Medicare Advantage Plan Quality Reporting: The Relationship Between Subjective and Objective Measures of Quality.

Michael Darden*

Very Preliminary and Incomplete

September 18, 2009

Abstract

The extent to which "report cards" improve overall welfare has been recently debated by economists and policy makers. Some have argued that report cards may alleviate information imperfections, common in both health services and health insurance markets, and thus improve market outcomes. Using data on Medicare enrollees and Medicare Advantage plan quality scores, this study enters the debate by investigating the correlation between subjective and objective measures of insurance plan quality. This paper also explores the extent to which these quality measures change plan enrollment patterns and alter insurance provider incentives. Preliminary findings suggest almost no correlation or negative correlation between subjective and objective measures of insurance plan quality. Furthermore, Medicare enrollees appear to respond to objective reports of quality only in areas with higher than average levels of educational attainment. Finally, the value of information, as measured by the variation in quality scores with a market area, appears to be increasing over time.

*Email:mdarden@email.unc.edu

Introduction

The well-known information imperfections of the health sector are argued to contribute to market inefficiencies with respect to the allocation of health care. The focus of this paper is on the specific quality information asymmetry faced by the Medicare enrollee in selecting a utility maximizing form of Medicare. How elderly consumers of health care learn about quality and, following learning, select a form of health care should be of interest to policy makers. Indeed, governments release quality "report cards" on numerous topics in an attempt to educate the public and alleviate information asymmetries. Governments thus face a fundamental question with respect to information: What is the most effective way to convey complex information to consumers such that, in the case of Medicare, consumers can select a utility maximizing form of care under the constraint of cognitive learning costs?

This paper considers the quality scores of Medicare Advantage plans and enters the debate by posing three main questions. First, when reporting quality information to consumers, what is the relationship between subjective and objective measures of quality in terms of both effectiveness in conveying information and of changing provider incentives? If subjective measures of quality simply reflect niceties enjoyed by the average consumer (e.g. better waiting rooms, closer parking lots, etc.) and if insurance benefits most those in the right tail of the expenditure distribution, overall welfare may diminish if plan market share increases dramatically in these scores. Second, what types of firms focus disproportionately on improving subjective versus objective measures of quality given the publication of scores? Third, what characteristics of individuals influence the response to subjective versus objective measures? Given recent findings that consumers only internalize subjective quality measures (Dafny and Dranove (2008)), this paper tests whether, for example, market areas with large populations of elderly women to respond more intensely to plans with high mammography scores.

Preliminary findings suggest that subjective measures of quality often do not correlation with objective counterparts. For example, subjective 2001 survey results of a plan's "Overall quality" on a 10 point scale negatively correlate with a plan's rate of diabetic eye exams, CABG procedures for men over the age of 85, and male and female hip replacements. Furthermore, the highest positive correlation between the above subjective measure and *any* objective score was 0.34. Also of note is the finding that while more highly educated markets (counties with high proportions of college graduates) do respond more to some objective measures of quality, markets with high proportions of elderly women do not appear to respond to high mammography rate scoring plans. While the evidence presented here is only suggestive at this stage of research, the findings do motivate further work.

I Background

I.1 Medicare Advantage Plans

Since the inception of Medicare in 1966, enrollees have had the option of traditional fee-for-service (FFS) Medicare.[1] With the rise of managed care agencies in the 1980s and 1990s, the choice set of enrollees with respect to health provision expanded greatly. In 1993, 1.8 million Medicare eligible consumers, roughly 5% of the total Medicare eligible population, were enrolled in a Medicare HMO. During the 1990s, that penetration rate increased considerably to a peak of around 16% in 1999. As of 2005 that percentage was roughly 12%.[2] As discussed below, much of the health plan learning literature suggests that the period of 1994-2005 was a period of significant potential for consumer learning, market-based or otherwise, with respect to plan quality.

Medicare Advantage (or Medicare Part C), allows enrollees to select an approved privately operated managed care plan, most notably a Medicare Health Maintenance Organization (HMO). Managed care plans wrap together the financing and provision of health care. Medicare reimburses Part C HMOs a fixed amount per enrollee that depends on age, gender, geographic location, and work status, among other factors. Plan enrollees are required to pay the supplemental Medicare part B insurance premium[3]. Medicare HMOs are allowed to charge additional premiums, although these premiums are highly regulated. Plans therefore compete on the basis of both quality and price. Proponents of the managed competition model of health care argue that price and quality competition lead to efficient market outcomes. Scanlon *et al.* (2002) note, "A cornerstone of the managed competition model is providing consumers with sufficient information about quality." Indeed, in the private market for health plans, consumers without an accurate quality assessment may use only price to evaluate the relative quality of health plans. Thus, in the absence of quality report cards, the much studied price elasticity of demand for health insurance may be smaller (Wedig and Tai-Seale (2002)). Royalty and Solomon (1999) show that price competition can only improve the efficiency of market outcomes with respect to health plans if consumers are well informed of plan quality.

[1]If available, Medicare enrollees have also had the option of privately operated managed care plans since 1966; however, Medicare managed care plan enrollees were initially a very small portion of overall Medicare Newhouse (2002).

[2]Author's tabulations from Medicare State/County/Plan enrollment files.

[3]$96.40/month as of 2009

I.2 Health Plan Choice

Scanlon *et al.* (2002) nicely summarizes the broad literature in health plan choice. The authors partition the determinants of choice into primary and secondary determinants. Primary determinants include specific plan characteristics: price (if applicable), quality, provider choice, benefits and coverage choice, plan convenience, etc. Secondary determinants are those that directly affect any of the primary determinants: demographic characteristics, health state, underlying economic conditions, etc. Much of the literature focuses on the primary determinants. Nearly all studies have found the expected inverse relationship between plan enrollment and price (Scanlon *et al.* (2002)). Estimates in the literature of the price elasticity of demand for health insurance range from 0.1 to 0.8 (Wedig and Tai-Seale (2002)). Estimates of the enrollment response to quality have been varied as quality is more difficult to measure. Chernew and Scanlon (1998) find that many quality determinants of plan choice are unobserved to the econometrician (plan processing time, ease of service, etc.). Controlling for unobservable quality determinants will prove important later in the paper.

Secondary factors can easily be incorporated into empirical frameworks using interaction terms. While fewer studies have concentrated on the secondary factors, Beaulieu (2002) studies secondary factors in the choice of employer provided health insurance plans. She hypothesizes that older workers may be less affected by changes in relative quality due to strong physician/provider relationships. Dafny and Dranove (2008) also suggest that older patients may also be less likely to learn from report cards and more likely to learn from "prior experience." The opposite argument may hold for new employees. New employees with the choice of employer provided health insurance may benefit greatly from report cards because new employees do not have the needed prior experience to make optimizing decisions. Finally, Beaulieu (2002) hypothesize that workers who provide the health insurance for their families may be less likely to respond to relative quality changes due to the high switching costs associated with changing family plans.

I.3 Learning and Plan Quality

How consumers learn about health plan quality can be partitioned into two main types: market-based and government sponsored. Report cards span both types of learning. An example of government sponsored report cards came with the 1997 Balanced Budget Act, requiring all Medicare managed care plans to report certain standardized quality measures to the Center for Medicare and Medicaid Studies (CMS). Standardized quality scores were developed by the Nation Consortium for Quality Assurance (NCQA). Quality scores were then distributed to all Medicare eligible consumers.

Market based learning mechanisms can be more broadly defined. Quality reports and ordinal rankings such as those published by the U.S. News and World Report are private market report cards that can be quite influential in consumer choice. Other types of market learning include "word-of-mouth" learning and prior experience learning. Being difficult to measure, few studies on health plan learning have considered such measures.[4]

Turning to the internalization of quality scores, consumers have often been thought to use "rules of thumb" when internalizing data. Pope (2006) tests whether consumers respond to U.S. News' rank changes in hospitals while controlling for the more complex (and also reported) underlying quality scores. He finds that consumers focus overwhelmingly on the ordinal ranks relative to the continuous measures. This result is common in the literature and report card distributors therefore often aggregate complex information into more simplistic rankings or ratings.

In an older work, O'Brien (1981) shows that there often exists a great deal of distortion when simplifying an underlying continuous distribution into categories. As an example, consider a consumer in the 2nd ranked health plan within their health market. Assume that the continuous quality score difference between the 2nd and 1st ranked plan is tiny. Our consumer may be induced to switch plans despite the fact that the switching cost could outweigh the benefit of the 1st relative to the 2nd plan. On the hand however, Israel (2005) finds that consumers are often reluctant to switch auto insurance companies even when faced with higher quality/lower price alternatives. Switching costs in health insurance are very probably as high or higher than those in auto insurance. Thus, governments, and more generally report card distributors, are faced with the trade-off between accessibility and accuracy when publishing report cards.

Historically the main empirical problem with measuring the effectiveness of report cards has the existence of unobservable (to the econometrician) quality measures that affect both a plan's rating and enrollment. Pope (2006) states, "a positive association between rank changes and consumer behavior will result if changes in rank simply confirm what consumers have already learned as opposed to providing new information." The endogenous nature of report card ratings has therefore prevented cross-sectional studies from identifying the causal relationship between ratings and enrollment.

Because of the endogeneity problem inherent in report card studies, much of the literature has focused on specific "natural" experiments within companies (e.g. General Motors, Scanlon *et al.* (2002)), universities (Harvard, Beaulieu (2002)), and federal employees Wedig and Tai-Seale (2002). Most of these studies cannot be generalized to a wider population. A majority of these studies have however found that managed care plan enrollment is positively

[4]The notable exception being Dafny and Dranove (2008)

related to high quality scores in report cards. The results are however far from uniform. Furthermore, most studies have not considered the possibility of other forms of learning. Federal employees, for example, may have a large amount of prior experience with health plan selection. Similarly, General Motors employees may select health plans based on communication with other employees. Report cards may simply reaffirm what workers already know.

Dafny and Dranove (2008) enters the literature most closely to the work of Scanlon *et al.* (2002) and Pope (2006). DD exploit the publication of the largest natural experiment with respect to report cards to date. Because the distribution of quality reports of health plans was due to the Balanced Budget Act of 1997, it can be justifiably argued that the decision to distribute scores to Medicare beneficiaries was uncorrelated with the quality score of any particular plan or type of plan. Furthermore, premiums and co-payments charged by Medicare HMOs are regulated as a function of the traditional Medicare provisions. Plans could therefore not,"price out" their quality differentials Newhouse (2002), Dafny and Dranove (2008). Dafny and Dranove (2008) examine only Medicare HMOs enrollment relative to traditional Medicare. With the inability to price out quality differentials, DD isolate how consumers learn about plan quality. Controlling for market forms of learning, Dafny and Dranove (2008) find evidence of learning, as measured by the within market Medicare Managed Care plan share of enrollees, from the Government published report cards. Learning was found to be greatest in markets with the highest variation in quality. Finally, Dafny and Dranove (2008) find that individuals responded only to subjective measures of quality.

While whether consumers pay attention to report cards can be debated, firms clearly take note. 16 out 17 of the top ranked hospitals in U.S. News have advertised their standing Pope (2006). Pope estimates that up to 750 million has been shifted from lower ranked hospitals to higher ranked hospitals from 1996-2006 as a result of U.S. News report cards. The incentive to maximize rank for a firm is clear. What is unclear is whether the quality "improvements" of firms are welfare improving. If quality information in report cards is presented in such a way that the cognitive costs of internalizing the information are too great for the average consumer, then report cards might not lead to any significant change in consumer behavior. Furthermore, if firms realize the importance specific quality measures (subjective or otherwise), the incentive structure maybe such that firms focus solely on a subset of quality measures, neglecting others. Firms may, in effect "teach-to-the-test". In a seminal 2003 paper, Dranove *et al.* (2005) argue that report cards on the surgeon and hospital mortality and morbidity rates associated with coronary artery bypass graft (CABG) surgery had a negative overall effect on patient welfare in New York and Pennsylvania in the early 1990s. They show that, rather than improving overall CABG surgery quality, hospitals simply limited access to CABG surgery for high-risk patients. The authors used a difference-in-differences approach to capture the effect of health report card implementation in New York (1991) and Pennsylvania (1993) relative to other states. Their study period ends however in 1994. It is reasonable to assume that improvements in the overall quality of CABG surgery would be more of a long run affect. An interesting study would measure the overall costs and benefits of CABG report cards over a longer period of time.

To summarize, Wedig and Tai-Seale (2002) conclude that report cards do indeed influence consumer choice; however, market efficiency is improved only if it can be shown that report cards provide valid measures of quality of care.

II Theoretical Motivation

Following Berry (1994), I now outline a simple discrete choice model of demand for Medicare services. I follow CMS and define the a county as the market area. Individuals in the model may choose among traditional FFS Medicare or any of the available Medicare Advantage plans in their market. Omitting both the market and time subscripts (to be added again later) for notational ease, consider a one-period model where consumer i purchases one unit of product j for $j \in \{0, 1, \ldots, J\}$. Define the utility associated with choice j as:

$$u_{ij} = x_j' \beta_i + \alpha p_j + \mu_j + \epsilon_{ij} \tag{1}$$

- x_j: A $kx1$ vector of observed product characteristics of j.
- p_j: The premium of plan j.
- μ_j: The mean, across consumers, utility valuation of some unobserved plan characteristic.
- ϵ_{ij}: An iid random error.

Given the random coefficients specification, we can decompose the vector $\tilde{\beta}_i$ as follows:

$$\tilde{\beta}_i = \beta + \sigma \zeta_i' \tag{2}$$

- β: The mean taste parameters across consumers $kx1$
- ζ_i: $kx1$ vector of iid shocks, each with a standard normal distribution

Distributing through we have:

$$u_{ij} = x_j' \beta + \mu_j + \alpha p_j + \nu_{ij} \tag{3}$$

Where:

$$\nu_{ij} = x_j' \sigma \zeta \tag{4}$$

which is clearly heteroskedastic. If we define mean utility as

$$\delta_j = x_j \beta + \alpha p_j + \mu_j \tag{5}$$

then we can rewrite the utility associated with choice j as:

$$u_{ij} = \delta_j + \nu_{ij} \tag{6}$$

An individual will choose choice j iff $u_{ij} \geq u_{ik}$ $\forall k \in \{0, 1, \ldots, J\}$ for $k \neq j$. Given some distribution for the joint unobservables, we can derive market share equations for plan j, s_j from the underlying utility specification. If we assume for the moment that $\beta_i = \beta$ $\forall i$, and if ϵ_{ij} is distributed extreme value type 1, then we have familiar logit equation for the market share equation:

$$s_j = \frac{exp(\delta_j)}{\sum_{k=0}^{J} exp(\delta_k)} \tag{7}$$

Furthermore, if we normalize the utility of the outside option (choice $j = 0$), then we can estimate the following using standard techniques:

$$ln(s_j) - ln(s_0) = \delta_j = x_j \beta + \alpha p_j + \mu_j \tag{8}$$

II.1 Nested Logit

The problem with the standard logit framework is how consumer tastes enter the model. This issue is the well-known independence of irrelevant alternatives assumption. Those consumers that actually chose plan j are likely to have higher than average values of β.[5] Given a change in quality such that consumers in plan j are willing to change plans, they are likely to shift into a similar plan (in style, mission, etc.) as j . In our example, it is natural to think that the substitution patterns between a Medicare HMO and traditional Medicare should differ from patterns within HMOs.

The nested logit model relaxes the I.I.A assumption in a very structured way. Berry (1994) suggests we group products (plans) into $G+1$ groups. In our case, within a market we can group all Medicare HMOs together and leave traditional Medicare for the other group. Here $G = 1$. The nested logit model assumes I.I.A within a group but not across groups. Let us redefine the utility function of consumer i as:

$$u_{ij} = x_j' \beta + \mu_j + \tau_{ig} + (1 - \sigma) \epsilon_{ij} \tag{9}$$

[5] By assuming $\beta_i = \beta$ $\forall i$, we implicitly assume that only mean utility differentiates plans. All consumers therefore have the same estimated value (the mean of the vector β) placed on all observed characteristics.

I omit the g subscript on utility because nest g is implied given plan j. This is the same as the simple logit model utility function except with the addition of τ_{ig}, consumer mean utility when enrolled in a plan in group g. The error term ϵ_{ij} remains an i.i.d. extreme value disturbance. σ represents the nesting parameter on the closed interval [0,1]. When σ approaches one, we say that the within nest correlation is high. As σ approaches 0, the within group correlation goes to zero. $\sigma = 0$ yields the standard logit model. Let $g = 1$ denote the nest of Medicare HMOs in the market and let $g = 0$ denote traditional Medicare. The market share of plan j within nest 1, that is, plan $j's$ share amongst Medicare managed care plans, after dividing through by $(1 - \sigma)$, is given by:

$$s_{j/1} = \frac{exp(\frac{x_j\beta+\alpha p_j+\mu_j}{1-\sigma})}{\sum_{k=1}^{N_1} exp(\frac{x_k\beta+\alpha p_k+\mu_k}{1-\sigma})} \tag{10}$$

Here, N_1 represents the number of Medicare HMOs in the market. Denote the denominator of this equation as $M_{g=1}$. We can define the share of all Medicare HMOs, that is, the total share of nest 1 in the market as:

$$s_1 = \frac{M_1^{1-\sigma}}{\sum_{g=0}^{1} M_g^{1-\sigma}} \tag{11}$$

Furthermore, the market share of traditional Medicare in this case is:

$$s_0 = \frac{1}{\sum_{g=0}^{1} M_g^{1-\sigma}} \tag{12}$$

The absolute market share of plan j is therefor given as:

$$s_j = s_{j/1} * s_1 = \frac{exp(\frac{x_j\beta+\alpha p_j+\mu_j}{1-\sigma})}{M_1^{\sigma} \sum_{g=0}^{1} M_g^{1-\sigma}} \tag{13}$$

We can formulate a similar estimating equation as in the simple logit by taking logs and differencing:

$$ln(s_j) - ln(s_0) = \frac{x_j\beta + \alpha p_j + \mu_j}{1-\sigma} - \sigma ln M_1 \tag{14}$$

Note however that:

$$\frac{1}{1-\sigma}\Big(ln\frac{M_1^{1-\sigma}}{\sum_{g=0}^{1} M_g^{1-\sigma}} - \frac{1}{\sum_{g=0}^{1} M_g^{1-\sigma}}\Big) = \frac{ln(s_j) + ln(s_0)}{1-\sigma} = ln M_1 \tag{15}$$

Rearranging and solving for mean utility yields the estimating equation:

$$ln(s_j) - ln(s_0) = x_j\beta + \alpha p_j + \sigma ln(s_{j/1}) + \mu_j \tag{16}$$

The term $ln(s_{j/1})$ is likely endogenous because μ_j measures unobserved product characteristics of plan j that are likely correlated with a plan's within nest share. For instruments, Berry (1994) suggests characteristics of other firms. Valid instruments will be correlated with plan $j's$ within nest market share but uncorrelated with $j's$ unobserved quality measures. I follow the literature in defining valid instruments as rival plan (i.e. those operating within the same market as plan j) characteristics. Specifically, I consider the following four variables as instruments: the presence of a rival nonprofit plan, the presence of a rival IPA plan within a specific county, and the mean and maximum CMS reimbursement rate within the county. I create indicator variables, taking the value 1 when, for example, there exists a plan within the same county as plan j that is nonprofit. All four instruments significantly predict the plan within nest share.

III Data

Following Dafny and Dranove (2008), I collect Medicare enrollment data from the Medicare Managed Care Quarterly State, County, Plan December Files 1993-2005.[6] Included in these data are Medicare eligibility counts for all counties in the United States as well as enrollment data at the Medicare managed care plan level. The files also include the base CMS payment rate for HMO enrollees in each county. After 1997, enrollment data were restricted to plans with ten or more enrollees; therefore, I restrict my analysis to such plans. Quality data were extracted from the HEDIS and CAHPS databases.[7] The HEDIS database contains a wealth of objective measures of quality for each Medicare Advantage plan with an enrollment over 10 enrollees. Data are available for years 1997 through 2005. The CAHPS data contain subjective measures of quality by plan but are only publicly available for 2005.[8] Luckily, Dafny provided me with certain CAHPS variables for 2001.

While I have a wealth of quality data, this study focuses on the specific measures studied in Dafny and Dranove (2008). As noted above, the 1997 Balanced Budget Act required CMS to collect and distribute quality information on Medicare Advantage Plans. In 2000, CMS distributed "Medicare & You" to over 40 million Medicare enrollees. This publication included the following quality measures: the extent to which individuals in a survey felt their doctors "communicated well" and a plan's mammography rate. The 2001 publication of "Medicare & You" replace the "communication" score with "Overall Plan Satisfaction" on a 10 point scale. Controlling for underlying quality, Dafny and Dranove (2008) also consider a plan's (unreported) diabetic eye exam rate and beta blocker rate. I therefore heavily consider these measures. Furthermore, Dafny and Dranove (2008) only consider enrollment trends through 2002. Given that I have quality data through 2005, I can extend their work

[6]These data are publicly available at CMS Health Plans, Reports, Files and Data.

[7]HEDIS are available at CMS HEDIS Public Use Files

[8]CAHPS data for 2005 are available at Medicare Download Database.

by asking: did Medicare eligibles respond to the publication of quality scores in 2000 and 2001 through 2005?

Merged together are therefore yearly enrollment data from 1997-2005, yearly objective quality measures from 1997-2005, and subjective quality measures for 2001 and 2005. Throughout the analysis, and indeed in much of literature, markets are defined at the county level. National HMO carriers often offer Medicare managed care plans at the county level. An observation in my data is therefore a plan/county/year. Table 1 provides observation counts by year. Table 2 provides plan characteristics. The average number plan in my data

Table 1: Plan/County Observations per Year

year	**N**
1997	644
1998	941
1999	1076
2000	1262
2001	1071
2002	1000
2003	900
2004	962
2005	1285
Total	9141

has 3228 enrollees. 79% of plans are organized as HMOs and 52% of plans are for-profit. The average annual reimbursement from CMS to plans is $564.29. In the spirit of also exploring

Table 2: Summary statistics: Plans

Variable	**Mean**	**Std. Dev.**
Enrolled	3227.94	9037.46
Penetration	0.05	6.94
Plan Share	0.50	0.42
For-Profit	0.52	0.50
IPA	0.57	0.50
HMO	0.79	0.41
CMS Rate	564.29	106.55
N	9141	

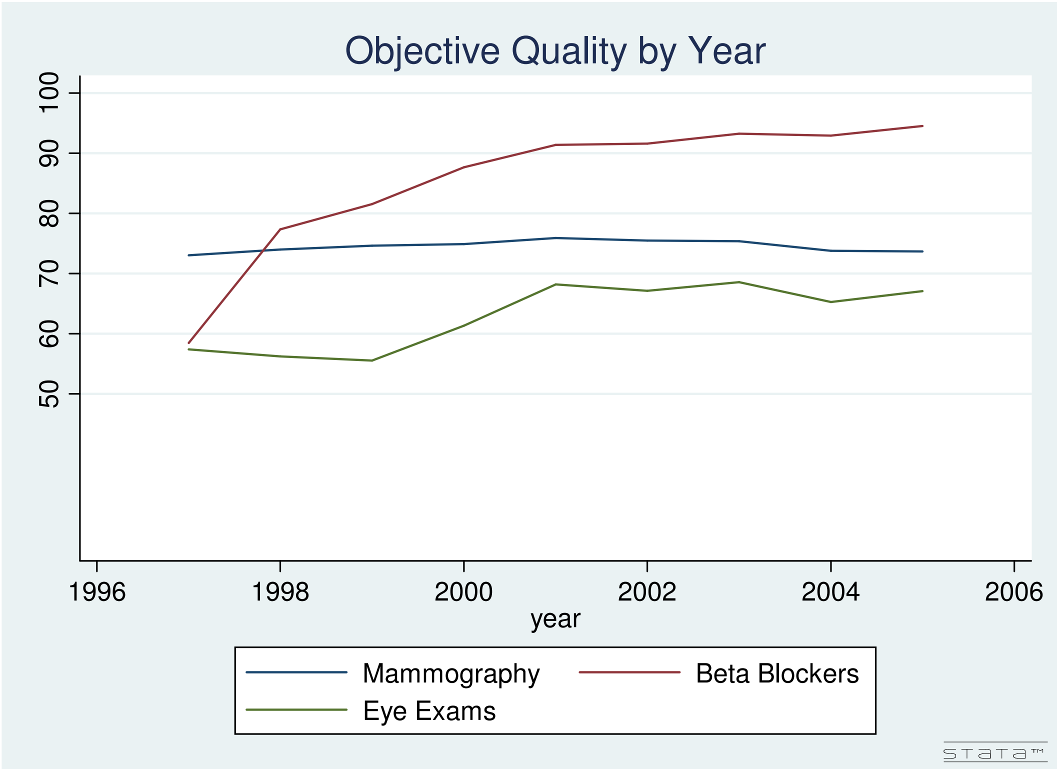
secondary characteristics of plan enrollment, I make a first attempt by merging county level data from the 2000 U.S. Census to my enrollment and quality data.[9] Table 3 provides a brief set of county characteristics. The average number of plans operating in a county was 3.55.

[9]Data are available at American Fact Finder.

Table 3: Summary statistics: Counties

Variable	**Mean**	**Std. Dev.**
# Plans	3.55	3.17
Population	440975.92	813369.78
College Graduates	0.23	0.10
Women $\geq$65	0.12	0.03
Men $\geq$65	0.09	0.03
N	9141	

Figure III shows the evolution of plan several quality scores by year. Table 4 details,


Objective Quality by Year
100
90
80
70
60
50
1996
1998
2000
2002
2004
2006
year
Mammography
Beta Blockers
Eye Exams

by year, my main subjective measure of quality: Best Care. Best Care is the percentage of respondents that indicated that their Medicare Advantage plan's overall rating was a "10/10". Meanwhile, table 5 gives summary statistics for other subjective measures. For these measures, I only have data for 2005.

Table 4: Overall Plan Rating

Year	**% 10/10**	**% ≥8/10**	**% ¡8/10**
2001	51.88	39.09	22.60
2005	38.24	38.96	22.77
Total	45.06	39.02	22.68

All summary statistics of quality measures detailed above are in percentages. To facilitate comparison across measures, I now standardize each measure, for each year available, using z-scores such that each measure/year has a standard normal distribution. Table 6 provides evidence of the potential for learning. Recall, the average number of plans competing within any given county/year is 3.55. Table 6 displays, by year, the average difference between the maximum and minimum z-score within a county/year. The final column displays the t-test significance level at which we can reject the null hypothesis that the average differences are equal. Note that for most quality scores, there exists at least a 1 standard deviation difference between the best and worst plan *within* a market. Furthermore, in many cases, that difference is increasing from 2001 to 2005.

Table 5: Other Subjective Measures

Measure	**% No Problem**	**% Small Problem**	**% Big Problem**
Seeing A Specialist	82.50	11.19	6.30
Overall Satisfaction with Health Care	45.86	39.76	14.29
Doctor Communication	68.76	24.60	6.61

Table 6: Differences in Max-Min Z-Score

Quality Measure	2001 Difference	2005 Difference	Significance
Best Care	1.02	1.45	***
Mammography	1.19	1.13	
Beta Blocker	0.94	1.06	**
Diabetic Eye Exam	1.20	1.31	**
CABG$\leq$65: Male	1.18	1.52	***
CABG 65–74: Male	1.18	1.50	***
CABG 75–84: Male	1.18	1.41	***
CABG$\geq$85: Male	1.20	1.17	
PTCA$\leq$65: Male	1.01	1.00	
PTCA 65–74: Male	1.15	1.29	***
PTCA 75–84: Male	1.22	1.14	**
PTCA $\geq$85: Male	1.17	1.51	***
Hip $\leq$65: Male	0.86	0.92	*
Hip 65–74: Male	1.18	1.47	***
Hip 75–84: Male	1.12	0.81	***
Hip $\geq$85: Male	0.73	0.75	
Prostate $\leq$65	1.14	1.33	***
Prostate 65–74	1.34	1.49	***
Prostate 75–84	1.11	1.50	***
Prostate $\geq$85	1.16	1.17	
CABG$\leq$65: Female	1.09	1.54	***
CABG 65–74: Female	1.17	1.25	
CABG 75–84: Female	1.06	1.25	***
CABG$\geq$85: Female	1.12	1.12	
PTCA$\leq$65: Female	1.00	1.07	*
PTCA 65–74: Female	1.06	0.87	***
PTCA 75–84: Female	1.22	1.25	
PTCA$\geq$85: Female	1.21	1.42	
Hip$\leq$65: Female	1.07	1.11	
Hip 65–74: Female	1.30	0.78	***
Hip 75–84: Female	1.33	0.78	***
Hip $\geq$85: Female	1.38	0.70	***

Significance Levels: ***1% Level,
**5% Level, *10% Level

IV Preliminary Results

While any results are quite preliminary and by no means authoritative, I present the following suggestive findings that motivate further work. First, table 7 is a cursory attempt at addressing whether subjective quality scores really signify better health outcomes. Furthermore, as motivated above, if subjective measures of quality do not reflect those in the right tail of the expenditure distribution, their inclusion in report cards may be ill advised. Table 7 shows the simple correlations between certain subjective and objective quality scores. For example, the first column presents the correlation between the 2001 plan best care score and a variety of 2001 objective quality measures. Column two presents the same correlations except where all objective measures are their 2005 counterpart. Note that most, if not all, correlations below show small positive or negative correlations. Furthermore, note that the correlation is often smaller or negative for an increasingly older population.

Table 8 shows the results of the following simple random-effects regression model:

$$score^{l}_{jm2005} = X_{jm2001}\beta_m + \mu_m + \epsilon_{mj} \tag{17}$$

Here m represents the county/state market area and j represents the specific plan. Furthermore, l, indexes the score mammography or best care. I run the model separately for each quality score. First note that the results strongly support a multi-level error structure. The estimate of the share of the total error variance that stems from the county-state specific term (rho) is significant at the 1% level in both models. As we might expect, having a high l score in 2001 is positively associated with a high l score in 2005. Interestingly, for-profit status correlates positively with 2005 scores where as IPA status does not. Finally, I estimate a set of models attempting to capture the effect of report card information on specific groups of individuals. Recall that Dafny and Dranove (2008) find that the publication of quality report cards in 2000/2001 did influence enrollment, but that only subjective scores like best care were influential. Dafny and Dranove (2008) do not however consider specific populations in this analysis. Furthermore, as motivated above, report card distributors often face a trade off between accuracy and accessibility. If Medicare Advantage quality report cards are complicated to the extent that the average individual does not internalize all or most of the information, individuals may revert to simple ordinal ranks or subjective scores to influence their decision.[10] In this section, I estimate models motivated by my theoretical results above. Specifically, I estimate the following equations:

$$ln(s_{jm2005}) - ln(s_{0m2001}) = x_{jm}\beta + \sigma ln(s_{jm2001/1}) + \mu_{jm} \tag{18}$$

[10] A simple Google search reveals numerous websites that decry the complexity of the Medicare and You publications.

Table 7: Correlation Matrix

Quality Measure	**Correlation Coefficient**	
	Best care 2001	Best care 2005
Best Care	1.00	1.00
Mammography	-0.01	0.29
Beta Blocker	0.02	0.03
Diabetic Eye Exam	-0.14	-0.00
CABG$\leq$65: Male	0.16	0.09
CABG 65–74: Male	0.28	-0.22
CABG 75–84: Male	0.30	-0.06
CABG$\geq$85: Male	-0.08	0.03
PTCA$\leq$65: Male	0.18	-0.08
PTCA 65–74: Male	0.24	-0.21
PTCA 75–84: Male	0.13	-0.14
PTCA $\geq$85: Male	-0.07	-0.17
Hip $\leq$65: Male	-0.04	-0.19
Hip 65–74: Male	0.09	-0.44
Hip 75–84: Male	-0.08	-0.23
Hip $\geq$85: Male	-0.20	-0.06
Prostate $\leq$65	0.24	0.02
Prostate 65–74	0.01	-0.10
Prostate 75–84	0.10	-0.29
Prostate $\geq$85	0.15	-0.11
CABG$\leq$65: Female	0.17	0.13
CABG 65–74: Female	0.24	-0.20
CABG 75–84: Female	0.30	-0.10
CABG$\geq$85: Female	0.34	-0.03
PTCA$\leq$65: Female	0.34	0.03
PTCA 65–74: Female	0.22	-0.01
PTCA 75–84: Female	0.13	-0.15
PTCA$\geq$85: Female	0.12	-0.11
Hip$\leq$65: Female	-0.13	-0.01
Hip 65–74: Female	-0.11	-0.21
Hip 75–84: Female	0.05	-0.12
Hip $\geq$85: Female	-0.30	-0.03

Table 8: Estimation results : xtreg

	Dependent Variable	
Covariate	**2005 Mammography Rate**	**(2005 Best Care)**
2001 HMO Share	-0.238	0.027
	(0.061)	(0.086)
For-Profit Dummy	0.224	0.100
	(0.047)	(0.067)
IPA Dummy	-0.198	-0.734
	(0.047)	(0.067)
2001 Reimbur. Rate	-0.003	-0.001
	(0.001)	(0.001)
2001 Mammography Rate	0.106	0.015
	(0.005)	(0.006)
2001 Beta Blocker	-0.006	-0.059
	(0.004)	(0.005)
2001 Diabetic Eye Exam	0.003	0.014
	(0.003)	(0.004)
2001 Best Care	0.021	0.025
	(0.004)	(0.005)
# 2001 Plans	0.038	-0.023
	(0.021)	(0.021)
Intercept	-7.175	2.770
	(0.533)	(0.663)
ρ	0.453	0.169

First, note the exclusion of the additional plan premium charged by the Medicare plan. Dafny and Dranove (2008) note that the median plan premium for all years between 1994-2002 was $0 and that the period saw extreme "premium compression". Plans effectively competed only on the basis of quality. Furthermore, those authors found no significant effect of plan premium on enrollment trends. Again, m denotes the county/state market. I therefore regress the log difference in market shares between plan j and traditional (FFS) Medicare in 2005 on a variety of *2001* covariates. Again, given the potential endogeneity of a plans within market share, I instrument for this variable with other characteristics of other plans *within* market m.[11] I report robust standard error that have been clustered at the market level. Initial results are presented in 9. Note that high 2001 best care and diabetic eye exam scores positively predict 2005 absolute market share. Furthermore, a high HMO Share in 2001 was positively associated with 2005 absolute market share. Table 10 presents estimates from

Table 9: Instrumental Variables Regression

Dependent Variable: $(ln(s_j) - ln(s_0))$		
Variable	**Coefficient**	**(Std. Err.)**
Intercept	-5.648	(0.622)
HMO Share	0.726	(0.077)
Best Care	0.298	(0.083)
Mammography Rate	0.110	(0.078)
Diabetic Eye Exam	0.251	(0.079)
Beta Blockers	-0.091	(0.079)
IPA	0.171	(0.134)
For-Profit	0.117	(0.145)
Reimb. Rate	0.004	(0.001)

the same model above except with quality score interactions with dummies for market areas that were in the 75th (column 1) and 90th (column 2) percentiles in college graduates in year 2000. Interestingly, it appears that counties in the 90th percentile in college graduates in 2000 had a significantly stronger positive reaction to high mammography rate scores than the average county. Furthermore, both counties at the 75th and 90th percentiles appeared to place less emphasis on best care scores.

Finally, table 11 presents results from the same model as above expect with interactions for the 75th and 90th percentile counties of populations of women over the age of 65. We might expect these counties to respond most heavily to the publication of high mammography scores. Interestingly however, high scoring mammography rate plans did not see an increase in market share in these counties.

[11] Section discussion in the theoretical section.

Table 10: Estimation results : ivreg

Dependent Variable: $(ln(s_j) - ln(s_0))$		
Variable	**Specification 1**	**Specification 2**
Intercept	-5.548	-5.628
	(0.605)	(0.621)
HMO Share	0.674	0.704
	(0.075)	(0.077)
Best Care	0.422	0.335
	(0.088)	(0.088)
Mammography Rate	0.094	0.082
	(0.092)	(0.083)
Diabetic Eye Exam	0.208	0.229
	(0.097)	(0.085)
Beta Blockers	-0.054	-0.073
	(0.087)	(0.081)
IPA	0.205	0.191
	(0.133)	(0.133)
For-Profit	0.102	0.121
	(0.145)	(0.146)
Reimb. Rate	0.004	0.004
	(0.001)	(0.001)
Mammography*College75	0.205	
	(0.158)	
Eye Exam*College75	0.170	
	(0.180)	
Beta Blocker*College75	-0.231	
	(0.183)	
Best Care*College75	-0.367	
	(0.170)	
Best Care*College90		-0.286
		(0.211)
Mammography*College90		0.567
		(0.229)
Beta Blocker*College90		0.158
		(0.255)
Eye Exam*College90		0.288
		(0.280)

Table 11: Estimation results : ivreg

Dependent Variable: $(ln(s_j) - ln(s_0))$		
Variable	**Specification 1**	**Specification 2**
Intercept	-5.506	-5.482
	(0.595)	(0.595)
HMO Share	0.695	0.714
	(0.074)	(0.074)
Best Care	0.143	0.236
	(0.096)	(0.089)
Mammography Rate	0.137	0.139
	(0.092)	(0.085)
Diabetic Eye Exam	0.218	0.237
	(0.088)	(0.083)
Beta Blockers	-0.141	-0.107
	(0.080)	(0.080)
IPA	0.191	0.180
	(0.133)	(0.133)
For-Profit	0.068	0.112
	(0.142)	(0.144)
Reimb. Rate	0.004	0.004
	(0.001)	(0.001)
Mammography*Elderly75	-0.010	
	(0.181)	
Eye Exam*Elderly75	-0.089	
	(0.238)	
Beta Blocker*Elderly75	0.158	
	(0.255)	
Best Care*Elderly75	0.578	
	(0.148)	
Best Care*Elderly90		0.656
		(0.170)
Mammography*Elderly90		-0.061
		(0.263)
Beta Blocker*Elderly90		-0.026
		(0.336)
Eye Exam*Elderly90		-0.135
		(0.337)

V Discussion: Where to go next?

I now briefly itemize the directions in which I would like to take this research:

- Obtain contemporaneous quality data, both objective and subjective from 2001 onwards. I could then follow the enrollment and market share responses after the publication of the "Medicare & You" report cards while also controlling for pre-report card quality.

- Obtain individual level data from the Medicare Beneficiary Survey. I could then explore the secondary characteristics of plan choice and the extent to which different populations respond to plan report cards.

- With the above data, I also would like to more richly model the supply side of the market. Specifically, I would like to estimate a model of HMO behavior and study the incentives generated by report cards.

References

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.

Chernew, M. and Scanlon, D. P. (1998). Health plan report cards and insurance choice. *Inquiry*, **31** (1), 9–22.

Dafny, L. and Dranove, D. (2008). Do report cards tell consumers anything they don't already know? the case of medicare hmos. *Rand Journal of Economics*, **39** (3), 790–821.

Dranove, D., Kessler, D., McClellan, M. and Satterthwaite, M. (2005). Is more information better? the effects of 'report cards' on health care providers. *Journal of Political Economy*, **111** (3), 555–588.

Israel, M. (2005). Services as experience goods: An empirical examination of consumer learning in automobile insurance. *American Economic Review*, **95** (5), 1444–1463.

Newhouse, J. P. (2002). *"Medicare" in American Economic Policy in the 1990s*. Cambridge: MIT Press.

O'Brien, R. M. (1981). Using rank category variables to represent continuous variables: Defects of common practice. *Social Forces*, **59** (4), 1149–1162.

Pope, D. G. (2006). Reacting to rankings: Evidence from "america's best hospitals and colleges", job Market Paper, Department of Economics, University of California, Berkeley.

Royalty, A. and Solomon, N. (1999). Health plan choice: Price elasticities in a managed competition setting. *Journal of Human Resources*, **43** (1), 1–41.

Scanlon, D. P., Chernew, M., McLaughlin, C. and Solon, G. (2002). The impact of health plan report cards on managed care enrollment. *Journal of Health Economics*, **21**, 19–41.

Wedig, G. and Tai-Seale, M. (2002). The effect of report cards on consumer choice in the health insurance market. *Journal of Health Economics*, **21**, 1031–1048.