/ HC condition	
	Main Effects	Interaction	Main Effects	Interaction	Main Effects	Interaction
Uninsured	-0.268	-0.457	-0.196	-0.536**	0.065	-0.057
	[0.296]	[0.420]	[0.154]	[0.222]	[0.209]	[0.318]
Public	0.334	0.343	0.062	0.076	0.569***	0.574***
	[0.249]	[0.250]	[0.142]	[0.142]	[0.183]	[0.183]
Chronic Condition	0.150	0.094	0.487***	0.392***	0.518***	0.479***
	[0.189]	[0.193]	[0.097]	[0.101]	[0.146]	[0.156]
Uninsured*Chronic Condition		0.388		0.632**		0.222
		[0.558]		[0.294]		[0.387]
Travel Time 15-29	0.413**	0.414**	0.278***	0.278***	0.253*	0.253*
	[0.188]	[0.188]	[0.099]	[0.099]	[0.150]	[0.150]
Travel Time 30+	0.552**	0.545**	0.453***	0.445***	0.624***	0.621***
	[0.246]	[0.246]	[0.142]	[0.142]	[0.178]	[0.178]
Rural Indicator	0.098	0.095	0.110	0.108	0.052	0.051
	[0.220]	[0.220]	[0.123]	[0.123]	[0.177]	[0.178]
Previous Use of Internet	0.169	0.167	0.304***	0.301***	0.317**	0.316**
	[0.204]	[0.205]	[0.106]	[0.106]	[0.159]	[0.159]
Education 12+ years	0.051	0.050	0.217**	0.218**	0.026	0.026
	[0.178]	[0.178]	[0.098]	[0.098]	[0.147]	[0.147]
Income 35-75K	-0.397*	-0.398*	-0.168	-0.171	-0.282*	-0.283*
	[0.206]	[0.206]	[0.109]	[0.109]	[0.164]	[0.164]
Income 75+K	-0.057	-0.060	0.037	0.030	-0.385*	-0.387*
	[0.250]	[0.250]	[0.137]	[0.137]	[0.215]	[0.215]
Male	0.105	0.108	-0.493***	-0.488***	-0.410***	-0.408***
	[0.176]	[0.176]	[0.093]	[0.093]	[0.136]	[0.135]
age3140	-0.529*	-0.533*	-0.511***	-0.523***	-0.316	-0.319
	[0.296]	[0.297]	[0.158]	[0.159]	[0.220]	[0.221]
age4150	-0.323	-0.325	-0.345**	-0.356**	-0.240	-0.242
	[0.258]	[0.258]	[0.146]	[0.147]	[0.208]	[0.208]
age5164	-0.512*	-0.507*	-0.336**	-0.333**	-0.434**	-0.431**
	[0.269]	[0.266]	[0.142]	[0.141]	[0.204]	[0.204]
Constant	-2.870***	-2.843***	-1.339***	-1.284***	-2.369***	-2.347***
	[0.328]	[0.332]	[0.180]	[0.181]	[0.252]	[0.259]
Observations	6262	6262	6261	6261	6253	6253

Logistic Regression - Coefficients shown

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5a: Magnitude of the Effects of Insurance Status and Health Status

	Insurance Status by Health Status						
	No Chronic		Chronic				
	Uninsured NC0	Insured NC1	Uninsured C0	Insured C1	Uninsured Effect NC0-NC1	Chronic Effect C1-NC1	Interaction Effect (C0-C1)-(NC0-NC1)
Use of the Internet for Health Information within the past year							
Ever Search for Information on the Internet	0.243	0.319	0.373	0.379	-0.076	0.060	0.070
Sometimes Search for Information on the Internet	0.076	0.113	0.188	0.153	-0.037	0.039	0.072
Use of the Internet or E-mail ever within the last year to:							
Communicate with a doctor or other health care provider	0.037	0.056	0.058	0.061	-0.020	0.005	0.016
Communicate with a family member or friend about health or health care	0.131	0.203	0.289	0.271	-0.072	0.068	0.090
Communicate with other people who have health conditions or concerns like mine	0.080	0.084	0.146	0.127	-0.004	0.044	0.023

Table 5b: Magnitude of the Effects of Travel Time

	Travel Time			
	low	medium	high	difference high to low
Use of the Internet for Health Information within the past year				
Ever Search for Information on the Internet	0.329	0.344	0.367	0.038
Sometimes Search for Information on the Internet	0.131	0.125	0.148	0.018
Use of the Internet or E-mail ever within the last year to:				
Communicate with a doctor or other health care provider	0.045	0.066	0.074	0.030
Communicate with a family member or friend about health or health care	0.201	0.247	0.278	0.077
Communicate with other people who have health conditions or concerns like mine	0.088	0.110	0.150	0.062

Table 6: Type of Use during last Visit Conditional on Any Use - Non-chronics

	Symptoms		Prevention		Providers		Lifestyle
Uninsured	0.228	**	0.008		-0.041	***	-0.043
	(0.104)		(0.043)		(0.012)		(0.089)
Public	-0.015		0.036		0.001		0.117
	(0.116)		(0.060)		(0.027)		(0.125)
Travel Time 15-29	0.033		0.009		0.018		-0.031
	(0.063)		(0.032)		(0.019)		(0.061)
Travel Time 30+	0.103		0.036		-0.006		-0.083
	(0.091)		(0.045)		(0.021)		(0.086)
Previous Use of Internet	-0.005		0.021		-0.018		-0.031
	(0.078)		(0.031)		(0.020)		(0.075)
Education 12+ years	0.066		-0.061	*	0.007		0.045
	(0.066)		(0.033)		(0.013)		(0.069)
Income 35-75K	-0.092		0.082	**	-0.012		0.030
	(0.067)		(0.042)		(0.016)		(0.069)
Income 75+K	-0.050		0.115	*	0.014		0.002
	(0.083)		(0.069)		(0.021)		(0.081)
Male	-0.142	**	0.029		0.032	*	0.059
	(0.057)		(0.029)		(0.019)		(0.059)
age4150	0.048		0.018		-0.009		0.029
	(0.071)		(0.036)		(0.017)		(0.069)
age5164	-0.042		0.069		0.000		0.045
	(0.071)		(0.045)		(0.018)		(0.070)
Observations							

Multinomial logistic regression - Marginal probabilities shown

Other types of information is the reference group

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Type of Use during last Visit Conditional on Any Use - Chronics

	Condition or Symptom		Prevention	Providers		Lifestyle	
Uninsured	0.167	**	0.029	-0.012	**	-0.069	
	(0.083)		(0.044)	(0.005)		(0.053)	
Public	-0.129	*	0.055	-0.002		-0.092	*
	(0.072)		(0.036)	(0.006)		(0.048)	
Travel Time 15-29	-0.056		0.005	0.018	*	0.085	*
	(0.059)		(0.021)	(0.010)		(0.051)	
Travel Time 30+	-0.257	***	0.050	0.018		0.009	
	(0.069)		(0.044)	(0.018)		(0.067)	
Rural	-0.043		-0.009	-0.005		-0.049	
	(0.075)		(0.025)	(0.009)		(0.051)	
Previous Use of Internet	0.030		-0.039	0.012	*	-0.055	
	(0.066)		(0.030)	(0.007)		(0.057)	
Education 12+ years	-0.112	*	0.030	-0.007		0.069	
	(0.059)		(0.019)	(0.008)		(0.046)	
Income 35-75K	0.002		0.010	0.030	**	-0.068	
	(0.062)		(0.019)	(0.014)		(0.045)	
Income 75+K	-0.033		-0.003	0.063	*	-0.004	
	(0.078)		(0.024)	(0.038)		(0.064)	
Male	0.068		-0.015	0.008		0.080	*
	(0.056)		(0.020)	(0.007)		(0.046)	
age4150	-0.021		0.017	-0.006		-0.031	
	(0.078)		(0.039)	(0.007)		(0.052)	
age5164	-0.018		0.016	0.006		-0.074	
	(0.066)		(0.027)	(0.011)		(0.048)	
Observations							

Multinomial logistic regression - Marginal probabilities shown

Other types of information is the reference group

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%