NBER WORKING PAPER SERIES

DO REPORT CARDS TELL CONSUMERS ANYTHING THEY DON'T ALREADY KNOW? THE CASE OF MEDICARE HMOs

Leemore S. Dafny
David Dranove

Working Paper 11420
http://www.nber.org/papers/w11420

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

We thank Tom Hubbard, Ilyana Kuziemko, Mara Lederman, Gautam Gowrisankaran, Phillip Leslie, Mike Mazzeo, Dennis Scanlon, Scott Stern, and participants at various seminars for helpful comments. We are grateful to Laurence Baker, Su Liu, and Robert Town for sharing their data with us, and to Shiko Maruyama and Subramaniam Ramanarayanan for outstanding research assistance. This research was made possible by a grant from The Searle Fund for Policy Research.The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Do Report Cards Tell Consumers Anything They Don't Already Know? The Case of Medicare HMOs

Leemore S. Dafny and David Dranove

NBER Working Paper No. 11420
June 2005, Revised February 2006
JEL No. D8, H4, I1

ABSTRACT

The use of government-mandated report cards to diminish uncertainty about the quality of various products and services is widespread. However, report cards will have little effect if they simply confirm consumers' prior beliefs. Moreover, documented "responses" to report cards may reflect learning about quality that would have occurred in their absence. Using panel data on Medicare HMO market shares between 1994 and 2002, we examine the relationship between enrollment and quality both before and after report cards were mailed to 40 million Medicare beneficiaries in 1999 and 2000. We find evidence for both market-based and report-card-induced learning. We estimate the report-card effect on enrollment in the 2 years following their release to be approximately equal to that of cumulative market learning between 1994 and 2002. The report-card effect is entirely due to beneficiaries' responses to consumer satisfaction scores; other reported quality measures such as the mammography rate did not affect enrollment.

Leemore S. Dafny
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2001
and NBER
l-dafny@kellogg.northwestern.edu

David Dranove
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2001
d-dranove@kellogg.northwestern.edu

Governments devote substantial resources to develop and disseminate quality report cards in a variety of settings, ranging from public schools to restaurants to airlines. The value of these interventions depends on the strength of market-based mechanisms for learning about quality. For example, the value of reports by the Department of Transportation on airline delays and lost luggage will be minimal if consumers can easily learn about performance along these dimensions through word-of-mouth, prior experience, or a scorecard created by a private company.

In this study we quantify the effect of the largest public report-card experiment to date, the release of HMO report cards in 1999 and 2000 to 40 million Medicare enrollees, on the subsequent healthplan choices of enrollees. We compare the magnitude of the learning induced by the report cards to that of ongoing, market-based learning, concluding that both played roughly equal roles in shifting enrollees to higher-quality healthplans during the study period, 1994-2002. We find the market-based learning curve is steepest in markets with private-sector report cards, which provides secondary evidence that report cards are an effective means of disseminating quality information, whether publicly or privately sponsored. Market-based learning is also stronger in areas with stable populations and low rates of prior experience with HMOs, suggesting that "word of mouth" and personal experience also facilitate learning. The report-card effect is wholly driven by enrollees' responses to customer satisfaction ratings; other reported measures did not affect subsequent enrollment.

Our parameter estimates, obtained from a model of healthplan choice, enable us to simulate the effects of market learning and report cards in a variety of scenarios. For example, in a market with three Medicare HMOs, two of which have average quality and one of which scores at the 80^{th} percentile, we estimate the report-card effect on enrollment within 2 years to be nearly as large as the effect of *cumulative* market learning between 1994 and 2002. Approximately 20

percent of enrollees in the average-quality HMOs and 1 percent of enrollees in traditional Medicare switch into the higher-quality HMO following the publication of the report card. These responses reveal a substantial effect of report cards on the information sets of HMO enrollees. They also suggest that quality reporting is unlikely to generate large increases in the HMO penetration rate among Medicare beneficiaries, one of the stated goals of the report-card intervention.

Our study complements recent work on the effects of restaurant hygiene report cards (Jin and Leslie 2003) and HMO report cards (e.g. Chernew et al. 2001, Scanlon et al. 2002, and Beaulieu 2002).[1] One of the primary advantages of the Medicare report-card experiment as compared to the HMO report-card experiments at General Motors (studied by Chernew et al. 2001 and Scanlon et al. 2002) and Harvard University (studied by Beaulieu 2002) is that the decision to release report cards is uncorrelated with the quality scores of any particular plan or set of plans. To the extent that employers who demand and distribute healthplan report cards are also more likely to offer high-quality plans, and healthplans that voluntarily gather and disclose data are more likely to be high-quality providers, consumer responses in prior studies may not be representative of responses in a typical setting. Second, it is much more difficult for Medicare HMOs to "price out" their quality levels, charging more if their scores are higher than competitors' scores. Pricing out will produce downward-biased estimates of the effects of quality data on enrollment unless data on employee-paid premiums (as well as appropriate instruments) are available. Third, the lengthy study period (1994-2002) allows us to distinguish between responses to report cards and market learning. Last, we also have data on *unreported*

[1] There is also a substantial literature on school responses to test-based "accountability" policies that link funding to student performance; see Jacob (2005) for an excellent summary.

quality scores, enabling us to compare enrollee responses to reported and unreported information.

The paper proceeds as follows. Section 1 provides background on Medicare HMOs and the report card mandate imposed as part of the Balanced Budget Act of 1997. Section 2 summarizes prior related research, and section 3 presents the data. Section 4 describes the main analysis and results, and Section 5 discusses extensions and robustness tests. Section 6 concludes.

1 Medicare HMOs and the Report Card Mandate

Although the vast majority of Medicare beneficiaries are enrolled in fee-for-service "traditional Medicare," the option of receiving coverage through participating, privately-managed HMOs has been available since the introduction of Medicare in 1966 (Newhouse 2002). Medicare enrollments in HMOs grew slowly at first, reaching just 1.8 million, or 5 percent of beneficiaries, by 1993. Between 1993 and 1998, enrollment in Medicare HMOs increased threefold, mirroring enrollment patterns among the privately-insured. Figure 1 graphs the HMO penetration rate for Medicare-eligibles and the privately insured between 1993 and 2001. HMO penetration in both populations peaked in 1999-2000 and has declined since.

Although there have been many changes in the statutes governing Medicare HMOs, throughout our study period (1994-2002) several key features remained intact. First, Medicare reimbursed participating HMOs a fixed amount per enrollee which varied by geographic area, gender, age, institutional and work status, and source of eligibility.[2] Second, plans faced statutory

[2] In 2001 and 2002, very small adjustments were also made for enrollees' health status. Between 1982 and 1997, the payment amount was 95 percent of the average cost for a traditional Medicare enrollee of the same age, gender,

limits on the premiums and copayments they could charge enrollees.[3] The result was substantial premium compression from above and below, which constrained the ability of plans to "price out" quality differentials. In every year in our study period, the median enrollee paid no premium at all, and the 75th percentile for monthly premiums ranged between $15 and $35.[4] Third, during the November "open enrollment" period, plans were required to accept new enrollees for the following January. Most plans also accepted enrollees throughout the year, at the start of each month. Enrollees were permitted to switch plans or return to traditional Medicare at the end of every month.

The Balanced Budget Act of 1997 (BBA 1997) required all managed care plans participating in the Medicare program to gather and disclose quality data to the Health Care Financing Agency, now known as The Centers for Medicare and Medicaid Services (CMS). Plans must report a set of standardized performance measures developed by the National Consortium for Quality Assurance (NCQA).[5] These measures are collectively called The Health Plan Employer Data and Information Set (HEDIS®).[6] Beginning in 1998, CMS began supplementing this data by conducting an independent annual survey of Medicare beneficiaries called the Consumer Assessment of Health Plans Study (CAHPS). Respondents are asked a

---

institutional status, and eligibility source, living in the same county. Following the BBA, payment rates were a blend of area costs and national costs (beginning with 90:10 and ending at 50:50 by 2003), subject to a minimum annual increase of 2 percent as well as an absolute floor (Newhouse 2002). CMS began implementing a risk adjustment formula in 2000, with transition to full risk-adjustment delayed to 2007 by the Benefits Improvement and Protection Act of 2000. Between 2001 and 2003, only 10 percent of the payment from the blend/floor formula was adjusted for health status, as determined by the enrollee's "worst principal inpatient diagnosis" to date, if any. As of 2004, CMS began implementing a risk-adjustment formula based on multiple sites of care (Pope et al. 2004, CMS 2004).

[3] The enrollee premium regulations, summarized in Newhouse (2002), limited premiums to "the actuarial value of the cost sharing provisions in traditional Medicare." Moreover, if the combination of Medicare and enrollee contributions exceeded the rate charged to commercial enrollees (adjusted for utilization factors), plans were required to add benefits, reduce premiums, or refund the difference to the government.

[4] Authors' tabulations using data described in Section 3.

[5] NCQA is a private not-for-profit organization whose mission is "to improve healthcare quality everywhere." In addition to collecting, standardizing, and releasing HEDIS data, NCQA uses this information to accredit health plans. Many employers refuse to contract with unaccredited plans.

[6] HEDIS consists of a broad range of measures covering areas such as patient access to care, quality of care as measured by "best practices," provider qualifications, and financial stability.

series of questions designed to assess their satisfaction with various aspects of their healthcare, including the communication skills of their physicians and the ease of obtaining care.

BBA 1997 also required CMS to provide Medicare beneficiaries with information about health plans and the enrollment process in November of each year (Goldstein et al. 2001). HMO quality measures were first published and distributed in November 1999, when Medicare mailed pamphlets entitled *Medicare & You 2000* to all beneficiaries. Both *Medicare & You 2000* and *Medicare & You 2001* (mailed in November 2000) contained selected HEDIS and CAHPS scores for most plans operating in the beneficiary's market area; plans with very low enrollments were exempted from reporting HEDIS data. Figure 2 presents an excerpt of the report card printed on pages 28-35 of the 73-page *Medicare & You 2001* booklet mailed to Illinois eligibles. The editions since 2001 refer readers interested in quality scores to the Medicare website and a toll-free number.

It is important to note that quality data was available via the Medicare website beginning in June 1998, and the toll-free Medicare helpline in March 1999. Due to the low rate of web access among Medicare beneficiaries in 1999 (21 percent according to Medicare surveys), as well as the low rate of requests for managed care information through the helpline, we consider the report card mailing to be the primary source of exposure to the quality data.

For the report cards to have a discernible effect on enrollee behavior, the following chain of events must transpire: (1) beneficiaries must read and comprehend the publications or communicate with someone who has done so; (2) beneficiaries must change their beliefs about plan quality in response to the reported scores; (3) these changes must be of sufficient magnitude to imply a change in the optimal plan for some enrollees; (4) some of these enrollees must take

actions to switch to their optimal plan. The enrollment changes we examine will only reveal the extent to which these requirements were collectively satisfied by *Medicare & You.*

There are several other formal and informal mechanisms for enrollees to learn about the quality of Medicare HMOs, including word of mouth, prior experience in a private-sector HMO offered by the same carrier, current experience in the Medicare HMO, information provided directly by the HMO, and publications of quality measures for a private-sector HMO offered by the same carrier. Some carriers made their HEDIS scores for private-sector enrollees available on NCQA's website. The popular magazine *U.S. News & World Report* published selected scores for all of these plans in their annual "America's Top HMOs" series from 1996-1998.

Of the 16 percent of beneficiaries who reported seeking managed care information in a nationwide survey conducted in 2001, the majority used non-CMS information sources. The most frequent sources cited were the managed care plans themselves, followed by physicians and their staff, and friends and family (Goldstein et al. 2001). These statistics suggest a substantial role for market learning, a hypothesis that is supported by the empirical results.

2 Prior Research

The few empirical papers on market learning focus on the ability of consumers to learn about the quality of so-called "experience goods" through personal experiences. They find rapid learning in markets with low switching costs (e.g. yogurt, Ackerberg 2002), but slower learning when switching costs are high (e.g. auto insurance, Israel 2005). Hubbard (2001) finds evidence that consumers also learn through the aggregate experiences of others: vehicle emissions inspectors with low aggregate failure rates enjoy more business, controlling for consumers' prior experience at these firms.

Recent studies suggest that report cards also facilitate consumer learning. Jin and Leslie (2003) find that restaurants posting an "A" grade enjoyed a 5-percent revenue boost relative to restaurants posting a "B." They find no evidence that revenues responded to changes in actual hygiene scores during the two years before grade cards were introduced. There are at least two reasons to expect more market-based learning about Medicare HMOs as compared to restaurants. First, in a broad class of learning models, learning will occur most rapidly in new markets, and the restaurant market is much more mature. Second, market-based mechanisms that facilitate learning are more likely to evolve in healthcare due to the magnitude of spending involved as well as the private incentives for large, private-sector buyers to assess quality.

A number of recent studies evaluate the impact of healthcare report cards on both consumers and providers. Most focus on the effects of hospital and surgeon report cards for cardiovascular care. These studies find a positive relationship between reported ratings and subsequent changes in market share (e.g. Mukamel and Mushlin 1988; Cutler et al. 2004), as well as an increase in provider selection behavior or "cream-skimming" (Dranove et al. 2003; Werner 2005). Most relevant to our study is the smaller body of research on the impacts of health*plan* report cards.

Early studies of the cross-sectional relationship between market share and quality ratings suffer from an inability to separate the effect of ratings from the effect of omitted variables that are correlated with both the ratings and enrollment (e.g. Chernew and Scanlon 1998). A similar problem prevents Jin (2002) from identifying a causal relationship between voluntarily-disclosed quality ratings and Medicare enrollments from 1993 to 1998.[7] A few recent studies have circumvented these endogeneity problems by focusing on employee responses to employer-

[7]The voluntarily-disclosed quality ratings were based on care provided to private-sector enrollees and were only available to Medicare eligibles through direct inquiries to NCQA and the *U.S. News* articles.

mandated report cards. Although these employers (and hence their choices of plans and employees) are not randomly selected, the release of report card data is arguably an exogenous shock to enrollees' information sets. Using this methodology, researchers find modest increases in the market share of highly-rated plans offered to employees of the federal government (Wedig and Tai-Seale 2002), Harvard University (Beaulieu 2002), and General Motors (Chernew et al. 2001 and Scanlon et al. 2002).[8] However, these increases could be wholly or partially due to market learning. Absent a longer panel of data, it is impossible to know whether enrollees were shifting over time to more highly-rated plans independently of the report card interventions.[9]

Our study complements existing research on consumer responses to healthplan report cards in several respects. We study a nationwide release of report cards to 39.6 million current and potential Medicare HMO enrollees, the largest such intervention to date. Although the results may not generalize to younger populations, the sample is a virtual census of the elderly and disabled, who currently account for nearly one-third of national healthcare expenditures.[10] In addition, the report cards provided quality ratings for over 250 plans, far broader than the range of plans offered by a single employer. The disclosure decision is uncorrelated with the

[8] The report card released to federal employees included six highly-correlated measures of enrollee satisfaction gathered through mailed survey responses. Wedig and Tai-Seale include two of these measure in their models: overall quality of care and plan coverage. The Harvard and GM report cards included HEDIS measures as well as patient satisfaction scores. Beaulieu (2002), Chernew et al. (2001), and Scanlon et al. (2002) use aggregations of all reported scores in logit models of plan choice.

[9] The results in Scanlon et al. (2002) are robust to including a term that reflects switching patterns among unionized employees during the same period (essentially an estimate of market learning). However, the authors note this control group is imperfect because unionized employees were not subjected to the same price changes as the treatment group, and enrollment data is only available at the state level (whereas the unit of observation for the treatment group is the market-coverage tier). In addition, we do not know if enrollment patterns for the two groups were similar prior to the report card intervention. Chernew et al. (2001) consider non-union employees at a Midwest-based Fortune 50 manufacturing company as an alternative control group. However, these employees did not choose from the same set of plans, so this specification can only control for the movement out of PPOs and into HMOs, and not for shifts across HMOs that might have occurred in the absence of report cards.

[10] CMS estimates that total healthcare expenditures for Medicare beneficiaries in 1999 was $385.2 billion, out of an estimated $1.2 trillion for the entire population (*Program Information on Medicare, Medicaid, SCHIP, and other programs of the Centers for Medicare & Medicaid Services*, June 2002; *U.S. Statistical Abstract 2004-2005*, Table 114, "National Health Expenditures.")

scores of any plan or set of plans, which cannot be assumed in the case of report cards released by individual employers. Because Medicare providers are constrained in their ability to pass along or "price out" the costs associated with better quality, this setting is also ideal for isolating the effect of learning. Previous studies have had to disentangle the effects of quality ratings and contemporaneous changes in price, which are typically correlated.[11] Finally, the length of our panel (1994-2002), as well as the availability of unreported quality measures, enables us to distinguish between responses to report cards and continuations of market learning.

3 Data

We use several datasets available online or through direct requests to CMS. We obtain enrollment data from the Medicare Managed Care Quarterly State/County/Plan Data Files for December of each year from 1994 to 2002.[12] Enrollment is available at the plan-county-year level, where "plan" refers to a unique contract number assigned by CMS.[13] Note that carriers may offer several different products within the same plan, such as a benefits package that includes prescription drug coverage and one that does not. Enrollment and benefits data is not available at this level of detail throughout the study period. However, the quality scores in *Medicare and You* were reported at the plan level, so combining enrollment across products within the same plan should not bias the results. Plan-county-year cells with fewer than 10 enrollees are not included in the data. The enrollment files also contain the base CMS payment

[11] In an unregulated market, high-quality sellers may boost their prices, resulting in market shares that do not reflect quality levels. (Compare the quality of a Lexus automobile, which is high relative to the typical passenger car, to its market share, which is very low.)

[12] http://cms.hhs.gov/healthplans/statistics/mpscpt.

[13] CMS assigns unique contract numbers to carriers (e.g. Aetna) for each geographic area they serve. Because these geographic areas are defined by the carriers and areas served by different carriers need not coincide, we follow CMS in considering the county as our market definition.

rate for HMO enrollees in each county, as well as the total number of Medicare eligibles in each county.[14]

The plan-level quality measures included in *Medicare & You 2000* and *2001* were extracted from the Medicare HEDIS files and the Medicare Compare Database.[15] Three measures were reported in each booklet: one from the HEDIS dataset, one from the CAHPS survey (included in the Medicare Compare Database), and the voluntary disenrollment rate.[16] The reported HEDIS measure in both years is *mammography,* the percent of women aged 50-69 who had a mammogram within the past 2 years. The CAHPS measure reported in *Medicare & You 2000* is *communicate,* the percent of enrollees who reported that the doctors in their plan always communicate well. *Medicare & You 2001* replaced *communicate* with *best care*, the percent of enrollees who rated their own care as the "best possible," a rating of 10 out of 10. The reported HEDIS scores were based on data gathered by plans 3 years prior, while the CAHPS scores and disenrollment rates were lagged 2 years. Appendix Table 1 provides detail on the sources and data years for reported scores.

Although *Medicare & You* reports the disenrollment rate for each plan, we do not include this measure in our analyses because it is a lagged component of the dependent variable (enrollment). The three reported scores we match to the enrollment data are therefore *mammography* from 2000 (which is highly correlated with reported 2001 scores),[17] *communicate* from 2000, and *best care* from 2001. To enable comparisons across all measures, the regressions use annual z-scores for each.

---

[14] The base payment rate is county and year-specific, and is adjusted to reflect enrollee characteristics. See footnote 3 for details.

[15] HEDIS data is available at http://cms/hhs.gov/healthplans/HEDIS/HEDISdwn.asp. CAHPS data is available from the Medicare Compare Database at http://www.medicare.gov/Download/DownloadDB.asp.

[16] Involuntary disenrollment is produced by plan exits. Participating plans must accept all Medicare beneficiaries desiring to enroll.

[17] The correlation coefficient for *mammography* reported in 2000 and *mammography* reported in 2001 is .86.

We also obtain from the HEDIS files the three measures that were audited by CMS but not included in the publications: *beta blocker* (the percent of enrollees aged 35+ receiving a beta blocker prescription upon discharge from the hospital after a heart attack), *ambulatory visit* (the percent of enrollees who had an ambulatory or preventive-care visit in the past year), and *diabetic eye exams* (the percent of diabetic enrollees aged 31+ who had a retinal examination in the past year. We use these measures to compute *unreported composite*, which is the average of a plan's z-scores on all three unreported measures.[18]

Most plans report a single set of quality measures pertaining to all of their enrollees. A small number of plans report data separately by submarket, e.g. San Francisco and Sacramento. These submarkets do not correspond to county boundaries, so we create enrollee-weighted average scores by plan in these cases, using enrollment data reported in the HEDIS files. For plans reporting CAHPS data separately by submarket, we create simple averages by plan because the CAHPS files do not include enrollments, and the CAHPS submarkets do not always correspond to the HEDIS submarkets.

Our sample includes plans with quality data for all six measures. Note that the quality data is measured at a single point in time, and it is matched to the panel data on plan enrollments. In Section 5, which addresses robustness, we describe and utilize the limited panel data available for some of the quality measures.

We obtain the minimum monthly enrollee premium for each plan and year from the December Medicare Coordinated Care Plans Monthly Report for 1994-1998, and directly from

[18] The unreported measures were obtained from the same source as *mammography* in 2000, and therefore pertain to data from 1996-97.

CMS for 2000-2002.[19] We estimate 1999 premiums using the average of each plan's 1998 and 2000 premiums, where available. We also construct an indicator variable that takes a value of 1 if a plan had an affiliate that was rated at least once by *U.S. News*. A plan is considered to have such an affiliate if both the Medicare plan and the plan appearing in *U.S. News* had a common carrier (e.g. CIGNA, Humana) and state; Medicare plans were not directly included in the *U.S. News* publications.[20]

Table 1 presents descriptive statistics for the complete plan-county-year dataset. During the study period, HMO enrollment averaged 3,557 per plan-county, or just under 5 percent percent of eligible enrollees in the county. Nearly two-thirds of the observations come from plans whose affiliates were rated by *U.S. News.* Table 2 provides additional detail regarding the number of competitors in each market and the variation in quality scores within markets. For markets with more than one HMO, we calculate the difference between the maximum and minimum reported (and normalized) quality scores in each market, and report the means in Table 2. For example, in markets with 2 competitors, the mean difference in mammography scores is .86 standard deviations. The table reveals substantial variation in quality within markets. Table 3 presents a correlation matrix for the quality scores. *Mammography* is highly correlated with *unreported composite*, but uncorrelated with *communicate* and *best care*, the correlated subjective measures from the CAHPS survey.

[19] The Medicare Coordinated Care Plans Monthly Reports are available at http://www.cms.hhs.gov/healthplans/statistics/monthly/. Many plans offer multiple products with varying benefits and premiums. We follow the literature and select the minimum premium.

[20] When the carrier name did not appear as part of the plan name, carrier identity was obtained by examining names in prior and subsequent years, performing literature searches, and searching the Interstudy database of publicly-reported data on HMOs. We do not incorporate the ratings measures reported by *U.S. News* due to the high number of missing values.

4 Analysis

The enrollment trends depicted in Figure 1 suggest that most enrollees were experiencing Medicare HMOs for the first time during the mid to late 1990s. This was therefore a period during which a great deal of market learning about HMO quality was potentially taking place, which might have caused enrollees to shift toward the highest-quality plans available. If the publication of the Medicare report cards had a separate impact on this learning process, we would expect discrete changes in enrollments following the publication, *controlling for previous trends.*

4.1 Methods

We estimate a discrete choice demand model in which each Medicare enrollee selects the option in her county that offers her the highest utility, including the "outside good" represented by traditional Medicare. As is well-known, the standard assumption of i.i.d. errors in consumer utility produces stringent restrictions on the substitution patterns across options. We estimate a nested model that allows substitution among HMOs to differ from substitution between HMOs and traditional Medicare. Adopting the notation of Cardell (1997), the utility consumer i obtains from selecting plan j in nest g is

$$u_{ij} = x_j\beta + \xi_j + \zeta_{ig} + (1-\sigma)\varepsilon_{ij}.$$

The x_j are observed plan characteristics, ξ_j represents the mean utility to consumers of unobserved plan characteristics, ζ_{ig} is the mean utility to consumer i of products in nest g, and ε is an i.i.d. extreme value random error term.[21] The parameter σ ranges between 0 and 1, with values closer to 1 indicating the within-nest correlation of utility levels is high and values closer

[21] For ease of exposition, we have omitted market subscripts.

to 0 indicating that substitution patterns do not differ across nests. The utility of traditional Medicare, denoted by j=0, is normalized to zero.

As compared to a reduced-form demand equation, the nested logit not only derives from a structural model of choice but also corrects for changes in the choice set, e.g. those caused by entry and exit. This model, which is widely used in the literature on healthplan choice, is particularly appropriate for our analysis because of the frequency of healthplan exit in the post-BBA era. It generates consistent utility parameters that do not depend on the specific competitors in a market. We can then use these parameters to measure the effects of report cards, abstracting away from entry and exit that independently affect enrollment. The model captures both movement across HMOs and movement between traditional Medicare and HMOs.

We follow the methodology introduced in Berry (1994) to estimate β using market share data and linear instrumental variables regressions of the following form:

$$(1) \qquad ln(s_{jc(s)t}) - ln(s_{0c(s)t}) = \sum_{l=1}^{3}\left[\beta_0^l score_j^l * f(year_t) + \beta_1^l score_j^l * post_t^l\right] + \omega_j + \kappa_{c(s)} + \psi_{st} + \sigma\, ln(s_{jc(s)t|gc(s)t}) + \zeta_{jc(s)t},$$

where $s_{jc(s)t}$ denotes absolute market share for plan j in county c (within state s) and year t, and $s_{jc(s)t/gc(s)t}$ denotes plan j's market share among HMO enrollees in county c and year t. ω_j, $\kappa_{c(s)}$, τ_t, and ψ_{st} are vectors of plan, county, year, and state-year fixed effects, respectively.

This specification includes a separate time trend for each score l ($\beta_0^l score_j^l * f(year_t)$) to capture learning about that score over time, as well as interactions between each score and a *post* dummy to capture deviations from trend following the publication of individual scores. For *mammography* and *communicate*, $post_t^l$ takes on a value of 1 beginning in 2000; for *best care*, $post_t^l$ equals 1 beginning in 2001. To determine the functional form for $f(year_t)$, we estimate a specification that interacts each score with dummies for every year (τ_t):

(2) $$ln(s_{jc(s)t}) - ln(s_{0c(s)t}) = \boldsymbol{\delta} score_j^l * \tau_t + \omega_j + \kappa_{c(s)} + \psi_{st} + \sigma ln(s_{jc(s)t|gc(s)t}) + \zeta_{jc(s)t}$$

This specification allows for a flexible learning pattern and easy detection of post-reporting deviations from trend. Unfortunately, the data cannot identify 24 parameters at once (3 scores*8 year dummies), so we use the results from these separate regressions to select $f(year_t)$ in specification (1), and to inform our discussion of the results.

In all specifications, we include plan fixed effects to capture time-invariant differences in the unobservable quality of plans (as perceived by consumers), and county fixed effects to capture time-invariant differences in consumer utility across markets. Such differences can be driven by local demographics, economic conditions, and market structure. For example, HMO penetration in the private sector is larger in urban counties and on the west coast. To the extent that Medicare HMO penetration tracks private sector penetration, county fixed effects will eliminate the time-invariant component of these differences across counties. The county fixed effects also imply that we are examining the relationship between *relative* quality scores within a county and plan market shares in that county. Because changes in national or state economic conditions and regulations may be correlated with quality levels and enrollment decisions, we also include state-year fixed effects. The standard errors are corrected for an arbitrary covariance structure across observations within the same plan and year.

The final term in equations (1) and (2) is the within-group share, i.e. plan j's enrollment in county c divided by total HMO enrollment in county c. As described above, the coefficient on this term (the "nesting parameter") reflects the extent to which HMOs are better substitutes for one another than is traditional Medicare. The within-group share is likely to be correlated with unobservable plan quality, $\xi_{jc(s)t}$. Ideally, we would like instruments that affect the number of competitors in a market (and therefore within-group market share), but are uncorrelated with the

unobservable characteristics of plan j. Some possibilities include the hospital occupancy rate and/or the number of hospitals per capita, as researchers have found that HMO entry is related to the local market structure of the hospital industry. However, the inclusion of county and state-year fixed effects in the regression leaves little variation in these measures to identify the nesting parameter. We therefore use the traditional instruments for this term, namely the characteristics of competing firms (Berry et al. 1995).

Competitor characteristics will be valid instruments if competitors do not *alter* their product characteristics in response to *changes* in plan j's unobserved quality, and if competitors' entry/exit decisions are uncorrelated with *changes* in plan j's unobserved quality.[22] We select product characteristics that are fairly immutable and unlikely to be affected by shocks to competitors' quality levels: indicator variables for not-for-profit ownership, chain membership, and whether the HMO is organized as an "Independent Practice Association (IPA)."[23] These variables are reported annually to CMS and are good individual predictors of $s_{jc(s)t|gc(s)t}$ in separate first-stage regressions, with coefficient estimates (standard errors) of -0.541 (.104), -.361 (.055), and -.515(.082), respectively. The assumption that entry/exit is also unaffected by changes in unobserved quality is tested formally in section 5.3.

4.2 Results

We begin by examining the results from specification (2), which is estimated separately for each reported score. Figure 3 plots the estimated coefficients on the $score_j^l * \tau_t$ interactions; the data

[22] The inclusion of plan fixed effects relaxes the usual assumptions substantially; rather than positing that observable competitor characteristics are uncorrelated with unobservable plan characteristics, we only require *changes* in observable competitor characteristics to be uncorrelated with *changes* in unobservable plan characteristics.
[23] IPA-model HMOs contract with independent physicians and groups of physicians, and they tend to offer a broader network of providers than "staff-model" or "group-model" HMOs, in which physicians are fully or mostly employed by the HMO.

for these figures is presented in Appendix Table 2. The vertical line in each graph signifies the start of the post period for each measure. We draw three conclusions from these graphs. First, mean utility for plans with higher scores is increasing throughout the study period. Second, the only measure that clearly deviates upward from trend during the post-period is *best care*. Prior to the report-card intervention, plans with high *best care* scores were generating more utility over time, but at a decreasing rate. In the first year after *best care* was reported, the effect of *best care* on utility increased more than it had over the three prior years combined. Third, it appears that a log time trend is more appropriate than a linear time trend for modeling the underlying increase in utility for plans with high scores.[24]

The first column in Table 4 presents results from specification (1). This model includes log trends for each reported score as well as interactions between each reported score and *post*. The reported coefficients are estimates of the mean marginal utilities associated with the corresponding scores in different years. Thus the positive coefficients on the trend variables imply that consumers value higher-quality plans more over time (at a decreasing rate). All are approximately the same size and statistically significant at p<.10. (Recall $score_j \sim N(0,1) \forall j$, so the coefficients are comparable). The *post* interactions reveal that plans with high *best care* scores generate even higher utility following the publication of their scores (p<.001). Publicizing the scores for *mammography* and *communicate* does not have a significant impact on utility.[25] The estimate of the nesting parameter, .739 (.056), strongly supports the use of a separate nest for HMOs.

[24]The concave trend is consistent with a learning model in which a decreasing percentage of the population learns each year.

[25]The negative coefficient estimate on *communicate*post*, though statistically insignificant, reflects the concave learning trend for communicate revealed in Figure 3.

Column 2 of Table 4 presents the results with the addition of *unreported composite*ln(year)* and *unreported composite*post* to the specification in column (1). Again, only *best care* deviates significantly from trend in the post-reporting period. Due to the high correlation between *mammography* and *unreported composite*, the learning trends for these scores are not individually significant. Given that the learning coefficients are all of the same magnitude, column 3 replaces the individual trends with a trend for *composite*, which is the average z-score across all 4 measures. This trend is statistically significant at $p<.001$, and the pattern of *post* interactions is unchanged.

The magnitudes of the coefficients reported in Table 4 are not readily interpretable. However, we can use them to simulate how enrollee choices change over time and in response to the publication of report cards. We use the results in column 3 to analyze two scenarios selected to highlight the relative importance of market versus report-card learning.[26] In both scenarios, we abstract away from time trends unrelated to consumer learning (the year and state*year fixed effects) and do not allow for entry and exit. Thus, we focus on how consumer demand changes over time as consumers learn about high-quality plans, either through market sources or the report card intervention. The two scenarios have the following shared features:

1) We suppose that there are three Medicare HMO plans in a hypothetical market throughout the time period 1994-2002.
2) In 1994, all three plans have identical shares of 2.70, which is the market share of the average plan in that year. This implies that the outside good (traditional Medicare) has a market share of 91.90.

[26] For computational methods and other examples of this type of simulation, see Stern (1996) or Lederman (2004).

3) There are no time trends in market share that are unrelated to learning about quality. Thus, if all the HMOs had average quality, their market shares would remain unchanged over time.
4) One of the three plans scores at the top 20^{th} percentile (i.e. the 80^{th} percentile) for all quality measures.

There are no report cards in the first scenario; thus, all consumer responses to quality derive from market learning. In the second scenario, we add the effect of the report cards; specifically, we consider how publicizing the quality scores affects consumer demand beyond what would be predicted by continued market learning.

Table 5 presents the simulation results, which are also graphed in Figure 4. The first three columns in Table 5 report market share trends for the first scenario. These trends correspond to the solid lines in Figure 4. Over time, the market share of the high-quality plan increases by 76 percent, reaching 4.74 percent of the market. Sixty percent of this increase comes from rival HMOs, with the remainder coming from traditional Medicare. The last three columns (dotted lines in Figure 4) add the effects of the report cards. The high-quality plan sees a substantial additional boost in market share. By 2002, its share reaches 6.46 percent. The report-card-related increase of 1.72 is the same order of magnitude as aggregate market learning between 1994 and 2002. Enrollees switching out of rival HMOs again account for roughly 60 percent of this increase.

The simulation illustrates how the report cards could have generated large swings in market shares for individual HMOs. The magnitude of these swings is partially driven by the low HMO penetration among Medicare beneficiaries; only 1.72 percent of all beneficiaries are

predicted to switch due to the report card, but this represents 40 percent of the HMO's market share in 1999 (the year before scores were published).[27]

5 Extensions and Robustness

Our main specification explores the relationship over time between quantity demanded (at a relatively-fixed price) and *reported* quality, which is measured at a single point in time. This model isolates the effect of the report cards by controlling for any movement toward highly-rated plans that might have occurred absent the report cards. The steady increase in consumers' valuation of quality over time is consistent with market learning, but there are other possible explanations. In this section, we describe a series of extensions and robustness checks we perform to evaluate alternative hypotheses.

5.1 Heterogeneity in Learning

We begin by considering the mechanisms through which market learning may be occurring. Our intepretation of the trend coefficients as evidence of market learning will be bolstered if the data are consistent with hypotheses about how this learning takes place. We examine three potential channels for market learning: friends and family (proxied by *stable population*, the share of the 1995 county population still living in the county in 2000); prior HMO experience (proxied by *HMO penetration*, the county Medicare HMO penetration rate in 1994, the start of the study period); and other published report cards (proxied by appearance of affiliated plans in the *U.S. News* "Best HMO" reports).[28] Descriptive statistics for these proxies are included in Table 1. If

[27] This simulation focuses exclusively on consumer reponses to report cards. To the extent the mandate stimulated quality improvements and/or exit by low-quality plans this scenario underestimates the total effect of report cards on the market share of high-quality plans.

[28] County demographic characteristics are from the U.S. Census Bureau and the 2002 *Area Resource File.*

"word of mouth" is a source of learning, we would expect a positive coefficient on the triple interaction term, *composite*ln(year)*stable population*, assuming population stability is correlated with the exchange of information among beneficiaries. If enrollees learn from prior HMO experience, we would expect diminished market learning during the study period and therefore a negative coefficient on *composite*ln(year)*HMO penetration*. Finally, if learning is facilitated by other sources of report-card data, we expect a positive coefficient on the interaction between *composite*ln(year)* and *U.S. News,* an indicator for whether all plans in a county have an affiliated plan that appeared at least once in the *U.S. News* publications.

Table 6 reports the results of adding each of these terms, first separately and then jointly, to the main nested logit specification. The baseline results are repeated in column 1, followed by estimates obtained when adding interactions with the county z-score for *stable population* (column 2), the county z-score for *HMO penetration* (column 3), the *U.S. News* indicator (column 4), and all three together (column 5). Note that main effects for the learning proxies are not needed due to the inclusion of county fixed effects in all specifications.

The data support all three mechanisms, with the strongest evidence for learning facilitated by other report cards. The magnitude of the learning coefficient in markets with complete *U.S. News* coverage is nearly twice that in markets with incomplete or no coverage.[29] A one-standard-deviation increase in *stable population* is associated with an increase of ~40 percent in the learning coefficient, while a one-standard-deviation *de*crease in prior HMO experience is associated with an increase of ~20 percent. The coefficients on the score**post* variables are unaffected by the inclusion of the new interactions.

[29] We do not incorporate the ratings measures reported by *U.S. News* due to the high number of missing values. The correlation between *composite* and the overall *U.S. News* rating (which ranges from 1 to 4 stars) is .64 for plans with data from both sources.

To further examine heterogeneity in market learning and report card effects, we also considered interactions with county-level demographic measures such as the fraction of college graduates and the share of women aged 65-74 (who may be particularly interested in mammography scores of Medicare HMOs). We find no significant relationships between these measures and the pace of learning.

5.2 Specification Checks Using Contemporaneous Quality

Our estimation strategy uses quality measured at a single point in time (1997 for *mammography,* 1998 for *communicate*, and 1999 for *best care.*) Given that these are the data actually reported in *Medicare & You,* panel data on quality throughout the study period is not necessary to examine the response to reported information. However, because enrollees are presumably learning about contemporaneous quality prior to the report-card release, it is useful to estimate our models using the subset of years for which contemporaneous quality data is available. Data on *mammography* is available for 1996-2001, on *communicate* from 1998-1999, and *best care* for 1998-2002.[30] Descriptive statistics for these data are presented in Appendix Table 4. Given the short time-series for *communicate*, we estimate the following specification using only mammography and best care:

$$(3)\quad ln(s_{jc(s)t}) - ln(s_{0c(s)t}) = \gamma_1 mammography_{j,t-1} + \gamma_2 bestcare_{j,t-1} + \phi_1 reported\ mammography_j * post_t^m + \phi_2 reported\ best\ care_j * post_t^b + \omega_j + \kappa_{c(s)} + \psi_{st} + \sigma\, ln(s_{jc(s)t|gc(s)t}) + \zeta_{jc(s)t},$$

As before, $post_t^m$ takes on a value of 1 in 2000-2002, and $post_t^b$ takes on a value of 1 in 2001-2002.[31] The $\hat{\gamma}$ will reflect enrollee responses to changes in contemporaneous plan quality, while

[30] *Unreported composite* is available for 1996-1998.
[31] Note that *reported mammography* and *reported best care* are the measures labeled as *mammography* and *best care* in the main specifications.

the $\hat{\phi}$ will capture responses to reported quality measures. Note that contemporaneous quality is lagged by one year due to the discrete nature of the data available; the earliest that beneficiaries can respond to quality measured during calendar year 1997 is 1998. Given the data limitations, the model can be estimated using only observations from 1999-2002.

The results, reported in Table 7, confirm our main findings. The coefficient estimates on both contemporaneous quality measures are positive and of the same order of magnitude as in the main models, and the *best care* estimate is statistically significant at $p<.05$. The reported value of best care continues to have a large, positive impact on utility in the post-period (2001-2002).[32]

5.3 Plan Benefits and Premiums

The identifying assumption of the main specification is that no omitted, plan-specific, time-varying factor is correlated with both reported quality and enrollments. Apart from changes in contemporaneous quality (addressed above), the most likely candidates for such factors are out-of-pocket premiums and plan benefits. If high-quality plans are more or less likely to increase premiums or benefits over time, the trend variable will reflect these characteristics as well as learning about quality. Similarly, if plans react to high reported scores by raising their premiums ("pricing out" quality) or lowering plan benefits, the *post* interaction terms will be downward-biased.

[32] Because specification 3 is limited to 1999-2002, we use the *mammography* data from 1996-2001 to further confirm that changes in contemporaneous plan quality are not producing the enrollment trend toward highly-rated plans. If plans with high initial quality are more likely to improve their benefits over time, consumers' valuation of these improvements will be captured in the market "learning" term. We therefore regress the change in *mammography* between 1996 and 2001 on *reported mammography* (which is measured in 1997). We obtain a coefficient estimate of 0.08 (0.12), providing little support for this alternative explanation.

Due to the premium compression described earlier, premiums are unlikely to generate much movement in enrollment during our study period. This conjecture is confirmed in column 4 of Table 4, which adds the minimum monthly premium (which varies by plan and year) to the main specification.[33] The coefficient estimates on the quality measures are virtually unchanged, while the effect of premium is small, negative, and statistically insignificant.

To examine the possibility that changes in benefits are biasing the coefficient estimates, we assemble panel data on prescription drug benefits offered by plans. Prescription drugs accounted for one-third of direct out-of-pocket spending by Medicare beneficiaries in 1999, and likely more for beneficiaries without supplemental insurance policies, the primary target market for Medicare HMOs.[34] Town and Liu (2003) estimate that 45 percent of the consumer surplus generated by the Medicare HMO program in 2000 was due to prescription drug coverage provided by (some of) the plans. Unfortunately, we lack a consistent definition of drug coverage during our study period. For 1994-2000, we have an indicator of drug coverage for the "base" option within each plan, provided by Town and Liu.[35] For 2000-2004, CMS provided us with indicators of drug coverage for all options within a plan, but we lack the "base" identifier included in the earlier data.[36] However, the median indicator for each plan in 2000 matches the

[33] Town and Liu (2003) point out that this premium should be expressed relative to the traditional FFS "premium," which can be viewed as the expected out-of-pocket costs associated with achieving the same benefits offered by an HMO while enrolling in traditional FFS Medicare. Town and Liu use Medigap premiums as an estimate of these costs. These premiums are only available at the state-year level, however, so they would not affect the premium coefficient in our models, which include state-year fixed effects.

[34] "Direct" out-of-pocket spending excludes premium payments for Medicare and supplemental insurance policies. (Source: "Program Information on Medicare, Medicaid, SCHIP, and other programs of the Centers for Medicare & Medicaid Services," Office of Research, Development, and Information, June 2002.)

[35] In 1999 and 2000, this indicator varies slightly across counties, so we use the maximum indicator for each plan-year. Town and Liu obtained the 1994-1998 data from the Medicare Coordinated Care Plans Monthly Reports, *ibid.* The 1999-2000 data is from an older version of the Medicare Compare Database and is not currently available online.

[36] We obtained detailed benefits data for all options offered by participating plans in 2000-2004 by direct request to CMS. The base plan is not identified, nor is enrollment data (which might be useful in identifying this plan) included.

base plan indicator in 2000 fairly well (sample mean of .85 vs. .83, respectively), so we use the median indicator for 2001-2004.[37]

Column 5 in Table 4 presents the results from the main specification with the addition of this drug coverage indicator. The coefficient estimate on drug coverage is positive but imprecisely estimated, and the magnitudes of the coefficients of interest are unchanged.[38] Although prescription drug coverage is but one of the unobserved plan characteristics in our models, this analysis suggests that unobserved changes in plan benefits are not driving the results.

We also utilize the panel data on drug benefits to test the assumption that our instruments for a plan's within-group share are uncorrelated with changes in the plan's unobserved quality. We regress each of the three instruments on all of the covariates in the main specification plus the indicator for drug benefits. The coefficients on the drug indicator are uniformly small and statistically insignificant. Again, while drug benefits are only one of many plan characteristics omitted from the main model, this test suggests that the instruments satisfy the exogeneity requirement.

Last, column 6 of Table 4 presents the main results with the addition of the base CMS payment rate, another factor that may affect unobserved, time-varying plan quality. The results are unaffected by the inclusion of this control.

[37] Note that any systematic change in the drug coverage indicator between 2000 and 2001 will be captured by the year dummies.

[38] Our estimated drug coefficient of .023 (.046) is the same order of magnitude as Town and Liu's estimate of .077 (.021).

6 Conclusions

Governments often evaluate the quality of various products and services and publish such information for consumers. The value of these initiatives depends, in part, on the pace at which consumers learn about quality in their absence. The health insurance market, through which nearly 15 percent of GDP flows, is perhaps the most important laboratory for these government initiatives.

Using panel data on Medicare HMOs and a nested logit model of demand, we examine whether and how Medicare enrollees learn about the quality of Medicare HMOs. We arrive at three main conclusions. First, between 1994 and 2002 Medicare enrollees were switching into higher-quality plans *independently* of the government report cards issued in 1999 and 2000, where quality is measured as a composite of the 6 available audited quality scores. This market learning attenuated over time, and was strongest in markets in which *U.S. News* provided report cards, and in which migration and prior HMO experience was relatively low. These findings suggest that market learning is facilitated through the private release of report cards, "word of mouth," and prior experience. The evidence for market learning implies that prior estimates of report card effects are likely biased upward.

Second, after controlling for market learning, we still find a response to the Medicare report cards. The report-card-induced enrollment changes are the same order of magnitude as the changes associated with market learning over the entire 8-year study period. The report-card effect is entirely due to beneficiaries' responses to consumer satisfaction scores; other reported quality measures such as the mammography rate did not affect enrollment. Given that public report cards are often justified on the grounds that individuals' subjective opinions are not good measures of the true quality of health care, it is surprising that satisfaction scores were included

at all, and potentially disconcerting that consumers ignored an alternative, objective measure of quality that was also provided. In our data, enrollee satisfaction is uncorrelated with the mammography rate as well as other measures that are believed to reflect best practices in disease screening and prevention. It can be affected by features that are not instrumental to producing better health, such as large parking lots and nice waiting rooms. These responses also create incentives for plans to maximize satisfaction ratings by directing resources toward "average" enrollees and away from outliers with catastrophic or expensive chronic conditions, precisely the individuals for whom insurance is most valuable.

Third, our estimates suggest that the report cards encouraged a substantial amount of switching among enrollees already in Medicare HMOs, but only drew a small fraction of enrollees in traditional Medicare into Medicare HMOs. This result is consistent with prior research in the private sector (using PPOs as the outside option), and suggests that quality report cards alone will be insufficient to convince Medicare enrollees to abandon traditional Medicare for the Medicare HMO program (currently known as Medicare Advantage).

Evaluating the aggregate welfare effects of government report cards requires estimates of supply-side responses. In the case of the *Medicare & You* report cards, examining plan responses is difficult due to the absence of pre-mandate quality data, as well as simultaneous changes in Medicare payment rates. Setting aside data and identification concerns, a priori there are several reasons to expect a poor response by plans during our study period. First, plans were required to report hundreds of measures, and CMS did not announce which would be publicized to enrollees. As compared to restaurant hygiene inspections, where a final summary grade is posted and the weights on the component scores are known, this report card design limits plans' incentives to improve along measured dimensions. Second, plans may not have anticipated a

significant enrollee response to the quality data, both because of assumptions about enrollee behavior and/or because they underestimated CMS' commitment to disseminating the data. As more recent data becomes available, it will be possible to see whether plans focus disproportionately on improving their scores on measures included in *Medicare & You.* Investigating the extent to which firms "teach to the test" and skimp on unreported quality is an important area for future research, and one of many inputs that will be needed to estimate the welfare effects of measuring and publicly disclosing quality information.

References

Ackerberg, D. A. (2002), "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination," *International Economic Review*, 44:1007-1040.

Beaulieu, N. D. (2002), "Quality Information and Consumer Health Plan Choices," *Journal of Health Economics*, 21:43-63.

Berry, S. T. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal of Economics*, 25(2): 242-262.

Cardell, N.S. (1997), "Variance Components Structures for the Extreme Value and Logistic Distributions," *Econometric Theory*, 13(2): 185-213.

Centers for Medicare and Medicaid Services (2004), "45 Day Notice for 2004 M+C Rates," www.cms.hhs.gov/healthplans/rates/2004/45day-section-a.asp.

Centers for Medicare and Medicaid Services (2002), "Program Information on Medicare, Medicaid, SCHIP, and other programs of the Centers for Medicare & Medicaid Services," http://www.cms.hhs.gov/charts/series/sec3-b3.pdf.

Chernew, M., Gowrisankaran, G., and D. Scanlon (2001), "Learning and the Value of Information: The Case of Healthplan Report Cards," *NBER Working Paper* 8589.

Chernew, M. and D. Scanlon (1998), "Health Plan Report Cards and Insurance Choice," *Inquiry*, 35(1): 9-22.

Cutler, D., Huckman, R., and M. Landrum (2004), "The Role of Information in Medical Markets: An Analysis of Publicly Reported Outcomes in Cardiac Surgery," *American Economic Review*, 94(2): 342-46.

Dranove, D., Kessler, D., McClellan, M., and M. Satterthwaite (2003), "Is More Information Better? The Effects of 'Report Cards' on Health Care Providers," *Journal of Political Economy*, 111(3): 555-588.

Goldstein, E., Teichman, L., Crawley, B., Gaumer, G., Joseph, C., and L. Reardon (2001), "Lessons Learned from the National Medicare & You Education Program," *Health Care Financing Review*, 23(1): 5-20.

Hubbard, Thomas N. (2002), "How Do Consumers Motivate Experts? Reputational Incentives in An Auto Repair Market," *Journal of Law and Economics*, 45: 437-468.

Israel, Mark (2005), "Services as Experience Goods: An Empirical Examination of Consumer Learning in Automobile Insurance," *American Economic Review* (forthcoming).

Jacob, Brian A. (2005), "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools," *Journal of Public Economics* 89: 761-796.

Jin, Ginger Z. (2002), "Did Quality Information Matter? Evidence from Medicare 1993-1998", *mimeo*, University of Maryland.

Jin, Ginger Z. and Phillip Leslie (2003), "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, 118(2): 409–51.

Lederman, Mara (2004), "Do Enhancements to Loyalty Programs Affect Demand? The Impact of International Frequent Flyer Partnerships on Domestic Airline Demand," *mimeo*, University of Toronto.

Mukamel, D. and A. Mushlin (1998), "Quality of Care Information Makes a Difference: An Analysis of Market Share and Price Changes after Publication of the New York State Cardiac Surgery Mortality Reports," *Medical Care* 36:945-54.

Newhouse, Joseph P. (2002), "Medicare," in American Economic Policy in the 1990s, eds., Jeffrey A. Frankel and Peter R. Orszag; Cambridge: MIT Press, 899-955.

Pope, Gregory C.; Kautter, John; Ellis, Randall P.; Ash, Arlene S.; Ayanian, John Z; Iezzoni, Lisa I.; Ingber, Melvin J.; Levy, Jesse M., and John Robst (2004), "Risk Adjustment of the Medicare Capitation Payments Using the CMS-HCC Model," *Health Care Financing Review*, 25(4): 119-141.

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

Stern, Scott (1996), "Product Demand in Pharmaceutical Markets," *mimeo*, Northwestern University.

Stigler, George J. (1961), "The Economics of Information," *Journal of Political Economy*, 61(3): 213-225.

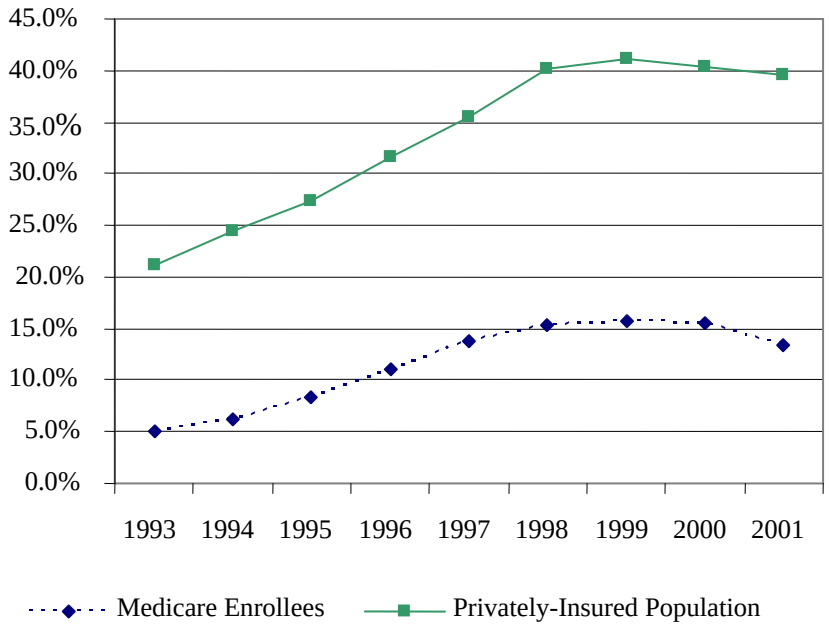
Town, Robert and Su Liu (2003), "The Welfare Impact of Medicare HMOs," *RAND Journal of Economics* 34(4): 719-736.

United States General Accounting Office (1996), "HCFA Should Release Data to Aid Consumers, Prompt Better HMO Performance", Report to Congressional Requesters, October 1996.

Wedig, Gerard J. and M. Tai-Seale (2002), "The Effect of Report Cards on Consumer Choice in the Health Insurance Market," *Journal of Health Economics* 21: 1031-1048.

Werner, Rachel (2005), "The Impact of Quality Report Cards on Health Outcomes and Racial Disparities," mimeo, University of Pennsylvania.

Figure 1. HMO Penetration Rates, 1993-2001



Sources: Author's tabulations from the Medicare Managed Care Quarterly State/County/Plan Data files for December 1993-2001; CMS Statistics, National Enrollment Trends; *U.S. Statistical Abstract*, various years.

Figure 2. Example of Medicare Report Card Appearing in *Medicare & You 2001*

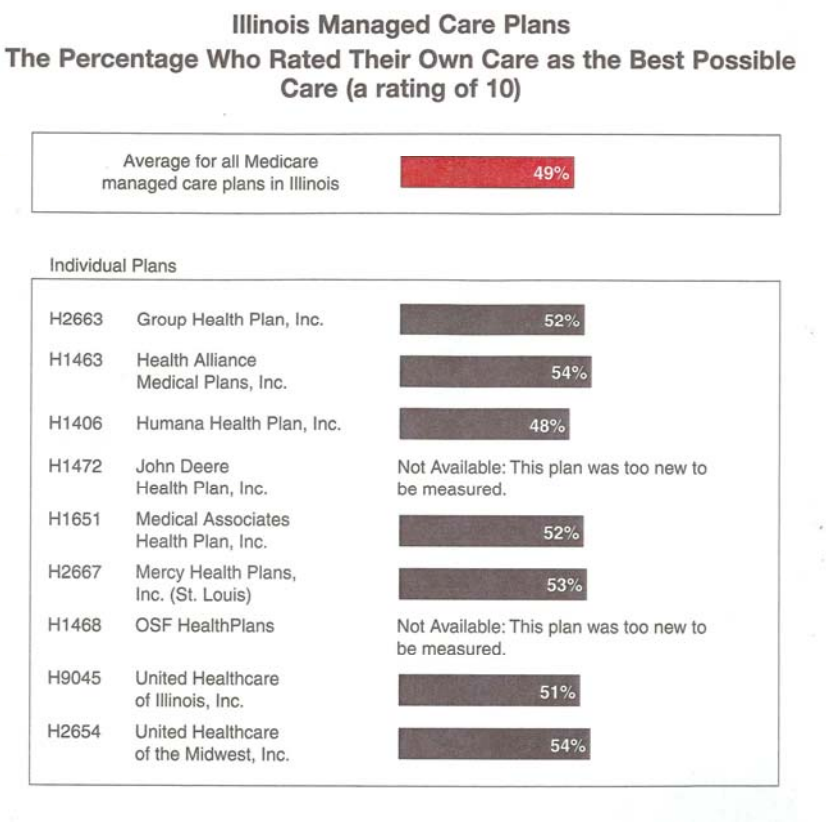
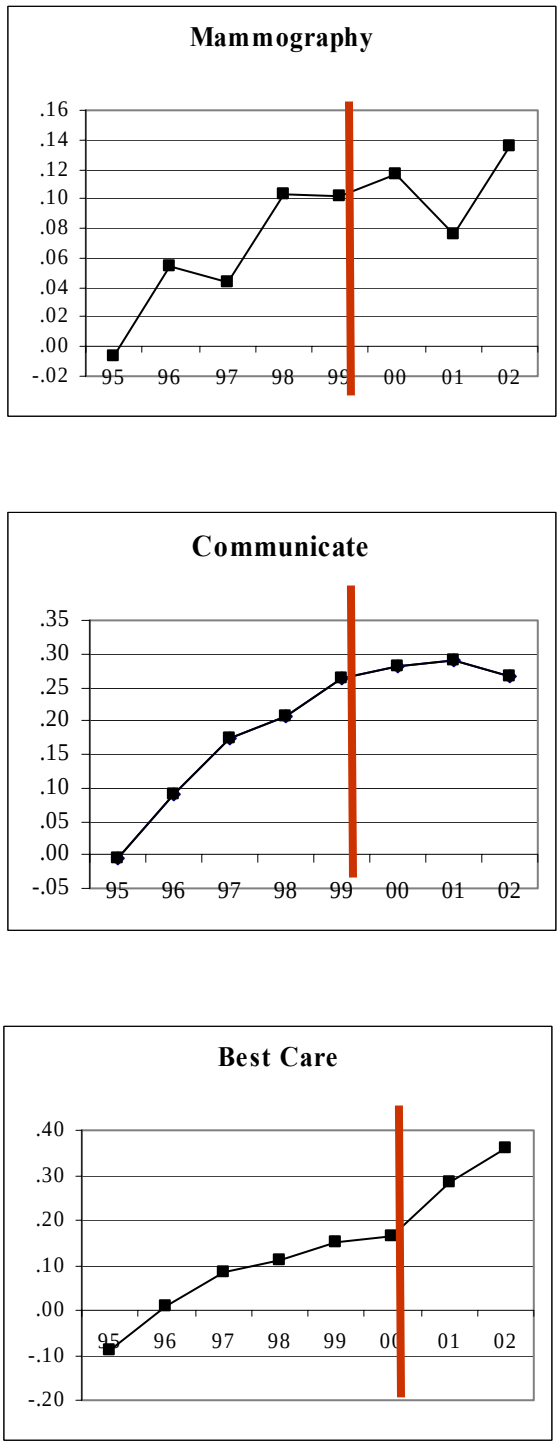
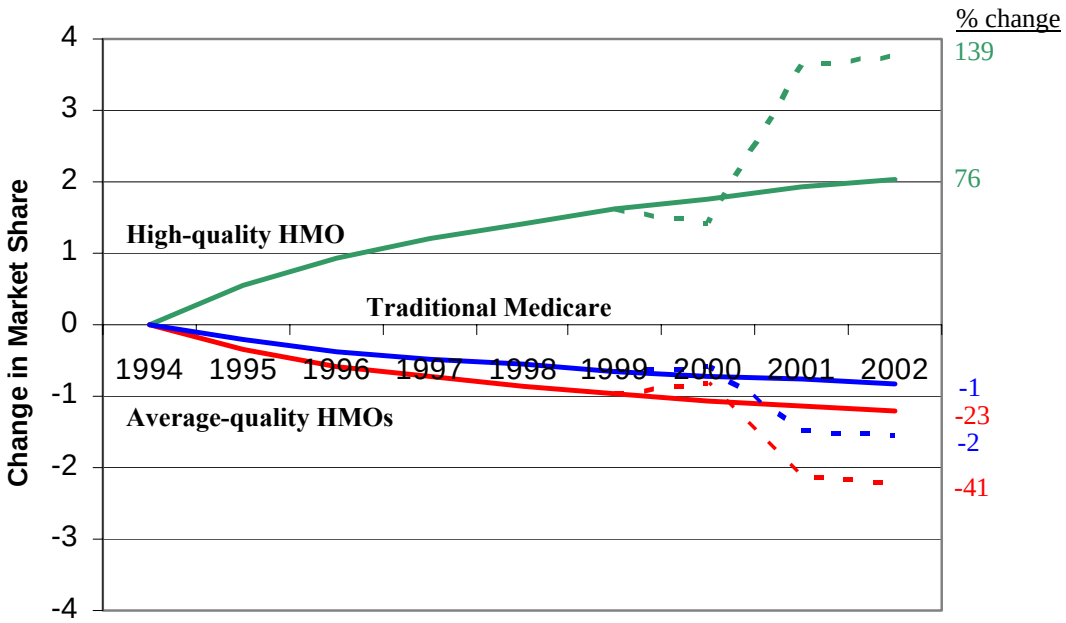


Figure 3. Estimated Coefficients on Score*Year Interactions



Notes: Coefficient estimates from specification (2) in the text. Data is reported in Appendix Table 1.

Figure 4. Effect of Quality on Predicted Plan Market Shares Over Time



Notes: Figure is based on simulation results in Table 5. Solid lines depict share changes associated with market learning only; dashed lines add report-card learning. Share changes for both average-quality HMOs are combined.

Table 1. Descriptive Statistics

	Mean	Std. Dev.
Plan Characteristics		
Enrollment	3,557	9,125
Share of county eligibles	4.58	6.59
Share of county HMO enrollment	49.98	41.66
Reported Quality Measures		
Mammography	75.69	7.41
Communicate	69.95	4.91
Best Care	49.53	6.22
Unreported Quality Measures		
Betablocker	81.20	12.69
Diabetic Eye Exams	59.06	12.75
Ambulatory Visit	88.73	8.03
Monthly premium ($)	16.56	24.73
Affiliate in *U.S. News*	66.61	47.16
CMS monthly payment rate ($)	495.42	96.62
Prescription drug coverage	70.17	45.40
Market Characteristics		
Number of rivals	2.24	2.36
Number of rivals belonging to a chain	1.70	1.97
Number of not-for-profit rivals	1.08	1.29
Number of IPA rivals	1.18	1.53
Stable population share (1995 to 2000)	79.71	9.40
HMO penetration rate (1994)	8.68	11.89
Percent of population aged 65-74 (2000)	7.10	2.22
Percent with college degree (2000)	15.78	6.51

Notes: N=8230 (8216 for monthly premium and prescription drug coverage). The unit of observation is the plan-county-year. Sample includes observations with 10 or more Medicare enrollees and nonmissing data for all quality measures. All data is reported in percentages unless otherwise indicated. Quality measures correspond to data reported in *Medicare & You 2000* (2001 for *best care*). *Stable population share* is the share of a county's 1995 population still living in the county in 2000.

Table 2. Market Characteristics

Number of Plans	Number of Markets	Mean of (Max-Min) z-scores Mammography	Communicate	Best care
1	240	-	-	-
2	114	.86	.59	.87
3+	117	1.48	1.49	1.64
Total	471	1.17	1.04	1.26

Notes: Sample includes all markets (=counties) in 2000. Quality measures correspond to data reported in *Medicare & You 2000* (2001 for *best care*).

Table 3. Correlation Matrix for Quality Scores

	Mammography	Communicate	Best Care	Unreported Composite
Mammography	1.00			
Communicate	0.10	1.00		
Best Care	0.02	0.82	1.00	
Unreported composite	0.73	0.17	0.05	1.00

Notes: Sample includes all markets (=counties) in 2000. Quality measures correspond to data reported in *Medicare & You 2000* (2001 for *best care*).

Table 4. Effect of Quality on Mean Plan Utility

Dependent Variable	Ln(HMO share of eligibles/traditional Medicare share of eligibles)					
	(1)	(2)	(3)	(4)	(5)	(6)
Market Learning						
Mammography*ln(year)	0.093*	0.059				
	(0.041)	(0.056)				
Communicate*ln(year)	0.107*	0.090†				
	(0.050)	(0.051)				
Best care*ln(year)	0.078†	0.094†				
	(0.049)	(0.049)				
Unrep. composite*ln(year)		0.070				
		(0.081)				
Composite*ln(year)			0.304***	0.314***	0.312***	0.301***
			(0.078)	(0.080)	(0.081)	(0.078)
Report Card Effect						
Mammography*post	0.009	-0.017	-0.009	-0.006	-0.003	-0.012
	(0.041)	(0.064)	(0.049)	(0.049)	(0.049)	(0.048)
Communicate*post	-0.070	-0.069	-0.023	-0.023	-0.017	-0.026
	(0.045)	(0.044)	(0.042)	(0.042)	(0.041)	(0.041)
Best care*post	0.174***	0.175***	0.201***	0.204***	0.194***	0.193***
	(0.054)	(0.054)	(0.051)	(0.053)	(0.050)	(0.051)
Unreported composite*post		0.060	-0.013	-0.013	-0.022	-0.014
		(0.098)	(0.081)	(0.082)	(0.081)	(0.079)
Price						
Monthly premium				-0.001		
				(0.001)		
Benefits						
Prescription drug coverage					0.023	
					(0.046)	
Medicare payment						
Base county rate						0.001
						(0.001)
Nesting parameter						
Ln(share of HMO enrollmen	0.739***	0.740***	0.745***	0.741***	0.740***	0.757***
	(0.056)	(0.056)	(0.055)	(0.055)	(0.055)	(0.055)
N	8230	8230	8230	8216	8216	8230
Adjusted R-squared	0.92	0.92	0.92	0.92	0.92	0.92

Notes: All specifications include plan, county, and state-year fixed effects, and are estimated by 2SLS, with the nesting parameter as the endogenous variable. Z-scores are used for all quality measures (*composite, mammography, communicate, best care, and unreported composite*). *Post* is an indicator variable that takes on a value of one beginning in 2000 (2001 for *best care* interactions). Robust standard errors clustered by plan-year are in parentheses.
† denotes p<.10, * denotes p<.05, ** denotes p<.01, *** denotes p<.001

Table 5. Effect of Quality on Predicted Plan Market Shares Over Time

	Market Learning			*Market Learning And Report Card*		
Year	High-quality HMO	Avg-quality HMO	Traditional Medicare	High-quality HMO	Other HMOs	Traditional Medicare
1994	2.70	2.70	91.90	2.70	2.70	91.90
1995	3.27	2.53	91.68	3.27	2.53	91.68
1996	3.64	2.41	91.54	3.64	2.41	91.54
1997	3.91	2.33	91.43	3.91	2.33	91.43
1998	4.13	2.27	91.33	4.13	2.27	91.33
1999	4.32	2.21	91.26	4.32	2.21	91.26
2000	4.48	2.17	91.19	4.11	2.28	91.33
2001	4.62	2.13	91.13	6.32	1.64	90.41
2002	4.74	2.09	91.07	6.46	1.60	90.33
Aggregate Change	2.04	-0.61	-0.83	3.76	-1.10	-1.57
Percentage Change	75.7%	-22.5%	-0.9%	139.2%	-40.6%	-1.7%

Notes: Predictions based on hypothetical county containing 3 Medicare HMOs with the national average market share in 1994. The high-quality HMO scores at the 80th percentile for all quality measures, while the two remaining HMOs have mean quality scores. Note the results for both average-quality HMOs are identical, so only one column is presented above. Market shares are calculated using coefficient estimates from column 3, Table 4. State-year effects are not included.

Table 6. Heterogeneity in Effects of Quality on Mean Plan Utility

Dependent Variable	Ln(HMO share of eligibles/traditional Medicare share of eligibles)				
	(1)	(2)	(3)	(4)	(5)
Market Learning					
Composite*ln(year)	0.304***	0.285***	0.350***	0.236**	0.265***
	(0.078)	(0.074)	(0.080)	(0.077)	(0.080)
Composite*ln(year)*stable population		0.114***			0.123***
		(0.028)			(0.027)
Composite*ln(year)*HMO penetration			-0.060**		-0.051*
			(0.023)		(0.022)
Composite*ln(year)**U.S. News*				0.208**	0.178*
				(0.072)	(0.072)
Report Card Effect					
Mammography*post	-0.009	-0.000	-0.001	-0.009	0.007
	(0.049)	(0.048)	(0.048)	(0.050)	(0.048)
Communicate*post	-0.023	-0.027	-0.025	-0.029	-0.034
	(0.042)	(0.041)	(0.041)	(0.042)	(0.040)
Best care*post	0.201***	0.197***	0.204***	0.199***	0.197***
	(0.051)	(0.050)	(0.051)	(0.051)	(0.051)
Unreported composite*post	-0.013	-0.021	-0.019	-0.021	-0.033
	(0.081)	(0.079)	(0.079)	(0.081)	(0.077)
Nesting parameter					
Ln(share of HMO enrollment)	0.745***	0.748***	0.744***	0.739***	0.743***
	(0.055)	(0.054)	(0.055)	(0.056)	(0.055)
N	8230	8230	8230	8230	8230
Adjusted R-squared	0.92	0.92	0.92	0.92	0.93

Notes: All specifications include plan, county, and state-year fixed effects, and are estimated by 2SLS, with the nesting parameter as the endogenous variable. Z-scores are used for all quality measures (*composite, mammography, communicate,* and *best care*), *stable population* and *HMO penetration*. *U.S. News* takes on a value of 1 if all plans in a county-year have affiliates that appeared in the *U.S. News* "Best HMOs" articles at least once, and 0 otherwise. *Post* is an indicator variable that takes on a value of one beginning in 2000 (2001 for *best care* interactions). Robust standard errors clustered by plan-year are in parentheses.

† denotes p<.10, * denotes p<.05, ** denotes p<.01, *** denotes p<.001

Table 7. Effect of Contemporaneous and Reported Quality on Mean Plan Utility

Dependent Variable	Ln(HMO share of eligibles/ traditional Medicare share of eligibles)
Market Learning	
(Lagged) contemporaneous mammography	0.050
	(.048)
(Lagged) contemporaneous best care	0.117*
	(0.046)
Report Card Effect	
(Reported) mammography*post	0.035
	(0.031)
(Reported) best care*post	0.124**
	(0.042)
Nesting Parameter	
Ln(share of HMO enrollment)	1.045***
	(0.063)
R-squared	.96
N	3409

Notes: This specification includes plan, county, and state-year fixed effects, and is estimated by 2SLS, with the nesting parameter as the endogenous variable. Z-scores are used for all quality measures. *Post* is an indicator variable that takes on a value of one beginning in 2000 for *mammography*, and 2001 for *best care*. Robust standard errors clustered by plan-year are in parentheses.
† denotes p<.10, * denotes p<.05, ** denotes p<.01, *** denotes p<.001

Appendix Table 1. Scores Reported in *Medicare & You*

Measure	Description	Data Years (Sources) for Scores Reported in 2000	Data Years (Sources) for Scores Reported in 2001
mammography	% of women 50-69 receiving a mammogram within past 2 years	1996-1997 (1998 HEDIS file)	1997-1998 (1999 HEDIS file)
communicate	% enrollees reporting the doctors in their plan "always communicate well"	1998 (2000 Medicare Compare Database)	Not reported
best care	% enrollees rating their own care a 10 out of 10	Not reported	1999 (2001 Medicare Compare Database)
disenrollment	% enrollees who voluntarily disenrolled	1998 (2001 Medicare Compare Database	1999 (2001 Medicare Compare Database)

Note: Measures matched to enrollment data are shaded in gray.

Appendix Table 2. Effect of Quality on Mean Plan Utility, *Specification 2*

Dependent Variable	Ln(HMO share of eligibles/trad. Medicare share of eligibles)		
Score	Mammography	Communicate	Best Care
Score*1995	-.007	-.006	-.088
	(.095)	(.080)	(.086)
Score*1996	.055	.091	.009
	(.077)	(.079)	(.081)
Score*1997	.044	.174*	.086
	(.074)	(.076)	(.079)
Score*1998	.103	.208**	.113
	(.073)	(.074)	(.077)
Score*1999	.102	.264***	.153
	(.076)	(.079)	(.078)
Score*2000	.117	.281***	.163*
	(.076)	(.081)	(.077)
Score*2001	.076	.291**	.286***
	(.090)	(.088)	(.085)
Score*2002	.135	.267*	.358***
	(.109)	(.107)	(.107)
Nesting parameter			
Ln(share of HMO enrollment)	.724***	.721***	.739***
	(.061)	(.060)	(.056)
N	8830	8230	8230
R-squared	.92	.92	.92

Notes: All specifications include plan, county, and state-year fixed effects, and are estimated by 2SLS, with the nesting parameter as the endogenous variable. Z-scores are used for all scores. Robust standard errors clustered by plan-year are in parentheses.
† denotes p<.10, * denotes p<.05, ** denotes p<.01, *** denotes p<.001

Appendix Table 3. Panel Quality Data

	1998	1999	2000	2001
Mammography	76.47	75.94	76.34	76.67
	(6.15)	(6.58)	(6.58)	(5.68)
Best Care	50.21	49.60	48.30	43.31
	(7.33)	(6.22)	(6.01)	(6.37)
N (number of plan-counties)	1124	1024	870	748

Notes: The unit of observation is the plan-county-year. Sample includes observations with 10 or more Medicare enrollees and nonmissing data for all *reported* quality measures. All data is reported in percentages unless otherwise indicated. Standard deviations are in parentheses.