

NBER WORKING PAPER SERIES

INFORMATION AND CONSUMER CHOICE:
THE VALUE OF PUBLICIZED HEALTH PLAN RATINGS

Ginger Zhe Jin
Alan T. Sorensen

Working Paper 11514
<http://www.nber.org/papers/w11514>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2005

Special thanks to Sarah Shih, Sarah Scholle, Greg Pawlson and James Bost at NCQA for providing assistance with the quality information reported to NCQA; to Bill Anderson, James Quayle, Agnes Kalland, Ralph Pierce, Ron Gresch, Leroy McKnight, and Ebbie Block in OPM for providing us individual records of FEHBP beneficiaries; and to the University of Maryland for financing the data cost. We are grateful to the editor and two anonymous referees for several helpful suggestions. Participants at the 2004 FTC Conference on Health Care also provided useful comments. Any errors are ours. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2005 by Ginger Zhe Jin and Alan T. Sorensen. 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.

Information and Consumer Choice: The Value of Publicized Health Plan Ratings
Ginger Zhe Jin and Alan T. Sorensen
NBER Working Paper No. 11514
July 2005
JEL No. I11, L15

ABSTRACT

We use data on the enrollment decisions of federal annuitants to estimate the influence of publicized ratings on health plan choice. We focus on the impact of ratings disseminated by the National Committee for Quality Assurance (NCQA), and use our estimates to calculate the value of the information. Our approach exploits a novel feature of the data—the availability of nonpublic plan ratings—to correct for a source of bias that is inherent in studies of consumer responsiveness to information on product quality: since publicized ratings are correlated with other quality signals known to consumers (but unobserved by researchers), the estimated influence of ratings is likely to be overstated. We control for this bias by comparing the estimated impact of publicized ratings to the estimated impact of ratings that were never disclosed. The results indicate that NCQA’s plan ratings had a meaningful influence on individuals’ choices, particularly for individuals choosing a plan for the first time. Although we estimate that a very small fraction of individual decisions were materially affected by the information, for those that were affected the implied utility gains are substantial.

Ginger Zhe Jin
University of Maryland
Department of Economics
3105 Tydings Hall
College Park, MD 20742-7211
and NBER
jin@econ.umd.edu

Alan Sorensen
Graduate School of Business
Stanford University
518 Memorial Way
Stanford, CA 94305-5015
and NBER
asorensen@stanford.edu

Information and Consumer Choice: The Value of Publicized Health Plan Ratings*

Ginger Zhe Jin
University of Maryland and NBER
jin@econ.umd.edu

Alan T. Sorensen
Stanford University and NBER
asorensen@stanford.edu

June 2005

Abstract

We use data on the enrollment decisions of federal annuitants to estimate the influence of publicized ratings on health plan choice. We focus on the impact of ratings disseminated by the National Committee for Quality Assurance (NCQA), and use our estimates to calculate the value of the information. Our approach exploits a novel feature of the data—the availability of nonpublic plan ratings—to correct for a source of bias that is inherent in studies of consumer responsiveness to information on product quality: since publicized ratings are correlated with other quality signals known to consumers (but unobserved by researchers), the estimated influence of ratings is likely to be overstated. We control for this bias by comparing the estimated impact of publicized ratings to the estimated impact of ratings that were never disclosed. The results indicate that NCQA’s plan ratings had a meaningful influence on individuals’ choices, particularly for individuals choosing a plan for the first time. Although we estimate that a very small fraction of individual decisions were materially affected by the information, for those that were affected the implied utility gains are substantial.

1 Introduction

Consumers often choose between products whose qualities or characteristics are *ex ante* uncertain. The reliability of a car, for example, cannot be fully known at the time of purchase, and the

*Special thanks to Sarah Shih, Sarah Scholle, Greg Pawlson and James Bost at NCQA for providing assistance with the quality information reported to NCQA; to Bill Anderson, James Quayle, Agnes Kalland, Ralph Pierce, Ron Gresch, Leroy McKnight, and Ebbie Block in OPM for providing us individual records of FEHBP beneficiaries; and to the University of Maryland for financing the data cost. We are grateful to the editor and two anonymous referees for several helpful suggestions. Participants at the 2004 FTC Conference on Health Care also provided useful comments. Any errors are ours.

quality of a restaurant isn't known until after one has dined there. For some products, the presence of such uncertainty has given rise to markets for information: for example, companies like Edmunds and Consumer Reports publish automobile ratings, and publications like Zagat provide restaurant reviews. The profitability of such information enterprises indicates that the information is valuable—but exactly how valuable is it? In some cases, government or nonprofit consumer-interest agencies have assumed the role of information provider. Does the information benefit consumers enough to warrant the costs of its acquisition and dissemination?

In this paper we attempt to answer these questions in the context of employee health plan decisions. For employees choosing from a menu of health plan options, how valuable are publicized plan ratings in making the decision? The National Committee for Quality Assurance (NCQA), an independent nonprofit agency that serves as a “watchdog for the managed care industry,” collects and actively disseminates information about health insurance plans. The ratings and statistics are used by large employers to inform their employees about plans, and for a period of time in the late 1990's summaries of the ratings were published in a major national news magazine, *U.S. News & World Report*. Using enrollment data for a large sample of U.S. government annuitants, we measure the impact of the health plan ratings on individuals' choices, and use our estimates to calculate the value (to consumers) of the information.

This study is not the first to analyze consumer responses to health plan quality information. Wedig and Tai-Seale (2002) examine data on federal employees in 1995 and 1996, when health plan report cards were first provided, and report evidence that plan ratings significantly influenced individuals' decisions. Beaulieu (2002) examined the health plan choices of Harvard University employees and found that the provision of quality information in 1997 had a small but significant effect. Scanlon et al (2002) use employee data from General Motors to study health plan decisions; they find that employees respond to overall quality indices but not to specific quality measures. Using the same dataset, Chernew et al (2004) estimate the parameters of a Bayesian learning model, in which reported plan ratings cause individuals to update their prior beliefs about plan quality. Their findings confirm the positive value of information, but also suggest that much of the information in reported ratings was already known.

Relative to the previous literature, the key contribution of this paper lies in its ability to correct for an important source of bias in estimating the quantitative impact of publicized ratings. A fundamental challenge for this study and all its predecessors is that published information may be highly correlated with quality signals consumers can obtain through informal channels. In general, this correlation between observed product ratings and unobserved product quality will cause econometric estimates of the ratings' impact to be biased. In the present context, if health plan ratings simply mirror information about plan quality that was already known to consumers (but not the econometrician), then high market shares for plans with high ratings could be incorrectly interpreted as evidence that the ratings were influential. This is obviously a critical concern if one's objective is to determine the value of information: if the health plan ratings merely reflect information that consumers already knew, then the ratings have no informational value at all.

Our ability to address this problem comes from a novel characteristic of the data. While most major health plans reported information to NCQA during our sample period, not all plans elected to allow public dissemination of their information. Essentially, plans reporting to NCQA had the option of being "public" or "nonpublic."¹ For the years 1998-1999, we have NCQA rating data for *both* types of plans. The reason this is valuable is that ratings for nonpublicly reporting plans by definition have no causal impact on choices: individuals cannot be influenced by information they do not have. Therefore, any measured effect of nonpublic ratings on plan choice must purely reflect the correlation of those ratings with unobserved plan quality. If the correlation between plan ratings and unobserved plan quality is the same whether a plan is public or nonpublic, the difference between the estimated impacts of public vs. nonpublic ratings provides a consistent measure of the public information's actual influence.

This approach is similar in spirit to that of Wedig and Tai-Seale (2002), who study consecutive years in which FEHBP provided health plan report cards to employees. In 1995 these report cards were distributed on a limited basis, whereas in 1996 the distribution was more widespread, so the authors compare the measured effects of the report cards in the consecutive years to test the hypothesis that report cards provided relevant information to consumers. Hence the underlying

¹See Jin (2004) for an empirical analysis that partially accounts for this disclosure decision. That paper pools "not reporting to NCQA" and "reporting to NCQA nonpublicly" as one group of non-disclosure.

idea—to control for the bias stemming from correlations with unmeasured quality by comparing the effects of information that is vs. is not publicized—is the same as the one motivating our study, with the primary difference being that we are able to make this comparison contemporaneously by comparing the impact of public vs. nonpublic information in the same year.

Our empirical strategy not only delivers a clean estimate of the informational value of publicized plan ratings, but also allows us to assess the overall importance of the correlation with unobservables problem. Previous authors have employed various approaches to circumvent or resolve the problem; having replicated those approaches, we can observe whether nonpublic plan ratings still appear to “influence” individuals’ choices. For example, Beaulieu (2002) emphasizes estimates of the impact of plan ratings on switching behavior rather than cross-sectional results relating plans’ market shares to their ratings, and Scanlon *et al* (2002) use a panel dataset to estimate a choice model with plan fixed effects, effectively eliminating any bias that would result if plans’ ratings are correlated with time-invariant unobservables. Our data enable us to employ similar controls, and then additionally compare the estimated effects of public vs. nonpublic ratings. If nonpublic ratings appear to be influential after controlling for time-invariant unobserved plan characteristics or individual enrollment history, then we could infer that such controls are insufficient—i.e., that plan ratings are likely correlated with unobservable, *time-varying* determinants of plan choice—and result in overstated estimates of the impact of publicized plan ratings.

In the following section we explain how health plan quality information is collected and disseminated to consumers, and then describe the features of our data. Section 3 presents empirical estimates of the influence of plan ratings. The results generally confirm that correlation of these ratings with unobservable plan characteristics is an important problem; however, we find that controlling for switching costs *and* time-invariant unobservables (e.g., by incorporating plan fixed effects or lagged plan market shares) mostly corrects the bias. In section 3.4 we use the empirical estimates to calculate the information value of publicized plan ratings. Although we estimate that a very small fraction of individual decisions were materially affected by the information, for those that were affected the implied utility gains are substantial. The overall implications of our findings are discussed in the concluding section.

2 Background & Data

2.1 How health plan ratings are constructed and publicized

NCQA programs

NCQA is the earliest and remains the most important verification agency for HMOs. Founded in 1979 by two major managed care trade associations, NCQA claimed independence in 1990 under the support of the Robert Wood Johnson Foundation along with matching funds from HMOs and purchaser contributions. NCQA offers a variety of health care quality evaluation services, including accreditation² and the collection, updating, and maintenance of the Health Plan Employer Data and Information Set (HEDIS[®]). HEDIS has been offered nationwide annually since August 1996. It permits HMOs to disclose their summary statistics on physician inputs and patient utilization of certain procedures. (Typical examples are breast cancer screening, diabetic eye exams, child immunization, and physician turnover rates.) These data are self-collected by HMOs and standardized subject to NCQA specifications. Every year since 1996, NCQA has also conducted independent consumer surveys for each participating HMO and reported satisfaction indices if it is permitted to do so. The survey was named the NCQA Member Satisfaction Survey (MSS) before 1998 and renamed the Consumer Assessment of Health Plan Satisfaction (CAHPS[®]) after 1998. Because HEDIS measures and consumer satisfaction data appear together in media publications, we group them as HEDIS/CAHPS.

Participation in both HEDIS and CAHPS is voluntary. Moreover, HEDIS/CAHPS participants may allow NCQA to publicize the collected data, or instruct NCQA to keep them nonpublic. For “public” plans, NCQA may include the plan-specific data in one comparison table and sell it to employers (typically at a cost of \$2,500-7,500 per copy), disclose the data to the media, make the data accessible through its own website, or provide the data to consumer groups and government agencies. For example, each year from 1996 to 1998, *U.S. News & World Report* used the HEDIS/MSS data to construct a comparison chart of HMOs, and assigned an overall rating from one to four stars. *Newsweek* also made a similar table in 1999. In principle, the participation and publication decisions can be made separately for HEDIS and CAHPS. In practice,

²For a detailed description of NCQA accreditation, see Jin (2004).

plans are more likely to report in HEDIS than in CAHPS. Participation in CAHPS is somewhat more costly because it is survey-based, and because NCQA requires the use of a licensed survey vendor following a strict protocol.

We define a plan to be “public” if it has an overall score in *U.S. News & World Report* or allows NCQA to publicize its data in both HEDIS and CAHPS. By this definition, throughout the whole HMO industry (with or without FEHBP contracts), 26.7% of operating HMOs participated in the public release of HEDIS/CAHPS in 1998, accounting for 49% of HMO patients. Another 18.1% reported nonpublic data, accounting for 29% of patients. The remaining 55.2% of HMOs (covering 22% of HMO patients) did not report to NCQA.

Because participation in the NCQA rating programs is voluntary, the decision to report publicly is itself potentially informative to consumers. In section 3.2 we discuss the endogeneity of disclosure status and how it affects our empirical analysis. Intuitively, one might expect that plans with the highest ratings are the ones who elect to report publicly; however, we discuss below that there are various idiosyncratic reasons why some plans choose not to report publicly, and we show that there is considerable overlap in the distribution of plan ratings for public vs. nonpublic plans.

FEHBP Guide

Operated by the Office of Personnel Management (OPM), FEHBP is the largest employer-sponsored health benefit program in the world, offering health insurance to roughly 1.7 million federal employees and 2.2 million annuitants (which include annuitants and employee survivors). Every year, it renegotiates contracts with 200-400 HMOs and 13-21 non-HMOs. In selecting plans to contract with, OPM says its most important concern is a plan’s financial stability, not its quality ratings. The rationale for including so many plans is to allow consumers to benefit from competition among health plans.

Strictly speaking, an “HMO” can be an HMO or a POS but cannot be both.³ For simplification, we group all HMO/POS plans as HMOs. All non-HMOs offer both PPO and non-PPO coverages, which are subject to different copay and coinsurance structures but in the data share the same plan ID, premium and quality ratings. If an individual chooses a non-HMO, she does not need to choose between PPO and non-PPO until the delivery of health services. For this reason, we refer to all

³FEHBP considers it unfair for a health plan to compete with itself.

non-HMOs as fee-for-service (FFS) and do not differentiate between PPO and non-PPO coverages.

Every year FEHBP publishes a guidebook right before the open enrollment season, November 1 to November 30. Plan choices made in the open enrollment season are valid for the duration of the next calendar year unless the beneficiary moves or exits the FEHBP system. The guidebook, disseminated in both hard copy and electronically on the internet, lists premium, main benefits (such as copays and prescription drug coverage), and some quality information for every plan. Premium is further broken down by self and family coverages, but the other variables do not vary by self/family. Detailed benefits are described in plan-specific brochures, available at local human resource offices and on the FEHBP website. Because there are few FFS plans and each FFS plan has two sets of copay and coinsurance structures for PPO and non-PPO, in the empirical analysis we define benefit variables for HMOs only.

FEHBP has long been praised as a model program for its provision of quality information. Its own consumer satisfaction survey was started in 1994 and the first set of survey results were disseminated to a limited subsample of beneficiaries in the 1995 open enrollment season. From 1996 on, the survey has been repeated every year and survey results have been published in the FEHBP guide for all beneficiaries. Unlike NCQA programs, FEHBP *requires* every plan with more than 300 enrollees in the previous year, including HMOs and non-HMOs, to participate in the FEHBP consumer survey. We focus on two measures of satisfaction: the percentage of enrollees who report being “very satisfied” or “extremely satisfied” overall, and the percentage of enrollees who report being satisfied with a plan’s coverage of services. These satisfaction rates, as reported in the FEHBP guide, are plan-specific and do not vary by self/family coverage, nor by respondent characteristics. The overall satisfaction rate is a time-consistent measure, since the survey question to which it corresponds did not change from year to year. The reporting format for plan coverage satisfaction changed in 1999, so in our empirical models we estimate coefficients for the coverage variable separately by year. FEHBP does not report HEDIS/CAHPS data, but they are available from other media reports as mentioned above.⁴

⁴FEHBP also began publishing NCQA accreditation status in 1996. We control for this information in the empirical analysis.

2.2 Data

The principal data sources for our analysis are the health plan enrollment records of all federal government annuitants. The complete data set, which was provided by the Office of Personnel Management (OPM), covers roughly 1.7 million employees and 2 million annuitants across the entire U.S. for the years 1995-2000. Current employees may be more sensitive to health plan information than annuitants, but OPM keeps employee records by working place instead of living place. Since an individual cannot enroll in an HMO unless she resides in one of the plan's serving counties (at the time of enrollment), the lack of residence information makes it difficult to define the choice set of health plans for each employee. We therefore focus the empirical analysis on annuitants, for whom this data problem does not exist because their residence information is recorded in the data. We use the term "annuitants" to include both retirees and surviving family members of deceased federal employees.⁵

Most of the annuitants in the data are old enough to qualify for Medicare. However, most still enroll in an FEHBP plan for supplementary coverage. For these individuals, Medicare serves as the primary insurer, and FEHBP plans adjust any payments so that they supplement rather than duplicate any Medicare benefits. Although the FEHBP plans are not Medigap plans (in the sense that they are not one of the ten standardized Medicare supplemental policies), annuitants have little reason to purchase Medigap policies when they can obtain similar supplemental coverage through FEHBP at heavily subsidized rates. FEHBP brochures explicitly discourage Medicare-eligible annuitants from purchasing both an FEHBP plan and a Medigap policy, since the combination of Medicare and FEHBP provides essentially comprehensive coverage. Because the majority of annuitants enroll in an FEHBP plan either for primary or supplementary coverage, publicized health plan information is directly relevant regardless of Medicare eligibility. However, it is possible that Medicare-eligible individuals respond to the information differently; we address this issue directly in the analysis of section 3.3.

For each annuitant in each year, we observe the health plan chosen and type of coverage (i.e.,

⁵Retirees entitled to annuitized retirement benefits are also eligible for FEHBP health insurance coverage if they were continuously enrolled in any FEHBP plan for the 5 years of service immediately prior to the start of their retirement annuity. Survivors are eligible if the employee was enrolled in family coverage at the time of death.

single or family),⁶ as well as individual characteristics such as age, income, sex, and county of residence. Table 1 summarizes the health plans offered in the FEHBP system. Because the FEHBP covers annuitants nationwide, there are hundreds of plans in the system—even though a given annuitant typically selects from a menu of roughly 5-10 HMO plans (in addition to all available FFS plans). As the table indicates, the total number of plans decreased dramatically during the sample period, and enrollment in HMO plans increased slightly.

In a typical year, about 26.5% of annuitants do not choose any FEHBP plan. This happens for two reasons: the annuitant is covered somewhere else outside the FEHBP system, or covered through his/her spouse who is eligible for the FEHBP benefits as well. Because we do not observe marital status, it is impossible to distinguish these two cases. We ignore these individuals who “opt out,” so all empirical analyses are conditional on having chosen an FEHBP plan.

To make the data more manageable and to focus attention on the data most relevant to our study, we narrowed the sample to 86 counties in which we observed the greatest number of public and nonpublic plans operating simultaneously. The main justification for doing this is that counties without publicly or nonpublicly reporting plans are essentially uninformative with respect to the key questions of our study. Moreover, although there are many counties outside our sample that have either public or nonpublic plans, we are most interested in observing counties where both types of plans are available. Table 2 lists the important characteristics of the counties in our selected sample, and compares them to the full sample. Not surprisingly, our sample is somewhat biased toward larger counties with more offered plans. However, we note that the sample covers a broad geographic and demographic range: every major region of the U.S. is represented, with large urban counties such as New York, NY included along with smaller and less populated counties such as Tuolumne, CA. A full list of the 86 selected counties is provided in the appendix.

In principle, the set of available health plans in each county can be easily reconstructed from the individual choices. However, while health insurance coverage extends from Jan. 1 to Dec. 31, the OPM data are based on coverage and residence information as of September 30. As a result,

⁶Medicare enrollment is individual-only, but Medicare-eligible annuitants may enroll in family coverage through FEHBP. In such cases, FEHBP serves as supplemental coverage for the Medicare-eligible annuitant, and as primary or supplemental coverage for the spouse depending on his/her Medicare eligibility.

some individuals may have moved before September 30 but kept the same health plan. Subject to plan portability, a plan may show up in a county (with an extremely low market share) even if it does not serve that county. To address this contamination, we cut off the choice sets by excluding plans with market shares below 0.5% and total enrollment less than 10. The numbers in table 2 reflect this definition of the choice sets. Overall, 126 (113) HMO plans and 8 (7) FFS plans are represented in the data in 1998 (1999). The FFS plans offered through FEHBP are national plans, while the HMO plans are city- or county-specific.

Data on health plan characteristics were obtained from FEHBP guidebooks for the relevant years. Some features of health plans, such as office or prescription drug copayments, exhibit little variation across plans. However, there is substantial variation in plan premiums. During our sample period, the government covered 71% of the listed premium, with the remainder being the responsibility of the enrollee. Premiums vary across plans—even within fee-for-service or HMO classes, some plans were considerably more expensive than others—and also (to a lesser extent) over time. The top panel of Table 3 shows summary statistics for plan characteristics, comparing the selected sample to the full sample. The bottom panels of the table summarize the individual characteristics of principal interest for the empirical analysis. Note that because the sample consists solely of annuitants, the average age is quite high and the average income is rather low. Our income measure reflects pension payments by the government, and will understate true income levels to the extent that individuals have income from other sources.

In analyzing health plan choice behavior it is useful to distinguish two specific types of individuals: “switchers” and “new enrollees.” We define a switcher as someone who switches plans in a given year without being compelled to do so by some change in her choice set. As indicated in table 3, the rate of plan switching is extremely low in our sample: in a given year, fewer than 2 percent of annuitants voluntarily switch to a different plan. In spite of how few they are in number, the switchers are very important in the empirical analysis: the impact of information on plan choice is identified partly from the prevalence of switching toward higher-rated plans. Table 4 shows some switching patterns in our sample, indicating that switches tend to be toward plans with higher satisfaction rates and higher NCQA scores. The bottom panel of the table shows similar pat-

terns for “new enrollees.” Unlike switchers, who change plans without being compelled to do so, new enrollees are annuitants whose circumstances essentially force them to select a new plan. The main cause for this in our data is a move: if an annuitant moves to a new county that is not served by her current health plan, she must choose a new plan from the set of available options in the new county. Additionally, in a small number of cases an annuitant’s previous plan was dropped from the FEHBP system, so the annuitant was required to select a new plan; these individuals are also coded as new for the respective year. Together, these new enrollees account for 6% of the sample. Given the low incidence of plan switching, new individuals’ choices are perhaps the most informative observations in the data for estimating the impact of plan characteristics and published information.

It is worth noting that the low switching rate is not unique to our data. Among the Harvard employees studied in Beaulieu (2002), only 5% voluntarily switched plans in 1997, and another 4% were forced to switch due to plan termination. The voluntary switch rates reported by Sorensen (2004) are also in the 5% range (for employees of the University of California). These percentages are not far from ours, and given the size of our sample, we observe a much greater number of voluntary and involuntary switches.

For the empirical analysis, we restrict our attention to the years 1998-1999. Health plan ratings based on HEDIS/CAHPS data were not widely available until 1997: although some information was being collected by NCQA, that information was not actively disseminated until late 1996. In September 1996, *U.S. News & World Report* released an issue devoted to health plan quality concerns, including a summary of the NCQA ratings for all publicly reporting U.S. health plans. This special issue was then repeated annually for the years 1997-1998. Due to the late publication dates, the ratings are relevant to the enrollment decisions for the following year, so 1997 was the first year in which annuitants’ plan choices could have been influenced by the published ratings. However, we are forced to ignore 1997 in the analysis because we do not have complete data from NCQA on health plan scores for 1997. Although in principle *U.S. News* scores could be used to impute the NCQA information (a process explained in more detail below), the imputation is problematic for 1997, since the weighting of various NCQA categories was dramatically different

in 1997 vs. 1998-1999 (the years for which we do have complete NCQA data). This essentially makes it difficult to calculate summary ratings that are comparable across years.⁷

The critical variables in this study are the public and nonpublic NCQA plan ratings. As mentioned in the previous section, NCQA collects health plan information in many categories as part of its HEDIS/CAHPS system. In order to summarize the various scores into a single composite rating, we make use of the ratings published by *U.S. News*. These ratings were probably the highest-profile publication of NCQA information, and they are convenient as summary measures of plan quality. However, the rating system differed slightly for 1998 vs. 1999, and the exact algorithm used to compute the summary scores was not available to us. We handle the first of these complications by standardizing the scores (subtracting the mean score and dividing by the standard deviation of scores for each year). For public plans, then, our NCQA score variable is simply the standardized *U.S. News* score: it reflects the number of standard deviations above or below the mean score in that year. For nonpublic plans, we impute the score variable using the estimated relationship between observed *U.S. News* ratings and NCQA scores. That is, we regress the *U.S. News* scores for public plans on all 39 NCQA variables (with another 39 dummies to indicate if the data for those variables were missing), and use the estimated coefficients to predict the (standardized) scores that the nonpublic plans would have received had they been published in *U.S. News*.⁸ The regressions were run separately for 1998 and 1999 since we were told that *U.S. News* had changed the relative weights of the NCQA categories in their calculation. Although this imputation procedure is not perfect, we are confident that the resulting score variable is internally consistent (i.e., comparable across years and across public vs. nonpublic status). The correlation between the imputed NCQA scores and the FEHBP satisfaction index is 0.43, suggesting that while both measures tend to identify similar plans as being good ones, they are not redundant (i.e., there is information content in

⁷The 1997 *U.S. News* ratings weighted preventive care categories much more heavily than in later years. In fact, the 1997 ratings were based solely on preventive measures, while 1998 and 1999 ratings included many non-preventive categories such as access to care, access to doctor, doctor credentials, and member satisfaction. When we tried imputing the 1997 scores and incorporating them in our analysis, our findings were consistent with this: annuitants, who are probably less concerned about preventive care than other populations, appeared to respond negatively to high *U.S. News* ratings. When we broke down overall *U.S. News* ratings into specific preventive measures, we found annuitants responding positively to cholesterol check, pap test and mammography, but negatively to child immunization and prenatal care.

⁸The R^2 for the prediction equation was .92 for 1998 and .91 for 1999.

both measures).

3 Empirical Model & Results

3.1 Model

Our objective is to measure individuals' responsiveness to publicized health plan ratings, and to use the resulting estimates to calculate the value of the information. We model consumers' health plan decisions in a traditional discrete choice framework, with expected plan utilities depending explicitly on the information available. In particular, we suppose that individuals' perceived payoffs from choosing each plan depend on a set of plan and individual characteristics, x , a set of information variables (quality signals), z , and an unobserved plan characteristic, w . (We will interpret w specifically as a quality signal that is observed by consumers but not by the econometrician.) In the application, x consists of plan characteristics such as premium, type (fee-for-service or HMO), emergency room and prescription drug copayments, etc., and interactions with individual characteristics such as age, income, and sex. The key information variables (z) in our study are the satisfaction index (from the FEHBP guidebook) and the NCQA scores (as described in section 2.2).

We assume that an individual's indirect utility from choosing a plan is a linear function of plan characteristics and plan quality, which we denote q , plus an idiosyncratic error. Of course, plan quality is uncertain at the time of enrollment, so individual i forms expectations about q using quality signals. Letting $u_{ijt}(\Omega_{ijt})$ denote individual i 's expected utility from choosing plan j in period t when the information available about plan j is Ω_{ijt} , we have

$$u_{ijt}(\Omega_{ijt}) = x_{ijt}\beta + E_i[q_{jt}] + \epsilon_{ijt} .$$

We consider two information sets: $\Omega_{ijt}^0 = \{x_{jt}, w_{ijt}\}$ when plan ratings are not available to the public, and $\Omega_{ijt}^1 = \{x_{jt}, z_{jt}, w_{ijt}\}$ when plan ratings get publicized. We also assume that the conditional expectation function for q is linear, so that:

$$E_i[q_{jt}|x_{jt}, w_{ijt}] = x_{jt}\delta_0 + w_{ijt}\theta_0$$

and

$$E_i[q_{jt}|x_{jt}, z_{jt}, w_{ijt}] = x_{jt}\delta_1 + z_{jt}\gamma_1 + w_{ijt}\theta_1 .$$

Expected plan utilities are then given by

$$u_{ijt}(\Omega_{ijt}^0) = x_{ijt}(\beta + \delta_0) + w_{ijt}\theta_0 + \epsilon_{ijt} \quad (1)$$

and

$$u_{ijt}(\Omega_{ijt}^1) = x_{ijt}(\beta + \delta_1) + z_{jt}\gamma_1 + w_{ijt}\theta_1 + \epsilon_{ijt} \quad (2)$$

The above expressions for plan utilities highlight two important difficulties in measuring consumers' responsiveness to information. First, since observable information variables are likely to be correlated with unobserved signals of plan quality, coefficients on those information variables will typically misrepresent the true impact of the information. In the present context, where the z 's are health plan ratings, we are concerned that $\hat{\gamma}_1$ will reflect more than just the impact of those ratings on consumers' choices: it will also reflect any correlation of plan ratings with unobserved plan quality. In the worst case, the plan ratings may simply serve as perfect proxies for w , so that even though they have no informational value they will be estimated to have a positive impact on consumers' choices. The second difficulty can be seen in the comparison of equations 1 and 2: when new information becomes available, other factors that were used in forming expectations about plan quality will be weighted differently. In other words, the coefficients on other variables are likely to change when z (the new information) is made available to consumers. For example, in the absence of published health plan ratings, consumers may rely on plan premiums as signals of quality. In that case, we would expect the estimated coefficient on premium to be different depending on whether the ratings were available.

The availability of nonpublic information in our data allows us to address these issues directly. To see the first point, let us ignore the second problem for the moment. If publicizing plan ratings does not entail changes in the other coefficients, γ should capture the full value of information and the only bias in $\hat{\gamma}$ arises from the correlation between plan ratings z and unobserved quality

signals w . Our data allow us to use the undisclosed health plan ratings to directly correct this bias. Because nonpublic plans' ratings were undisclosed, they have no direct impact on consumers' choices: i.e., the coefficient on nonpublic plan scores must be zero in the choice model. Therefore, if we include nonpublic plan scores in the empirical model, then the estimated coefficient reflects *only* the correlation of those scores with unobserved quality. If we also assume that the correlation of plan scores with unobserved plan quality is the same regardless of whether the scores were made public or not, then the difference between the coefficients for public vs. non-public plan scores identifies the direct impact of the information.

If publicizing plan ratings leads consumers to reweight other quality signals, the above logic must be modified. The change in perceived plan utilities resulting from new information is obtained by differencing equations 1 and 2:

$$u_{ijt}(\Omega_{ijt}^1) - u_{ijt}(\Omega_{ijt}^0) = x_{ijt}(\delta_1 - \delta_0) + z_{jt}\gamma_1 + w_{ijt}(\theta_1 - \theta_0). \quad (3)$$

This expression suggests that γ_1 is not the only parameter of interest: introducing new information changes the weights on other quality signals, and this indirect effect should be taken into account in computing the value of the information. For the x variables we can address this issue directly, by allowing their coefficients to differ across information regimes (i.e., public vs. nonpublic plans); however, since w_{ijt} is unobserved, we cannot directly measure changes in its impact on perceived utilities. In what follows, we report results from estimating the model under the assumption that $\delta_0 = \delta_1$ and $\theta_0 = \theta_1$, and also consider specifications in which δ_0 and δ_1 are estimated separately. If these assumptions are wrong, our estimate of γ_1 will reflect a combination of the pure impact of the information and changes in the influences of other quality signals. However, while this suggests a bias in $\hat{\gamma}_1$ in the technical sense, it is arguably advantageous for calculating the value of information: to the extent possible, we want our calculations to incorporate all the changes in consumers' choices that result from the new information.

The public-nonpublic comparison we propose above is complementary to various empirical strategies utilized in the existing literature. When data from multiple time periods are available, one approach is to assume that unobserved quality is time-invariant and common to all individuals

(i.e., $w_{ijt} = w_j$). In that case, the w_j 's can be estimated as fixed effects. This is a way of controlling for unobserved plan quality, and it is essentially the approach adopted by Chernew et al (2004). However, if in fact there are time- or individual-specific quality signals observed by consumers but not by econometricians—i.e., $w_{ijt} = w_j + \eta_{ijt}$ —then the endogeneity problem is still present to the extent that $score_{jt}$ is correlated with the specific quality signal η_{ijt} .

Another approach to control for unobserved quality signals is to directly account for individual specific enrollment history. For example, Beaulieu (2002) estimates plan choice (and the probability of switching) as a function of the number of years an individual has been enrolled in a specific plan. This allows each individual to have better knowledge about the plan she chose before, but still assumes that the unobserved quality signal is orthogonal to observed plan ratings for all the plans she did not choose. Apparently, this approach is complementary to the fixed-effects one: enrollment history controls for individual- and time-specific knowledge of plan quality, while plan fixed effects control for time-invariant information that is common to every individual.

Because our data include multiple years *and* some nonpublic scores, we are able to augment the previous approaches with the public-nonpublic comparison outlined above. Rather than estimate fixed effects, however, we instead consider specifications in which each plan's market share from the previous year is included as a control for w_j . Lagged market shares essentially absorb the unobserved heterogeneity in the same way fixed effects would, while still allowing for some movement in unobserved plan quality over time.⁹ More importantly, the public-nonpublic comparison allows us to assess how effectively the existing approaches address the problem of unobserved quality signals: if nonpublic ratings still appear to have a significant impact after we include a specific control, then that control must not fully correct the bias from the unobservables.

3.2 Endogeneity of Disclosure Decisions

A potential concern with our empirical analysis is that health plans' disclosure decisions are endogenous. Participation in the NCQA rating programs is voluntary, so plans choose whether to report and, conditional on reporting, whether to allow the information to be publicized. Plans that

⁹In what follows we report results from specifications with lagged market shares included as controls; specifications with plan fixed effects yielded qualitatively and quantitatively similar results.

elect not to report to NCQA apparently do so either because it is too costly to construct data in compliance with NCQA specifications or because they face little market pressure to do so. Also, some plans do not meet minimum size requirements for reporting clinical indicators. Some plans capitate most doctors and it is difficult to synchronize different databases kept in each individual doctor's office. In the October 13, 1997 issue of *U.S. News*, the editor wrote "Don't assume that (an HMO not listed in the comparison table) is a subpar performer. Many managed-care plans in this fast-changing sector are being absorbed into others or are in the process of restructuring themselves, and yours simply may be too new to have the several years' worth of data required for listing. Others lack the information-gathering abilities to collect vast quantities of data. And HMOs with minimal competition may feel little pressure to generate numbers if local employers do not ask for them."

It is more puzzling why a plan would report its data to NCQA but choose to keep the data nonpublic. One reason cited by NCQA staff is concerns about heterogeneous auditing practices. Although NCQA required submitted data to be audited by a third party beginning in 1998, prior to that many HMOs submitted unaudited data. Some plans felt it was unfair for their audited data to be compared with others' unaudited data, but still wanted their numbers to enter in the calculation of national averages, since the latter may be used in shaping future public policies for the managed care industry.

However, this does not explain why nonpublic reporting still exists after 1998. Economic theory suggests that zero-cost disclosure should be 100% unraveling: as long as consumers assume non-disclosing firms have the worst quality, every firm above the worst quality has an incentive to disclose, and therefore disclosure and non-disclosure are equally revealing for the worst ones.¹⁰ We suspect the unraveling theory fails in the context of health plans because consumers do not necessarily know what the worst quality is. Obviously, if the data turn out to be worse than consumer beliefs, the plan may choose to hide the data. Another possibility is that plans may not believe NCQA specifications provide an unbiased description of HMO quality. For example, some plans are good in dental and vision care, but NCQA measures do not capture this competitive advantage. Reporting in the publicized dataset may mislead consumers to cherry-pick plans that are only

¹⁰See, e.g., Milgrom (1981), Grossman (1981), and Jovanovic (1982).

strong in dimensions reported by NCQA.

Although modeling plans' disclosure decisions is beyond the scope of this paper, we address the issue indirectly by including an indicator for publicly reporting plans as a covariate in the empirical models we estimate below. To the extent that a plan's reporting status is informative about its quality, consumers' mean response to that information should be captured in the coefficient on this dummy variable.

Given the empirical strategy of the paper—to compare the estimated impacts on consumer decisions of public vs. nonpublic ratings—it is important that there be some overlap in the distributions of scores for public and nonpublic plans. Implicitly, our analysis asks what the market shares of nonpublic plans would have been if their scores had been public. But we intuitively expect nonpublic plans to have lower scores than public plans—since low-scoring plans are precisely the ones with an incentive to withhold the information—and comparisons of nonpublic plans with public plans would be difficult if the former were of uniformly lower quality. However, as mentioned above, there are other idiosyncratic reasons why plans might elect to be nonpublic, and in fact we find that many nonpublic plans have relatively high ratings. Figure 1 shows the score distributions for public and nonpublic plans, both for the full sample and for the selected sample. As expected, public plans clearly have higher average scores, but the overlap in the distributions is still substantial. This overlap suggests (and our conversations with researchers at NCQA have tended to confirm) that plans' reasons for selecting nonpublic status are often somewhat idiosyncratic, and not merely a function of plan quality.

An additional concern is that the endogeneity of plans' disclosure decisions may undermine our assumption that the correlation between NCQA scores and unobserved quality signals is the same for public and nonpublic plans. However, we see no reason that endogenous selection of public vs. nonpublic plans would imply differences in this correlation. Suppose, for example, that plans choose to allow public disclosure if their ratings are better than the quality signal observed by consumers. Although this would truncate the distribution of ratings that consumers get to observe, it would not necessarily induce a difference in the correlation between ratings and quality signals for public vs. nonpublic plans.

3.3 Estimation & Results

We estimate the choice model in equations (1) and (2) assuming the distribution of the errors, ϵ_{ijt} , is type I extreme value, so that the choice probabilities have the convenient logit form. Our x variables include plan premium (interacted with age, sex, and income), a fee-for-service indicator (also interacted with age, sex, and income), and copays for outpatient visits, emergency room visits, and prescription drugs (separately for branded and generic).¹¹ Summary statistics for these variables are reported in Table 3. Our information variables, z , are the satisfaction index from the FEHBP guidebook and the NCQA plan score.¹² As described in section 2.2, the fraction of individuals switching plans in a given year is typically quite low. To accommodate this feature of the data, we include an additional covariate to represent switching costs. In particular, we define an indicator $switch_{ijt} = 1$ if individual i did *not* choose plan j in period $t - 1$. In other words, in any given period, this indicator is turned on for every plan except the one chosen last period.

Table 5 reports maximum likelihood estimates for four specifications of the model.¹³ Each of the four columns in the table represents different sets of controls, to facilitate comparisons of the corresponding estimates of plan ratings' impact. The first column controls for neither switching costs nor lagged market shares; essentially, it estimates the impact of plan ratings purely from the cross-sectional patterns in market shares. The second and third columns control for switching costs and lagged market shares (respectively), and the fourth column—our preferred specification—controls for both.

The estimated coefficients on plan characteristics are mostly consistent with basic intuition about health plan choice.¹⁴ The coefficient on premium is negative and statistically significant, and

¹¹Some additional variables are included in the estimated models to serve as controls, but not reported in the tables in order to save space. These include indicators for type of coverage (self or family), Medicare eligibility, NCQA accreditation status, and dummy variables to indicate whether the values of certain variables were missing.

¹²The NCQA scores apply only to HMO plans. For FFS plans, we set the scores to zero and include an FFS dummy to control for the “missing” scores.

¹³The large size of our sample makes the model computationally burdensome to estimate. To reduce the burden, we randomly sampled non-new and non-switching enrollees for inclusion in the estimation: the sample used for estimation includes all new enrollees, all enrollees who switch plans, and a random 5% sample of all other enrollees. The observations are then weighted so that the estimated coefficients accurately reflect the effects that would be obtained from the full sample.

¹⁴Several plan characteristics are included in the model as controls but omitted from the table to save space. These include HMO age, model type, an indicator for whether the HMO belongs to a national managed care chain, the

it differs meaningfully depending on individual characteristics: e.g., price sensitivity decreases with income. The large negative coefficient on the switch indicator reflects the persistence of health plan enrollment decisions: people very rarely switch away from their chosen plan. Females are less likely to switch plans than males, and switching is more likely among younger annuitants with higher incomes.¹⁵ The importance of allowing for this inertia is evident from the dramatic improvement in fit when the switch indicators are included (the pseudo- R^2 increases from 0.18 to over 0.90). The coefficients on both the switch variables and the lagged market shares reflect a general persistence in plans' market shares, which could result from reluctance to switch plans at the individual level, or from small net changes in enrollment shares in spite of high turnover at the individual level. Including the switch variables in the model allows the data to distinguish these two explanations (in favor of the former), and also yields more reliable estimates of the premium coefficients.

In general, the coefficients on the information variables are sensible, indicating that annuitants prefer plans with higher published ratings. However, the important pattern shown in the tables is that the magnitudes of the coefficients shrink when switching costs and lagged market shares are included as controls. This is especially true for the satisfaction rate. The pattern is suggestive of the central concern this paper addresses: that reported plan ratings are correlated with unobserved plan quality. Note that if the impact of plan ratings is estimated purely from the cross-section (column I), the estimate is quite large with a very small standard error. However, it is clear from table 5 that much of this estimated effect must be spurious: the coefficient on nonpublic NCQA scores is also large and statistically significant, which would imply that individuals were affected by information they never received. Our ability to observe "responsiveness" to ratings that were never disclosed allows us to assess the effectiveness of various controls at mitigating the bias. The estimates reported in column IV of table 5 suggest that switch costs and lagged market shares help control for unobserved plan quality, as the coefficient on nonpublic scores is closer to zero. This is

number of hospitals under contract with the HMO, and the number of physicians under contract with the HMO.

¹⁵Note that although it is tempting to convert the switch coefficient to dollars (using the premium coefficient) and interpret it as a literal switching cost, this would ignore other plausible explanations for the observed inertia. For instance, the negative coefficient could reflect persistence in preferences: consumers tend to stay in the same plan not because it is costly to switch, but because their idiosyncratic preferences for plans are highly stable over time (e.g., the ϵ_{ijt} 's could be serially correlated).

somewhat encouraging, since in general it is easier to obtain panel data and control for plan effects than it is to obtain nonpublic data. However, it appears that neither control is sufficient to entirely resolve the problematic correlation between published information and unobserved quality.

The positive difference in the estimated coefficients for public vs. nonpublic plan ratings reveals clearly that published ratings indeed contain information that is acted upon. If these coefficients were equal, we would be forced to conclude that published scores contained no new information—i.e., they simply mirrored information that was already known. However, the estimated difference is statistically and economically significant. If interpreted literally as the causal effect of the ratings on perceived plan qualities, the difference in the public vs. nonpublic coefficients (0.292 in the preferred specification in column IV) is large enough to suggest the published scores have a meaningful impact on choice: if a plan’s score increased by one standard deviation, this would increase a typical individual’s likelihood of choosing that plan by 2.63 percentage points. To put this in context, an equivalent increase in the likelihood of selection would result from a \$759 decrease in the annual gross premium (which would mean \$220 less per year out of the annuitant’s pocket). The quantitative estimates of individuals’ responsiveness to information will figure prominently in the analysis of section 3.4, where we use the results to compute the dollar value of the information.

Table 6 presents estimates of the choice model based on two divisions of the sample. The first two columns compare the results for Medicare-eligible vs. non-Medicare-eligible individuals. As mentioned above, Medicare-eligible annuitants purchase FEHBP plans as supplemental coverage, and therefore may respond differently to published information than individuals for whom the FEHBP plan is the primary insurer. However, the results indicate that choice behavior is quite similar between the two groups. Medicare-eligible annuitants are more price-sensitive and less responsive to NCQA ratings, but the differences are small. This is encouraging, since it suggests that our findings aren’t peculiar to the Medicare subpopulation.

Perhaps the most informative comparison afforded by the data is that of “old” vs. “new” enrollees. As is evident from the estimated coefficients on the switch indicator, annuitants tend to choose a plan and then stick with it. The choices of individuals who were already enrolled in a certain plan are in a way less informative than the choices of “new” individuals whose circum-

stances required a fresh decision. Given that the new annuitants we identify are primarily those moving to a new county (with a different set of available health plans), we expect them to have less prior information about the options, and to therefore be more responsive to published ratings. Columns III and IV of Table 6 present estimates of the choice model separately for new vs. old enrollees. As expected, individuals making new decisions appear to be much more sensitive to published information about plan quality. For annuitants who had the option of simply staying in their previous plan, the difference between the coefficients on public vs. nonpublic scores is in fact slightly negative (though the difference is not statistically significant). For new enrollees, on the other hand, the difference is large: a one standard-deviation increase in a public plan's score would increase its probability of selection by approximately 3.89 percentage points for an individual with mean characteristics.

As suggested in section 3.1, it is potentially important to allow the coefficients on the x variables to differ depending on whether the information variables were available. In principle, individuals forming expectations about plan quality should weight plan characteristics differently for public vs. nonpublic plans. To account for this possibility, we estimated the model allowing the coefficients on the x variables to be different for public vs. nonpublic plans. As discussed in section 3.1, this gives us some explicit idea about how consumers reweight the observable x 's as quality signals. Table 7 reports results from this alternative specification. The estimated coefficients on the NCQA score variables are very similar to those reported in Tables 5 and 6, but some of the x -interactions are significant. One notable difference is that individuals appear to be more price-sensitive for public plans, which is consistent with the idea that premiums could be used as signals of quality in the absence of other direct information about quality. (For nonpublic plans, higher premiums might be taken to imply higher quality, which would tend to make demand appear price-inelastic.) Individuals also appear to be less responsive to lagged market shares for public plans, and less averse to switching to public plans, which again could be taken to imply that plan information from alternative channels is relied on more heavily when publicized ratings are unavailable. However, the coefficients on the Coverage \times Public interactions are an exception to this pattern: individuals appear to be more responsive to coverage ratings for public plans than for

nonpublic plans.

Collectively, the results in tables 5-7 tell a plausible story about consumer choice behavior and responsiveness to information, and the pattern of coefficients on the public and nonpublic NCQA scores is fairly robust. Several additional controls were included in the estimation but omitted from the tables to save space; in general the coefficients on these controls are consistent with intuition, but a few of the coefficients have unexpected signs. For example, the coefficient on a plan's copay rate for branded drugs is positive,¹⁶ which could mean that higher copays for branded drugs are correlated with more generous formularies (about which our data have no information). Because these few anomalous coefficients suggested we may have some important omitted variables, we re-estimated the models to see if omitting different sets of plan characteristics would have an appreciable impact on our main results. The results are qualitatively and quantitatively similar, though the coefficients on the NCQA score variables change slightly when copay variables and/or contracting variables are omitted.¹⁷ We prefer the reported specification because copays are clearly printed in the FEHBP brochures and contracting variables were found to be important in numerous previous studies; including these variables as controls should capture (to the extent our data allow) the other information that enrollees observe beyond the NCQA scores.

One potential concern with our empirical results is that premiums may be endogenous in the sense of being correlated with unobserved plan quality. This is a classic problem with price coefficients in discrete choice models, and generally tends to bias the estimates toward finding relatively inelastic demand. This is very important in our case, since in order to calculate a meaningful estimate of the value of information, we need to use the estimated price coefficient to convert numbers to dollars. If our estimate of individuals' price sensitivity is too close to zero (too inelastic), when we divide by it we will get an upward-biased estimate of the dollar value of information.

The magnitudes of the premium coefficients indeed suggest that annuitants are not very sensitive to price. For example, for a plan with a market share of 10%, a \$300 increase in the gross annual premium would reduce its probability of selection by only 1.04 percentage points for a

¹⁶We also control for the copay of generic drugs. Its coefficient is negative, as expected.

¹⁷The coefficient on the public NCQA scores increases slightly, and the coefficient on the non-public NCQA scores tends to decrease slightly when omitting observable plan characteristics. Details of these sensitivity tests (along with tables listing the full sets of estimated coefficients) are available from the authors.

typical annuitant (based on the estimates in column IV of table 5). This relative unresponsiveness to price is due in part to employer cost-sharing—the \$300 increase in gross premium requires only \$100 in additional payments by the enrollee—but we note that it may reflect the endogeneity problem described above. Nevertheless, controlling for lagged market share should alleviate much of the problem: any bias that remains must reflect correlation between premiums and period-specific quality shocks. Also, comparisons with findings from other studies are reassuring. Wedig and Tai-Seale (2002) estimate price sensitivities very close to ours (based on the specifications in column IV of Table 5 or column I of Table 7), and the estimates of Chernew et al (2004) are actually smaller (implying even less price sensitivity than what we find).

3.4 Calculating the value of information

The estimates reported in tables 5-7 can be used to derive the value to consumers of the NCQA scores. Changing notation slightly from the previous section, let $u_i(j)$ denote individual i 's indirect utility from plan j . If a_i is the plan that individual i would choose if the information is available, and b_i the plan that would be chosen in the absence of the information, then the dollar value of the information to individual i is $[u_i(a_i) - u_i(b_i)]/\alpha_i$, where α_i is the marginal utility of a dollar. Under this definition, information is valueless (*ex post*) if it doesn't change the individual's choice—i.e., if $a_i = b_i$. Note that in principle one could allow the utility function itself to change depending on the information available, in which case the information could have value even if $a_i = b_i$. For example, people could derive utility from being told that their chosen plan is highly rated even if that information would not have affected their choice. In what follows, we instead choose to hold the utility function fixed, and calculate the *ex post* value of information by determining (a) whether the information affected the individual's choice and (b) if so, what the resulting gain in utility was.

To generate numerical estimates of the value of information, we consider two counterfactual experiments. First, we ask how utility would have been affected if the public NCQA scores had been nonpublic—i.e., we compare the actual information regime to one with no publicized information at all. Second, we ask how utility would have changed if the nonpublic NCQA scores had been publicized. In each case, we estimate the value of information for each individual as

$$\hat{v}_i = \frac{\hat{u}_i(\hat{a}_i) - \hat{u}_i(\hat{b}_i)}{\hat{\alpha}_i},$$

where \hat{u}_i is the estimated utility function from the “informed” regime, $\hat{\alpha}_i$ is the estimated marginal utility of a dollar for individual i ,¹⁸ and \hat{a}_i (\hat{b}_i) is the model’s prediction of individual i ’s choice with (without) information. To be specific, we obtain estimates from the following simulation procedure: (1) calculate plan utilities for each individual under each regime (with vs. without information), using random draws from the extreme value distribution for the ϵ ’s; (2) obtain \hat{a}_i (\hat{b}_i) by determining which plan yields the maximal predicted utility in the informed (uninformed) model; (3) calculate the difference in utilities from \hat{a}_i and \hat{b}_i using the utility function from the informed regime; (4) divide that difference by $\hat{\alpha}_i$ to convert the number to dollars per year.¹⁹

Table 8 shows average estimates from 100 replications of the above simulation procedure. Based on the specification in column I of table 7, we estimate that the choices of 3,539 individuals were directly affected by the publicized NCQA scores, in the sense that they would have chosen a different plan if they hadn’t had the information. However, this is a very small percentage of the total: 99.3% of the enrollment choices in our sample would have been the same with or without the information. This is primarily due to the considerable inertia in health plan enrollment decisions: for the vast majority of individuals, the switching costs swamp any effect the information might have. For those individuals whose decisions were materially affected by the ratings, the value of the information is estimated to be substantial: \$160 per person per year. Averaged over all individuals in the sample, the value of the published scores is only \$1.11 per person. Interestingly, the second column of the table indicates that publicizing the nonpublic scores (the second counterfactual experiment) would have had a slightly lower informational value, presumably because nonpublic plans tend to have lower market shares than public plans and there are fewer nonpublic plans overall.

The relatively small fraction of decisions for which the ratings are pivotal suggests separating the estimates for new vs. old individuals: the choices of annuitants whose circumstances necessitate

¹⁸This varies across individuals because the coefficient on premium in the choice model depends on the individual’s age, sex, and income.

¹⁹In this case, $\hat{\alpha}_i$ is obtained as (coefficient on premium for i) $\times 26 \times 0.29$, since payments are biweekly and the annuitants’ share of the cost is 29 percent.

a new health plan decision are presumably more sensitive to the ratings, and we should therefore expect the information to be more valuable to them. The second panel of table 8 presents calculated values based on separately estimated models for old and new annuitants. The value of publicized NCQA scores is clearly smaller for individuals who have the option of maintaining enrollment in their previous plan: less than half of one percent of these individuals are estimated to have made different choices than they would have in the absence of the publicized information. In contrast, the information appears to be much more influential for individuals making new choices: over 8% of these annuitants were affected by the publicized ratings, and the average value of the information among those whose choices “changed” was \$588.67 (and \$47.31 among all new enrollees). Since new enrollees comprise roughly 6.2% of the sample, our estimate of the overall average value of the information for the entire sample is $(47.31)(0.062)+(588.67)(0.0938)=\3.39 per person per year.²⁰

These estimates can be used to make a rough calculation of the overall gain in consumer welfare derived from the information provided by NCQA. Multiplied by the 1.7 million individuals in the full sample who chose an FEHBP plan, the \$3.39 per person implies a \$5.76 million overall benefit per year to annuitants of the U.S. government. This number is likely to overstate the true benefit, however, since plans with publicized ratings are less prevalent in general than in the specific subsample we selected for the analysis. However, even if there were no benefits whatsoever to individuals outside the selected sample, the estimated overall benefit to those 250,000 individuals who are in the sample would be \$0.85 million per year.

Dynamic considerations also suggest these numbers may understate the true value of information. We calculate the expected dollar value for a given year, but the effects of information are not confined to the current period: costs of bad choices continue over each additional period in which the incorrect choice persists. Also, switching costs are incurred when an individual ultimately switches out of a bad choice. This suggests that the present value of information should incorporate not just the benefits of correct choices in the current year, but also the implicit benefits in future years of avoiding the costs associated with bad choices.

²⁰In order to put a confidence interval around this number, we performed a parametric bootstrap procedure in which we resampled the parameter vector from the distribution implied by our estimates, and recalculated (using the simulation method described above) the value of information number for each resampled value of the parameter. Using 30 repetitions we get a 95% confidence interval of [\$3.14,\$3.64].

4 Discussion and conclusions

Our empirical results clearly indicate that publicized ratings have a direct impact on individuals' choices of health plans—a finding that is consistent with other recent studies of information and health plan choice. More importantly, our ability to compare the effects of publicized and non-publicized plan ratings allows us to assess the significance of a potential source of bias: if publicized ratings largely reflect information that is already known to consumers (but not econometricians)—that is, if ratings are correlated with unobserved plan quality—then naive estimators will overstate the actual impact of the information. Not surprisingly, this problem is most severe when the effect of information variables is estimated from a simple cross-section: in that case, even information that was unavailable to consumers appears to have had a substantial impact.

We suspect nonpublic data will rarely be accessible to researchers, so comparisons like the ones made here will not typically be feasible. Nevertheless, our findings about the effectiveness of various control methods are somewhat encouraging. In particular, using panel data and incorporating plan fixed effects or lagged market shares appears to be helpful: in our application, including lagged market shares and controlling for switching costs eliminate some (but not all) of the bias.

Our estimates imply the average dollar value of the publicized plan information is modest (\$3.39 per person per year when averaged over all individuals in our sample), but the value is much more substantial for the small fraction of individuals whose decisions are affected by the information. In any case, even this seemingly modest estimate implies benefits that far outweigh the costs of acquiring and distributing the information. If the \$3.39 per household is crudely extrapolated to every U.S. household enrolled in a health plan, we would arrive at a total benefit easily exceeding \$200 million per year. Though such a number is obviously speculative, it seems clear that the overall benefits to consumers far outweigh the costs. (NCQA's annual operating expenses were roughly \$20 million for the period we studied, and only a fraction of those expenses were related to the HEDIS and CAHPS programs.)

Note also that the benefits we have calculated are the *direct* benefits to consumers; we have made no attempt to estimate the indirect benefits that may accrue from health plans' increased incentives to provide quality. If active publication of health plan quality reports causes plans to

improve their services and obtain better health outcomes, the consequent benefits to consumers could conceivably swamp the benefits we've estimated here. Analyzing health plans' responses to the publication of information is a difficult but important issue to explore.²¹

Finally, our results suggest open questions about the format in which information is delivered to consumers. Would the impact of the publicized plan ratings be greater if the ratings were reported differently? How much more influential is the information if it is delivered directly to consumers in a brochure instead of indirectly via news publications or the internet? Also, for which categories of care does information about plan quality appear to be highly valued? Our findings confirm that quality information is valued by consumers in a general sense, leaving these (and other) questions about the nuances of publicizing quality ratings for future research.

²¹Jin (2004) is one example of a paper that focuses on supply-side responses to the provision of information.

References

- Beaulieu, N. D. (2002, January). Quality Information and Consumer Health Plan Choices. *Journal of Health Economics* 21(1), 43–63.
- Chernew, M., G. Gowrisankaran, and D. P. Scanlon (2004, February). Learning and the Value of Information: The Case of Health Plan Report Cards.
- Grossman, S. (1981). The Informational Role of Warranties and Private Disclosure about Product Quality. *Journal of Law and Economics* 24, 461–489.
- Jin, G. Z. (2004). Competition and Disclosure Incentives: An Empirical Study of HMOs. forthcoming, *RAND Journal of Economics*.
- Jovanovic, B. (1982). Truthful Disclosure of Information. *Bell Journal of Economics* 13, 36–44.
- Milgrom, P. (1981). Good News and Bad News: Representation Theorems and Applications. *Bell Journal of Economics*, 380–391.
- Scanlon, D. P., M. E. Chernew, C. G. McLaughlin, and G. Solon (2002, January). The Impact of Health Plan Report Cards on Managed Care Enrollment. *Journal of Health Economics* 21(1), 19–41.
- Sorensen, A. (2004). Social Learning and Health Plan Choice. Stanford University working paper.
- Wedig, G. J. and M. Tai-Seale (2002, November). The Effect of Report Cards on Consumer Choice in the Health Insurance Market. *Journal of Health Economics* 21(6), 1031–1048.

Appendix A: Counties in selected sample

County	# annuitants	County	# annuitants
Alameda CA	10479	Merced CA	681
Albany NY	1936	Montgomery NY	301
Allegheny PA	8093	Nash NC	305
Amador CA	318	Nassau NY	5480
Ashtabula OH	256	Nevada CA	755
Beaver PA	518	New York NY	3776
Boulder CO	1652	Nye NV	427
Bronx NY	4046	Orange CA	7051
Butler OH	613	Orange NY	1762
Butler PA	817	Parker TX	472
Butte CA	1283	Putnam NY	264
Calaveras CA	346	Queens NY	8471
Cape May NJ	943	Rensselaer NY	1093
Cattaraugus NY	227	Riverside CA	5413
Chautauqua NY	362	Robeson NC	369
Clark NV	7879	Rockland NY	774
Cook IL	19084	Sacramento CA	14512
Cowlitz WA	352	San Bernardino CA	6842
Cumberland NJ	282	San Diego CA	18188
Cumberland PA	4214	San Francisco CA	5147
Custer CO	29	San Mateo CA	3943
Denver CO	4875	Santa Clara CA	4629
El Paso CO	4319	Santa Cruz CA	676
Essex NY	176	Saratoga NY	916
Fremont CO	467	Schenectady NY	1052
Genesee NY	270	Schoharie NY	107
Gloucester NJ	1659	Sonoma CA	2704
Graham AZ	166	Stanislaus CA	1374
Greene NY	201	Suffolk NY	5551
Hamilton OH	4708	Sullivan NY	296
Highland OH	262	Teller CO	149
Jefferson CO	5089	Tulare CA	726
Johnston NC	392	Tuolumne CA	434
Kings CA	369	Ventura CA	5138
Kings NY	7282	Walworth WI	195
Lawrence PA	300	Warren NY	213
Lewis WA	455	Warren OH	404
Los Angeles CA	21741	Washington NY	179
Louisa VA	258	Washington PA	703
Madera CA	403	Wayne MI	6442
Madison NY	282	Westchester NY	3102
Maricopa AZ	15087	Wyoming NY	119
Marin CA	1495	Yamhill OR	412

Table 1: Summary of health plans

	1998	1999
Total # of health plans	442	362
# FFS plans	14	13
% enrollment in FFS	60.95	60.94
# HMO plans	428	349
% enrollment in HMO	12.38	12.49
# Publicly reporting plans	175	133
% enrollment in public	7.25	7.10
# Nonpublic plans	72	95
% enrollment in nonpublic	2.64	3.33

The percentages enrolled in FFS and HMO plans do not add to one because some annuitants opt out, choosing not to enroll in any FEHBP plan. The percentages enrolled in public and nonpublic do not add to the percentage enrolled in HMOs because some plans are non-reporting (i.e., neither public nor nonpublic).

Table 2: Summary statistics for counties: selected vs. full sample

	Full Sample		Selected Sample	
	1998	1999	1998	1999
# of counties	3180	3220	86	86
Median # annuitants per county	106	106	736	741
Mean # annuitants per county	530	523	3000	2971
Median # plans per county	6	6	11	10
Median # HMOs per county	2	2	5	5
Median # reporting plans	2	2	3	3
Median # public per county	1	1	2	2
Median max# public per county	1	1	3	3
Median # nonpublic per county	1	1	1	2
Median max# nonpublic per county	1	1	2	2

Table 3: Summary of plan and individual characteristics (means and standard deviations)

	Full sample	Selected sample
Plan Characteristics:		
Gross Premium (\$)	83.71 (14.96)	85.14 (17.89)
Office copay (\$)	6.39 (3.12)	6.06 (3.35)
Emergency copay (\$)	38.85 (15.70)	34.65 (16.53)
Prescription copay (\$)	7.96 (3.74)	7.48 (3.79)
Generic copay (\$)	5.49 (2.21)	5.30 (2.19)
Satisfaction rate (%)	57.68 (9.29)	56.82 (9.00)
NCQA score	-0.20 (1.11)	-0.08 (1.03)
NCQA accredited? (0/1)	0.54 (0.50)	0.66 (0.48)
Individual characteristics:		
Age	70.82 (10.38)	71.61 (10.46)
Female	0.43 (0.49)	0.44 (0.50)
Income	20315.70 (12337.95)	19822.66 (11720.29)
Number of individuals	3,427,163	509,139
Percent new	6.35	6.00
Percent switchers	2.10	1.74
Percent in FFS plans	82.91	65.60
Percent w/ family coverage	46.02	40.60
Percent eligible for Medicare	66.47	67.84

The table lists means with standard deviations in parentheses below. Gross premium is biweekly; the annuitants' share of the payment is 29%. Copay numbers are for HMO plans only. "Generic copay" is the required copayment for generic prescriptions. The satisfaction rate is the percentage of customers who reported being "very satisfied" or "extremely satisfied," as opposed to "not satisfied" or "somewhat satisfied." Income is annual government pension payments, and does not include outside income. The number of individuals reflects the number who enrolled in some FEHBP plan—i.e., those that did not opt out. "New" individuals are those whose circumstances required them to choose a new plan (e.g. because they moved). Switchers are those who were not new (by the above definition) and who switched plans in the corresponding year.

Table 4: Choice patterns for switchers and new enrollees

	Switchers		
	Prev. plan	New plan	Other plans
Percent FFS	64.92	60.56	56.47
Percent Public	46.85	49.26	39.85
Percent Nonpublic	32.20	28.18	27.04
Mean satisfaction rate	61.54	63.41	60.79
Mean NCQA score	-0.18	0.04	-0.32
Mean previous-year market share	11.84	19.48	6.24

	New enrollees		
	Prev. plan	New plan	Other plans
Percent FFS	–	48.94	57.52
Percent Public	–	54.03	37.88
Percent Nonpublic	–	23.57	27.95
Mean satisfaction rate	–	62.23	60.13
Mean NCQA score	–	0.21	-0.57
Mean previous-year market share	–	23.10	5.75

Table 5: Estimates of plan choice model

	I	II	III	IV
Premium	-0.0316 (0.0001)	-0.0290 (0.0003)	-0.0015 (0.0001)	-0.0100 (0.0004)
Premium×Female	-0.0136 (0.0002)	-0.0051 (0.0005)	-0.0003 (0.0002)	0.0040 (0.0006)
Premium×Age	-0.0001 (0.0000)	-0.0002 (0.0000)	0.0001 (0.0000)	0.0001 (0.0000)
Premium×Income	-0.0010 (0.0001)	0.0004 (0.0002)	-0.0018 (0.0001)	0.0008 (0.0002)
Switch	—	-6.6811 (0.0140)	—	-6.5029 (0.0141)
Switch×Female	—	-0.3392 (0.0268)	—	-0.3372 (0.0270)
Switch×Age	—	-0.0495 (0.0011)	—	-0.0455 (0.0011)
Switch×Income	—	0.1360 (0.0102)	—	0.1284 (0.0103)
Lag Market Share	—	—	6.1489 (0.0113)	4.8167 (0.0351)
Satisfaction rate×FFS	0.1131 (0.0004)	0.0829 (0.0012)	0.0193 (0.0004)	0.0146 (0.0014)
Satisfaction rate×HMO	0.0074 (0.0006)	-0.0003 (0.0015)	-0.0076 (0.0006)	-0.0092 (0.0016)
Coverage	0.0434 (0.0009)	0.0786 (0.0025)	0.0006 (0.0008)	0.0518 (0.0024)
Coverage×1999	-0.5171 (0.0050)	-0.2053 (0.0160)	0.0292 (0.0060)	0.1593 (0.0180)
Public	0.1259 (0.0153)	0.0537 (0.0315)	-0.0487 (0.0147)	-0.0367 (0.0318)
NCQA Score×Public	0.3781 (0.0081)	0.4152 (0.0173)	0.2900 (0.0080)	0.3747 (0.0175)
NCQA Score×Nonpublic	0.0951 (0.0062)	0.1331 (0.0132)	0.1082 (0.0061)	0.0828 (0.0135)
<i>N</i>	153,405	153,405	153,405	153,405
Average Log-likelihood	-7.1596	-0.8331	-5.9692	-0.7664
Pseudo- R^2	.1805	.9046	.3168	.9123

Standard errors in parentheses. Premium is biweekly gross premium; “Switch” is an indicator equal to zero for the previous year’s plan, and equal to one for all others; lag market share is the plan’s market share in the individual’s county in the previous year. “Coverage” refers to the FEHBP satisfaction rate for the plan’s coverage of services; we estimate a separate coefficient on this variable for 1999 because the survey question was altered between 1998-1999. Several additional controls were included in the estimation, but to save space are not reported here. (A table listing the full set of coefficients is available from the authors.)

Table 6: Medicare beneficiaries and new enrollees

	Medicare	Non-Medicare	New enrollees	Non-new enrollees
Premium	-0.0152 (0.0006)	-0.0055 (0.0006)	-0.0048 (0.0006)	-0.0158 (0.0006)
Switch	-6.5554 (0.0184)	-6.3284 (0.0286)	—	-6.5330 (0.0154)
Lag Market Share	4.4888 (0.0580)	5.0153 (0.0443)	5.5374 (0.0462)	3.8004 (0.0572)
Satisfaction rate×FFS	0.0240 (0.0022)	0.0092 (0.0017)	0.0037 (0.0019)	0.0325 (0.0021)
Satisfaction rate×HMO	-0.0143 (0.0028)	-0.0071 (0.0019)	-0.0149 (0.0020)	0.0004 (0.0025)
Coverage	0.0712 (0.0040)	0.0411 (0.0031)	0.0220 (0.0032)	0.0924 (0.0036)
Coverage×1999	0.1117 (0.0297)	0.1803 (0.0227)	0.2707 (0.0244)	0.0176 (0.0277)
Public	-0.1182 (0.0583)	0.0007 (0.0379)	-0.0389 (0.0415)	-0.0841 (0.0504)
NCQA Score×Public	0.3436 (0.0310)	0.3823 (0.0212)	0.4459 (0.0237)	0.1995 (0.0282)
NCQA Score×Nonpublic	0.0871 (0.0240)	0.0794 (0.0163)	0.0137 (0.0169)	0.2670 (0.0225)
<i>N</i>	74,156	79,259	31,756	121,649
Average Log-likelihood	-0.6139	-0.9054	-1.9333	-0.4486
Pseudo- R^2	.9495	.8362	.2397	.9567

New enrollees are annuitants whose circumstances required choosing a new health plan in the given year—e.g., because they moved to a new county or because their previous plan was dropped from the FEHBP. All variable definitions are the same as in Table 5. Interactions of annuitant characteristics (age, income, sex) were included in the estimation, but the coefficients are omitted to save space. The reported coefficients on Premium and Switch reflect effects for males with average age and income.

Table 7: Model estimates: including interactions with reporting status

	All enrollees	Non-new enrollees	New enrollees
Premium	-0.0106 (0.0004)	-0.0162 (0.0006)	-0.0047 (0.0006)
Premium×Public	-0.0008 (0.0002)	-0.0007 (0.0003)	-0.0009 (0.0003)
Switch	-6.5346 (0.0154)	-6.5316 (0.0176)	—
Switch×Public	0.1846 (0.0347)	-0.0529 (0.0442)	—
Lag Market Share	4.8392 (0.0403)	3.7365 (0.0634)	5.7200 (0.0545)
Lag Market Share×Public	-0.1469 (0.0934)	0.1887 (0.1656)	-0.6188 (0.1204)
Satisfaction rate×FFS	0.0195 (0.0015)	0.0366 (0.0022)	0.0071 (0.0021)
Satisfaction rate×HMO	-0.0041 (0.0018)	0.0070 (0.0028)	-0.0160 (0.0024)
Satisfaction×HMO×Public	-0.0128 (0.0024)	-0.0197 (0.0039)	-0.0050 (0.0034)
Coverage	0.0485 (0.0025)	0.0961 (0.0037)	0.0110 (0.0033)
Coverage×Public	0.0063 (0.0013)	0.0097 (0.0022)	0.0054 (0.0017)
Coverage×1999	0.0568 (0.0214)	-0.0704 (0.0315)	0.1625 (0.0301)
Coverage×1999×Public	0.3073 (0.0466)	0.4441 (0.0775)	0.2397 (0.0616)
Public	0.3164 (0.1115)	0.4779 (0.1652)	0.0786 (0.1537)
NCQA Score×Public	0.3740 (0.0185)	0.2345 (0.0305)	0.4394 (0.0258)
NCQA Score×Nonpublic	0.0916 (0.0140)	0.2126 (0.0231)	0.0587 (0.0178)
<i>N</i>	153,405	121,649	31,756
Average Log-likelihood	-0.7656	-0.4479	-1.9623
Pseudo- <i>R</i> ²	0.9124	0.9567	0.2425

All variable definitions are the same as in Tables 5 and 6. Standard errors in parentheses.

Table 8: Value of information estimates

Specification	Counterfactual	
	Public scores withheld	Nonpublic scores publicized
Old/new annuitants pooled		
(Table 7, column I)		
Number of choices changed	3,539	2,078
Percent of choices changed	0.69	0.41
Average value of information:		
Choices changed	\$160.37	\$218.46
Overall	\$1.11	\$0.88
Old and new separately		
(Table 7, columns II-III)		
OLD:		
Number of choices changed	2,240	96
Percent of choices changed	0.47	0.02
Average value of information:		
Choices changed	\$105.45	\$10.50
Overall	\$0.49	\$0.002
NEW:		
Number of choices changed	2,552	1,259
Percent of choices changed	8.04	3.97
Average value of information:		
Choices changed	\$588.67	\$644.36
Overall	\$47.31	\$25.55

Figure 1: Distributions of NCQA scores for public and nonpublic plans

