NBER WORKING PAPER SERIES

LEARNING AND THE VALUE OF INFORMATION: EVIDENCE FROM HEALTH PLAN REPORT CARDS

Michael Chernew
Gautam Gowrisankaran
Dennis P. Scanlon

Working Paper 8589
http://www.nber.org/papers/w8589

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2001

This work was supported by a grant from the Agency for Healthcare Research and Quality (AHRQ), grant # 1-R01-HS10050. We are grateful to Tom Cragg and Bruce Bradley for providing the data for this study. We also acknowledge comments received from Dan Ackerberg, Scott Cardell, Tom Holmes, Phillip Leslie, Andrea Moro, Rob Porter, Gary Solon and seminar participants at the Federal Reserve Bank of San Francisco, UCLA, the University of Minnesota and IHEA 2001 in York, UK. Finally, we appreciate the capable programming assistance of Joe Vasey. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research, the Federal Reserve Bank of San Francisco or any other institution with which the authors are affiliated.

Learning and the Value of Information: Evidence From Health Plan Report Cards
Michael Chernew, Gautam Gowrisankaran and Dennis P. Scanlon
NBER Working Paper No. 8589
November 2001, Revised September 2006
JEL No. I11, D83, D12

ABSTRACT

We estimate a Bayesian learning model in order to assess the value of health plan performance information and the extent to which the explicit provision of information about product quality alters consumer behavior. We take advantage of a natural experiment in which health plan performance information for HMOs was released to employees of a Fortune 50 company for the first time. Our empirical work indicates that the release of information affected health plan choices. Consumers were willing to pay an extra $276 per year per below average rating avoided, and the average value of the information per employee was $22 per year. The priors on quality and the quality ratings have a correlation of 0.14 that is statistically significant. The results suggest that despite the existence of a variety of informal mechanisms to convey information, including reputation, consumers may value formally constructed performance measures.

Michael Chernew
Department of Health Management & Policy
Department of Economics
Department of Internal Medicine
The University of Michigan
109 S. Observatory
Ann Arbor, MI 48109-2029
and NBER

Gautam Gowrisankaran
Department of Economics
University of Minnesota,
Federal Reserve Bank of San Francisco,
and NBER

Dennis P. Scanlon
Department of Health Policy & Administration
Center for Health Policy Research
The Pennsylvania State University

Section 1: Introduction

In many markets, products vary substantively in terms of quality. However, quality is often not readily observable. Failure to capture full information can result in a lack of equilibrium or incomplete markets (Akerlof, 1970; Rothschild and Stiglitz, 1976) and may diminish welfare in a variety of ways (Stiglitz, 1989). Certainly markets capture some information through informal mechanisms such as reputation, but it is uncertain how well these mechanisms work. In particular, it is often hard to develop markets for information because information is hard to value before it is known and often has characteristics of a public good.[1] For these reasons, economists have long been interested in understanding the impact of information in markets with products of heterogeneous quality.

This paper estimates the value and impact of report card information in the market for health insurance plans. Our analysis is based on a report card dissemination effort in which the General Motors Corporation (GM) started distributing formal ratings of health maintenance organization (HMO) health plans to its non-union employees for the 1997 open enrollment period. GM has been a leader in creating health plan performance measures and was one of the first companies to provide such measures directly to employees. For each offered HMO, the GM ratings listed the performance in a variety of dimensions as one of four levels: superior, average, and below expected performance, and no data (which indicates that the HMO did not provide the information necessary to assess performance). Our data include employee plan choice before and after the release of the report card (i.e., from 1996 and 1997) and thus can explain the extent to which information affects choice.

[1] For example, Arrow (1963) comments on the "elusive character" of information as a commodity.

We develop a formal Bayesian learning model of health plan quality, estimate the parameters of the model with simulated maximum likelihood, and use the estimated model model's estimates to quantify the value of the report card information. In our model, each employee makes a discrete choice from one of the offered health plans each year in order to maximize her expected utility. Expected utility is a function of plan price, benefits, perceived quality, and idiosyncratic unobserved components. In 1996, employees have priors regarding plan quality; they use these priors and the signal from the ratings to form posterior distributions of quality in 1997.

We specify two different functional forms for the learning process: a specification with continuous quality levels that uses Gaussian priors and signals and another with discrete quality levels that uses Beta priors and Binomial signals. We model prior mean quality levels via fixed effects for each plan in each market, and we examine the impact of ratings using a variety of different specifications and for different subgroups. These methods allow us to evaluate the robustness of our findings to functional form and to obtain results that are consistent with heterogeneous priors and responses. As GM is a national employer, our data contain over 100 HMOs and approximately 70,000 employees observed over two years across many different markets. This provides us with a large amount of variation in ratings, plan attributes, and plan choices that are useful in identifying the values of different types of information.[2]

This paper contributes to two related literatures. First, a recent literature has examined the impact of report cards on managed care health plan enrollment (Beaulieu, 2002; Chernew et al., 2004; Dafny and Dranove, 2006; Jin and Sorenson, forthcoming; Sorensen, forthcoming;

[2] One limitation of the study design is that everyone in our sample received ratings in 1997. Because the U.S. experienced a general trend towards HMOs in this period (see InterStudy, 1996 and 1997), we would not want to attribute any trend towards HMOs at GM solely to the release of ratings. As we detail in Section 3, we use supplementary data to control for this limitation.

Scanlon et al., 2002; Wedig and Tai-Seale, 2002). These papers all find that consumers respond to ratings, in particular to measures such as patient satisfaction. None of these papers estimate a formal learning model, and hence none of them can quantify the value of the report card information or assess the relation between the report card ratings and consumers' prior information. Understanding the role of information in the health insurance market is important since the market is notoriously plagued by a variety of information imperfections (Arrow, 1963). Information about the quality of managed care health plans is particularly relevant since these plans provide a mechanism for individuals to commit to a package of benefits and style of care before an illness. To the extent that increased information about managed care plans will increase enrollment, report cards can lessen the moral hazard problems of traditional health insurance.[3]

Scanlon et al. (2002) deserves particular mention since it is based on an evaluation of the same report card release, though it uses less comprehensive data.[4] They find that consumers respond to ratings primarily by shifting away from plans with below average ratings. They did not seek to understand whether information could affect the set of people choosing HMOs. This study builds on Scanlon et al. (2002) by estimating formal Bayesian learning models that quantify the value of information and also by evaluating the response to ratings in more detail, in particular, by allowing for heterogeneity in responses based on observable and unobservable factors.

Second, another literature has estimated the extent to which consumers learn from information for goods ranging from yogurt to prescription drugs (Ackerberg, 2003; Crawford

[3] In contrast, Dranove et al. (2002) show that incomplete report cards can lower welfare by creating adverse selection incentives.
[4] Scanlon et al. (2002) do not include employees who chose plans other than HMOs in the sample nor does it model the "no data" rating.

and Shum, 2005; Erdem and Keane, 1996; Jin and Leslie, 2003; Milyo and Waldfogel, 1999). The first three of these papers estimate formal Bayesian learning models. We contribute to the literature on Bayesian learning models in two ways. First, we show how to apply a Bayesian learning model to a study design that exploits a policy intervention using detailed panel data and fixed effects. Second, we estimate a specification with Beta priors and Binomial signals which is, to our knowledge, the first estimation of this type of learning process. This type of specification may be useful in other studies where the set of possible quality levels is discrete.

The remainder of this paper proceeds as follows. Section 2 describes the data. Section 3 specifies the model and estimation. Section 4 provides results. Section 5 concludes.

Section 2: Data

Sample

During the late 1990s, GM provided health insurance and benefits for over 1.6 million active employees, retirees, and dependents in the U.S. Our analysis is based on the 1996 and 1997 health plan enrollment decisions for the approximately 70,000 active, non-union U.S. GM employees.[5]

Employees could choose from four different coverage tiers: single, employee and spouse, employee and children, and employee and family. In addition to the coverage tier, employees could choose from a menu of different health plans. In both periods, all employees could choose from fee-for-service basic (FFSB) and fee-for-service enhanced (FFSE) plans, with additional HMO and preferred provider organization (PPO) options depending on employees' zipcodes of

[5] We did not analyze dependents separately because they almost always made the same choice as the employee. We excluded retirees because they are frequently eligible for Medicare, making the nature of plan choice different than for the non-Medicare population. We excluded union employees because we lacked detailed enrollment data for them.

residence. The set of available plans was very similar across the two years. Benefits were standardized within each of the four plan types, although they varied across types. In addition to plan choice and coverage tier, our data include age of employee, tenure at GM, and ages and relations of dependents.

We divided zipcodes into geographic areas, where every zipcode in a geographic area contains the same set of offered HMOs and PPOs. We define a plan to be offered in a zipcode if it was chosen by at least one person in that zipcode in both years. While geographic areas are mutually exclusive, plans may serve multiple geographic areas. To create our final sample, we dropped employee/year observations with missing or obviously incorrect zipcode information, observations where plan and zipcode were never observed for the other year of data,[6] observations with missing price or ratings data, and observations in zipcodes for which no one chose an HMO or PPO.

We define a market as a particular geographic area/coverage category combination. We excluded markets with less than 5 employees in either year. The GM data contain 150,089 employee-year observations, and our final estimation sample contains 133,383 observations (about 89 percent), 437 markets and 1,964 plan-market pairs. Hence, our sample includes the vast majority of the employees. Table 1 details the number of employees by coverage category and plan type kept in our sample for both years. About 37.6% of employees chose HMOs in 1996, a number that rises to 40.7% in 1997. In 1997, HMOs were the most popular type of plan for employees with coverage for children, while FFS plans dominated for employees without coverage for children.

[6] In most cases, this would occur when the plan was not a realistic choice for the employees largely because geographic mobility resulted in a plan choice that was not consistent with the listed zipcode.

Report cards ratings and prices

We now summarize the report card ratings and prices; details are included in Scanlon et al. (2002). The set of health insurance plans from which employees could choose, as well as the prices employees were charged for each plan, were determined by GM. During the open enrollment period for 1997, which occurred in the Fall of 1996, non-union GM employees were given report cards with ratings for each of the HMOs in their choice set. Ratings covered all HMOs but not FFS or PPO plans because the measures used to construct the ratings are only collected for HMOs. GM did not distribute report cards to union employees.

Figure 1 provides a simulated sample report card. HMOs were rated along six domains: operational performance, preventive health care services, medical and surgical care, women's health issues, access to care, and patient satisfaction. In each domain. an HMO could obtain one of four ratings: below expected performance, average performance, superior performance, or no data. Employees were informed that the plans with "no data" ratings did not provide sufficient information and hence we treat a "no data" rating differently from no rating.

The performance ratings were mostly based on data from the Health Plan Employer Data and Information Set (HEDIS), developed by an independent and impartial data source, the National Committee for Quality Assurance (NCQA), and aggregated and compiled by GM. GM picked a subset of the HEDIS measures that were generally accepted to be important, and then aggregated them using standard statistical techniques. Two measures, operational performance and patient satisfaction, were constructed by GM from site visits to HMOs and surveys, respectively. The underlying HEDIS data relate to rates of utilization of selected services, survey responses regarding satisfaction, rates of medically appropriate procedures (e.g. mammographies, cardiac catheterizations and prenatal visits, as appropriate), and measures of access to physicians. The ratings did not include any outcomes data. The report cards also indicated whether the plan was accredited by the NCQA and whether GM designated the plan as a "benchmark" HMO (a positive designation) based on quantitative data and a qualitative assessment. We do not use the benchmark designation in our specifications since Scanlon et al. (2002) found that it had almost no impact on choice and since it only applies to a small number of plans.

The employees paid for health plans using "flex dollars" that could be allocated across several benefit categories (e.g., health insurance, life insurance, disability insurance, and dental insurance) as well as out-of-pocket pre-tax dollars. The price for every health plan was at least as high as the amount of flexible benefit dollars received, which implies that the marginal contribution for health coverage came from out-of-pocket expenses. We define price as the difference between the annual out-of-pocket price and the allotted flex dollars.

Table 2 provides summary statistics on health plan prices by coverage tier and ratings during our two year period. Although the mean out-of-pocket prices for plans stayed relatively constant from 1996 to 1997, there is substantial variation in the change in price between 1996 and 1997; for instance, for Tier 4 (family) coverage, the standard deviation of the price difference is $432 relative to a mean price of $1,312. According to GM benefit managers, changes in prices between 1996 and 1997 were chosen largely to be correlated with observed quality measures, in order to steer employees to high quality plans.

Section 3: Model and Estimation

Model

We consider a Bayesian learning model where individual i resides in market m at time period t, and must choose among a set of plans j.[7] Individuals care about the perceived quality of care that the plan will provide them and other plan attributes. We assume that individuals are well informed about the price that they pay for the plan, as well as general plan coverage characteristics, such as copays and deductibles, but that they may lack information about the quality of care that they would receive from the plan, which we denote q_{ijm}. For instance, individuals may not know how easy it is to find a specialist that will accept new patients; they may not know whether the health plan and its physicians are good at recommending medically appropriate treatments ranging from diagnostic procedures such as mammographies to invasive surgeries; they may not know the extent to which a serious illness would be accompanied by long waits to see physicians; and they may not know the quality of surgical care.

We specify the expected utility function for the individual as:

$$(1) \qquad u_{ijmt} = E_t\left[q_{ijm}\right] - \alpha_i P_{jmt} + \delta_{ijmt} + \varepsilon_{ijmt},$$

where E_t is a conditional expectation at time t, P_{jmt} is price,[8] α_i are parameters,[9] δ_{ijmt} are other plan attributes, and ε_{ijmt} is a component of utility that is not systematically related to plan quality and is unobservable to the econometrician.[10]

[7] Our plan choice model builds on a number of recent papers that have estimated the impact of price (though not quality) on the choice of health plans (Buchmueller and Feldstein, 1997; Cutler and Reber, 1998; Royalty and Solomon, 1999).

[8] Since (1) includes price, it is an indirect utility function. The underlying direct utility function that generate this would specify overall utility to be the sum of the utility from the health plan and from some numeraire good, which costs $1 per unit and gives a constant utility α_i per unit.

[9] We index all parameters by "i" because some specifications allow for heterogeneous responses to information across different consumers, a topic we return to below.

[10] Although consumers may learn about plan quality from experiences while enrolled in the plan, we assume there is sufficient noise in the learning process that consumers do not consider the value of learning when choosing a health plan. With this assumption, consumers will choose the health plans that maximize their current expected utilities (1). We believe that "sampling" plans is very uncommon, and therefore that this assumption is reasonable.

Following Cardell (1997) and Berry (1994), we assume a nested logit error structure for ε_{ijmt} which allows for correlated unobservables within a plan type. Specifically, we let

(2) $$\varepsilon_{ijmt} = \varepsilon'_{ig(j)mt} + \lambda_i \varepsilon''_{ijmt},$$

where ε' and ε'' are independent, λ_i are parameters, $g(j)$ indexes the type of health plan j (i.e., HMO, PPO, or FFS), ε'' is distributed extreme value, and $\varepsilon' \sim C(\lambda)$, defined as the unique h_{2i} distribution that makes ε extreme value given λ and the distribution of ε''. If $\lambda_i = 1$, then the model will be identical to the logit model and the unobservables will be i.i.d., while if $\lambda_i = 0$, the unobservables will be perfectly correlated within a group. We estimate a nested logit because this specification provides a natural way to estimate the extent to which consumers are willing to switch between types of plans.

We consider individuals at two time periods, 0 and 1 (i.e., 1996 and 1997, respectively). Signals, in the form of health plan report cards for HMOs, are given to individuals immediately before they make their choice of health plan at time 1. The conditional distribution of quality at time 0 (i.e., the prior) is a function of reputation and experience, while the conditional distribution of quality at time 1 (i.e., the posterior) is a function of both the prior and the signal. We estimate two specifications for the learning model, one with continuous quality levels and the other with discrete quality levels. These specifications will approximate the true, unknown, densities in different ways, and thus add to the robustness of our findings. We now discuss both of these specifications in turn.

Continuous quality levels

This specification assumes that the support of q_{ijm} is continuous with Gaussian priors and signals, specifically that the prior is distributed $N\left(\bar{q}_{ijm}, h_{1i}^{-1}\right)$ and the report card signal, s_{ijm}, is distributed $N\left(q_{ijm}, h_{2i}^{-1}\right)$, where $\bar{q}_{ijm}$ are parameters, and h_{1i} and h_{2i} are precisions of the priors and signal respectively. We assume that the priors and signals are uncorrelated across plans in a market. We let s_{ijm} be related to the published ratings r_j as

(3) $$s_{ijm} = \tilde{\beta}_i r_j + \tilde{\sigma}_i \nu_{ijm},$$

where $\tilde{\beta}_i$ and $\tilde{\sigma}_i$ are parameters and $\nu_{ijm} \sim N(0,1)$ captures other sources of health plan information obtained during period 0, e.g., media coverage. We include this term to make the signal more continuous, in keeping with the assumption that its distribution is Gaussian.

In this specification, the prior mean quality is $E_0\left[q_{ijm}\right] = \bar{q}_{ijm}$. Using (3) in conjunction with standard Bayesian updating formulas, the posterior mean quality is

(4) $$E_1\left[q_{ijm}\right] = \frac{h_{1i}\bar{q}_{ijm} + h_{2i}\left(\tilde{\beta}_i r_j + \tilde{\sigma}_i \nu_{ijm}\right)}{h_{1i} + h_{2i}}$$

for plans which receive ratings.

We require certain normalizations in order to identify our model. In particular, since utility is not observable, we normalize the fee–for–service basic (FFSB) plan to have expected prior quality 0 for every market. We normalize FFSB because it does not have published ratings, is homogeneous and is offered in every market. We also cannot jointly identify the precisions, h_{1i} and h_{2i}, since they are collinear, as can be seen from (4). We estimate instead $h_i \equiv h_{1i} / \left(h_{1i} + h_{2i}\right)$. Defining $\sigma_i = \tilde{\sigma}_i h_{2i} / \left(h_{1i} + h_{2i}\right)$ and $\beta_i = \tilde{\beta}_i h_{2i} / \left(h_{1i} + h_{2i}\right)$, expected utility for a rated plan at time 1 can then be expressed as

(5) $$u_{ijm1} = h_i\bar{q}_{ijm} + \beta_i r_j + \sigma_i \nu_{ijm} - \alpha_i P_{jm1} + \varepsilon_{ijm1}.$$

Discrete quality levels

This specification assumes that the support of q_{ijm} is discrete with mass on two points, v_{il} (low quality) and v_{ih} (high quality).[11] We assume that the prior density of the probability that q_{ijm} is v_{ih} is distributed $\text{Beta}\left(a_{ijm}, b_{ijm}\right)$.[12] Thus, the expected prior probability that q_{ijm} is v_{ih} is $a_{ijm}/\left(a_{ijm}+b_{ijm}\right)$. The interpretation of the Beta distribution is that a_{ijm} is the number of high quality draws and b_{ijm} is the number of low quality draws. Expected prior quality then becomes:

(6) $$E_0\left[q_{ijm}\right] = v_{il}\frac{b_{ijm}}{a_{ijm}+b_{ijm}} + v_{ih}\frac{a_{ijm}}{a_{ijm}+b_{ijm}}.$$

We assume each report card rating is a Binomial signal of either v_{il} or v_{ih}. Let R_{jl} and R_{jh} denote the number of low and high quality ratings for plan j, respectively. Using standard Bayesian updating formulas, the posterior density of the probability that q_{ijm} is v_{ih} is distributed $\text{Beta}\left(a_{ijm}+R_{jh}, b_{ijm}+R_{jl}\right)$ and hence the expected posterior probability that q_{ijm} is v_{ih} is $\left(a_{ijm}+R_{jh}\right)/\left(a_{ijm}+b_{ijm}+R_{jh}+R_{jl}\right)$.

As with the continuous case, we require normalizations to identify the parameters. We cannot identify both a_{ijm} and b_{ijm} for each plan in each market, because the two parameters

[11] Note that we could specify a Dirichlet prior and a multinomial signal and expand our specification to allow for four values for quality (instead of two) to fully exploit the fact that there are four ratings. While it is straightforward to evaluate the posterior for this model, we still cannot identify more than one coefficient implying the need for more normalizations, many of which might be unintuitive.

would be predicting market share in a collinear manner. Accordingly, we estimate the a_{ijm} parameters and one parameter $info_i \equiv a_{ijm} + b_{ijm}$ in place of all the b_{ijm} parameters. This normalization fixes the prior *total* number of draws that people receive from each plan, while letting the number of positive draws vary across plans and markets. Similar to the continuous model, we normalize the FFSB plan to have prior $a_{i,FFSB,m} = v_{il} \times info_i / (v_{il} - v_{ih})$, which implies (from (6)) that the expected prior quality for this plan is 0. Analogous to (5), expected utility for a rated plan at time 1 can be expressed as

$$(7) \qquad u_{ijm1} = v_{ih} \frac{a_{ijm} + R_{jh}}{info_i + R_{jh} + R_{jl}} + v_{il} \frac{info_i - a_{ijm} + R_{jl}}{info_i + R_{jh} + R_{jl}} + \delta_{ijmt} - \alpha_i P_{jm1} + \varepsilon_{ijm1}.$$

Parameterization

We allow prior mean quality to differ across markets and plans because of the local nature of information. Thus we estimate $\bar{q}_{ijm}$ or a_{ijm} (for the continuous and discrete specifications respectively) as a separate parameter for each plan j and market m for a given set of consumers i. Note that this assumption is similar to allowing plan-market fixed effects in a linear specification.

We specify several different functional forms for ratings. Our base specification for the continuous model assumes that the response to each of the six performance domains is the same and allows for four ratings (superior, average and no data, with below average excluded) and a dummy for whether or not the plan was accredited by the NCQA. We use this specification since consumers often use decision rules such as selecting plans with the most superior ratings or

[12] It is standard to define a Binomial on the set {0,1} and a Beta over the interval [0,1]. We renormalize to v_l and

fewest below average ratings (Hibbard *et al.*, 1997) and evidence from laboratory settings is consistent with such decision rules (Hibbard *et al.*, 2000). Other specifications for the continuous model allow for variation in the ratings coefficients across performance domains. Our discrete model is limited to two signals. Based on evidence from the continuous model below, we group superior with average and no data with below average.

We also cannot identify non time-varying components of δ_{ijt} from choice data (since we estimate plan-market fixed effects) and so we only consider time-varying components. We include three plan-type interactions for time 1, $\delta_{i,FFSE,1}$, $\delta_{i,PPO,1}$, and $\delta_{i,HMO,1}$, designed to capture shifts in acceptance for different plan types over time; all are relative to the FFSB time trend.

These variables, particularly $\delta_{i,HMO,1}$, are very relevant since U.S. HMO enrollment increased substantially between 1996 and 1997,[13] likely because of a relative increase in the value of HMO services,[14] and we would not want to attribute an increase in GM HMO enrollment solely to ratings. Unfortunately, since every employee received ratings in 1997 for every HMO, $\delta_{i,HMO,1}$ is collinear with ratings, and hence we cannot estimate it. However, we obtained aggregate data from a similar Midwest-based Fortune 50 manufacturing company that did not distribute ratings. That firm experienced an increase in HMO enrollment of 1.99 percentage points (from 40.78% to 42.77%) among its non-union employees between 1996 and 1997. Thus, we choose $\delta_{i,HMO,1}$ to be the value that would have caused a 1.99 percentage point increase in GM HMO enrollment between 1996 and 1997 at the estimated parameters in the

---

v_h respectively, because this fits better with our utility framework.

[13] InterStudy (1996, 1997) reports that the number of "pure HMO" enrollees in the U.S. increased from 52.5 million to 58.8 million people during 1996.

[14] For instance, drug treatments over this era were becoming increasingly effective and expensive (Lichtenberg, 2001) and HMOs generally provide lower copays for drug and other treatments than other plans.

absence of ratings or any price or sample change. We also experimented with other values of $\delta_{i,HMO,1}$ and found similar results for nearby values.

Thus far we have indexed all parameters with an "i" to indicate potential variation across consumers. Our base model assumes that the parameters are the same across individuals; in the interest of clarity we suppress the "i" subscript when discussing these specifications. However, we also examine several alternate specifications which generalize this assumption. In particular, in some specifications we define subgroups based on observable characteristics (e.g., gender, presence of young children) and allow all the parameters to vary across subgroups. In addition, for some specifications of the continuous model, we allow for random coefficients for the ratings. For these specifications, we let the coefficients on the ratings be distributed around some mean $\bar{\beta}$, i.e., $\beta_i = \bar{\beta} + \bar{\sigma}_i \bar{v}_{ijm}$ with $\bar{\sigma}_i$ being a parameter and $\bar{v}_{ijm} \sim N(0,1)$.

Identification

We first consider the identification of the coefficients on ratings (β_i for the continuous specification and v_{il} and v_{ih} for the discrete specification) and price (α_i). We treat both these variables as exogenous, and now explain why. Since we include a fixed effect for the prior quality of each plan in each market, endogeneity would occur only if particular ratings or changes in prices are correlated with changes in unobservable plan characteristics that might change market shares even in the absence of the changes in price or ratings.

We believe that endogeneity is unlikely for ratings because it is unlikely that particular ratings would change unobserved plan characteristics or vice versa. Specifically, ratings were provided only to non-union GM employees who formed a small subset of the enrollment base for any given health plan, suggesting that it is unlikely that plans would react to ratings by changing their unobserved characteristics. Moreover, the ratings, which were released in 1997, were based on 1995 plan performance, when most plans would not have anticipated the construction and release of the report card, suggesting that plans could not have endogenously influenced the ratings based on changes in their unobserved characteristics between 1996 and 1997. In addition, there is more direct evidence against endogeneity (or omitted variable bias) from Scanlon et al. (2002). This study included share among GM unionized employees as a control group, albeit at a more aggregate level,[15] and found virtually identical results. Since the union employees did not receive the report card information, this further suggests that any changes in enrollment among non-union workers that correlates with ratings is caused by the ratings.

We treat price as exogenous for similar reasons. Our prices are based on out-of-pocket costs charged to employees. We do not observe premiums charged to GM, which might be endogenous in a market setting, varying positively with quality. In contrast, out-of-pocket prices were set by GM and, as noted in Section 2, managers report that changes in prices were chosen largely to be correlated with observed (but not unobserved) performance measures. Moreover, as with ratings, Scanlon et al. (2002) find that the coefficient on price remains very similar when using union employees, who did not experience price changes, as a control group. Last, unlike ratings, several studies have measured the effect of price on health insurance plan market shares, and, as we show in Section 4 below, our figures are similar to those in the literature.

Other parameters, including the parameters that are specific to the two learning models, are similarly identified from intuitive variation in the data. One parameter of note is the nested logit correlation parameter, λ. In the context of a fixed effects model, this parameter will be

[15] We could not use union employees directly as a control group since the only available data is aggregate plan market shares by state.

identified from changes in the attributes of the choices over time within a market. Since our data contain many such changes, they are useful in identifying this parameter.

Estimation and simulation

We estimate the parameters of the models using a maximum likelihood. Each enrollee at each time period constitutes one observation. The likelihood for the observation is the probability that the chosen plan was selected, given the parameter vector. For the continuous specification, we simulate unobservables ν_{ijm} and $\bar{\nu}_{ijm}$ for the random effects specifications, and hence use simulated maximum likelihood.

To define the likelihood for the specifications where parameters do not vary by person "i", let y_{imt} denote the chosen plan for individual i in market m at time t, and let x_{mt} denote the exogenous variables in market m at time t, which include ratings, prices, and plan identities. Let θ denote the parameters: $\theta = \left(\bar{q}_{j,m}\ \forall j \text{ and } m, h, \beta, \sigma_v, \alpha, \lambda, \delta_{PPO,1}, \delta_{FFSE,1}\right)$ for the continuous specification and $\theta = \left(a_{j,m}\ \forall j \text{ and } m, v_l, v_h, \text{info}, \alpha, \lambda, \delta_{PPO,1}, \delta_{FFSE,1}\right)$ for the discrete specification. Then, the log likelihood for an individual i for the continuous specification satisfies

(8) $$\ln L\left(\theta \middle| y, x\right) = \sum_{i,m,t} \ln\left(\tfrac{1}{NS}\sum_{s} \Pr\left(\text{Choice for enrollee } i,m,t \text{ is } y_{imt} \middle| \theta, x_{mt}, \nu_{ijms}\right)\right),$$

where NS is the number of simulation draws per individual, ν_{ijms} is one simulation draw, and the probabilities of the observed choices are calculated using the nested logit model applied to the utility function specified by (5) and a simpler utility function without ratings.

For the discrete specification, the log likelihood is analogous to (8) but uses (7) in place of (5), and does not include simulations over ν_{ijms}. The likelihood for specifications with different subgroups based on observed consumer types is also similar but includes separate parameters by consumer type (we generally estimate one subgroup at a time). The likelihood for the random coefficients models is similar, but includes the parameter $\bar{\sigma}$ and involves simulation over $\bar{v}_{ijms}$.

We mention a couple of details about the estimation process. We estimate the model using a Newton-Raphson search. This derivative based search converges reasonably quickly, which is necessary given that each estimation includes over 1,500 parameters. We set NS to 20, and our conclusions are insensitive to estimates computed with 40 draws. As is generally done for simulated likelihood estimators, we use the same draws across parameter values.

Using our estimated parameters, one of our main goals is to measure the value of information. This is different than measuring the value of other product attributes (e.g., gas mileage for automobiles) since good and bad information are both valuable to the extent that they cause consumers to alter their behavior.[16] The textbook measure of the value of information when faced with subsequent decisions is given by DeGroot (1970, p. 197). To use this measure in our context, let $\Im_t$ denote the information set at time t, $Y_{im1}(\Im_t)$ denote the optimal choice for person i in market m given plan attributes at time 1 and an information set $\Im_t$, $U_{im1}(\Im_t, Y)$ denote expected utility given plan attributes at time 1, information set $\Im_t$ and choice Y, and $f_t(\Im_1)$ be the density over information sets $(\Im_1)$ at time t. Then, the aggregate value of the information, expressed in utility units, is

[16] Information may affect the behavior of health care providers or employers, which we do not account for. In addition, information may affect utility even if it does not alter behavior because it can reassure, or worry, consumers independent of any effects on plan choice. We follow the statistical literature and focus only on the portion of value generated as a result of behavior changes.

(9) $$V = \sum_m \sum_i \int \left[U_{im1}\left(\Im_1, Y_{im1}\left(\Im_1\right)\right) - U_{im1}\left(\Im_1, Y_{im1}\left(\Im_0\right)\right) \right] f_0\left(\Im_1\right) d\Im_1 ,$$

or in words, the probability with which the choice with information ($Y_{im1}\left(\Im_1\right)$) is different than the choice without information ($Y_{im1}\left(\Im_0\right)$), times the difference in expected utilities for the informed conditional on being in this set.

The value of information described in (9) is based on the expected distribution of information, $f_0\left(\Im_1\right)$, which we do not observe and hence cannot directly compute. Thus, we make one further assumption, that the ex-post distribution of signals was equal to the ex-ante distribution. While this is a strong assumption that effectively results in an ex-post valuation measure, our large sample of plans spread across many markets renders this assumption less problematic relative to papers that are based on fewer plans and fewer markets. With this assumption, we obtain

(10) $$V = \sum_m \sum_i \int \left[U_{im1}\left(\Im_1, Y_{im1}\left(\Im_1\right)\right) - U_{im1}\left(\Im_1, Y_{im1}\left(\Im_0\right)\right) \right] f_1\left(\Im_1\right) d\Im_1 ,$$

which can be computed by using the actual distribution of signals. We are interested in finding the per-capita value of information in dollar terms, which we obtain by dividing the value in utility units by the marginal utility of money, α_i, and the number of people.

Section 4: Results

Results from base continuous specification: Specification 1

This section details the estimates and implications of the model developed in Section 3. As discussed in Section 3, our base specification, Specification 1 in Table 3, groups ratings across performance domains. This specification reveals a coefficient on price that is negative and statistically significant. We cannot evaluate the economic magnitude of this coefficient using a price elasticity, since a large and unknown portion of the price is paid by the employer. We instead evaluate the semi-elasticity of price, defined to be the average percent change in the probability of choosing a plan given a \$100 increase in the annual price. We find that the \$100 increase in price would result in a reduction in plan share of 2.7% on average across plans. The literature on health plan choice finds values ranging from 2.5 percent to 4 percent, which is consistent with our value.[17]

We find that superior and average ratings are both significantly positive and similar in magnitude. A "no data" rating is significantly worse than below average, though smaller in magnitude than the other two ratings. The implication is that consumers react to ratings primarily by staying away from plans with below average scores or no data. The table, which reports magnitudes of the coefficients in dollar units by dividing the ratings coefficient by the coefficient on price, shows that one extra average rating in place of a below average rating would increase the willingness to pay for one year of plan coverage for a given plan by \$332.

We estimate the nested logit parameter, λ, to be .330 with a small standard error of .030. The standard error allows us to easily reject the logit model, which imposes $\lambda = 1$, and thus, we do not present results from the logit model. Nonetheless, we estimated the logit model and obtained similar results to our base specification. The estimated value suggests that there are

[17] Cutler and Reber (1998) find an elasticity of –2 for Harvard employees, which is equivalent to a semi-elasticity of 4% per \$100 increase given that the average gross premium is roughly \$5000 in their study. Royalty and Solomon (1999) report price elasticities of –1 to –1.8 for Stanford employees. Using the midpoint of –1.4 and noting that their average gross premium is roughly \$4,000, this implies a semi-elasticity of 3.5% per \$100 price change. Buchmueller and Feldstein (1997) report that an increase in net price from \$120 to \$240 reduced plan share by 4% for University of California employees, and that a further \$120 increase reduced the plan share by 3%. Scaling these down to \$100 increments yields semi-elasticities of between 2.5% and 3.3% per \$100 price change. Because they allow a discrete jump in response associated with any positive change in price, Buchmueller and Feldstein (1997) find much larger price elasticities, which we do not replicate, when the price changes from \$0 to \$120.

substantial correlations in preferences, in the sense that people with a high unobserved affinity for a PPO (for example) are likely to have a high unobserved affinity for another PPO.

This specification includes 1,527 plan-market prior dummies, as do all specifications that use the full data set. In the interest of brevity, we do not list these coefficients. However, their magnitudes are much larger than the magnitudes of the ratings coefficients: the absolute value of these variables has a mean of .774 and a standard deviation of .568.

We estimate the prior weight coefficient, h, to be .929 and significantly different from both 0 and 1. This implies that the posterior precision of plan quality is only about 8% higher than the prior precision. The estimated values of h and the plan-market prior dummies together imply that prior information is much more important than the signal in determining the posterior.

We estimate a value of the standard deviation for the unobserved shock in the signal, σ, that is small (e.g., less than half the magnitude of any ratings coefficient) and statistically insignificant. Recall that σ indicates the magnitude of the information that consumers obtain during the first period from sources other than the report card. Thus, this suggests that most of the *learning* about plan quality during 1996 came from the report card.

Our model includes three plan type-year interaction variables for 1997, all relative to FFSB. The estimated $\delta_{FFSE,1}$ and $\delta_{PPO,1}$ coefficients are both positive and significant. FFSE differed from FFSB only in that it had lower copays and deductibles, and thus the positive sign on $\delta_{FFSE,1}$ must be due to an increase in value from these features. We believe that the reasons that $\delta_{PPO,1}$ is positive are similar to the reasons why HMO market share was increasing over time nationally, noted in Section 3.

As discussed in Section 3, the HMO-time interaction term, $\delta_{HMO,1}$, cannot be estimated since ratings are distributed to all employees for all HMOs in 1997, but rather is chosen to generate an increase in HMO market share of 1.99 percentage points between 1996 and 1997 to match an aggregate control group at the estimated parameters. In keeping with the increase in market share, we find a positive value of $\delta_{HMO,1}$ that is larger than either the PPO or FFSE interactions. We cannot obtain a standard error for the parameter. Note that $\delta_{HMO,1}$ is perfectly collinear with the "rated" parameter and hence its value will not affect any of the other parameter estimates. However, a higher value of $\delta_{HMO,1}$ will result in a lower value of "rated" which will then attribute more of the 1997 increase in market share for HMOs to ratings and less to plan acceptance. This will in turn affect the value of information. The sign of this latter effect is not clear, since both good and bad information is useful. In practice, we found that reasonable values of $\delta_{HMO,1}$ gave very similar numbers for the value of information.

Using our estimated parameters and equation (10), we compute the value of the information contained in the report card. We find a reasonably modest value of information, an average of $19 per consumer.[18] We believe that the evidence that the impact is modest is well-substantiated in the data: the report cards did not get too many people to switch plans. In particular, only 12.4% of employees in our sample in both years switched health plans between 1996 and 1997. Some of that is due to ratings and some to other factors, such as price changes, changes in geographic location, and changes in unobserved components. Our base specification finds that ratings caused only 3.89% of employees to switch plans.[19] Moreover, the HMO market

[18] Note that one could bootstrap from the variance/covariance matrix of the parameter estimates in order to obtain a confidence interval for this figure.
[19] We compute this figure by simulation using 1997 plan attributes.

share increased by a net of only 3.1 percentage points between 1996 and 1997. Our model attributes that 1.0 percentage points of that to ratings, and the rest to greater HMO acceptance and changes in pricing and other plan attributes.

Our modest value of information occurs in spite of the reasonably large willingness to pay to avoid below average or no data ratings. The substantiation in the data for this dichotomy is that people did not often switch plans because of either price changes or ratings, and the willingness to pay figures are essentially a ratio of how willing people are to switch plans for better ratings to how willing they are to switch plans because of a lower price. This is also consistent with our finding that plan priors are more important than either ratings or prices. Note that among the 3.89% of employees who switched plans as a result of ratings, ratings were worth an average of $488 ex-post.

Our evidence that ratings have an impact on choice is consistent with survey data that suggest that measures such as these are salient for potential health plan enrollees (see Hibbard and Jewett, 1996 and Tumlinson et al., 1997). Our willingness-to-pay figures are also consistent with Scanlon et al. (2002) who find comparable numbers using similar data but a different model. Our results on employee switching and the value of information are also broadly consistent with other studies (see Beaulieu, 2002, for Harvard University employees, Jin and Sorensen, forthcoming, for federal employees, and Dafny and Dranove, 2006, for Medicare beneficiaries) who all find a small, but significant, amount of consumer switching resulting from report cards.[20]

[20] Jin and Sorensen (forthcoming) and Dafny and Dranove (2006) report smaller effects of switching than we do. However, there is no reason to expect the magnitudes to be the same since the value of information and extent of switching behavior is dependent on the type of ratings information, prior knowledge, and choice sets, all of which vary between our study and these studies.

Impact of discrete learning process: Specifications 2 and 3

We next examine the discrete learning specification, Specification 3, also in Table 3. Recall that we assume a two-point support for the distribution of quality and group together superior and average ratings and no data and below average ratings, because of the similarity of these coefficients in Specification 1. We use the six performance domains as the sources of information for this specification, and do not include accreditation. For comparison purposes, Specification 2 (also in Table 3) provides estimates for the continuous model with the ratings aggregated into two groups as in the discrete specification.

We find that the discrete learning specification provides very similar results to the continuous specification to the extent that they are comparable. In particular, the value of information, willingness to pay to avoid low ratings, the price coefficient, nested logit correlation and time interactions are almost identical across the two specifications. These results should add evidence that the results from the continuous model are not largely driven by functional form.

The discrete model also shows that prior information is very important relative to the signal from the report card ratings. In particular, we estimate the parameter "info" to be 86.0. This suggests that prior information about plan quality was equivalent to 86 ratings measures, some good and some bad. In contrast, the report card information contained only 6 measures, and hence contributed much less to the posterior.

Effect of specific performance domains: Specifications 4 and 5

In order to understand further which performance domains contribute value, Table 4 presents specifications where the signal from the report card is allowed to vary across domains. We use only continuous specifications here since our discrete model restricts the ratings to take one of two values. We estimate a specification (Specification 4) where we allow each of the 19 individual ratings to have a separate coefficient, and one where we allow for variation in the coefficients across performance domains but group together superior and average ratings and no data and below average ratings, as in Specification 2.

Specification 4 generally results in ratings coefficients that are not very precisely estimated and do not have a consistent pattern. We believe that the reason for this is that we are trying to estimate 19 ratings coefficients from data on only 105 plans, and hence there is not enough variation in the ratings to identify these coefficients. Indeed, one of the domains, operational performance, has no plans with a "no data" rating, and hence this parameter is excluded.

In contrast, Specification 5 shows a pattern that is more internally consistent and also consistent with Table 3. In particular, consumers value average or above average ratings for 5 of the 6 domains positively, and in 4 of these 5 cases, the coefficients are statistically significant. Moreover, a likelihood ratio test would allow us to reject the hypothesis that individuals respond equally to all ratings. It is useful to analyze responses to specific performance domains. However, we do this with the caveat that the probability that every conclusion below is accurate is less than the probability of any one of them being accurate.

We find that people value patient satisfaction and access to care measures, which is consistent with evidence from Chernew et al. (2004) and Dafny and Dranove (2006) for employers and Medicare beneficiaries respectively. However, the strongest response is to the medical and surgical care rating. This is intriguing because these measures are so imprecisely measured to not even include outcomes, except for one readmission rate. The fact that employees respond to even imprecise information along this dimension suggests to us that there is much uncertainty about the quality of medical and surgical care and employees may trust these measures more than informed observers might. Nevertheless, the result suggests that there may be considerable value in creating better measures.[21] In contrast, the coefficients on preventive care and women's health measures were smaller (also consistent with the two studies above), perhaps because there are less information problems for these domains. We are unsure what to make of the negative response to better operational performance. Perhaps employees view plans as achieving operational performance at the expense of quality care (e.g., employees do not have to wait to see a doctor, but the doctor spends only five minutes with each of them). Or perhaps, they were simply unsure about the meaning of this measure.

Note that the estimated values of information for these specifications are somewhat higher than in Specification 1, which occurs because the point estimates for certain individual ratings are larger in magnitude than the base point estimates, suggesting more value from switching plans in response to ratings. Indeed, we find that 4.03% of employees switch plans as a result of ratings in Specification 4, as compared to the 3.89% figure from Specification 1.

Heterogeneity in responses across employees: Specifications 6-11

Specifications 6-9 in Table 5 examine the extent to which there is a heterogeneous impact of ratings on different subgroups. Specification 6 presents results from the sample of employees with covered women (i.e., employees who were female or who had a covered female spouse). We allowed for the full 19 ratings as in Specification 4, but we report only the coefficients for the women's health performance domain. We find no evidence that women value this domain. Indeed, the point estimates for superior and average ratings for this domain are negative here as

[21] See, for instance, Geweke, Gowrisankaran and Town (2003) for an example of a study that attempts to create better measures of hospital quality.

in Specification 4, though somewhat less so. Thus, there is no evidence of heterogeneity along this domain.

Specification 7 reports the same model as in Specification 1, but for the sample of patients over 50. Older people have higher mortality and morbidity rates, have lower managed care enrollment rates than younger people, and may have other reasons to value ratings more. We find that the ratings coefficients for this group are somewhat larger than in Specification 1 but that the price coefficient is also somewhat larger. Overall, this yields a slightly larger willingness-to-pay to avoid below average ratings ($384 vs. $332) and a slightly smaller value of information. The coefficients here are much less precisely estimated than in the base sample.

Specifications 8 and 9 consider the same model as in Specification 1 but for employees with a covered child 12 years or younger and ones whose tenure at GM is less than 5 years, respectively. The coefficient estimates are generally similar to Specification 1, though with much less precision. The price coefficient for people with children is smaller in magnitude than in the base specification and not significant. People with children may have a lower income per person, suggesting more elastic demand. However, they may also be more likely to use healthcare, suggesting less elastic demand and hence a coefficient that is smaller in magnitude. The value of information for this group is higher than for the base specification, but these differences are not significant, since the price elasticities are not significantly different from 0.

Table 6 examines the extent to which there is a heterogeneous impact of ratings based on unobservable factors, by estimating a random coefficients specification. Specifications 10 and 11 duplicate Specifications 1 and 2 with the addition of random coefficients for all the ratings, respectively. Our findings reveal generally small point estimates on the standard deviations of the ratings coefficients. Indeed, of the 7 standard deviation parameters across the two specifications, only 1 is statistically significant. All the other parameter estimates are similar to the base specifications, although we estimate a somewhat higher value of information with this specification. Thus, we find no compelling evidence of heterogeneity based on unobservables, and it appears that whatever heterogeneity exists does not affect our conclusions very much.

Section 5: Conclusions

This paper assesses the value and impact of information on health insurance plans by applying a Bayesian learning model to a study design that includes panel data and fixed effects and that exploits a policy intervention (i.e., GM non-union employees were given health plan report cards). We find that information affects health plan choice in that consumers have a moderately large willingness to pay to avoid plans with bad ratings. Only about 3% of people switch plans as a result of the ratings, implying a moderate per capita value of the report card at about $20. The results are robust across discrete and continuous specifications for the learning process. We find evidence of heterogeneity in responses across performance measures, with people valuing medical and surgical care quality, and satisfaction and access measures, the most. In contrast, we find no significant evidence of heterogeneity in responses across different employee groups.

While our model cannot provide definitive answers as to why the impact of the ratings was modest, it does allow us to draw some inferences. One possible explanation is that people already are fully informed about health plan quality. However, this is contradicted by the fact that individuals report that they would like to see ratings information (see Hibbard and Jewett, 1996). Another explanation is self-selection, i.e., that people are already fairly satisfied with their plans. The fact that people do not often switch health plans for either price or ratings reasons suggests some validity of this second explanation.

A final explanation is that the GM ratings are not fully informative, as suggested by our finding that the signals are imprecise relative to prior information. For instance, there are few indicators in the ratings about the quality of the covered physicians and hospitals. In contrast, the ratings include measures such as the utilization rates for recommended age or gender specific preventive care or cancer screenings, but it is not clear that these ratings should influence one's choice of health plan, since the guidelines for this type of care is fairly straightforward (e.g., women over age 40, etc.). This is supported by the findings that people react to performance domains such as patient satisfaction which would not suffer from the imprecisions noted above. It is also supported by studies that find that consumers do not feel fully informed as a result of ratings.[22]

Our results also suggests that consumers might value other, more directly pertinent, ratings information much more strongly. To provide a more definitive answer as to the types of report card information that would add value, it ultimately might be necessary to understand which information impacts medical costs and medical utilization rates and through that employees' health. While we lack this type of data in this study, we feel that this is an important topic for future research.

[22] See Hibbard and Jewett, 1996; Hibbard et al., 2000; Robinson and Brodie, 1997; and Tumlinson et al., 1997.

References

Ackerberg, Daniel A., 2003. "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination," International Economic Review 44, 1007-1040

Advertising, Learning and Consumer Choice in Experience Good Markets: An Empirical Examination." Mimeo, UCLA.

Akerlof, George A., 1970. The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. Quarterly Journal of Economics 84: 488-500

Arrow, K., 1963. Uncertainty and the Welfare Economics of Medical Care. American Economic Review. 53(5). 941-973.

Beaulieu, Nancy (2002). Quality information and consumer health plan choices. Journal of Health Economics, 21(1), 43-63.

Berry, S.T., 1994. Estimating discrete-choice models of product differentiation. RAND Journal of Economics, 25, 242-262.

Buchmueller, T.C., and P. Feldstein, 1997. The effect of price on switching among health plans. Journal of Health Economics, 16, 231-247.

Cardell, N.S., 1997. "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," Econometric Theory, 13, 185-213.

Crawford, Gregory S. and Matthew Shum, 2005. "Uncertainty and Learning in Pharmaceutical Demand." Econometrica, 73: 1137-1174.

Chernew, Michael, Gautam Gowrisankaran, Catherine McLaughlin and Teresa Gibson, 2004. "Quality and Employers' Choice of Health Plan," Journal of Health Economics 23: 471–92.

Chernew, M.E., Scanlon, D., and R. Hayward, 1998. "Insurance Type and Choice of Hospital for Coronary Bypass Graft Surgery," Health Services Research, 33(3): 447-466.

Cutler, David M. and Sarah J. Reber. (1998) "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection," Quarterly Journal of Economics 113(2): 438-66.

Dafny, Leemore and David Dranove (2006). "Do Report Cards Tell Consumers Anything They Don't Already Know? The Case of Medicare HMOs." Mimeo, Northwestern University.

DeGroot, Morris H., 1970. Optimal Statistical Decisions, New York: McGraw–Hill.

Dranove, D., D. Kessler, M. McClellan, M. Satterthwaite, 2002. Is More Information Better? The Effects of 'Report Cards' on Health Care Providers. NBER Working Paper 8697.

Erdem, T. and M. Keane, 1996. "Decision Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." Marketing Science 15: 1-20.

Geweke, John, Gautam Gowrisankaran and Robert J. Town (2003). "Bayesian Inference For Hospital Quality in a Selection Model." *Econometrica* 71: 1215 – 1238.

Hibbard, J.H., L. Harris-Kojetin, P. Mullin, J. Lubalin, and S. Garfinkel, 2000. Increasing the impact of health plan report cards by addressing consumers' concerns. Health Affairs, 19, 138-143.

Hibbard, J.H., and J.J. Jewett, 1996. What type of quality information do consumers want in a health care report card? Medical Care Research and Review, 53, 28-47.

Hibbard, J.H., P. Slovic, and J.J. Jewett, 1997. Informing consumer decisions in health care: implications from decision-making research. Milbank Quarterly, 75, 395-414.

InterStudy (1996, 1997). The Competitive Edge. St. Paul, MN: InterStudy Publications.

Irwin, D. and P. Klenow, 1994. Learning-by-Doing Spillovers in the Semiconductor Industry. Journal of Political Economy, 102, 1200-1227.

Jin, Ginger Z. and Philip Leslie, 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," Quarterly Journal of Economics 118: 409-51.

Jin, Ginger Z. and Alan Sorensen, forthcoming. "Information and Consumer Choice: The Value of Publicized Health Plan Ratings," Journal of Health Economics, forthcoming.

Lichtenberg, F.R. (2001). "Are the benefits of newer drugs worth their cost? Evidence from the 1996 MEPS." *Health Affairs*, 20(5): 241-251.

Luft, H. S., D. H. Garnick, C. S., Mark, D. H., Peltzman, D.J., Phibbs, C.S., Lichtenberg, E., McPhee, S.J. June 6, 1990. "Does Quality Influence Choice of Hospital?" Journal of The American Medical Association 263(21): 2899-2906.

Mennemeyer ST. Morrisey MA. Howard LZ, 1997. Death and reputation: how consumers acted upon HCFA mortality information. *Inquiry*. 34:117-28.

Milyo, Jeffrey and Joel Waldfogel, 1999. The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart. American Economic Review 89: 1081-96.

Robinson, S., and M. Brodie, 1997. Understanding the quality challenge for health consumers: the Kaiser/AHCPR survey. Journal on Quality Improvement, 23, 239-244.

Rothschild, Michael and Joseph E. Stiglitz, 1976. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. Quarterly Journal of Economics 90: 630-49.

Royalty, A.B., and N. Solomon, 1999. Health plan choice: price elasticities in a managed competition setting. Journal of Human Resources, 34, 1-41.

Scanlon D.P., Chernew, M.E., McLaughlin, C.G., Solon, G., 2002. "The Impact of Health Plan Report Cards on Managed Care Enrollment," Journal of Health Economics, 21, 19-41.

Sorensen, A., forthcoming. Social Learning in the Demand for Employer-Sponsored Health Insurance. RAND Journal of Economics.

Stiglitz, Joseph E., 1989. "Imperfect Information in the Product Market.". In Richard Schmalensee and Robert D. Willig (ed.), Handbook of Industrial Organization: Volume 1, Amsterdam: North-Holland.

Tumlinson, A., H. Bottigheimer, P. Mahoney, E.M. Stone, and A. Hendricks, 1997. Choosing a health plan: what information will consumers use? Health Affairs, 16, 229-238.

Wedig, Gerard J. and Ming Tai-Seale, 2002. "The Effect of Report Cards on Consumer Choice in the Health Insurance Market." Journal of Health Economics 21, 1031-48.

Figure 1

Example information sheet

COMPARING YOUR 1997 GM MEDICAL OPTIONS

The following table shows the rating of the HMO option(s) available in eight selected quality measures. The ratings are based on historical data and therefore may not necessarily represent the quality of care you will receive in the future. GM does not endorse or recommend any particular medical plan option. The medical plan you elect is your personal decision.

For a more complete description of the eight selected quality measures, see the *GM Medical Plan Guide.*

	NCQA Accredited?	Benchmark HMO?	Opera-tional Performance	Preventive Care	Medical/ Surgical Care	Women's Health	Access to Care	Patient Satisfaction
0001 Basic Medical Plan	Information Currently Not Available							
0002 Enhanced Medical Plan	Information Currently Not Available							
PPO 2190 Blue Preferred Plus	Information Currently Not Available							
HMO 2101 Care Choices HMO	Yes	No	▲▲	▲▲▲	▲▲	▲▲▲	▲▲	▲▲▲
HMO 2108 BCN G Lakes SW MI	Yes	No	▲▲	▲	▲	▲	▲	▲▲
HMO 2113 BCN Health Central	Yes	No	▲	▲▲▲	▲▲▲	▲▲▲	▲	▲▲
HMO 2116 Priority Health	Yes	No	▲▲	▲	▲	▲▲	▲▲▲	▲▲▲
HMO 2117 Care Choices HMO W MI	Yes	No	▲▲	▲▲▲	▲▲	▲▲▲	▲▲	▲▲▲
HMO 2118 BCN West Michigan	Yes	No	▲▲	▲	▲	▲	▲	▲▲

Key: ▲ =below expected performance ▲▲ =average performance ▲▲▲=superior performance
ND=no data was available from this plan

Michigan - Lansing and West

Table 1
Number and percent of employees by coverage category and plan type

	HMO	PPO	FFS	Total
1996	25,275 (37.6)	10,768 (16.0)	31,204 (46.4)	67,247 (100)
1997	26,903 (40.7)	10,110 (15.3)	29,123 (44.0)	66,136 (100)
Tier 1 (Employee)	11,295 (33.8)	5,100 (15.3)	17,002 (50.9)	33,397 (100)
Tier 2 (Emp./Spouse)	11,213 (34.5)	5,876 (18.1)	15,448 (47.5)	32,537 (100)
Tier 3 (Emp./Child)	3,780 (44.1)	1,685 (19.7)	3,103 (36.2)	8,568 (100)
Tier 4 (Family)	25,890 (44.0)	8,217 (14.0)	24,774 (42.1)	58,881 (100)

Note: The universe is all active non-union employees kept in sample. Percentage of row in cells are in parentheses below the numbers.

Table 2
Summary of price and ratings characteristics

	All Plans: (HMO/PPO/FFS)				
	N	Mean	Std. Dev.	Min	Max
1996 annual Tier 1 (employee) price	133	$481	$179	$0	$708
1997 annual Tier 1 price	133	$476	$193	$0	$732
1996 annual Tier 4 (family) price	133	$1,325	$494	$0	$1,956
1997 annual Tier 4 price	133	$1,312	$528	$0	$2,004
Difference between Tier 1 prices, 1997-1996	133	-$4	$137	-$468	$252
Difference between Tier 4 prices, 1997-1996	133	-$13	$432	-$1,608	$960
	HMO Plans				
	N	Mean	Std. Dev	Min	Max
Number of superior ratings	105	2.18	1.79	0	6
Number of average ratings	105	1.91	1.27	0	5
Number of below average ratings	105	1.41	1.31	0	5
Number of no data ratings	105	0.50	1.09	0	5
	N	Yes		No	
Accreditation	105	74 (70%)		31 (30%)	
Benchmark plan	105	15 (14%)		90 (86%)	

Note: annual prices reflect the difference between the GM price-tag and the allotted flex dollars.

Table 3
Base coefficient estimates and estimated value of information

	Continuous quality, four ratings (1)	Continuous quality, two ratings (2)	Discrete quality, two ratings (3)
Rated (base: below avg.; (2) below avg. and no data)	–.091** (.023)	–.140** (.034)	
# superior ratings	.040** (.005)		
# average ratings	.047** (.006)		
# no data ratings	–.034** (.006)		
# average / superior		.053** (.011)	
Not accredited	.041* (.017)		
Prior weight (h)	.929** (.012)	.935** (.012)	
Std. dev. param. (σ)	.015 (.013)	.016 (.014)	
Utility from avg./sup. (v_h)			2.75** (.708)
Util. below avg./no data (v_l)			–2.15** (.671)
Prior draws (info)			86.0** (5.18)
Price (thousands per year)	–$.141** (.024)	–$.124** (.031)	–$.125** (.031)
Nested logit param. (λ)	.330** (.030)	.348** (.070)	.349** (.070)
PPO–year 1 dummy ($\delta_{PPO,1}$)	.036* (.018)	.037* (.019)	.037* (.019)
FFSE–year 1 dummy ($\delta_{FFSE,1}$)	.027** (.008)	.028** (.010)	.028** (.010)
HMO–year 1 dummy ($\delta_{HMO,1}$)	.127	.128	.128
Log likelihood	–183,641	–183,667	–183,665
Willingness to pay per below avg. rating changed to average	$332	$428	$458
Estimated average value of information per employee	$19	$22	$21

Note: Standard errors in parentheses. All specifications include 1,527 plan-market prior dummies. The symbols "*" and "**" indicate significance at the 5% and 1% levels respectively.

Table 4
Estimates with heterogeneity across performance domains

	Four ratings (4)	Two ratings (5)
Rated	–.025 (.028)	–.008** (.021)
Operational performance superior	–.027 (.021)	
Op. perf. avg.; (5) avg. and superior	–.048** (.017)	–.031** (.012)
Operational performance no data		
Preventive care superior	.076* (.035)	
Prev. care avg.; (5) avg. and superior	.027 (.023)	.032** (.012)
Preventive care no data	–.007 (.026)	
Medical/surgical care superior	.077** (.024)	
Med./surg. avg.; (5) avg. and superior	.119** (.029)	.112** (.021)
Medical/surgical care no data	–.083* (.034)	
Women's health superior	–.035 (.036)	
Women's avg.; (5) avg. and superior	–.020 (.016)	.011 (.010)
Women's health no data	.131* (.063)	
Access to care superior	.028 (.017)	
Access avg.; (5) avg. and superior	.034 (.018)	.046** (.013)
Access to care no data	.014 (.029)	
Patient satisfaction superior	.028 (.017)	
Pat. sat. avg.; (5) avg. and superior	.032 (.020)	.052** (.012)
Patient satisfaction no data	–.007 (.026)	
Not accredited	–.010 (.019)	
Prior weight (h)	.933** (.013)	.940** (.013)
Std. dev. param. (σ)	.011 (.010)	.011 (.010)
Price (thousands per year)	–$.098** (.029)	–$.096** (.023)
Nested logit param. (λ)	.247** (.052)	.235** (.044)
PPO–year 1 dummy ($\delta_{PPO,1}$)	.029 (.018)	.029 (.018)
FFSE–year 1 dummy ($\delta_{FFSE,1}$)	.020** (.007)	.019** (.006)
HMO–year 1 dummy ($\delta_{HMO,1}$)	.121	.120
Log likelihood	–183,567	–183,604
Estimated average value of information per employee	$29	$26

Note: Standard errors in parentheses. All specifications include 1,527 plan-market prior dummies. The symbols "*" and "**" indicate significance at the 5% and 1% levels respectively.

Table 5
Estimates with heterogeneous responses across groups

	Employees with covered women (6)	Employees over age 50 (7)	Employees with child age 12 or under (8)	Employees at GM less than 5 years (9)
Rated (base: below avg.)	.017 (.035)	–.212* (.086)	.018 (.030)	–.098 (.083)
# superior ratings ((6): women's health)	–.002 (.042)	.065** (.024)	.014 (.011)	.045 (.026)
# average ratings ((6): women's health)	–.005 (.019)	.079** (.026)	.020 (.015)	.043 (.026)
# no data ratings ((6): women's health)	.146 (.078)	–.055* (.028)	–.025 (.019)	–.025 (.017)
Not accredited	–.007 (.023)	.130* (.066)	.027 (.024)	–.011 (.042)
Prior weight (h_i)	.930** (.015)	.939** (.022)	.878** (.021)	.905 (.034)
Std. dev. param. (σ_i)	.005 (.012)	.082 (.059)	.020 (.017)	.020 (.027)
Price (thousands per year)	–$.132 (.039)	–$.206 (.110)	–$.060 (.046)	–$.101 (.086)
Nested logit param. (λ_i)	.264** (.062)	.721** (.209)	.137 (.100)	.239 (.134)
PPO–year 1 dummy	.042* (.021)	.056 (.048)	.035 (.035)	.014 (.026)
FFSE–year 1 dummy	.029** (.011)	.067 (.042)	.009 (.010)	.021 (.017)
HMO–year 1 dummy	.133	.182	.109	.109
Number of employee/year observations	103,989	38,804	39,184	15,395
Log likelihood	–139,539	-48,479	–54,906	–22,888
Willing. to pay per below avg. rating to average	n/a	$384	$331	$428
Estimated average value of information per employee	$18	$16	$41	$54

Note: Standard errors in parentheses. All specifications include plan-market prior dummies. Specification (6) includes dummies for all other ratings as in specification (3). The symbols "*" and "**" indicate significance at the 5% and 1% levels respectively.

Table 6
Estimates with unobserved heterogeneity in responses to ratings

	Random effects with four ratings (10)	Random effects with two ratings (11)
Mean: rated	–.084** (.027)	–.140** (.034)
Standard deviation: rated	.004 (.013)	.006 (.014)
Mean: # superior ratings	.037** (.008)	
Std. dev.: # superior ratings	.002 (.004)	
Mean: # average ratings ((11): avg. and superior)	.044** (.010)	.053** (.011)
Std. dev.: # average ratings ((11): avg. and superior)	.005 (.007)	.006 (.014)
Mean: # no data ratings	–.032** (.008)	
Std. dev.: # no data ratings	.009 (.015)	
Mean: not accredited	.030 (.017)	
Std. dev.: not accredited	.125** (.040)	
Prior weight (h)	.931** (.012)	.935** (.012)
Std. dev. param. (σ)	.013 (.013)	.016 (.014)
Price (thousands per year)	–$.131** (.034)	–$.124** (.031)
Nested logit param. (λ)	.306** (.064)	.347** (.070)
PPO–year 1 dummy ($\delta_{PPO,1}$)	.034 (.019)	.037* (.019)
FFSE–year 1 dummy ($\delta_{FFSE,1}$)	.025** (.009)	.028** (.010)
HMO–year 1 dummy ($\delta_{HMO,1}$)	.125	.128
Log likelihood	183,637	183,665
Estimated average value of information per employee	$25	$25

Note: Standard errors in parentheses. All specifications include 1,527 plan-market prior dummies. The symbols "*" and "**" indicate significance at the 5% and 1% levels respectively.