

HEDG Working Paper 10/18

Information and Quality when Motivation is
Intrinsic: Evidence from Surgeon Report Cards

Jonathan Kolstad

July 2010

Information and Quality when Motivation is Intrinsic: Evidence from Surgeon Report Cards*

Jonathan T. Kolstad**
April 2010

Abstract: If profit maximization is the objective of a firm, new information about quality should affect firm behavior only through its effects on market demand. I consider an alternate model in which suppliers are motivated by a desire to perform well in addition to profit. A simple model demonstrates the way in which performance data can alter both pecuniary incentives (i.e. extrinsic motivation) and incentives unrelated to profit (i.e. intrinsic motivation). The introduction of quality “report cards” for cardiac surgery in Pennsylvania provides an empirical setting to test for an effect of new information on quality (mortality) and to isolate the relative role of extrinsic (demand side) and intrinsic (supply side) incentives in determining surgeon response to new information. Using a structural demand system, I estimate the profit incentives facing each surgeon from the introduction of report cards. Extrinsic incentives due to quality reporting led to a .09 percentage point (three percent) decline in mortality. Consistent with a mixed model of objectives, information on performance that was new to surgeons and unrelated to patient demand led to an intrinsic response three times as large as surgeon response to profit incentives.

Key Words: demand estimation; intrinsic motivation; health care; quality competition

* I would especially like to thank my thesis committee: Susan Athey, David Cutler and Robert Huckman. In addition, I am grateful to Amitabh Chandra, Mike Chernew, Richard Frank, Amanda Kowalski, Tom McGuire, Joe Newhouse, Heidi Williams, Frank Wolak and seminar participants at the ASHE Biennial Meeting, Stanford, Emory, RAND, Wharton, Cornell, Washington University, CBO and Harvard for many helpful comments. Support from the National Institute on Aging (Grant P01 AG005842) is gratefully acknowledged. All errors are my own.

The data used in this analysis were obtained from the Pennsylvania Health Care Cost Containment Council (PHC4), which requests the following disclaimer: The Pennsylvania Health Care Cost Containment Council (PHC4) is an independent state agency responsible for addressing the problem of escalating health costs, ensuring the quality of health care, and increasing access to health care for all citizens regardless of ability to pay. PHC4 has provided data to this entity in an effort to further PHC4's mission of educating the public and containing health care costs in Pennsylvania. PHC4, its agents and staff, have made no representation, guarantee, or warranty, expressed or implied, that the data -- financial, patient, payor, and physician specific information -- provided to this entity, are error-free, or that the use of the data will avoid differences of opinion or interpretation. This analysis was not prepared by PHC4. This analysis was done by Jonathan T. Kolstad. PHC4, its agents and staff, bear no responsibility or liability for the results of the analysis, which are solely the opinion of the authors.

** Wharton School, University of Pennsylvania. Contact: 108 Colonial Penn Center, 3641 Locust Walk, Philadelphia PA 19104. (215) 573-9075 or jkolstad@wharton.upenn.edu

1. Introduction

A defining feature of health care markets is imperfect information (Arrow, 1963). Standard models show that in settings where consumers are poorly informed about product quality there are welfare losses due to the less-than-optimal supply of costly quality (Dranove and Satterthwaite, 1992) or the absence of markets for products and services that consumers otherwise value (Akerlof, 1970). Losses generally stem from profit maximizing suppliers who are able to profit from information to which they are privy but cannot be obtained (or verified) by the demand side (e.g. Arrow, 1963; Gaynor, 2006; Salop and Stiglitz, 1977). Consequently, policies to correct market failures due to asymmetric quality information focus on demand. In health care markets, where quality is widely believed to be suboptimal, this is the rationale for efforts to gather performance information, such as mortality, and offer it directly to consumers. Viewed in the standard framework, however, the existing evidence on such quality reporting in health care is paradoxical. Studies generally find improvements in measured quality but little evidence for corresponding changes in consumer demand (Epstein, 2006; Steinbrook, 2006).

To better explain the observed behavior of surgeons facing public quality reporting and to explore the nature of incentives among firms and individuals receiving public ratings, I consider the role of information in determining market outcomes when suppliers also have non-pecuniary incentives. I use the term “intrinsic motivation” to refer to incentives unrelated to profit and model it as a function not only of quality itself but of the ability to observably perform well relative to a reference group. In this context, information enters profit motives and alters intrinsic incentives when collecting and disseminating information provides the individual with a better sense of his or her own quality compared to peers.

The empirical setting for this study is the introduction of quality “report cards” for surgeons performing Coronary Artery Bypass Graft (CABG) surgery in Pennsylvania. Utilizing a detailed panel of data on surgeons and patients, I explore the effects of quality report card release on subsequent surgeon performance. Exploiting the information contained in the report card’s risk adjustment scheme, I model surgeons’ prior and posterior beliefs about market quality levels (both their own and that of their peers). I find

an impact of this information on changes in surgeon quality but not on demand— evidence for the presence of non-pecuniary incentives resulting from quality reporting.

To incorporate profit incentives, I estimate a structural model of consumer demand for surgeons. Consumer utility is modeled as a function of the detailed set of individual patient and surgeon observables. In addition, I account for unobserved (to the econometrician) influences on choice, including the role of physician agency, using a random coefficients demand model (Berry, *et al.*, 1995; Train, 2003). Simulations, relying on the estimated demand parameters, produce a measure of the additional market rewards for quality due to reporting. Variation in the *ex ante* distribution of patient demand for quality and the competitive structure of the markets leads to large differences in extrinsic incentives between surgeons. Individuals facing stronger profit incentives following the release of quality report cards show greater improvements in performance. This effect, however, is relatively small. Extrinsic incentives led to an additional 3 percent decline in the statewide risk adjusted mortality rate (RAMR) between the pre- and post-report card periods. Incorporating estimates of the intrinsic response to information predicts changes in surgeon quality that accords well with the observed response to reporting. The intrinsic response to quality reporting is about three times as large as the response due to profit incentives induced by changes in demand for quality. Limited sensitivity analysis finds that the observed intrinsic response cannot be explained by dynamic pecuniary incentives or inter-temporal technological change.

This paper contributes to the debate in economics on the merit of reporting schemes. Quality reporting and disclosure policies have been implemented across a variety of industries to address a multitude of market failures (Chatterji and Toffel, 2007). Reviewing the evidence, Fung *et al.* (2007) find the welfare impact of information based policy interventions varies dramatically depending on the market and regulatory environment. Their assessment, however, appraises disclosure through the lens of the standard profit maximizing model. The same approach also underlies much of the evaluation of reporting for CABG surgery. Despite finding that additional quality improvement followed the introduction of reporting programs (Epstein, 2006; Ghali *et al.*, 1997; Peterson, *et al.*, 1998; Hannan, *et al.*, 2003), the debate on and analysis of reporting efforts has focused on consumers' ability to interpret and respond to the

information supplied. If quality report cards deliver information on and to suppliers who care about performance intrinsically then their impact (both positive and normative) is not solely mediated through changes in demand.

Decisions about the value and type of quality information that should be measured and publicly reported depend critically on the model of supplier behavior. If suppliers operate under a standard profit model, demand side incentives can produce quality improvements. On the other hand, to the extent that information about peers alters surgeons' intrinsic incentives, public release is of less relevance. In fact, contrary to current efforts to simplify provider report cards, it may be preferable to deliver data with more clinical detail.

This paper also contributes to a broad literature in economics on information and incentives. Work in behavioral game theory and experimental economics has demonstrated a potential role for reference-based utility in individual behavior and incentives (Fehr and Schmidt, 2006; Heffetz and Frank, 2008). To date, however, relatively few empirical studies have documented such incentives in practice (Sauer mann and Cohen, 2008). This is due in part to the difficulty of empirically identifying changes in information in a market. The CABG setting allows me to overcome this problem by observing an exogenous and measurable change— the release of quality report cards.

The paper proceeds as follows. Section 2 develops a model of quality choice with intrinsic motivation. Section 3 discusses the data and setting. Section 4 presents the econometric specifications and results. Section 5 provides discussion and section 6 concludes.

2. A Model of Surgeon Objectives with Uncertainty

I begin by considering the equilibrium quality choices of surgeons who gain utility from income (profit) and from performing well. Because quality is also valued independently of earnings, surgeons are willing to forgo some profit to enhance quality. This willingness is a function of the ability to observe performance— determined by the information structure of the market. The intuition of the model is that a surgeon with little information on his own performance and that of his peers is unable to accurately observe both static levels of quality and improvements. In this way, increased uncertainty dilutes

the intrinsic incentive for quality improvement because surgeons do not see the result of effort (or may believe, in the absence of information, that they are performing as well they should be). This information also impacts surgeon incentives through the standard channel- quality information informs consumers and, subsequently, determines quality elasticity of demand (and profit).

Consider the quality choice of a surgeon in a monopolistically competitive market with regulated prices above the marginal cost of production. Prices are fixed at a regulated level, p_{reg} and surgeons maximize utility by selecting a quality level θ_i subject to a convex production technology, $c(\theta_i, q_i)$. To incorporate preferences that include both profit and intrinsic incentives, I express supplier utility as:

$$U_i = \Pi_i(\theta_i, \theta_{-i}, \Omega) + \Gamma_i(\theta_i, \theta_j, \Omega) \quad (1)$$

Allowing profit and intrinsic utility from quality to enter as additively separable terms can incorporate a range of intrinsic preferences and is a common feature of models of physician behavior (Fehr and Schmidt, 2006; Harsanyi, 1955; McGuire, 2000; Segal and Sobel, 2004).¹ The term $\Gamma_i(\theta_i, \theta_j, \Omega)$ captures individual i 's intrinsic utility from quality relative to the reference group $j \in J$.² Firm demand, $q_i(\theta_i, \theta_{-i}, \Omega)$, is determined by the quality of surgeon i as well as the quality choices of all competing surgeons. The reference group J is not necessarily the same as the set of all competing surgeons, indexed by $-i$. Information in the market is captured by the variable Ω , the effect of which I return to below. Surgeon i solves the following problem:

¹ Segal and Sobel (2004) take an axiomatic approach to show that a utility function of this sort incorporates a wide range of the models for other-regarding utility functions.

² Intrinsic motivation has a strict definition in the psychology literature: utility from the activity must be derived from a stimulus within the individual (Sauermaun and Cohen (2008) provide a useful synopsis). The model here captures intrinsic incentives under this definition as well as in a more general sense (typically used by economists). It fits a strict interpretation of intrinsic utility if reference utility is derived internally but, due to uncertainty, is altered by the outside information. However, if surgeons care directly about a ranking that is provided by the report card this is (strictly speaking) an extrinsic motive because it is initiated by an outside stimulus. For simplicity, in this paper I refer to all non-pecuniary rewards as intrinsic though I acknowledge this may not adhere to convention in some fields.

$$\max_{\theta_i} U_i = \Pi_i(\theta_i, \theta_{-i}, \Omega) + \Gamma_i(\theta_i, \theta_j, \Omega) = q_i(\theta_i, \theta_{-i} | \Omega) p_{reg} - c(\theta_i, q_i) + \Gamma_i(\theta_i, \theta_j | \Omega) \quad (2)$$

The argument that maximizes (2) (optimum quality) is reached when:

$$\left(\frac{\partial q_i(\theta_i, \theta_{-i}^* | \Omega)}{\partial \theta_i} \right) p_{reg} + \frac{\partial \Gamma_i(\theta_i, \theta_j^* | \Omega)}{\partial \theta_i} = \frac{\partial c(\theta_i, \theta_j^*)}{\partial \theta_i} \quad (3)$$

where all partial derivatives are taken with respect to own quality, taking the best response of other surgeons as given. The optimum is simply the point at which the sum of the marginal revenue (determined by the price and demand elasticity of quality) and marginal intrinsic utility is equal to the marginal cost of quality.³ Implicitly differentiating (3), optimal quality increases in demand for quality (determined jointly by consumers' willingness-to-pay for quality and their ability observe it (Dranove and Satterthwaite, 1992; Gaynor, 2006)) and intrinsic utility from quality improvement. Quality declines with the marginal cost of quality. In the standard model (where a surgeon cares only about pecuniary rewards) $\frac{\partial \Gamma_i(\theta_i, \theta_j^* | \Omega)}{\partial \theta_i} = 0$ and the equilibrium condition is reduced to setting marginal cost equal to marginal revenue.⁴ Optimum quality is an increasing function of the residual elasticity of demand for quality.⁵

Comparative statics incorporating intrinsic incentives require additional structure on the utility function. I assume that a surgeon determines his or her own quality relative to a reference group $\tilde{\theta}_j \in (\underline{\theta}_j, \bar{\theta}_j)$. Intrinsic utility from quality is captured in the model as a function that maps the deviation between an individual's quality and their reference point to a change in utility: $\Gamma_i(\theta_i^* - \tilde{\theta}_j)$. The precise relationship depends on the reference group and the shape of the intrinsic loss function.

Up to this point, I have not explicitly considered the role of quality reporting. The information structure of the market captures the impact of quality reporting. Information

³ One could also move the marginal intrinsic incentive in equation (3) to the right hand side. In this case the intrinsic incentive enters as a "reduction" in marginal cost. This interpretation is developed by Gaynor (2006) to model supplier quality choice with not-for-profit incentives.

⁴ Because prices are set by a regulator (Medicare), demand is equal to marginal revenue provided quality elasticity of demand is not a function of patient cost (i.e. raising quality does not lead to differential increases in demand from the most severe patients).

⁵ I do not consider explicitly whether demand is "correct" in a normative sense. This depends on the regulated price and whether quality elasticity of demand reflects some socially desirable taste for lower mortality. Despite this, incorporating a social marginal revenue curve into this framework that would allow theoretical conclusions as to whether changes in incentives induced by changes in information are optimal.

in period t is indexed by the set $\Omega^t \in \{\mu, \varepsilon\}$ containing two elements. The first term, μ , captures information on the relative location of a surgeon in the distribution. The second term, ε , measures the “quality” of information or the precision of a surgeon’s beliefs about the distribution of reference qualities.⁶ From equation (1) and the first-order condition in (3), it is clear that changes to the information structure can alter both components of surgeon utility. First, improved information allows consumers to more easily observe the quality of their full choice set of surgeons. This change in demand alters the pecuniary returns to quality improvement. Second, a better signal provides surgeons with more precise information on the performance of the set J in the reference group. Improved knowledge about the reference group alters the shape of the intrinsic utility function because changes in performance are more easily measured and produce utility gains.

Figures 1 and 2 present a graphical example. In Figure 1 each surgeon has a convex marginal cost curve (labeled MC) and a marginal revenue curve that is increasing in quality (labeled MR).⁷ In this standard model a surgeon is solely extrinsically motivated. He selects the point at which the marginal cost of quality improvement equals the marginal pecuniary benefit (MR). Introducing information on performance alters quality by changing the slope of the marginal revenue from improving quality. This can be seen in the rotation from MR0 to MR1. Quality reporting leads to improvement in performance from θ^* to θ^{**} by altering demand for quality and, thus, the pecuniary reward.

Figure 2 introduces a mixed surgeon utility function. The upper panel contains two possible intrinsic utility functions. Two surgeons (labeled 1 and 2) have the same concave intrinsic utility function but compare themselves to the lowest and highest

⁶ In a general model, surgeons are Bayesian learners and new information induces a posterior distribution of reference quality. Updating alters both mean beliefs about relative quality and the precision of the posterior.

⁷ The marginal revenue curve is determined both by demand elasticity for quality as well as the marginal profitability of a patient. I assume payments are sufficient to make marginal patients profitable. In cardiac surgery this condition is likely to hold. Huckman (2006) finds that cardiac surgical DRGs are profitable on average and at the margin. Chernew, *et al.* (1998) also find evidence that reimbursement for cardiac surgery is greater than cost (though the degree of profitability varies by payer).

surgeon in their reference set respectively.⁸ Note that both curves are monotonically increasing. For any combination of a utility function and reference point leading to monotonically increasing returns, the marginal intrinsic incentive for quality is positive and bounded below by zero. The deviation between profit and utility maximizing quality is determined by the shape of the MC, MR and MB curves. For any monotonically increasing intrinsic utility function the marginal benefit curve (MB) is higher than MR. Graphically this is captured in the increase in equilibrium quality from θ^{**} to θ' or θ'' , depending on surgeon's intrinsic utility function and reference point.⁹

To see the impact of information on equilibrium quality, consider a surgeon choosing quality prior to report cards. Prior to quality reporting, without information on peers, any change in quality is indistinguishable from noise. In the top panel of Figure 1 this is the flat intrinsic utility function $\Gamma(\theta_i - \underline{\theta} | \Omega^0)$. Without information, improving quality does not increase utility because it cannot be observed. Utility from quality need not, however, be set at zero if surgeons gain some level of static intrinsic utility – the “warm glow” from being a cardiac surgeon.

After report cards are released new information is provided with the signal $\Omega^1 \in \{\mu'', \varepsilon''\}$. Information that alters a surgeon's perceived relative quality changes his utility either positively or negatively (e.g. they learn they are better or worse than expected). Marginal incentives are also altered by the quality of the signal, ε , and the ensuing shape of the utility function. In this case, because there was no information prior to quality reporting, the prior slope of intrinsic utility is zero so any signal that provides new information will unambiguously increase the slope of the intrinsic utility curve resulting in increased intrinsic incentives for all physicians.¹⁰ This need not be true,

⁸ Note that the curve for surgeon 2 is convex as presented due to the reference point. Utility, however, is still concave in that returns are diminishing for changes further from the reference quality, the “best” surgeon.

⁹ In Figure 2, equilibrium quality is relatively high even for a profit maximizing supplier so for surgeon 1 (where the reference point is defined by the lowest quality in the group) intrinsic incentives increase quality to θ' , a small effect. On the other hand, for surgeon 2, who aspires to be “the best”, intrinsic incentives are relatively strong, leading to an equilibrium quality choice of θ'' .

¹⁰ This also underscores the fact that for any intrinsic utility function unrelated to information, quality reporting should not alter non-pecuniary incentives for quality. Even if surgeons are purely altruistic and perfect agents for patients, quality reporting would not alter quality levels unless information enters utility (the argument parallels perfect agency in Ellis and McGuire (1986)). This is particularly relevant in thinking about the importance of suppliers with non-pecuniary incentives (i.e. not-for-profits, etc.) in

however. If surgeons have some information on the distribution of reference quality prior to formal reporting then the slope of the intrinsic utility function will not initially be zero and some surgeons can receive new information that diminishes or leaves incentives unchanged. In general, any signal that increases the slope of the intrinsic utility curve at the posterior level of quality relative to the slope at the prior quality level increases the incentive to improve performance- regardless of its effect on demand (MR).

A measure of extrinsic and intrinsic utility can also be computed in this framework. In Figure 2, the area below the marginal cost curve between A and B and above the marginal revenue curve between A and C measures the additional cost in excess of revenue a surgeon is willing to expend in order to improve performance solely to gain intrinsic utility.¹¹ That is, for every quality investment beyond θ^{**} a surgeon loses money at the margin. Willingness to undertake such investments captures a non-pecuniary incentive in the utility function related to costly quality.

The primary goal of the remainder of the paper is to measure empirically the relative contribution of each of these incentive components in determining surgeon response to quality reporting in Pennsylvania. Viewed in Figure 2, this effort reduces to decomposing the observed quality improvement from θ^* to θ'' into the share due to the move from MR0 to MR1 and the share due to changes in intrinsic incentives from $\Gamma(\theta_i - \bar{\theta} | \Omega^0)$ to $\Gamma(\theta_i - \bar{\theta} | \Omega^1)$.

3. Background and Setting

3.1 Quality Reporting in Health Care

Quality reporting programs have been implemented in many forms across a variety of markets for health insurance and for providers (see Kolstad and Chernew (2009) for a review of the evidence to date). As of 2006, forty seven states had some form of quality reporting system in place for health care providers (thirty seven are

determining market outcomes in health care. Not-for-profit providers are assumed to more accurately reflect social preferences in their strategies but, in general, models have not considered the knowledge they have about performance or how these incentives interact with quality reporting policies (Arrow, 1963; Newhouse, 1970; Sloan, 2000).

¹¹ For any cost function and demand curve the dollar value of intrinsic utility is $\int_{\theta^{**}}^{\theta''} MC(\theta)d\theta - \int_{\theta^{**}}^{\theta''} MR(\theta)d\theta$.

mandatory and ten are voluntary) (Steinbrook, 2006). The most studied within the provider context have been the CABG report card programs in New York and Pennsylvania.¹² Reporting of surgeon and hospitals' risk adjusted mortality rate (RAMR) for CABG began in 1989 in New York State. Pennsylvania's experiment followed shortly thereafter and was led by the Pennsylvania Health Care Cost Containment Council (PHC4), a public/private partnership. They began collecting discharge data on outcomes and patient comorbidities in 1990. The first widely available report card was released in May of 1998 and included data from 1994-95.¹³

Some of the earliest evidence on these policies comes from surveys of market participants and suggests that quality reporting did not significantly alter consumer choice. Schneider and Epstein (1998) survey patients who received CABG surgery in Pennsylvania following the release of the initial rounds of report cards (1990-1993). They find that roughly 20 percent of patients were aware of the report cards and only 12 percent knew about the report cards before receiving surgery. The same authors (Schneider and Epstein, 1996) also survey cardiac surgeons and cardiologists and find that report card measures were "very important" in surgeon selection for only 10 percent of cardiologists. Hannan *et al.* (1997) conduct a similar survey of cardiologists in New York State and find that 67 percent of cardiologists thought the reports were very or somewhat useful and 38 percent said the data had "somewhat/very much" affected their referral patterns.¹⁴

Studies that rely on observed consumer behavior find more evidence for an effect of quality reporting on aggregate market share (Cutler, *et al.*, 2004; Dranove and Sfekas, 2008; Mukamel and Mushlin, 1998). Mukamel and Mushlin (1998) find that hospitals and surgeons in the New York market with lower RAMR for CABG surgery had increased growth in market share after report cards were released. Cutler *et al.* (2004)

¹² Similar report card programs for cardiac surgery are now in use in many states including California, Massachusetts, Florida and New Jersey as well as at the country level in the United Kingdom (Steinbrook 2006).

¹³ Reports based on 1990-1993 data were constructed and released between 1992 and 1995. However, these reports are no longer available and discussions with experts suggest that these data and the reports were not widely distributed. Furthermore, the risk-adjustment measures differed from later reports.

¹⁴ It should be noted that these studies were conducted in the early years of both efforts and may suffer from some reporting bias as a result. The introduction of quality reporting was highly contentious in New York and Pennsylvania (Bumiller, 1995). To the extent that surgeons or cardiologists had an opinion on the topic there is a concern that their answers reflected their preference for or against quality reporting.

find a statistically significant reduction of 5 surgeries per month (10 percent of the average hospital's volume) following a low quality indication. Dranove and Sfekas (2008) estimate a discrete choice model that controls for consumers' beliefs about provider quality prior to the release of report cards. The authors find a significant effect of new information on hospital market share.

Another line of research on supplier response to quality reporting (particularly in economics) has focused not on quality improvement (effort, investments, etc.) but on selection against sicker patients. In their survey, Schneider and Epstein (1998) found that 63 percent of surgeons report reduced willingness to operate on severe patients and 59 percent of cardiologists report having difficulty finding a surgeon for their more severe patients. Dranove *et al.* (2003) compare outcomes for cardiac patients in the Medicare population in New York and Pennsylvania with those in locations without report cards. They find patients had better matches with providers— a gain from the release of information— as well as selection by surgeons against sicker patients, higher resources use and worse outcomes.¹⁵ Fong (2008) considers similar selection behavior in a theoretical setting. She proposes a set of optimal rules to minimize selection while still gaining value from quality reporting. The presence of selection behavior (both adverse selection and moral hazard) means the socially optimal mechanism (scoring rule) relies on quality reporting as well as tenure and probation for low performance. While this paper does not bear on the existence of such effects, her findings underscore the important role supplier objectives should play in determining optimal policy.¹⁶

Taken together the evidence to date is paradoxical in the context of the standard model. Survey evidence and, to a slightly lesser extent, revealed preference suggest little demand response. On the other hand, a review of the medical literature finds, "...there is evidence that the public disclosure of death rates associated with surgery in New York and other states has contributed to reductions in operative mortality..." (Steinbrook,

¹⁵ The aggregate welfare effect of report cards during their study period was negative.

¹⁶ It should be noted that studies of selection were conducted in the early years of quality reporting and, as such, may not capture the long run impact of quality reporting on selection incentives. This is particularly relevant if selection is driven by mistrust of the risk-adjustment system. As physicians become more familiar with quality reporting and risk adjustment the incentives or willingness to select against sicker patients may decline.

2006). Applying the mixed incentive model in this paper can explain these paradoxical findings.

3.2 CABG Surgery

CABG surgery is one of a range of possible treatments for coronary artery disease, a condition in which a patient's blood flow to the heart is compromised by narrowing of the coronary arteries. The severity and symptoms of the disease vary with the degree of obstruction. Cardiac catheterization, a process that allows a cardiologist to image the blockage(s), is used to assess the extent of the disease and determine the appropriate treatment regime. Patients can be managed medically using drugs (beta blockers, Aspirin, ace-inhibitors, etc.) or surgically with either PTCA or CABG.

Diagnosis and treatment of a patient with coronary disease is often an integrated effort requiring both a primary care physician and a cardiologist. If a surgical intervention is decided upon the patient must then choose between angioplasty and CABG and select a cardiac surgeon. All of these choices depend not only on patient characteristics but also on the incentives facing their agent in the choice, the cardiologist.¹⁷ Angioplasty, a process by which a balloon tipped catheter is used to clear blockages, was invented in the 1970s but began to see more widespread use in the 1990s, in part because it is less invasive than bypass and relatively well reimbursed (Cutler and Huckman, 2003). Cardiac bypass surgery is the most invasive treatment for cardiovascular disease. After opening the chest wall, the surgeon creates a bypass around the blocked coronary artery using either internal mammary arteries or arteries from the leg. Patients are supported during the procedure by a heart and lung bypass machine.¹⁸ Recovery after CABG often requires a stay of several days in the hospital Intensive Care Unit (ICU).

¹⁷ A number of papers have focused on positive and normative implications of physician norms in treatment choices (see Frank and Zeckhauser (2008) for a recent discussion). Chandra and Staiger (2007) estimate a Roy model for geographical specialization in the treatment choice between medical management and technologically intensive approaches to treatment for acute myocardial infarction (AMI) and show that local patterns influence the decisions and appropriateness of treatment. Afendulis and Kessler (2007) show that vertical integration of cardiologists into invasive procedures (the rise of invasive cardiologists who can perform angioplasties) leads to changes in treatment decisions at the margin between medical treatment and angioplasty as well as between angioplasty and CABG.

¹⁸ Recent innovation has led to "off pump" surgeries in which the procedure is performed while the heart is still beating.

The production function for CABG surgery is complex and determined not only by the attending surgeon but also by a team of physicians and support staff (Edmondson, *et al.*, 2001). The physicians required for a CABG procedure include a cardiac anesthesiologist and the operating surgeon. In addition, a perfusionist, nursing and other support staff play an important role in the procedure itself and the follow up care as the patient recovers. A well studied and widely documented effect in this market is the presence of a volume-outcome relationship. This is generally attributed to learning-by-doing, though the endogeneity of volume raises the alternate mechanism of selective referral (Arrow, 1962; Gowrisankaran *et al.*, 2006; Gaynor, *et al.*, 2005; Huckman and Pisano, 2006; Ramanarayanan, 2007).

3.3 Data

Data were obtained from the Pennsylvania Health Care Cost Containment Council (PHC4) and contain observations of for 89,406 CABG surgeries performed in the state of Pennsylvania in 1994-95, 2000 and 2002-03 (PHC4, 1994, 1995, 1999, 2002, 2003). Each observation includes information on the surgeon performing the surgery, the hospital at which the surgery was performed, patient demographics, a set of patient comorbidities, the patient's home zip code, data on the payer type, and a set of outcome variables.¹⁹ The outcome of interest in this paper is inpatient mortality. In addition, I merge data on surgeon tenure in the Pennsylvania market.²⁰ This is intended to capture the life-cycle nature of returns to quality as well as costs that are associated with age of the surgeon. To account for the fact that academic surgeons may differ in their incentives, I also gather data on the number of publications for each surgeon by 1995 and the number of citations to those publications.

I compute a measure of risk adjusted mortality to capture surgeon performance using the standard approach used by PHC4. Each observation includes a dummy variable equal to 1 if a patient died in the hospital during or immediately following surgery. The log probability of death is computed as follows:

¹⁹ Patient characteristics include age, indicators for cardiogenic shock, concurrent angioplasty, complicated hypertension, dialysis, female gender, heart failure, and prior CABG or valve surgery.

²⁰ This is computed based on the date the surgeon was first licensed in the State. These data are gathered by matching surgeon names with the state database of license information available at <http://www.dos.state.pa.us/bpoa/cwp/view.asp?a=1104&q=432785>.

$$\ln\left(\frac{\Pr(MORT_{i,s,h} = 1 | x_i)}{1 - \Pr(MORT_{i,s,h} = 1 | x_i)}\right) = \beta_0 + \beta_1 \cdot X_i + \varepsilon_{i,s,h} \quad (4)$$

where i indexes patient, s surgeon and h hospital. $MORT$ is the indicator variable that equals 1 if the patient died in the hospital. This model is estimated for each report card period (1994-95, 2000, 2002 and 2003). The fitted values are obtained for each patient to form a predicted probability of death– the Expected Mortality Rate (EMR). For each surgeon I then compute a measure of risk adjusted performance (RAMR):

$$RAMR_{s,h} = \left(\frac{OMR_{s,h}}{EMR_{s,h}}\right) OMR_{PA} \quad (5)$$

where the risk adjusted, expected and observed mortality rates for each surgeon s or hospital h are $RAMR$, EMR and OMR respectively. Risk adjustment is accomplished by dividing the actual number of fatalities by the expected number of deaths conditioning on the actual patients selecting surgeon s or hospital h . This ratio is then normalized by multiplying it by the statewide average mortality rate.

Table 1 contains summary statistics for key variables in the data. As angioplasty gains market share, substituting for CABG, the statewide volume declines (this also occurred in neighboring states, see Cutler, *et al.*, 2010). The performance improvement over time is also apparent in the reductions in the mean $RAMR$. Between 1994 and 2003 the mean surgeon $RAMR$ dropped 42 percent from a rate of 3.42 percent to 2 percent. Figure 3 plots quarterly mean $RAMR$ and unadjusted mortality highlighting the drop following the release of quality report cards in 1998. Comparing performance in 1994-95 to the period between 2000 and the second quarter of 2002 (quarters 24 to 33 in the graph) shows a decline in quarterly average $RAMR$ and OMR , though these results are noisy from quarter to quarter.

Table 2 presents a matrix of transition probabilities between quintile measures of surgeon quality (5 indexes the highest quality and 1 the lowest in each period) in 1995 and 2000. The evidence suggests that it is feasible for surgeons to improve performance regardless of baseline quality. Of the worst performing surgeons in 1995, 8 percent were in the top 20 percent in 2000 while 19 percent remained in the lowest performing group. Surgeons in the 3rd quintile of performance in 1995 were equally likely to be in the highest quality as the lowest quality quintile by 2000 (16 percent in both cases).

Transition probabilities are computed using the full sample of surgeons. Looking at the right most column of Table 2 it is also clear that exit is substantial, particularly in the highest and lowest quality (quintiles 5 and 1 respectively).²¹ For this reason subsequent analyses are limited to surgeons observed in both the pre- and post- reporting period.

4. Econometric Models and Results

4.1 Identifying Intrinsic Incentives Using Information

To measure the effect of new information on surgeon choices due to intrinsic incentives, I construct an estimate for the *new* information provided by quality reporting that was available only to surgeons but not observed by consumers. The empirical question is whether surgeons who receive new information that is unrelated to demand differ in their subsequent quality improvement.

To construct the information measures I assume that, in the absence of data, a surgeon forms beliefs about performance by observing successes and failures— his or her own inpatient mortality rate. Surgeons do not, however, know with great certainty whether a patient is likely to have died due to underlying severity. The introduction of risk adjustment provides information to surgeons on the true difficulty of their cases— the expected outcome had the average surgeon in the state handled the case. The degree to which this confirms or differs from surgeons’ priors determines the amount of new information contained in the report for each surgeon.

I approximate this empirically as a non-linear function of the difference between a surgeon’s mean pre-report card RAMR and OMR: $f(RAMR_{pre} - OMR_{pre})$. A surgeon with a larger difference between his RAMR and his observed mortality, regardless of the level of each, is provided more information by the introduction of quality reporting. Figure 4 presents a histogram of the frequency of this measure. Information contained in the risk-adjustment appears to be roughly normally distributed around zero. A substantial

²¹ Glazer and McGuire (2008) develop a theory of surgeon response to quality reporting and argue that for some surgeons the introduction may not satisfy their participation constraint, leading them to leave the market. This is consistent with the high exit among poor performers but it less clear why it should affect the highest quintile of quality. Future work should consider the impact of reporting on exit and the possibility that exit choices can be used to recover structural parameters of surgeon utility.

share of surgeons have a RAMR that differs from their OMR, providing support for the empirical strategy.

I begin the analysis by plotting the difference between surgeon RAMR and OMR for 1994-95 against both a measure of extrinsic rewards from this new information (the percentage change in surgeon volume from 1995 to 2000) and the percentage change in RAMR over the same period. These results are presented in Figure 5. I fit the data using a Kernel-weighted local polynomial smoother (Fan, 1996) to estimate a nonparametric surgeon response to new information.

The lower curve plots the relationship of the difference between surgeon RAMR and OMR in 1994-95 and the change in surgeon volume between 1995 and 2000. The curve is remarkably flat, suggesting that this measure of new information did not lead to large changes in demand. The upper curve in Figure 5 plots the same information measure against surgeons' change in RAMR following reporting (computed by taking the difference relative to 1995 levels so positive values represent quality improvement). The curve is u-shaped centered near zero. These results suggest that RAMR remained roughly the same between 1995 and 2000 for surgeons receiving no new information. Moving away from zero in either direction we see that increased information led to larger improvements in quality.

The effect increases monotonically in each direction over most information ranges. Enhanced incentives for surgeons whose performance was worse than expected ($\text{RAMR} > \text{OMR}$) fits easily into the model described in section 2. Performance improvement among surgeons learning they were at high than expected quality is less obvious, but also consistent with predictions. If the shape of the intrinsic utility function is such that the posterior slope is greater than the prior, then improved information can enhance incentives, regardless of the sign of the change.²² An example of this type of effect is a surgeon who desires to be the best. Learning that he or she is better than expected and closer to his/her objective increases the incentive to improve.

To implement this identification strategy in a testable model requires additional parametric assumptions. I saturate the model by separating the magnitude of new

²² While I do not formally rely on this as a test, responding to information that performance is better than expected by improving performance is inconsistent with the behavior predicted by a standard model of profit maximization.

information into quintiles of the difference between RAMR and OMR, indexed by n . Surgeons in the bottom two quintiles (groups 1 and 2) receive information that they are worse than they thought they were (i.e. RAMR is higher than their OMR). Surgeons in the middle quintile (percentiles 40 to 59 of the distribution and the omitted category in estimation) receive no new information (RAMR roughly equal to OMR) and the surgeons in the top two quintiles (groups 4 and 5) receive new information that suggests they are better than expected. To establish that this information measure does not alter demand for surgeons, I estimate the following model among the surgeons observed both before and after the release of quality report cards:

$$\Delta q_s = \alpha + \beta_n \sum_{n=1}^5 I_{n,s} (RAMR_{pre} - OMR_{pre}) + X_s + X_h + \varepsilon_{s,h} \quad (6)$$

where the dependent variable is the change in surgeon s ' volume between the pre- and post-report card period. X_s is a set of surgeon level observables, X_h is a set of hospital level controls and ε_s is an i.i.d., mean zero error term. The coefficient β_n captures the effect of differing levels of report card information (indexed by n) on the outcome variable. If this measure is unrelated to extrinsic rewards I expect that it has no effect on demand, captured in the joint hypothesis $H_0 : \beta_1 = \beta_2 = \beta_4 = \beta_5 = 0$.

Results from estimating equation (6) using data from 1994-95 and 2000 are presented in Table 3. The measure of information gained from quality reporting does not lead to a significant change in volume. For each specification I cannot reject the joint hypothesis H_0 that all of the information variables are jointly equal to zero.

These results suggest that the information contained in the risk adjustment underlying quality reporting was unrelated to consumer response. Because risk adjustment does, however, provide information to surgeons on their own performance and that of their peers I use this to test for an intrinsic response. The primary estimating equation is as follows:

$$\Delta \theta_s = \alpha + \lambda \Delta \Pi_s + \xi_n \sum_{n=1}^5 I_{n,s} (RAMR_{pre} - OMR_{pre}) + X_s + X_h + \varepsilon_{s,h} \quad (7)$$

where the dependent variable is the change in surgeon quality (RAMR) between the pre- and post-reporting periods. Information, surgeon and hospital observables are captured as

above. If information alters quality due to intrinsic incentives, additional data from report cards should produce performance improvement. The hypothesis $H_1 : \xi_n > \xi_3$ is a test for intrinsic incentives associated with the information contained in each group n . With quality as the dependent variable, changes in profit incentives due to quality reporting also enter the model, captured by $\Delta \Pi_s$. Estimating (7) thus requires constructing a measure of the change in extrinsic incentives (the relative slope of MR in Figure 1) induced by quality reporting. This is the focus of the next section of the paper.

One important concern in estimating (7) is the potential for mean reversion that could bias the estimated effect of both intrinsic and extrinsic incentives on changes in surgeon quality. Suppose that high RAMR surgeons revert to lower RAMR by 2000 solely due to statistical error or other unobserved factors. If this is true, surgeons who appear to gain more information from quality reporting and to be worse than expected will also be observed to have lower RAMR in 2000 solely due to measurement error. Alternately, surgeons with very high RAMR in the pre-period could be predicted to have large changes in profit incentives. If high RAMR reverts to the mean, this would also bias estimates of λ . To account both of these possibilities, I include a surgeon's average RAMR in 1994-95 in the vector X_s . This eliminates mean reversion in the estimated effect of information and controls for the mechanical relationship between a surgeon's baseline quality and the implied returns to quality improvement.²³ Despite these empirical advantages, I also present results without the pre-reporting RAMR because any correlation between RAMR and the intrinsic information measure will bias the β_n coefficients towards zero.

Finally, technological change over time could affect the model. I account for this in two ways. First, risk adjustment is computed for each period and, subsequently, incorporates changes in the average ability to treat a patient with a given comorbidity. That is, a surgeon's risk adjusted performance in period t is defined in terms of the

²³ I also note that, even in the absence of 1994-95 RAMR control, if I have accurate measures of quality concerns about mean reversion will be minimized (i.e. the error term on the estimated quality in each period is small relative to the treatment effect). Because I have observations from a relatively long pre- and post- period (I observe an average of 240 surgeries per surgeon in the pre- and 160 in the post-reporting period), I expect the variance in estimated quality in each period to be small relative to the magnitude of the effect I am trying to identify (RAMR change).

expected mortality conditional on patient covariates given the current period technology. This controls for the majority of inter-temporal improvement in technology. A second concern is that technological change alters quality due to unobserved factors uncorrelated with the risk adjusters. I assume that this effect is equal across the panel of surgeons, conditional on surgeon and market observables. Under this assumption, the intercept α captures any additional unobserved technological change. Furthermore, specifications that include a surgeon's baseline RAMR allow the impact of technology to enter flexibly across the *ex ante* quality distribution. If, for example, a new technique or device were introduced that is more effective for lower performing surgeons (or they adopt this later than surgeons observed to be high quality) it is captured in the relationship between mean 1994-95 RAMR and changes in quality.

4.2 Surgeon Quality and Patient Demand

In order to more precisely relate demand side factors to surgeon quality choices following reporting, I estimate a structural model of consumer demand. Relying on parameter estimates for patients' utility, I can simulate alternate information environments. I model patients' discrete selection of a surgeon allowing for factors that are unobserved (to the econometrician) but alter choice. These enter as random coefficients. (Berry, Levinson and Pakes, 1995; Nevo, 2000; Train 2003). I present a brief discussion of the key variables and refer the reader to Kolstad (2009) for a detailed description of the structural model.

Each patient selects from the set of surgeons in his or her Hospital Referral Region (HRR). The utility for patient i from choosing a given surgeon s is a function of cost (both monetary and time costs), expected health improvement (capturing all components of quality) and an error term. Indirect utility to consumer i who selects surgeon s is:

$$u_{i,s,h} = g(X_i, \eta_i, Z_{s,h}, \theta_{s,h}; \rho) + \varepsilon_{i,s,h} \quad (8)$$

where X_i and η_i are vectors of observed and unobserved patient characteristics all of which lead to differences in taste. $Z_{s,h}$ is a K-dimensional vector of hospital and surgeon characteristics not directly related to expected health. $\theta_{s,h}$ is the expected quality (gains

in health) for surgeon s at hospital h . Finally, $\varepsilon_{i,s,h}$ is an iid error term with a type-1 extreme value distribution and ρ is a vector of parameters. A patient selects surgeon s at hospital h if and only if $u_{i,s} > u_{i,-s} \forall -s \neq s$. Indirect utility in (8) is derived directly from a quasilinear utility function without wealth effects or prices. Prices do not enter choice because patients are generally well insured (55 percent by Medicare) so the out-of-pocket cost is unlikely to vary in any meaningful way between surgeons.²⁴

Information (and quality reporting) enters the model by altering beliefs about expected health gains from choosing a given surgeon, s . Patients are assumed to develop beliefs based on all available information on surgeon performance (from both formal and informal information sources). Because the choice is likely influenced by an experienced agent (the patient's cardiologist), I control for a measure of the patient-surgeon match assumed to be mediated through agency. This is captured by the deviation between a patient's severity and the lagged average severity seen by surgeon s in the prior quarter. The expected quality if patient i chooses surgeon s at hospital h is:

$$\theta_{i,s,h} = X_i RC_{s,h} + \eta_i RC_{s,h} \quad (9)$$

RC is a vector of surgeon and surgeon-hospital characteristics observed by the patient. Included in RC is a continuous measure of surgeon performance— prior quarter RAMR— and dummies for discrete quality ratings included in the report card (i.e. better, worse or as expected mortality based on patient severity). Substituting back into (8) the patient's utility function is:

$$u_{i,s,h} = X_i Z_{s,h} + X_i RC_{s,h} + \eta_i RC_{s,h} + \varepsilon_{i,s,h} \quad (10)$$

Individual choice is thus a function of observed (to the econometrician) patient and surgeon attributes as well as unobserved factors that alter patient response to quality information.²⁵ The probability that patient i chooses surgeon s at hospital h is:

$$P_{i,s,h} = \int \frac{e^{X_i Z_{s,h} + X_i RC_{s,h} + \eta_i RC_{s,h}}}{\sum_{-s \in S} e^{X_i Z_{-s,h} + X_i RC_{-s,h} + \eta_i RC_{-s,h}}} \phi(\eta_i | b, W) d\eta_i \quad (11)$$

²⁴ To test this assumption I estimate a specification of the choice model using data only from Medicare and Private fee-for-service patients (those who are known not to face any constraints on choice). This has little effect on first stage parameter estimates for demand or second stage estimates of surgeon response.

²⁵ Contained in such unobserved influences are agency, insurance network constraints and patient-surgeon matching. Kolstad (2008) considers the implications of these factors for the positive and normative choices of surgeon by patients with and without information.

where the unobserved components of utility are distributed according to the distribution $\phi(\eta_i|b,W)$ that is known up to a mean and covariance, b and W , to be estimated. Using this expression, I fit the data using simulated maximum likelihood (Train, 2003).

Estimates for the demand system are computed by solving analytically for the logit choice probabilities and integrating out the random taste distribution by taking draws from the joint distribution of unobserved terms.

Results are presented in Table 4. I estimate the model with time period interactions for the period before (1994-95) and following the release of the 1998 report card but prior to the release of the ensuing report cards (2000 and the first two quarters of 2002). The estimates are generally consistent with expectations. Travel distance enters choice significantly, as does surgeon quality. The effect of the introduction of quality reporting on demand can be seen in the interactions of the quality variables with the dummy for post-quality reporting. Columns 3 and 4 include controls for agency induced surgeon-patient matching. Because these effects are significant in all specifications, I include these controls in all subsequent analysis.

4.3 Calibrating Report Card Related Extrinsic Incentives

To calibrate the magnitude of the demand side incentives facing each surgeon, I simulate a series of counter-factual scenarios. Extrinsic incentives are captured by the measured returns to quality identified by the demand system.²⁶

The impact of information on profit is measured empirically by computing a predicted choice probability under alternate information scenarios— with and without quality reporting. The estimate for the report card induced change in profit is:

$$\begin{aligned} \Delta \Pi_s = & \lambda_1 \left[\sum_{i=1}^I \hat{P}_{i,s,h,t}(X_i^t, Z_{s,h}^t, \theta_{s,t} | \Omega^{pre}) - \sum_{i=1}^I \hat{P}_{i,s,h}(X_i^t, Z_{s,h}^t, \theta_{s,t} | \Omega^{post}) \right] + \lambda_2 I_s^{Dem\uparrow} \\ & + \lambda_3 I_s^{Dem\uparrow} * \left[\sum_{i=1}^I \hat{P}_{i,s,h,t}(X_i^t, Z_{s,h}^t, \theta_{s,t} | \Omega^{pre}) - \sum_{i=1}^I \hat{P}_{i,s,h}(X_i^t, Z_{s,h}^t, \theta_{s,t} | \Omega^{post}) \right] \\ & + I_{hrr} + \varepsilon_{i,s,h,t,hrr} \end{aligned} \quad (12)$$

²⁶ A concern in this approach is that with a structural interpretation any misspecification will alter results. In this setting, however, it would be difficult to measure extrinsic returns without doing this. I also estimate the model with alternate demand specifications that relax a number of the assumptions and find little effect on the results.

where $\sum_{i=1}^I \hat{P}_{i,s,h}(X_i^t, Z_{s,h}^t, \theta_{s,t} | \Omega^{RC})$ is the sum of the fitted choice probabilities for all patients i receiving surgery in the HRR in which surgeon s practiced in period t , a measure of expected demand. Ω^{RC} indexes the information environment (pre- or post-report cards) that define the consumer utility parameters. $I_s^{Dem\uparrow}$ is a dummy variable equal to 1 if surgeon s has an expected demand using post-reporting demand parameters that is greater than expected volume using pre-report card demand (a windfall profit from quality reporting).²⁷ Including $I_s^{Dem\uparrow}$ and interacting it with the change in predicted volume allows the model to flexibly capture potentially discontinuous changes in incentives at a surgeon's current volume due to income effects (an issue discussed in more detail below). I also include market fixed effects to control for all time invariant market level factors that influence the profitability of quality. In this model, differences in the profitability of quality for surgeons are identified by differences in geographic locations, the distribution of patients' tastes, the competitive structure of the market and (in models without a control for 1994-95 RAMR) baseline quality. For example, a surgeon who is at a low quality level in the pre-reporting period and faces few high quality competitors and patients who (given their locations and attributes) respond strongly to quality reporting has a greater return to improving performance.

Surgeon response to pecuniary rewards for quality depends the strength of income and substitution effects (McGuire and Pauly, 1991). If income effects dominate, reductions in demand at a surgeon's current level of quality should alter extrinsic incentives more than the opportunity to gain additional patients by improving performance. Losing volume (profit) increases the marginal utility of an additional patient, mediated through the marginal utility of income.²⁸ On the other hand, if income

²⁷ The difference in volume is measured with respect to the predicted volume pre-quality reporting.

Surgeons expecting to *gain* volume after quality reporting ($I_s^{Dem\uparrow} = 1$) have *negative* estimates for the volume difference ($\sum_{i=1}^I \hat{P}_{i,s,h,t}(X_i^t, Z_{s,h}^t, \theta_{s,t} | \Omega^{pre}) - \sum_{i=1}^I \hat{P}_{i,s,h}(X_i^t, Z_{s,h}^t, \theta_{s,t} | \Omega^{post}) < 0$)

²⁸ The extreme version of this argument is consistent with the classic target income hypothesis (see McGuire (2000) for a review of a long literature). If surgeons are at an equilibrium volume (income) prior to quality reporting expecting to drop below that volume should lead to large utility losses and enhanced incentives to make investments to return to the target level. On the other hand, returns to volume growth beyond their current level are relatively low.

effects are not sufficiently strong, demand side incentives are better measured by computing the *ceteris paribus* returns to a reduction in RAMR. Equation (12) models both effects. The income effect is captured by the coefficient λ_2 . When estimating (7) including controls for baseline RAMR, differences in expected gains or losses in volume are identified solely by demand side responsiveness. A significant quality improvement by surgeons expecting a one patient *increase* in volume beyond their current level provides an estimate for substitution of costly effort towards reducing RAMR due to the increased marginal profitability of quality after reporting. This is captured by the coefficient λ_3 .

Using the approach in (12) and the full first stage demand system (including agency controls), the sample mean predicted quarterly change in volume is -.14 with a standard deviation of 3.09. The average surgeon could expect a very slight increase in demand due to quality reporting (changes are estimated with respect to the pre-reporting predicted demand so a negative number means the predicted volume using post-reporting parameter estimates is larger). The relatively small average effect is offset by a large variance across surgeons. Figure 6 plots the frequency of quarterly differences between surgeons' pre- and post-report card predicted volumes. Breaking the impact down by positive and negative demand effects, the mean predicted change in demand is 1.3 patients for surgeons losing volume in a given quarter and a gain of 2.5 patients for surgeons gaining volume. Compared to the average quarterly volume of 30 surgeries in the sample these effects are economically significant. Huckman (2006) finds that the marginal profit from a CABG surgery is \$6900.²⁹ Thus surgeons facing the average decline in demand expected a reduction in annual profit (joint with the hospital at which they practice) of \$35,000 and surgeons gaining volume expected an annual profit increase of \$69,000.

4.4 Mixed Incentives and Quality Improvement

Using this measure of profit incentives and the information contained in the risk adjustment, I estimate equation (7). The coefficients λ and ξ_n capture the effect of a

²⁹ This is computed using data from New York and is a measure of total marginal profit to the hospital and the surgeon. Average profit is \$2000 for a Medicare beneficiary and \$3900 across all payer types.

change in profit and a change in information respectively on the change in surgeon RAMR between 1994-95 and 2000. I account for measurement error from the first stage demand estimates by using bootstrap simulation to compute the standard errors.

Table 5 presents results. Parameter estimates for ξ_n are positive for all surgeon information groups. The effects are large and significant for information groups 1 and 2 (surgeons learning their performance was lower than expected). I fail to reject the hypothesis H_1 — that information increased intrinsic incentives for quality— for surgeons in groups 1 and 2. In addition, without controlling for surgeons' mean 1994-95 RAMR (columns 1 and 3), the coefficient estimates for group 5 are highly significant— surgeons learning they were much *better* than expected responded by improving performance.

Interpreting the coefficient estimates in Tables 5 suggests that, after controlling for demand side incentives and mean reversion, surgeons who learn significantly more about their own quality and that of their peers and find that they are not performing as well as expected (group 1) improve quality by an average of .57 to .62 percentage points more than surgeons receiving no information from reporting. In columns 1 and 3, those without controls for 1994-95 RAMR, the impact of learning is much larger for the same group, around 2 percentage points. The same learning, but for those only slightly worse than expected, produced an average improvement of between .48 and .52 percentage points of RAMR. The magnitude of the response is relatively similar including a control for pre-reporting quality. Finally, surgeons learning they were much better than expected improved by an average additional .76 to .79 percentage points compared to surgeons gaining no new information, though this effect is ameliorated by including pre-reporting average RAMR.

I next turn to the estimates of the λ coefficients that capture surgeon response to profit incentives. Profit incentives enter with the “right” sign— surgeons facing more responsive patients improve quality by more. The coefficient estimates of λ_1 suggest that for every additional patient a surgeon expected to lose per quarter under quality reporting (given *ex ante* surgeon, competitor and consumer characteristics) they reduced RAMR by an average of .043 to .046 additional percentage points. Estimates of λ_2 are insignificant, suggesting that the response does not drop discontinuously when surgeons expect to

maintain their pre-report card volume (e.g. income effects are relatively small). Consistent with a stronger substitution effect, surgeons expecting to gain patients are also observed to respond to greater report card induced demand response. This can be seen in the significant estimates of λ_3 . In Table 5 the extrinsic response effects are similar even after controlling for baseline quality, suggesting the differences are driven by surgeons responding to patients who are more quality elastic, not merely by the mechanical relationship between current quality and demand. Overall, surgeons expecting to gain more patients due to quality reporting improve by more, regardless of their current volume.

5. Discussion

5.1 Comparing Intrinsic and Extrinsic Response to Quality Reporting

The evidence thus far is consistent with a role for both extrinsic and intrinsic incentives in determining surgeons' response to quality reporting. I use the estimates from equation (7) to calibrate the relative magnitudes of these incentives by fitting the model under alternative incentive scenarios (i.e. with and without extrinsic and intrinsic incentives). For brevity, I only report results from the analysis using full controls (the relevant coefficient estimates are in column 4 of Table 6), though I note that the findings are robust to alternate specifications.

Overall, the model predicts changes in RAMR relatively well. The coefficient of correlation between the predicted and actual change in RAMR for each surgeon is .89. However, there is a tendency to over-predict quality improvement. This is due in part to very large predicted quality improvements based on observations in the tails. The full model predicts a mean improvement of 1.1 percentage points of RAMR compared to the actual average of .72; the median of .67 is closer to the median observed response of .79. To account for this bias, I use predicted improvement to compare the role of extrinsic and intrinsic incentives. As long as the error is not correlated with the intrinsic or extrinsic incentive measures, the predicted changes are consistent.

I begin by computing the predicted change in quality had report cards not been implemented. Constraining $\lambda = 0$ and $\xi_n = \xi_3$ the average predicted change in RAMR between 1994-95 and 2000 is .742 percentage points (median=.31). Allowing for

extrinsic incentives only (incorporating λ and continuing to set $\xi_n = \xi_3$), the predicted quality improvement is .841 percentage points of RAMR (median=.41). Surgeons' response to profit incentives led to an additional decline in RAMR of .09 percentage points. Compared to the mean statewide RAMR in 1994-95 (3.42 percent), this constitutes an additional 2.6 percent decline in mortality.

Allowing only information-induced intrinsic incentives to enter the model (i.e. setting $\lambda = 0$) the model predicts a market wide average RAMR that is an average of 1.04 percentage points lower due to quality reporting (median=.66). The intrinsic response to quality reporting alone led to an additional .3 percentage point decline in RAMR or 8.7 percent lower statewide RAMR. Comparing the two effects, the impact of quality reporting mediated through intrinsic response is roughly three times as large as the response to profit incentives.

5.2 Sensitivity and Alternate Explanations

Perhaps the most relevant threat to the validity of the model is the potential for dynamic pecuniary returns to quality improvement. If current period investments are related to future demand, intrinsic incentives may in fact be quality improvement in response to expected future compensation. To provide a simple test for such long run pecuniary incentives, I estimate the following model:

$$q_{s,t} = \alpha + \Delta\theta_{s,t-y,t-n} + q_{s,t-z} + \theta_{s,t-n} + X_s + X_h + \varepsilon_{s,t} \quad (13)$$

where $q_{s,t}$ is the quantity of patients seen by surgeon s in period t and $\Delta\theta_{s,t-y,t-n}$ is the change in surgeon quality between period $t-n$ and $t-y$. z , n and y are positive integers satisfying: $n > y$. Including the quantity of surgeries by surgeon s in all periods, $t-z$, controls for the full stock of demand and quality effects prior to period t . Because I am primarily interested in surgeons' response to quality reporting introduced in 1998, I estimate (13) using surgeon volume in 2003 as the dependent variable and measure the change in quality as the difference between a surgeon's 1994-95 and 2000 RAMR. I control for baseline quality and surgeon volume in 1994, 1995, 2000 and 2002. The estimated coefficient on $\Delta\theta_{s,2000,1994-95}$ is -1.98 (s.e.=1.52). The negative and statistically insignificant estimate suggests that surgeons were unlikely to expect changes in quality to

produce pecuniary returns in the future. If anything, it appears that greater quality improvement led to a slightly smaller number of patients in the long run. I also estimate (13) with the change in quality between 1994 and 1995 as an independent variable. The estimated coefficient on $\Delta\theta_{s,1995,1994}$ is .43 (s.e.=.6), again an insignificant effect of the change in quality on subsequent annual volume, in this case 8 years hence.

I turn next to a sensitivity test for the specific model that information alters surgeons' behavior by altering their knowledge about a reference group. The base specification (equation (7)) models surgeons' learning using the risk adjustment based on statewide performance. If surgeon incentives are related to their knowledge about the performance of their peers I also expect that new learning about other peer groups will alter the intrinsic incentives. The most easily defined peer group is the set of other surgeons practicing at the same hospital. To test for an effect of learning about within-hospital peer performance, I re-estimate (7) and include a measure of the difference between surgeon s ' RAMR in the pre-reporting period and the top (lowest RAMR) and bottom (highest RAMR) surgeon at the hospital at which they practiced in 1994-95. The model is:

$$\Delta\theta_s = \alpha + \lambda\Delta\Pi_s + \xi_n \sum_{n=1}^5 I_{n,s} (RAMR_{pre} - OMR_{pre}) + \eta(RAMR_{pre} - \min RAMR_{h,s,pre}) + \mu(RAMR_{pre} - \max RAMR_{h,s,pre}) + X_s + X_h + \varepsilon_{s,h} \quad (14)$$

Results are presented in Table 8 with and without market fixed effects. In both models, coefficient estimates for μ are negative and significant, suggesting a surgeon learning he is further from the worst surgeon at his hospital improves RAMR by relatively more. The coefficient estimates for η are not significant but the sign is also consistent with enhanced incentives for surgeons learning they are closer to being the top performer within their hospital. These results support the earlier findings, that report card learning alters incentives by improving knowledge about a reference group. Moreover, adding these controls produces larger coefficient estimates for ξ_n , suggesting that surgeons compare performance both to their peers at their own hospital and to the full statewide population.

6. Conclusion

The impact of information on equilibrium quality is mediated through demand insofar as suppliers choose quality levels to maximize profit. In this paper I present an alternate model in which reference intrinsic utility also determines surgeons' willingness to make costly quality improvements. Information (quality reporting) alters a surgeon's beliefs about their own quality level as well as the full distribution of the reference set. If intrinsically motivated suppliers update their beliefs in a Bayesian fashion, the improved posterior beliefs alter the shape of the intrinsic utility function and thus a surgeon's marginal incentive to improve quality.

The risk-adjustment model that underlies quality report cards provides a simple way of identifying the magnitude of the new information provided to surgeons and its effect on performance. Surgeons who gain more information about their performance relative to their peers (from risk adjustment) improve significantly more. A structural patient choice model allows me to estimate the profit incentives from quality reporting. Conditioning on extrinsic incentives, the intrinsic response to information leads to significant declines in RAMR and is large relative to the response to profit motives.

The results of this paper add to the literature on the behavioral response to improved information in market. However, I note some shortcomings and directions for future work. I consider a specific setting – cardiac surgery – that may not generalize to other markets or professions. The field of medicine, perhaps more than any other, relies on not-for-profit incentives to correct market failures. As trainees, physicians also become accustomed to evaluation mechanisms based on relative performance (grades and MCAT for medical school, board scores for residency and licensure, etc). For these reasons, the large intrinsic effects (at least compared to the response to profit incentives) I find among cardiac surgeons is best interpreted as an upper bound across all fields.

Future work that empirically models mixed incentives in other fields or that explicitly considers dynamic pecuniary rewards will build on these findings. The results do, however, provide an empirical first step and guidance on the potential role for mixed incentives in determining skilled professionals' effort and investment. Further, they contribute to our understanding of the effects of quality reporting in health care and inform policy making in this market.

References

- Akerlof, G.A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*. 84(3), 488-500.
- Akerlof, G.A. and Kranton, R.E. (2005). Identity and the Economics of Organizations. *Journal of Economic Perspectives*. 19(1), 9-32.
- Arrow, K. (1963). Uncertainty and the Welfare Economics of Medical Care. *American Economic Review*, 53, 941-73.
- Berry, S., Levinsohn, J. and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63 (4), 841-890.
- Besley, T. and Ghatak, M. (2008). Status Incentives. *American Economic Review: Papers and Proceedings*. 98(2), 206-211.
- Bolton, G.E. and Ockenfels, A. (2000). A Theory of Equity, Reciprocity and Competition. *American Economic Review*. 81, 1096-1136.
- Bumiller, E. (1995). Death Rate Rankings Shake New York Cardiac Surgeons. *New York Times*, September 6.
- Burns, L.R., and Wholey, D.R. (1992). The Impact of Physician Characteristics in Conditional Choice Models for Hospitals. *Journal of Health Economics*, 11 (1), 43-62.
- Chatterji, A.K. and Toffel, M.W. (2007). Shamed and Able: How Firms Respond to Information Disclosure. Working Paper, Harvard Business School.
- Charness, G. and Rabin, M. (2002). Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics*. 117, 817-869.
- Cutler, D.M., Huckman, R.S. and Kolstad, J.T. (2010). Input Constraints and the Efficiency of Entry: Lessons from Cardiac Surgery. *American Economic Journal: Economic Policy*. 2(1), 51-76.
- Cutler, D.M., Huckman, R.S. and Landrum, M.B. (2004). The Role of Information in Medical Markets: An Analysis of Publicly Reported Outcomes in Cardiac Surgery. *American Economic Review*, 94 (2), 342-346.
- Chernew, M.E., Gowrisankaran, G. and Scanlon D.P. (2001). Learning and the Value of Information: Evidence from Health Plan Report Cards. *NBER Working Paper*, #8589.
- Chernew, M.E., Scanlon, D. and Hoy, R. (1998). Insurance Type and Choice of Hospital for Coronary Artery Bypass Graft Surgery. *Health Services Research*, 33 (3), 447-466.

- Dafny, L. and Dranove, D. (2005). Do Report Cards Tell People Anything They Don't Already Know? The Case of Medicare HMOs. *NBER Working Paper*, #1142.
- Dranove, D., Kessler, D., McClellan, M. and Satterthwaite, M. (2003). Is More Information Better? The Effect of "Report Cards" on Health Care Providers. *Journal of Political Economy*, 111(3), 555-588.
- Dranove, D. and Satterthwaite, M. (1992). Monopolistic Competition When Price and Quality and Imperfectly Observable. *RAND Journal of Economics*. 23(4), 518-534.
- Dranove, D. and Sfekas, A. (2008) Start Spreading the News: A Structural Estimate of the Effects of New York Hospital Report Cards. *Journal of Health Economics*. 27(5), 1201-1207.
- Edmondson, A., Bohmer, R. and Pisano, G. (2001). Disrupted Routines: Team Learning and New Technology Implementation in Hospitals. *Administrative Science Quarterly*. 46(4), 685-716.
- Epstein, A. (1998). Rolling Down the Runway: The Challenges Ahead for Quality Report Cards. *Journal of the American Medical Association*, 279, 1691-1696.
- Fan, J. (1996). Local Polynomial Modeling and Its Applications. London: Chapman and Hall.
- Fehr, E. and Schmidt, K.M. (1999) A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics*. 114, 817-868.
- Fehr, E. and Schmidt, K.M. (2006). The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories. In Kolm and Ythier (eds.) *Handbook of the Economics of Giving, Altruism and Reciprocity*. Amsterdam: Elsevier.
- Fong, K.G. (2008). Evaluating Skilled Experts: Optimal Scoring Rules for Surgeons. Mimeo, Stanford University.
- Fung, A., Graham, M. and Weil, D. *Full Disclosure: The Perils and Promise of Transparency*. Cambridge: Cambridge University Press, 2007.
- Gaynor, M. 2006. What Do We Know About Competition and Quality in Health Care Markets? *NBER Working Paper* 12301.
- Gowrisankarin, G., Chernew, ME, and Scanlon, D. (forthcoming). Learning and the Value of Information: Evidence from Health Plan Report Cards. *Journal of Econometrics*.

- Harsanyi, J. (1955). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy*. 98, 1119-58.
- Ho, K. (Forthcoming). The Welfare Effects of Restricted Hospital Choice. *Journal of Applied Econometrics*.
- Jin, G.Z. and Sorensen, A.T. (2006). Information and consumer choice: the value of publicized health plan ratings. *Journal of Health Economics*, 25 (2), 248-75.
- Kolstad, J.T. and Chernew, M.E. (2009). Consumer Decision Making in the Market for Health Insurance and Health Care Services. *Medical Care Research and Review*. 66 (1), 28S-52S.
- Kolstad, J.T. (2009). Essays on Information, Quality and Competition in Health Care Provider Markets. Dissertation, Harvard University.
- Luft, H.S., Garnick, D.H., Mark, C.S., Peltzman, D.H., Phibbs, C.S., Lichtenber, E. *et al.* (1990). Does Quality Influence Choice of Hospital? *Journal of the American Medical Association*, 263 (11), 2899-906.
- McGuire, T.G. (2000). Physician Agency. In Newhouse and Culyer (Eds.) *Handbook of Health Economics*. Amsterdam: Elsevier.
- McGuire, T.G. and Pauly, M.V. (1991). Physician Response to Fee Changes with Multiple Payers. *Journal of Health Economics*. 10(4), 385-410.
- Mukamel, D.B and Mushlin A.I. (1998). An Analysis of Market Share and Price Change After Publication of the New York State Cardiac Surgery Mortality Reports. *Medical Care*, 35 (1), 945-54.
- Mukamel, D.B and Mushlin A.I. (2001). The Impact of Quality Report Cards on Choice of Physician, Hospitals and HMOs: A Midcourse Evaluation. *Joint Commission on Accreditation of Healthcare Organizations*.
- Mukamel, D.B., Weimer, D.L., Zwanziger, J., Gorthy, S.H. and Mushlin, A.I. (2004/2005) Quality Report Cards, Selection of Cardiac Surgeons, and Racial Disparities: A Study of the Publication of the New York State Cardiac Surgery Reports. *Inquiry-Excellus Health Plan*, 41 (4), 435-446.
- Mukamel, D.B., Weimer, D.L, Gorthy, S.H. and Mushlin, A.I. (2002) Quality of Cardiac Surgeons and Managed Care Contracting Practices. *Health Services Research*, 37(5).
- Murdock, K. (2002) Intrinsic Motivation and Optimal Incentive Contracts. *RAND Journal of Economics*. 33(4), 650-671.

- Nevo, A. (2001). Measuring Market Power in the Ready-to-eat Cereal Industry. *Econometrica*, 69 (2).
- Newhouse, Joseph P. (1970) Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital. *American Economic Review*, 60(1): 65-71.
- Pauly, M.V. (2004) Competition in Medical Services and the Quality of Care: Concepts and History. *International Journal of Health Care Finance and Economics*, 4(2): 113-130.
- Pennsylvania Health Care Cost Containment Council. (1998). *Coronary Artery Bypass Graft Surgery—1994-95 Data: Research Methods and Results*.
- Pennsylvania Health Care Cost Containment Council. (2002). *Coronary Artery Bypass Graft Surgery—2000 Data: Research Methods and Results*.
- Pennsylvania Health Care Cost Containment Council. (2004). *Coronary Artery Bypass Graft Surgery—2002 Data: Research Methods and Results*.
- Pennsylvania Health Care Cost Containment Council. (2005). *Coronary Artery Bypass Graft Surgery—2003 Data: Research Methods and Results*.
- Salop, S. and Stiglitz, J. (1977) Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion. *Review of Economic Studies*. 44(3), 493-510.
- Sauermann, H. and Cohen, W.M. (2008) What Makes Them Tick? Employee Motives and Firm Innovation.
- Segal, U. and Sobel, J. (2007). Tit for Tat: Foundations of Preferences for Reciprocity in Strategic Settings. *Journal of Economic Theory*. 136(1), 197-216.
- Schneider, E.C. and Epstein, A.M. (1996), Influence of Cardiac-Surgery Performance Reports on Referral Practices and Access to Care-A survey of Cardiovascular Specialists. *New England Journal of Medicine*. 335(4), 251-256.
- Schneider, E.C. and Epstein, A.M. (1998), Use of Public Performance Reports: A Survey of Patients Undergoing Cardiac Surgery. *Journal of the American Medical Association*. 279(20), 1638-1642.
- Sloan, F.A., (2000). Not-for-profit Ownership and Hospital Behavior. In Newhouse and Culyer (Eds.) *Handbook of Health Economics*. Amsterdam: Elsevier.
- Spence, M.A. 1980. "Product Selection, Fixed Costs, and Monopolistic Competition." *Review of Economic Studies*, 43(2): 217-235.

Steinbrook, R. (2006), Public Report Cards- Beyond Cardiac Surgery. *New England Journal of Medicine*. 355(18), 1847-1849.

Wennberg, J.E., Cooper, M.M. and Dartmouth Atlas of Health Care Working Group. 1999. *The Dartmouth Atlas of Health Care 1999*. Chicago. American Hospital Publishing.

Table 1: Descriptive Statistics by Year

Year	Observations	Surgeons	Hospitals	Mean RAMR*	Mean OMR*
1994-95	18,351	201	43	3.42	3.23
2000	19,594	182	55	2.38	2.20
2002	15,999	187	62	2.02	1.82
2003	15,157	183	63	2.00	1.85

*Surgeon weighted average

Table 2: Transition Matrix for RAMR Quintile Between 1995 and 2000

		2000 RAMR Quintile					
		1	2	3	4	5	Exit
1995 RAMR Quintile	1	0.08	0.11	0.03	0.03	0.13	0.63
	2	0.08	0.24	0.26	0.16	0.08	0.18
	3	0.16	0.13	0.11	0.29	0.16	0.16
	4	0.16	0.11	0.13	0.08	0.18	0.34
	5	0.19	0.05	0.08	0.05	0.16	0.46

Table 3: New Information and Change in RAMR and Volume

	Dependent Variable: Change Volume s 1995 to 2000			
1994-95 Report Card Info (RAMR-OMR) Group				
Much Better than Expected (0-20%)	1.519	(1.315)	1.622	(1.306)
Slightly Better than Expected (20-40%)	1.758	(1.307)	1.700	(1.208)
Slightly Worse than Expected (60-80%)	1.827	(1.307)	1.758	(1.191)
Much Worse than Expected (80-100%)	1.531	(1.312)	1.633	(1.291)
Surgeon License Year (PA)			36.167	(30.214)
Surgeon License Year (PA) Squared			-0.009	(0.008)
Publications			0.035	(0.047)
Change in Hospital RAMR 1995-2000				0.193 (0.208)
R Squared	0.1191		0.1793	0.1988
Observations (surgeon/quarter)	122		122	104

, and * denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Random Coefficients Demand System Parameter Estimates

		Dependent Variable: Log Probability Patient i Selects Surgeon s							
		Mean		S.D.		Mean		S.D.	
Mean Surgeon RAMR s,t-1		-0.001	(0.002)	-0.001	(0.003)	-0.001	(0.002)	0.000	(0.003)
Mean Surgeon RAMR s,t-1*Post RC		-0.008	(0.004) **	-0.001	(0.008)	-0.008	(0.004) **	-0.001	(0.008)
Distance (Miles) i,s		-0.216	(0.006) ***	0.014	(0.011)	-0.212	(0.006) ***	-0.007	(0.005)
Distance (Miles) i,s*Post RC		0.072	(0.010) ***	0.007	(0.013)	0.073	(0.010) ***	-0.004	(0.008)
Distance Squared (Miles) i,s		-0.001	(0.000) ***	0.005	(0.000) ***	-0.001	(0.000) ***	0.005	(0.000) ***
Distance Squared (Miles) i,s*Post RC		-0.002	(0.000) ***	0.000	(0.000) *	-0.002	(0.000) ***	0.000	(0.000)
EMR i-EMR s,t-1						1.084	(0.516) **	0.092	(0.886)
EMR i-EMR s,t-1*Post RC						-3.616	(1.173) ***	0.208	(2.145)
EMR i-EMR s,t-1 Squared						-17.322	(3.483) ***	-0.308	(4.919)
EMR i-EMR s,t-1 Squared*Post RC						-17.107	(9.839) *	-5.032	(14.717)
Low Quality RC Score 1994-95 s		-1.059	(0.121) ***	1.540	(0.154) ***	-0.499	(0.075) ***	0.601	(0.188) ***
Low Quality RC Score 1994-95 s*Post RC		-0.165	(0.087) *	0.238	(0.225)	-0.002	(0.068)	-0.046	(0.238)
High Quality RC Score 1994-95 s		0.283	(0.024) ***	0.017	(0.112)	0.271	(0.024) ***	-0.033	(0.095)
High Quality RC Score 1994-95 s		-0.130	(0.105)	0.277	(0.511)	-0.097	(0.055) *	-0.087	(0.377)
Not in Report Card 1994-95 s		0.137	(0.023) ***	0.043	(0.091)	0.131	(0.023) ***	0.022	(0.090)
Not in Report Card 1994-95 s*Post RC		0.079	(0.042) *	0.064	(0.192)	0.067	(0.049)	0.193	(0.243)
Observations		720,364				720,364			
Log Likelihood		-86606.20				-86570.20			
Sample		1994-95, 2000, 2002 (Q1-2)				1994-95, 2000, 2002 (Q1-2)			

*, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Effect of Report Card Induced Information and Demand on Surgeon Quality

	Dependent Variable: Change RAMR s 1994-95 to 2000 ¹			
	(1)	(2)	(3)	(4)
Intrinsic Incentives				
1994-95 Report Card Info (RAMR-OMR) Group				
Much Better than Expected (0-20%)	0.790 (0.190) ***	0.211 (0.199)	0.761 (0.244) ***	0.251 (0.206)
Slightly Better than Expected (20-40%)	0.245 (0.235)	0.105 (0.162)	0.243 (0.241)	0.121 (0.230)
Slightly Worse than Expected (60-80%)	0.568 (0.235) **	0.482 (0.162) ***	0.591 (0.233) **	0.523 (0.232) **
Much Worse than Expected (80-100%)	1.976 (0.264) ***	0.574 (0.250) **	2.045 (0.252) ***	0.624 (0.285) **
Extrinsic Incentives				
Pred Vol No RC-Pred Vol RC	0.062 (0.026) **	0.045 (0.021) **	0.064 (0.026) **	0.043 (0.021) **
I[RCDem>0]	-0.258 (0.210)	-0.145 (0.172)	-0.284 (0.215)	-0.255 (0.195)
I[RCDem>0]*Pred Vol No RC-Pred Vol RC	-0.049 (0.062)	-0.142 (0.061) **	-0.055 (0.064)	-0.166 (0.045) ***
Controls				
Mean RAMR 1994-95		0.715 (0.050) ***		0.712 (0.051) ***
Surgeon License Year (PA)	-42.030 (4.965) ***	-37.772 (3.815) ***	-42.412 (5.858) ***	-38.497 (5.352) ***
Surgeon License Year (PA) Squared	0.011 (0.001) ***	-0.203 (0.040) ***	0.011 (0.001) ***	0.010 (0.001) ***
Publications	-0.258 (0.047) ***	0.008 (0.001) ***	-0.273 (0.050) ***	-0.215 (0.047) ***
Market Fixed Effects?	No	No	Yes	Yes
Observations (surgeon/quarter)	920	920	920	920
R Squared	0.1919	0.3654	0.2154	0.3724

, * and * denote statistical significance at the 10%, 5%, and 1% levels, respectively. Bootstrap standard errors are clustered at the surgeons
 1. Note.: Changes are computed with respect to 1994-95 RAMR so positive values represent quality improvement.

Table 6: Effect of Report Card Induced Information and Demand on Surgeon Quality Including Learning about Within Hospital Reference Surgeons

Dependent Variable:	Change RAMR s 1994-95 to 2000 ¹	
	(1)	(2)
Intrinsic Incentives		
1994-95 Report Card Info (RAMR-OMR) Group		
Much Better than Expected (0-20%)	0.256 (0.154) *	0.272 (0.158) *
Slightly Better than Expected (20-40%)	0.231 (0.196)	0.240 (0.172)
Slightly Worse than Expected (60-80%)	0.453 (0.167) ***	0.482 (0.166) ***
Much Worse than Expected (80-100%)	0.728 (0.244) ***	0.709 (0.256) ***
RAMR s-Best RAMR h	0.110 (0.088)	0.125 (0.089)
RAMRs-Worst RAMR h	-0.048 (0.008) ***	-0.051 (0.008) ***
Extrinsic Incentives		
Pred Vol No RC-Pred Vol RC	0.099 (0.029) ***	0.106 (0.034) ***
I[RCDem>0]	0.214 (0.140)	0.251 (0.171)
I[RCDem>0]*Pred Vol No RC-Pred Vol RC	-0.051 (0.055)	-0.055 (0.052)
Controls		
Mean RAMR 1994-95	0.637 (0.096) ***	0.632 (0.105) ***
Surgeon License Year (PA)	-36.221 (5.137) ***	-37.056 (4.859) ***
Surgeon License Year (PA) Squared	0.009 (0.001) ***	0.009 (0.001) ***
Publications	-0.205 (0.040) ***	-0.204 (0.047) ***
Market Fixed Effects?	No	Yes
Observations (surgeon/quarter)	920	920
R Squared	0.3893	0.3801

, * and * denote statistical significance at the 10%, 5%, and 1% levels, respectively. Bootstrap standard errors are clustered at the surgeon level.
 1. Note.: Changes are computed with respect to 1994-95 RAMR so positive values represent quality improvement.

Figure 1: Profit Maximizing Quality Choice

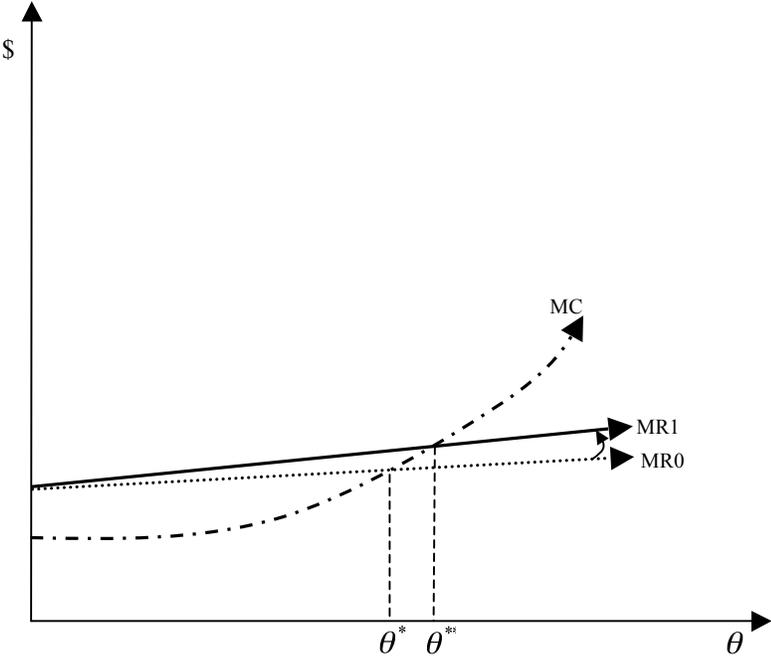


Figure 2: Intrinsic Utility and Quality Choice

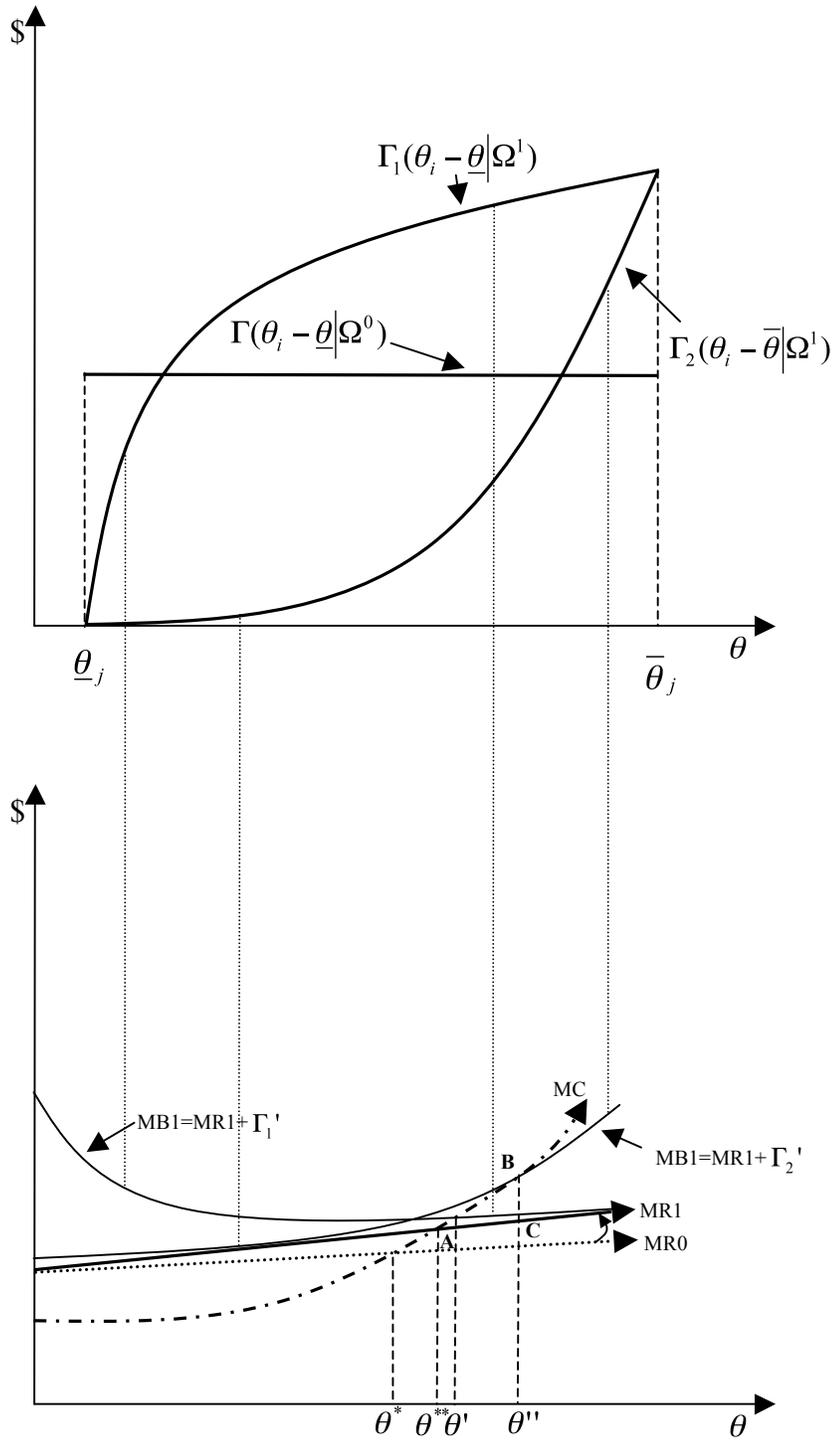


Figure 3: Mean Quarterly Performance

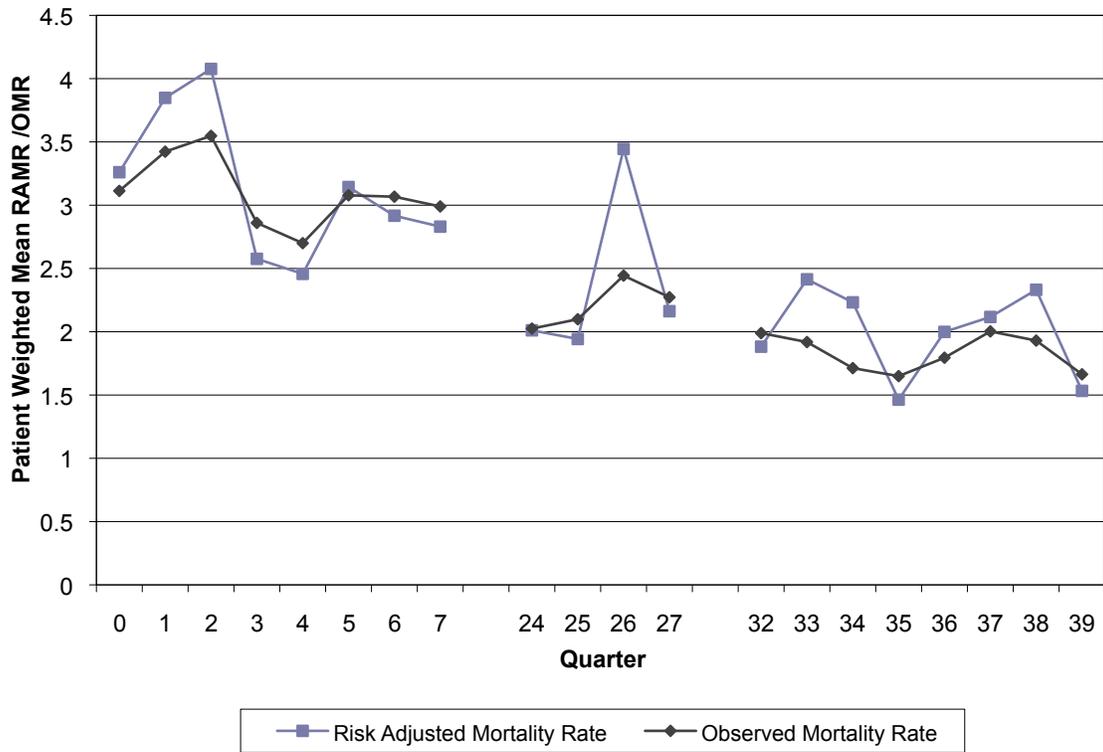


Figure 4: Frequency of New Information Provided by the 1994-95 Report Card

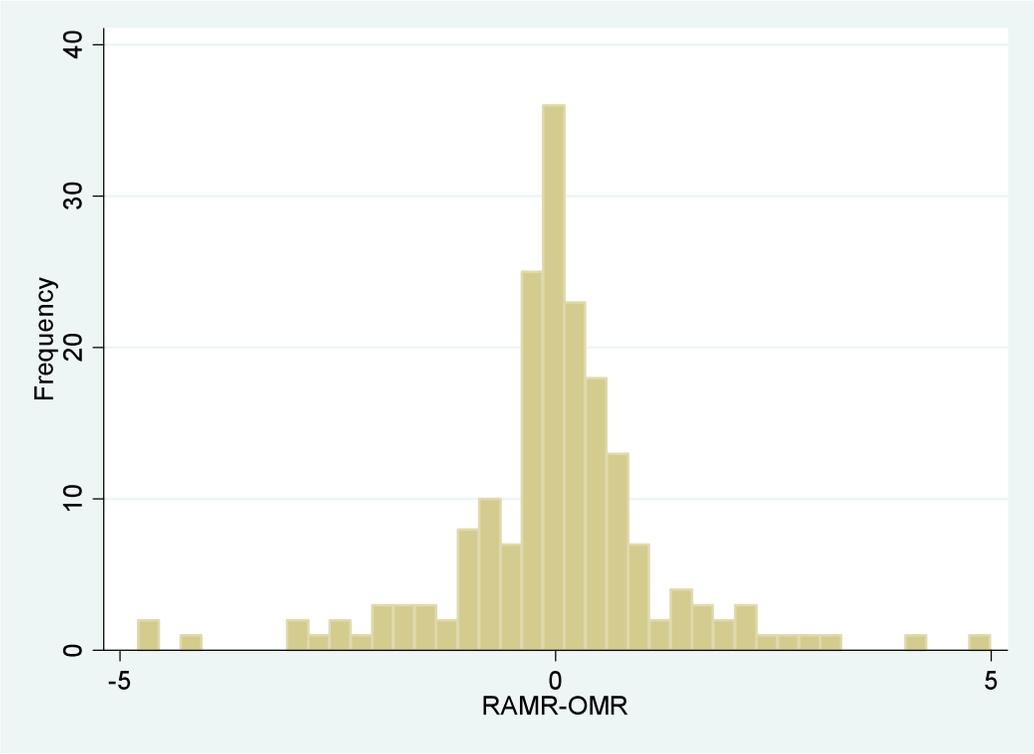


Figure 5: Local Polynomial Smoothed Estimates for the Relationship Between New Information and Changes in Volume and Quality

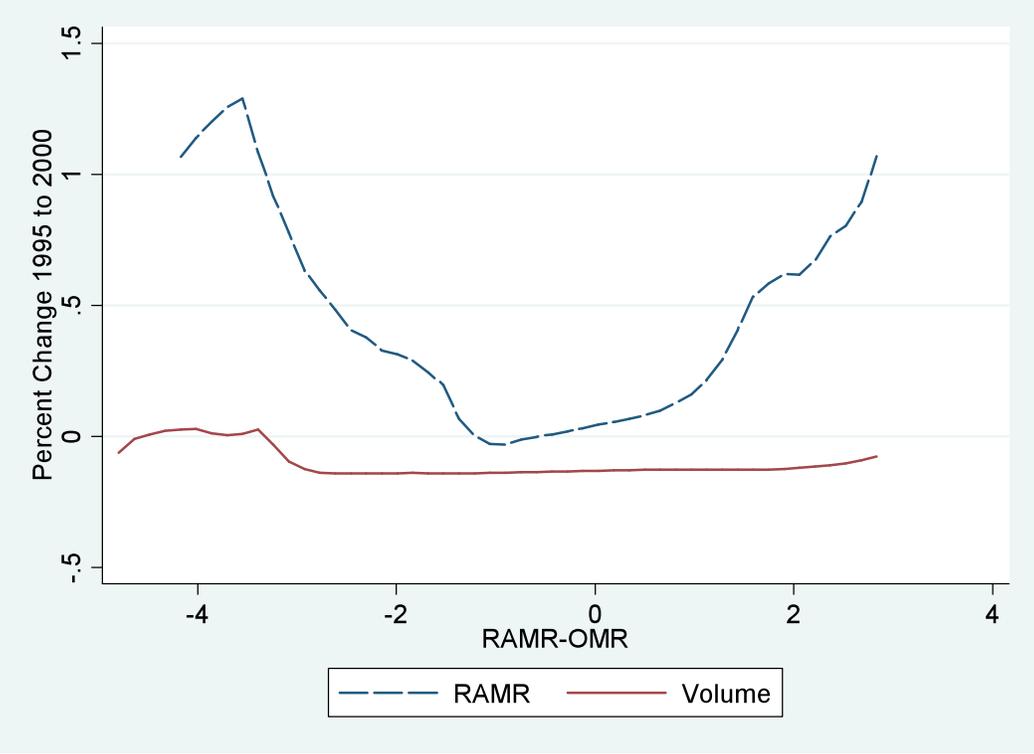


Figure 6: Frequency of Quarterly Differences Between Surgeons' Pre- and Post-Report Card Predicted Demand

