

NBER WORKING PAPER SERIES

HEALTH INSURANCE ON THE INTERNET AND
THE ECONOMICS OF SEARCH

Mark V. Pauly
Bradley Herring
David Song

Working Paper 9299
<http://www.nber.org/papers/w9299>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2002

This research was supported, in part, by the Robert Wood Johnson Foundation's Changes In Health Care Financing and Organization (HCFO) Initiative. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2002 by Mark V. Pauly, Bradley Herring, and David Song. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Health Insurance on the Internet and the Economics of Search
Mark V. Pauly, Bradley Herring, and David Song
NBER Working Paper No. 9299
October 2002
JEL No. D83, G22, I10

ABSTRACT

This paper explores the level and dispersion of premiums paid for individual health insurance by comparing “asking price” data posted on an electronic insurance exchange with survey data on premiums actually paid in the period just before the advent of electronic exchanges. The primary theoretical question is whether the pattern of differences between asking prices and transactions prices can be explained using a simple search theory. We hypothesize, following suggestions of Stigler and Rothschild, that higher risks who expect to pay higher premiums for a given policy will engage in more intensive search than lower risks, given the same distribution of asking prices. As a result, for a given distribution of asking prices, the dispersion of premiums actually paid (transactions prices) will be smaller for higher risks. Therefore, the introduction of an electronic exchange should have a larger potential influence on the dispersion and level of premiums paid for lower risks than for higher risks. We find evidence consistent with each of these hypotheses.

Mark V. Pauly, Ph.D.
Chair, Health Care Systems Department
The Wharton School
University of Pennsylvania
208 Colonial Penn Center
3641 Locust Walk
Philadelphia, PA 19104-6218
and NBER
pauly@wharton.upenn.edu

Bradley Herring, Ph.D.
Department of Health Policy & Management
Emory University’s Rollins School of Public Health
1518 Clifton Road, NE. Atlanta, GA 30322
bjherri@sph.emory.edu

David Song
Health Care Systems Department
The Wharton School, University of Pennsylvania
204 Colonial Penn Center
3641 Locust Walk
Philadelphia, PA 19104-6218
dasong@wharton.upenn.edu

I. Introduction

Insurance markets of all types have become available to consumers on the Internet. Electronic methods offer the possibility of reducing the cost to the consumer of searching among policies, companies, and premiums virtually to zero. Before the advent of Internet-based exchanges, the great bulk of consumers used insurance brokers or agents to assist them in their search. Agents are widely thought to have provided information of the type now available on the Internet, at a higher cost. The impact of substituting one technology for another, both overall and for different types of consumers, is not yet fully clear.

Some people obtain health insurance as individuals, in the same way as they obtain auto, homeowners, or most other insurances. At present, this fraction is less than 10 percent of total private insurance since the vast majority of private health insurance is employment-based. Recent policy proposals envision offering the currently uninsured subsidies of various types that they can use to obtain insurance in the individual market, which could greatly increase the size of this market. But how does (or could) this market perform? In this paper we investigate search behavior in this market, using the advent of Internet-based exchanges as an instrument to do so. The individual health insurance market appears to have a competitive market structure, and to earn only competitive rates of return. Yet premiums charged and paid in that market for a given insurance product appear to vary a great deal. The precise nature of premium variation, the contribution of search (or non-search) behavior to that variation, and whether the variation is greater in some circumstances than in others, are not well known. The advent of electronic exchanges for health insurance offers the promise of substantial improvements in the way health insurance markets function. We use data from an exchange to simulate the search process both

before and after the arrival of Internet-based exchanges. Our analysis of the search process tests hypotheses that have implications in other markets as well.

Empirical evidence on the effect of online price information for markets in general is mixed. The advent of life insurance quotes on the Internet appears to have reduced premiums in both Internet and conventional markets (Brown and Goolsbee, 2002). Online auto retailers appear also to have lowered prices and price dispersion of cars (Scott-Morton, Zettelmeyer, and Risso, 2000). However, the advent of the Internet in the book industry appears to have lowered neither prices nor price dispersion (Clay, Krishnan, and Wolff, 2001). In this paper we examine the individual health insurance market in which there has historically been both substantial price dispersion and price inflation.

There is a strong belief that the individual health insurance market has been subject to both high average administrative or “loading” costs and a wide dispersion of premiums for the same product, partly because of costly search and substantial buyer ignorance (Pauly and Herring, 1999). However, virtually all our information on individual-market premiums is based on data about the “posted” or proposed premiums listed by a few large or selected insurance firms; see, for example, Frogue (2001). Information on the level of dispersion of “transactions” prices actually paid to all the firms buyers use in this market is scarce. Since no or few transactions may be expected to occur at unusually high posted prices, a better understanding of what actually happens would come from looking at premiums consumers actually pay, and comparing them with premiums firms propose to charge, rather than focusing entirely on the latter. Such a comparison will tell us both about dispersion of actual realized market prices, and about the search behavior of consumers of different types that contributes to this outcome.

Finally, there is evidence that older people are more likely than young people to obtain insurance in the individual market, other things equal, but no explanation of their behavior currently exists (Nichols and Pauly, forthcoming). Since older people pay higher premiums, why are they more likely to buy? Our analysis provides a potential explanation for this phenomenon that does not rely on an assumption of misperception of risk among young adults. Rather than assume that young people mistakenly believe that they will almost never use medical care, we suggest, and provide evidence for, the hypothesis that search is relatively less efficient for younger people. They respond by buying less frequently. However, the advent of the Internet could in principle change this age bias and result in a reduction in the number of young uninsured persons.

Overview

There exists, to our knowledge, no database that combines detailed information on posted individual insurance prices with information on transactions prices. Instead, most information has come from interviews with firms about their “pricing policies” (Chollet and Kirk, 1998) or from premium quotes for hypothetical customers (Pollitz, Sorian, and Thomas, 2001). However, in this study we match premium quotations from a particular Internet exchange with data from the large sample 1996-1997 Community Tracking Survey’s (CTS) Household Survey that records premiums actually paid.

The CTS data provides information on actual transactions for a relatively large and nationally-representative sample of individual insurance buyers in the period just before the advent of Internet-based health insurance exchanges. The Internet data is provided by the site for

eHealthinsurance.com. This website is the largest health insurance “exchange,” and provides a set of premium “quotes” for a buyer who provides information on date of birth, gender, smoking status, and location. We compare the inflation-adjusted CTS data for 438 individual purchases from 1996-1997 with data extracted from the website in the spring of 2001, when nationwide quotes became common. (We make a similar comparison for family coverage using website data extracted in the summer of 2002.)

This exercise allows us both to judge how premiums paid compare with premiums posted and to investigate whether use of what is arguably a more efficient way to search will in fact provide substantial benefits to different kinds of buyers. In principle, the easy availability of premium quotations from multiple firms for many policies should allow buyers to more easily identify sellers with low premiums and to select products more closely matched to their preferences. Before Internet-based exchanges were available, most buyers used agents or brokers to perform these tasks. Would people be able to do better online than they had been doing with using agents? And would the Internet provide larger potential gains to some buyers than others, either compared to what they would have paid in the pre-Internet era or compared to the insurance benefits they would expect to receive?

Because of state premium-setting regulations, insurers are only allowed to post on the Internet the same premiums (plus a service fee) they would have charged to other buyers (i.e., no “Internet specials”) and generally all of the listing firms’ premiums were posted. While there may have been self-selection of firms that chose to list on exchanges, the comparison itself will tell us whether firms that chose to list were relatively low priced. Buyers as well are self-selected; we do not know the transaction prices that would have been paid by buyers who searched and then decided not to buy, and some consumers never search at all. Since searchers

(or non-searchers) who did not buy might well have paid higher prices than those who did, our results may underestimate the benefit from using the Internet for the average person who searches for individual health insurance.

II. Modeling the Search Process

The model of the search process for our empirical work is based on Stigler's (1961) classic insight that consumers decide how much to search for lower prices by comparing the cost of additional search with the expected benefit (in terms of discovering a yet lower price) of more search. McCall (1965) modified and extended this theory to the more appropriate case of sequential search. A main result Rothschild (1973) draws in his survey of search models is that, if one assumes that all consumers have the same average cost of search, those consumers buying higher priced products will search more and, as a result, the dispersion of final prices paid will be higher for low priced products (Rothschild uses screwdrivers as an example) than for big ticket items (like automobiles).¹

A complete model of the search process and market equilibrium intended to explain the prices paid by different kinds of consumers is best constructed by assuming that the key distinguishing characteristic, for buyers facing the same search cost per seller, is the typical quantity each buyer purchases. The gain from discovering a lower price is obviously proportional to the amount that would be purchased at that price (disregarding negligible second-

¹ Stigler actually said that dispersion would be lower for commodities that form a larger share of total household consumption spending, whether or not the per-unit price was high or low. Sorensen (2000) more recently has argued that "search-induced price variation will be independent of scale" because "dispersion is a function of search costs, which are generally modeled as independent of prices." (p. 838). We reconcile Rothschild's conjecture with Sorensen's proposition in our discussion.

order effects). While we are unaware of a formal model that includes variation in intended quantity, it seems clear that, *ceteris paribus*, prices are likely to be lower relative to cost and less dispersed the larger is the average or typical amount purchased per year. Prices for white bread are probably closer to competitive than prices for screwdrivers. Also, those who typically purchase above-average quantities would be expected to engage in above-average levels of search, and those who typically purchase below-average quantities in below-average levels of search; these behaviors should affect the level and dispersion of prices each type of buyer actually samples.

For health insurance (viewed as a financial instrument) the relevant measure of the price is the loading: the amount which must be paid to the insurer to deliver a dollar of expected benefits, or the percentage of the premium which is not returned to the insured as expected benefits.² For health insurance, there is an identifiable source of variation in the amount of benefits sought: the individual's expected medical expense, or risk level. Higher risk people buy more dollars of benefits than do lower risk people, and the total premium, equaling the price (loading per dollar of expected benefits) times the quantity (dollars of expected benefits), that they pay reflects this variation. For example, premiums are typically two or more times higher for middle-aged people than for younger people.

In effect, a high-risk person buys more "insured dollars" per year than does a low-risk person³ so that, following Stigler's logic, such a person (sampling from the same distribution of offering prices) should on average search more and pay less. (We examine empirically whether

² The Phelps (1997) textbook notes that "the price of insurance is L, the 'loading fee' of the insurance above expected benefits" (p. 343).

³ Thus, in Sorensen's (2000) terminology, although search is assumed to be independent of the loading (as a price), it is dependent on the premium (as a measure of total annual spending, or price times quantity).

the distribution of *quotes* varies by risk level.) Stigler's theory would therefore predict that the dispersion of transactions prices should be lower for individual insurance actually purchased by older persons compared to that for younger persons. It would therefore also suggest that transaction prices should be lower relative to posted or quoted prices for older buyers, if the distribution of posted prices is the same at all risk levels.

To sum up: Figure 1 shows the relationship between the number of searches per high-risk person (who buys more and therefore more costly insurance) and the number per low-risk person. The marginal benefit (MB) is the expected savings from conducting an additional unit of search, and thus both high-risk and low-risk buyers will in equilibrium purchase at the marginal benefit level equal to the search cost. Since $MB_H > MB_L$, high-risk (older) buyers will engage in a greater number of searches, all other things equal. Specifically, if the marginal cost (MC) of search is assumed to be positive and uniform at $\$C$, the number of searches made by high-risk individuals (S_H) will be greater than the number of searches by low-risk individuals (S_L). Hence, we would expect the dispersion of transaction prices for older buyers of health insurance to be lower than that for younger buyers.

A second issue is the relationship of transactions premiums to expected insurance benefit payments. If high-risk buyers engage in more search, the profit margin on their business will be lower than for low-risk buyers engaged in less search. Facing a higher markup (even though a lower premium), low-risk (young) buyers will be less likely to buy insurance. If insurers find it advantageous to offer insurance products to all (profitable) buyers at all risk levels, and if there are some fixed costs at the firm level, there can be a market-level equilibrium in which all firms earn the same profits, prices are dispersed, and yet premiums are closer to expected expenses for high-risk than low-risk buyers (Stigler, 1961; Salop and Stiglitz, 1977; Salop and Stiglitz, 1982).

These considerations suggest our first two hypotheses:

- 1) The dispersion of transactions prices (relative to the dispersion of asking prices) before the advent of Internet-based exchanges will be lower for individual insurance sold to high-risk buyers than for individual health insurance sold to low-risk buyers.
- 2) The ratio of transactions prices to expected expenses will fall as the risk level increases.

The Effect of the Internet

If the marginal search cost falls to zero (i.e., MC falls from \$C to 0), then people of both risk levels end up making the same number of searches--given that they face the same distribution of loadings (or prices per dollar of expected insurance benefit). This number of searches is shown as S* in Figure 1. It therefore follows that the gain (in terms of expected premium reduction) from reducing search costs will be larger for low-risk individuals than for high-risk individuals.

Presumably, the advent of Internet-based exchanges permits a much lower cost of search. The expected impact on any individual depends on the individual-specific change in the cost of search, and the shape of what Stigler describes as the “marginal benefit schedule.” Because high-risk (e.g., older) buyers will wish to search more than low-risk (e.g., younger) buyers, we can immediately conclude that high-risk buyers will be more likely to use the Internet than low-risk buyers, other things equal. If the Internet reduces the cost per search by the same amount for all buyers, the change in the amount of search depends on the slope of the *marginal* search benefit schedule at a relative *high* value for low-risk buyers, versus that at a relatively *low* value for high-risk buyers. High-risk buyers will have already exhausted the gains from search, and

therefore will have less to gain (relative to premiums they have already discovered offline) in terms of finding more favorable prices from using the Internet than will low-risk buyers who previously searched little. This argument therefore implies our third hypothesis:

- 3) The difference between “pre-Internet” premiums and (measures of the magnitude of) “post-Internet” premiums will be smaller for high-risk buyers than for low-risk buyers, other things (including the distribution of quoted premiums) held constant.

All three of these conjectures depend on the similarity of shape of the distributions of quoted premiums by risk level, but our main point here is that the gain in terms of lower prices from shifting from the use of more costly search methods to a much less costly Internet-based method will be smaller for those who previously searched more.

III. Previous Research, Our Methodology, and the Data Used

Previous research provides some evidence in support of the suggested hypotheses. We found, using the 1987 National Medical Expenditure Survey (NMES) and 1996-1997 CTS data, that individual insurance premiums actually paid rose with risk, but at a much less than proportional rate (Pauly and Herring, 1999; Herring and Pauly, 2001). That is, the ratio between premiums paid and expected benefits (the loading) declined as expected benefits rose, just as would be expected if higher risks searched more. Even for easily observable demographic characteristics such as age and gender, the elasticity of transaction premiums with respect to expected expense was significantly less than unity, usually on the order of 0.5. Including all of the high-risk conditions such as diabetes and cancer that could be identified in the data (some of which may not have been known to the insurer at the time insurance was sold), the analogous

elasticity was approximately 0.2. However, neither of those studies contained information on the premiums insurers actually proposed to charge, so the hypothesis that the firms set *posted* premiums closer to expected expense for high risks could not be tested.

The primary data innovation in this paper is the generation of a large number of Internet premium “quotations” for self-only and family health insurance coverage for the sample of individual insurance purchasers in the 1996-1997 CTS Household Survey. We analyze variations within these data for several purposes. We first examine whether the shape of the distribution of quoted premiums varies by risk level. We then test whether the dispersion of paid premiums is lower for older, high-risk individuals than that for younger, low-risk individuals (Hypothesis #1). We then examine the relationship between premiums and risk (for both quoted premiums from the website and paid premiums for the CTS respondents) to test whether the ratio of paid premiums to expected expense decreases as risk increases (Hypothesis #2). Also of interest is whether the ratio of quoted premiums with respect to expected expense is less than unity. Finally, we compare the quoted premiums from the website with the paid premiums from the CTS to examine the potential savings from using the Internet for buyers previously engaging in some other type of costly search. This allows us to test the hypothesis that an Internet-related reduction in search costs would provide a larger benefit to low-risk consumers than to high-risk consumers (Hypothesis #3).

The CTS Household Survey was undertaken in late 1996 and early 1997, just before Internet-based insurance exchanges became available. Even at that time, however, potential insurance buyers had several ways to shop for their individual health insurance. They could use insurance brokers as “proxy shoppers”: while consumers might not be familiar with the distribution of premium quotations because they undertook search relatively infrequently,

insurance brokers in contrast arrange such transactions all the time and are thought to be knowledgeable about the premiums (and other contract terms) available from different insurers. Alternatively, individual buyers could obtain much of the premium information that brokers knew about by spending a modest amount on an information service such as Quotesmith. For about \$15, anyone could have the same kind of report about premiums for a set of policies for an individual with given characteristics as was available to brokers.

Generating Internet Based Comparisons

Our empirical analysis began with the sample of 438 persons who reported purchasing self-only individual (or “nongroup”) insurance in response to the CTS survey, who provided information on the type of insurance plan and deductible, and who could be subsequently matched with our website data. We constructed a computer script (or “automated agent”) that collects pricing information from the public website for the largest Internet exchange, www.ehealthinsurance.com; for more detail, see Pauly, Herring, and Song (2002). A visitor to the website can query for individual insurance plans and premium quotations after entering his or her age, gender, zip code, and smoking habits. Since each respondent to the 1996-1997 CTS Household Survey was asked to provide information on age, gender, smoking habits, location of residence, premium paid, and some characteristics of the insurance bought, our analysis essentially replicates an online query for each of these individuals (upon generating a random zip code from within the county identified in the CTS). This computer script therefore can generate (in several different ways to be described below) relevant sets of website premium quotations for

people of various demographic characteristics. The single policyholders in the CTS survey were associated with 10,650 website premium quotations.

About half of this sample was located in an area where the policy identified in the CTS could be exactly matched with one or more policies listed on the website, based on the plan type (HMO, PPO, etc.) and size of the deductible. We call this the “perfect match” subsample. To use the full 438-observation sample, we estimated a “hedonic” premium regression using the premiums in the website data, relating the website premiums to plan and buyer characteristics. We then used this regression’s results to generate predicted “composite” website quotations to match with the actual CTS premiums; we call this the “composite premium” sample. None of the reasons for missing data seem likely to be related to the behavior we are trying to describe.

Finally, the CTS premium was inflated by a health premium cost index to be consistent with the time period of the Internet data, and the premium quotes and paid premiums were then compared for people with the same characteristics buying the same policies. Although we show results for the expected or mean premium within the distribution of premium quotes, no rational buyer, confronted with the range of website prices from low to high, would choose to pay the average price. Instead, the person would choose to pay a low price—perhaps not the absolute lowest, because that price may have special restrictions, but a price in the lower part of the distribution of quoted prices. Accordingly, we also show comparisons of paid premiums to the 10th decile website premium quote.⁴

⁴ For the “composite premium” sample of website purchase, this estimate was constructed by selecting the premiums whose deviations from the predicted premium were in the bottom decile of all premiums. Then a regression was run on the observations to estimate a “10th decile” premium for each CTS observation.

Generating Estimates of Risk

We want to examine how paid premiums from the CTS sample vary with risk relative to how quoted premiums from the website vary with risk. Following Herring and Pauly (2001), we first generate an estimate of expected expense by regressing actual medical expenses on several age-gender dummy variables and geographic indicators for insured individuals in the 1996 Medical Expenditure Panel Survey (MEPS). We use the coefficients from this regression to “sum up” expected expenses for the sample of individual insurance purchasers in the CTS sample. We then use this predicted value as an explanatory variable in a premium regression, using either the paid premium or premium quotes as the dependent variable; we use the log of the premium and the log of expected expense in the second-stage premium regression to provide an appropriate elasticity estimate with respect to risk (Pauly and Herring, 1999; Herring and Pauly, 2001).

Whatever is the case for consumers in general, our main interest is how the difference between premiums paid in the CTS and quoted premiums from the website varies by consumer risk level. Since by far the most important and consistent determinants of risk variation are age and gender, we also calculate the mean difference between premium quotes and paid premiums by five ten-year intervals for males and females separately. In some cases, the sample size within these ten age-gender cells is too small, so we instead make comparisons simply between individuals aged 18 to 44 and individuals aged 45 to 64.

IV. Results

Dispersion of Premiums (Hypothesis #1)

Recall that we use two methods of calculating the dispersion of premiums. First, for the “perfect match” sample in which each CTS observation was linked to policies from the website with the same plan characteristics, we calculate the mean website premium for each observation in the CTS data. We calculate the standard deviation about the mean for this observation, and then calculate the average standard deviation across all matched observations. By this calculation, the average coefficient of variation for quoted premiums in 2001 was 0.35 for persons aged 18 to 44, and a very similar 0.33 for persons aged 45 to 64. In effect, the distribution of prices posted on the Internet does not seem to vary with the level of risk (as would be expected if relative premiums are highly correlated within a firm that applies the same loading charge to all products).

To use the larger “composite premium” sample of observations where the CTS person can be matched only to a website that does not contain an exactly-matched policy, we estimated our “hedonic” premium regressions using all premium quotations for the CTS sample observation separately for the younger low-risk (aged 18 to 44) and older high-risk (aged 45 to 64) groups. We then calculate the standard errors of each regression and divide them by the mean. The measures of variation for premium quotes in 2001 calculated in this way were 0.31 for the younger sample and 0.30 for the older sample. Based on these comparisons of these two distributions of quoted or posted premiums, we conclude that the distributions about the mean of posted premiums are virtually the same for both risk groups.

For the sample of paid premiums from the CTS data, we do not have direct measures of the distribution or dispersion of premiums paid in the sample of transactions prices because

virtually every observation is a unique combination of age, gender, location, and plan type. Instead we follow the same procedure outlined in the previous paragraph, estimating a hedonic premium regression for the two subsamples. The measures of variation for CTS *transactions* premiums do differ between the younger low-risk sample and older high-risk subsamples, thereby confirming Hypothesis #1. The measure of variation was 0.74 for the low-risk subsample, but only 0.56 for the high-risk subsample.

In short, while the dispersion of asking prices does not appear to vary with risk, the dispersion of the transactions prices people choose to pay from those similar distributions is lower for high-risk individuals than for low-risk individuals. Given similar patterns of asking prices, high-risk individuals end up paying transactions prices which are less dispersed than those paid by low-risk individuals. It is, however, somewhat puzzling that the distribution of posted premiums displays lower variation than the distribution of paid premiums. This is perhaps both due to the premiums in the CTS being self-reported and thus containing random noise and due to the fact that only a fraction of insurers list on these sites.

Relationship of Premiums to Risk (Hypothesis #2)

The next question is how premiums (both quoted premiums and paid premiums) vary with a measure of an individual's risk. (Recall that this estimate of risk is constructed from regressing medical expenditures from the MEPS survey on age, gender, and location.) Specifically, we regress both quoted premiums from the website and paid premiums from the CTS on the measure of risk (and a constant term) using a log specification. We also estimate these models for the full sample, a non-HMO sample, an HMO sample, and the "perfect match"

sample.⁵ The results are shown in Table 1. First, the estimated elasticities are all significantly less than unity. Second, and most importantly, they are all lower for the paid premiums in the CTS than for the quoted premiums from the website; the difference is significant for each comparison, except for the non-HMO subsamples. Note that, because the paid premiums reflect actual approved applicants who have been through underwriting, the insurers would have been able to query buyers about the health status when they applied for insurance (but obviously, neither insurer nor buyer would know about changes in health status that occurred after purchase). We obtain similar results when we use the smaller “perfect match” sample.

These results for paid premiums suggest that the greater search effort higher risks undertake offsets some of the effect of higher expected expenses. It also means that insurance actually sold to high-risk individuals is less profitable on average than that sold to low-risk individuals. (This would still be consistent with individual firms charging similar premiums (relative to those of other firms) at all risk levels.)

The Effect of the Internet on Premium Savings (Hypothesis #3)

Can buyers expect to pay less on the Internet than they actually paid before the Internet-based exchanges were available? Table 2 shows the difference between the median quoted premium from the website (using the “composite” predicted website premium methodology) and the paid premium from the CTS, by different age-gender intervals. The table shows both the absolute difference in premiums and the relative difference in premiums. The difference is always positive: average website premiums are higher, for most of this sample of purchasers,

⁵ Herring and Pauly (2001) noted different elasticities for premiums with respect to risk based on whether the plan was an HMO.

than what they paid before the availability of Internet-based exchanges. Randomly selecting a website premium would have led most of these buyers to pay significantly higher premiums.

However, since we expect consumers to select a low premium—perhaps not the absolute lowest, because that plan may have special restrictions, but a premium in the lower part of the distribution of quotes, we show in Table 3 the comparison of paid CTS premiums to the 10th decile “composite” website premium. The data suggests that, before the advent of Internet-based exchanges, young people paid transactions prices that were higher than the “low” price they could find by going on the website in 2001, whereas older people (specifically, men over age 55 and women over age 45) still would pay less than they would have paid by searching on the Internet. While the sign of the comparisons differs depending on whether the median or the 10th percentile website premium is used, there is a consistent pattern: hypothesized website transactions prices increase relative to pre-Internet transactions prices as risk increases.

Table 4 shows the same type of comparisons between hypothesized website transactions and pre-Internet paid premiums, based on the smaller subsample of perfect matches. The website premium here was the lowest one among those plans that matched to the CTS plan. The premiums on the website are higher than the CTS premiums, but the difference is much smaller for low-risk persons. As before, high-risk individuals had previously obtained much lower transactions prices relative to Internet prices.

So does the Internet save money for individual purchasers? The answer depends in part on who the buyer is and on how the person is assumed to use the Internet to search. Specifically, if we use the average website premium as the benchmark, before the Internet was available all types of purchasers actually paid less than the website price, but the difference was largest by far for older people. (Remember that insurers did not post lower premiums on the website than they

charged before; indeed, usually they added a small amount to the premium they would have charged in order to cover the cost of using the site.) If, instead of paying the average, buyers choose a low price posted on the Internet, younger people might gain (or lose only a little) but older people will still lose substantially from using the Internet. In short, it is possible to save money by using the Internet, but there are no guarantees. The evidence here therefore suggests even more strongly that younger low-risk people were previously less aggressive searchers (and therefore had more to gain or less (in terms of prices) to lose from going to the more convenient posted prices on the web) than were older high-risk people.

However, because the Internet saves on search costs, it may still be rational for older people to use it even if their final premiums were modestly increased. Both sets of results are consistent with the view that high-risk individuals had formerly secured transactions prices that, compared to low-risk individuals, were more advantageous relative to expected benefits. Measured solely in terms of premium savings, high-risk individuals therefore have less to gain from using the Internet than low-risk individuals. However, this smaller gain (or loss) for high-risk individuals must be set against their presumably greater expenditure of effort before the Internet became available for searching to find these more favorable prices.

We also generated a similar sample of family premiums in 2002. We used only a “perfect match” sample here because a regression model did not deal well with the variety of different kinds of families (mature adults and children). Family premiums are higher than individual premiums for either young or older (higher risk) individuals because the expected expense for a family is typically larger than that for an individual. As in the case of high-risk individuals, the median Internet premium for family coverage was found to be higher than the “inflated” CTS premium. The median difference for family premiums was \$867, compared to a

median difference for individual premiums of \$294. This result is consistent with the hypothesis that family purchasers, like high risk individual purchasers, had already searched effectively for the best premiums in the market before the Internet became available, so the premiums they could find on the Internet were not more advantageous than what they had already discovered. That is, as our theory suggests, the results for family coverage are similar to those for high-risk individuals.

Age and the Use of the Web

Are there other theories that could explain why younger people gain more (or lose less) from the use of the Internet than do older people? An alternative (though not mutually exclusive) hypothesis suggested by the life insurance analysis of Brown and Goolsbee (2002) might assume that the implicit cost to using the Internet is lower for younger people. Being more familiar with the use of the Internet, they use it more frequently. However, Brown and Goolsbee implicitly assumed that the amount of search in the pre-Internet period was independent of the use of the Internet. Moreover, it seems implausible that, given the still-rare current use of Internet-based health insurance exchanges, health insurers would have adjusted their posted premiums. In contrast to life insurance pricing discussed by Brown and Goolsbee, health insurance premiums posted on the Internet had to be the same as the prices charged to others, the use of the Internet-based exchanges remains relatively rare, and premiums have been rising.

It is possible that those insurers who chose to list on the Internet might have been those with a premium schedule skewed to attract younger people. Further research is needed to separate this hypothesis from the one we have proposed.

V. Conclusion

The results in this paper generally support the hypothesis that the level and dispersion of premiums paid for individual health insurance both depend on buyer search behavior. High-risk buyers traditionally have searched more, and consequently paid premiums which (compared to low-risk buyers) are more favorably related both to benefits received and to average posted or quoted premiums. Their transactions prices also display less relative variation. These results somewhat paradoxically imply that low-risk buyers have more to gain from using Internet-based exchanges than do high-risk buyers. However, in the current situation, the main advantage of using the Internet may not be so much the ability to secure lower prices as it is a much lower amount of effort required to find whatever price is finally paid.

This analysis is not able to address two additional questions which should be answered in future work. First, we had no data on the search behavior of people who ultimately decide not to buy insurance. It is possible, for example, that website premiums might be much lower than the best potential transaction premium for those who previously stopped searching and did not buy. Second, we know nothing about the underwriting behavior of insurers, either those who list on the website or those who actually sold coverage to those in the CTS sample. We do know what insurers are willing to sell to older people (so they do not “try to avoid high risks” entirely), but we do not know how they deal with potential buyers with less than perfect health.

From a policy viewpoint, these results help to explain the paradox in which older persons charged higher premiums are nevertheless more willing than younger people to obtain insurance: older people search more until they find better deals. They also imply that such ways of

reducing search costs as the Internet may be useful in inducing younger lower risks to pay more attention to the need for insurance.

Figure 1

Health Insurance Search

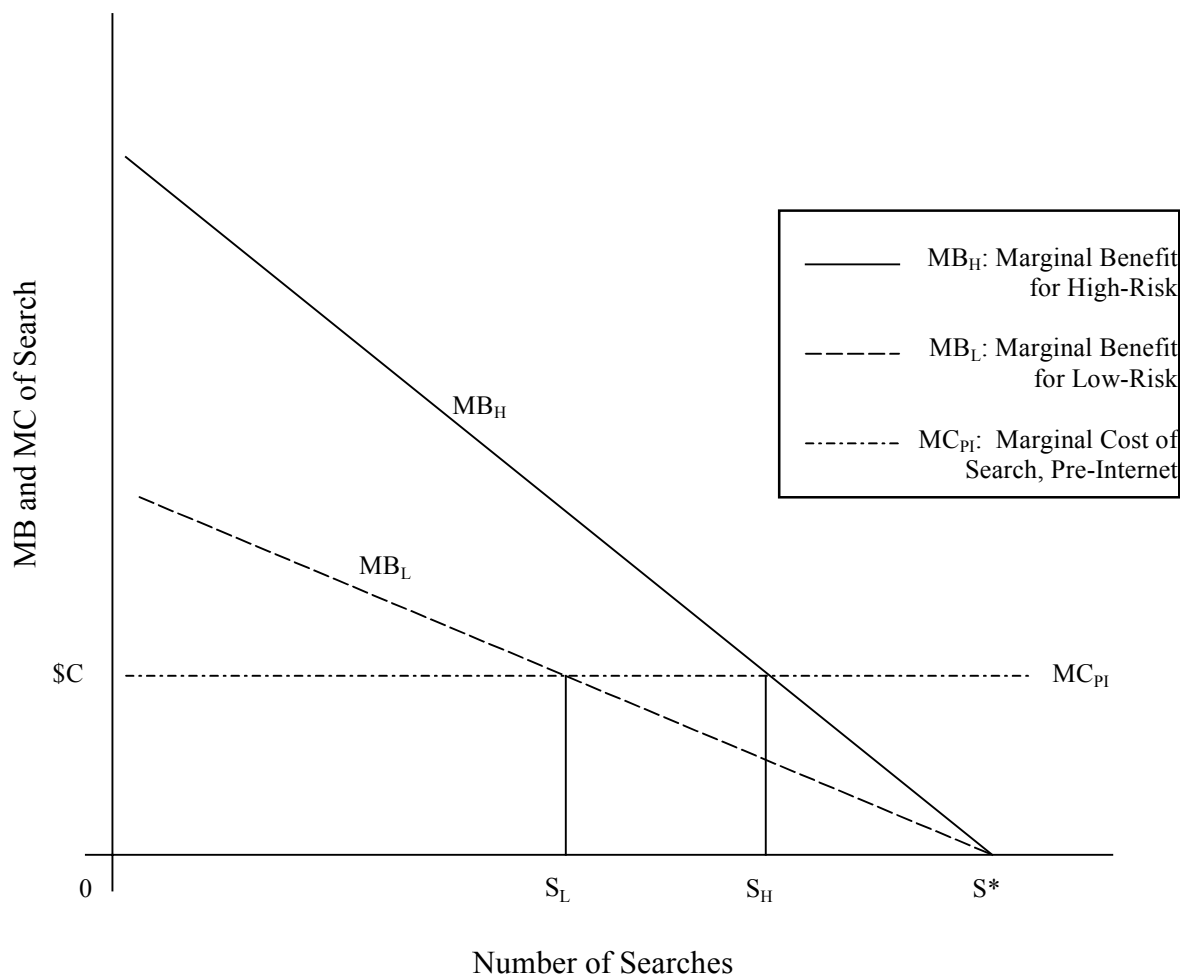


Table 1
Elasticities of Premiums with Respect to Expected Expense
(Standard Errors in Parentheses)

	All	Non-HMO	HMO	Perfect Matches
Premium Quotes* (Website)	0.561	0.565	0.484	0.640
	(.02)	(.03)	(.04)	(.06)
Paid Premiums (CTS)	0.459	0.501	0.376	0.437
	(.05)	(.07)	(.09)	(.07)
N	436	293	143	214

* Predicted premiums matched to the CTS respondent's characteristics (including plan preferences) were estimated from a hedonic regression of Internet quotes on Internet plan characteristics and risk variables.

Table 2

Value of Predicted Website Premium Minus Paid CTS Premium:
Medians by Age-Gender Category

(“Composite Premium” Sample Constructed from the Mean Predicted Website Premiums)

Age	Males			Females		
	Absolute Difference*	Relative Difference**	N	Absolute Difference	Relative Difference	N
18-24	575	0.57	26	445	0.46	15
25-34	501	0.45	59	410	0.37	50
35-44	388	0.35	41	614	0.42	34
45-54	593	0.38	29	1013	0.63	75
55-64	1311	0.68	25	1473	0.65	84

* *Absolute Difference* = *Website Premium* – *CTS Premium*

** *Relative Difference* = (*Website Premium* – *CTS Premium*) / *CTS Premium*

Table 3

Value of Predicted Website Premium Minus Paid CTS Premium:
Medians by Age-Gender Category*

*(“Composite Premium” Sample Constructed from the Lowest Decile of Predicted Website Premiums)***

<i>Age</i>	<i>Males</i>	<i>Females</i>
18-24	-262	-265
25-34	-370	-378
35-44	-406	-171
45-54	-175	83
55-64	253	243

* *Absolute Difference = Website Premium – CTS Premium*

** For this set of premium comparisons, we estimated a hedonic regression of website premiums at the lowest decile, then computed predicted website premiums from this model.

Table 4

Value of Predicted Website Premium Minus Paid CTS Premium:
Medians by Age

(“Perfect Match” Sample, Median Website Premium)

Age	Absolute Difference*	Relative Difference**
Less than 45	153	0.10
45 or older	1249	0.33

* *Absolute Difference = Matched Website Premium – CTS Premium*

** *Relative Difference = (Matched Website Premium – CTS Premium) / CTS Premium*

References

- Brown, Jeffrey R., and Goolsbee, Austan. "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry." *J.P.E.* 110 (June 2002): 481-507.
- Chollet, Deborah J., and Kirk, Adele. *Understanding Individual Health Insurance Markets*. Menlo Park, CA: The Henry J. Kaiser Family Foundation, 1998.
- Clay, Karen; Krishnan, Ramayya; and Wolff, Eric. "Prices and Price Dispersion on the Web: Evidence from the Online Book Industry." Working Paper no. 8271. Cambridge, Mass.: NBER, May 2001.
- Froque, James. "Recent Survey Points to Affordable Individual Health Insurance." Executive Memorandum no. 740. Washington, D.C.: The Heritage Foundation, April 2001.
- Herring, Bradley, and Pauly, Mark V. "Premium Variation in the Individual Insurance Market." *International Journal of Health Care Finance and Economics* 1(March 2001): 43-58.
- McCall, John J. "The Economics of Information and Optimal Stopping Rules." *Journal of Business* 38 (July 1965): 300-317.
- Nichols, Len, and Pauly, Mark V. "The Non-Group Health Insurance Market: Short on Facts, Long on Opinions and Policy Disputes." *Health Affairs Web Exclusive*, forthcoming.
- Pauly, Mark V., and Herring, Bradley. *Pooling Health Insurance Risks*. Washington, D.C.: AEI Press, 1999.
- Pauly, Mark V.; Herring, Bradley; and Song, David. "Tax Credits, the Distribution of Subsidized Health Insurance Premiums, and the Uninsured." *Frontiers in Health Policy Research* 5, Garber, Alan M, ed.: 103-122.
- Phelps, Charles E. *Health Economics*, 2nd Edition. Reading, Mass.: Addison-Wesley, 1997.
- Pollitz, Karen; Sorian, Richard; and Thomas, Kathy. *How Accessible is Insurance for Consumers in Less-Than-Perfect Health?* Menlo Park, CA: The Henry J. Kaiser Family Foundation, 2001.
- Rothschild, Michael. "Models of Market Organization With Imperfect Information: A Survey." *J.P.E.* 81 (November-December 1973): 1283-1308.
- Salop, Steven, and Stiglitz, Joseph. "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion." *Rev. Econ. Studies* 44 (October 1977): 493-510.
- Salop, S., and Stiglitz, J.E. "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents." *A.E.R.* 72 (December 1982): 1121-1130.
- Scott-Morton, Fiona S.; Zettelmeyer, Florian; and Silva-Risso, Jorge. "Internet Car Retailing." Working Paper no. 7961. Cambridge, Mass.: NBER, October 2000.
- Sorensen, Alan T. "Equilibrium Price Dispersion In Retail Markets for Prescription Drugs." *J.P.E.* 108 (August 2000): 833-850.
- Stigler, George J. "The Economics of Information." *J.P.E.* 69 (October 1962): 213-225.