

Are Two Report Cards Better than One? The Case of CABG Surgery and Patient Sorting

Yang Zhang*

December 29, 2011

Abstract

Public reporting of information regarding quality may encourage sellers to improve product quality. In this paper, I study how quality disclosure affect health care providers' behavior. In particular, I examine the impact of mortality report cards on the degree of within-hospital sorting using hospital discharge data on coronary artery bypass graft (CABG) surgery in New York State from 1986 to 1993. Hospital report cards may encourage surgeons to allocate patients in more efficient ways to boost hospital scores. However, surgeon report cards may provide incentives for surgeons to avoid risky patients. I find that the level of within-hospital sorting increased immediately following the publication of hospital report cards but fell after the publication of surgeon report cards for hospitals in Manhattan, but not hospitals elsewhere. This phenomenon may have been driven by the intense competition faced by hospitals and surgeons in Manhattan.

*Paul Merage School of Business, University of California, Irvine. yang.zhang@uci.edu. I am indebted to David Dranove for his continuous guidance and support. I am also grateful to Bernard Black, Leemore Dafny, and Craig Garthwaite for their advice. I would like to thank Jennifer Brown, Cory Capps, Kitt Carpenter, Steven Farmer, Jie Gong, Mark Satterthwaite, Todd Sarver, Tuan-Hwee Sng, and Bruce Spencer for valuable discussions. Conversations with Robert Bonow, Angelo Costas, and Steven Farmer vastly improved my understanding of medical and institutional background information. Chieko Maene provided generous GIS support. This work has also benefited from the comments of seminar participants at Kellogg School of Management, Institute for Policy Research at Northwestern University, Department of Economics at National University of Singapore, Bates White, Paul Merage School of Business at University of California, Irvine, Acumen LLC., and Cornerstone Research. All errors are mine.

1 Introduction

Is more information always better? The benefits and hazards of quality disclosure have long been debated. During the past few decades, quality report cards have become increasingly popular, especially in areas such as health care, education, and finance. The underlying rationale for these report cards is that disclosing quality information can help consumers make better choices and encourage sellers to improve product quality.

A large recent literature documents the impact of report cards across a wide range of industries. Some examples include restaurant hygiene report cards (e.g., Jin and Leslie, 2003), school report cards (e.g. Figlio and Lucas, 2004), and a number of disclosure programs in the health care industry—coronary artery bypass graft (CABG) surgery mortality report cards (e.g., Dranove, Kessler, McClellan, and Satterthwaite, 2003), health plan report cards (e.g., Dafny and Dranove, 2008), hospital rankings (e.g., Pope, 2009), nursing homes report cards (e.g., Lu, 2011), fertility clinics report cards (e.g., Bundorf, Chun, Goda, and Kessler, 2009), and hospital infection rates report cards (e.g., Kim and Black, 2011). Some of the existing studies provide evidence in support of the use of report cards—consumers use report cards to select better-rated sellers and sellers respond by improving product quality, but others have raised concerns by showing that report cards may induce sellers to game the system in ways that hurt consumers.

Most of the existing literature is concerned with whether quality information should be disclosed. Much less attention has been given to *how* such information should be disclosed. In fact, report-card programs in practice display a surprising amount of heterogeneity in the ways they are designed. Some disclosure programs report detailed quality information (e.g., risk-adjusted mortality rates, birth rates by age groups) while others use coarse performance measures (e.g., Healthgrades.com uses a 5-star reporting system and Leapfrog group reports hospital quality based on a 4-bar standard). Some emphasize on rankings (e.g. US News and World Reports) while others highlight quality outliers (e.g., OSHPD ICU mortality rates report cards). In particular, disclosure programs often differ in their levels of reporting. For example, CABG surgery mortality rates are usually reported at hospital and surgeon levels while fertility clinics report cards only disclose clinic-level data. Even for the same type of report cards, there could be cross-sectional variations in the levels of reporting—Pennsylvania reports hospital infection rates at hospital level while New York publishes similar report cards at ICU level.

What is the best way to report quality information? At what level should quality be evaluated and reported? These questions cannot be separated from the larger question of whether quality information should be disclosed, yet they remain largely unexplored, both

theoretically and empirically.¹ This paper is an attempt to investigate whether reporting quality at different levels have different consequences.

I exploit a unique policy change—New York State’s sequential publication of hospital- and surgeon-level CABG surgery mortality report cards, to study the different incentives for health care providers created by hospital level and surgeon level report cards. In 1989, New York State took steps to introduce public reporting of hospital-level risk-adjusted mortality rates for CABG surgeries. Two years later, state authorities were instructed by the court to release surgeon-level report cards following a lawsuit.

Specifically, I investigate whether these report cards encouraged hospitals and surgeons to improve within-hospital sorting between patients and cardiac surgeons. Matching heterogeneous consumers with appropriate sellers is an important issue in economics. In the health care market, sorting may occur at different levels and take various forms. For example, one can study the sorting between patients and hospitals or the sorting between patients and physicians. At the physician level, one can examine whether patients with a particular risk factor are matched with physicians specializing in that condition or whether patients who prefer a particular treatment approach are treated by physicians who use that approach. The type of sorting that I am interested in involves the riskiest patients being matched with the best physicians. In my context, I define within-hospital sorting as the phenomenon where sicker patients are matched with better surgeons within the same hospital. This scenario is welfare improving if the severity of patients’ conditions and the quality of the physicians are complements in the health care production function.²

Ideally, report cards should provide incentives for both hospitals and surgeons to improve their performance. Hospital-level report cards should also encourage individual hospitals to assign risky patients to their best surgeons—i.e., to improve within-hospital sorting—so as to minimize mortality rates. Suppose this were indeed the case; the distribution of risky patients in each hospital would then become more skewed towards good surgeons upon the introduction of report cards. This effect would perhaps be stronger for hospitals experiencing more competitive pressure.

However, the publication of surgeon report cards could create incentives for the best surgeons to oppose within-hospital sorting to avoid risky patients, and surgeons facing more com-

¹A few exceptions on the theory side include Chen (2011), Epstein (2004) and Fong (2009). Some empirical works also have implications for the design of report cards. For example, Cutler, Huckman, and Landrum (2004) and Wang, Hockenberry, Chou, and Yang (2011) find reduced patient volume for hospitals and surgeons with high-mortality flags but no effect for low-mortality hospitals and surgeons, indicating that report cards should focus on singling out the worst providers.

²Zhang (2011) finds that a marginal increase in the level of within-hospital sorting improves the treatment outcomes for risky patients and does not affect non-risky patients, indicating that severity and quality are indeed complements in the production function for CABG surgeries.

petition may have more incentives to do so. In other words, the two components of the public reporting system may have contradicted each other in the case of New York State. While hospital-level report cards provided incentives for hospitals to encourage sorting, surgeon-level report cards possibly created reasons to discourage it. On the other hand, if low-skilled surgeons have more incentives to avoid risky patients, the publication of surgeon report cards may actually strengthen within-hospital sorting. All in all, it is not particularly clear whether within-hospital sorting would improve with the introduction of report cards in New York, especially with the publication of surgeon report cards. And this is precisely the issue I seek to clarify.

Though many studies have examined the consequences of New York State's CABG report cards, to my knowledge, no one has yet analyzed the potentially different incentives resulting from hospital report cards and surgeon report cards. Some of the existing research has also examined sorting as a possible mechanism for improving quality and some of them have shown that report cards promoted the matching of patients to hospitals (Dranove et al., 2003) and nursing homes (Werner, Konetzka, Stuart, and Polsky, 2011), very little has been written about sorting within organizations.

I combine the data sets from the 1986-1993 New York State hospital discharge database and the CABG report-card information from the New York State Department of Health to test the impact of report cards on within-hospital sorting. The first data set provides extensive information at the patient level. The second data set contains details from hospital- and surgeon-level report cards. I develop a measure for patient-condition severity using diagnosis codes, and use the surgeon-level report-card scores as a measure of surgeon quality. To investigate the effect of report cards on health care providers' behavior, I further limit the sample to patients admitted to hospitals through the emergency department following a heart attack to focus on how hospitals and surgeons allocate patients. I calculate the hospital-specific HHI as well as the HHI of each geographic market to quantify the degree of competition.

The specific questions I ask are the following.

1. Did the publication of hospital report cards encourage hospitals and surgeons to improve within-hospital sorting?
2. Was the effect of report cards different after the publication of surgeon report cards (compared to the preceding period when only hospital report cards were issued)?
3. Was the impact of report cards on within-hospital sorting stronger for hospitals facing more competition?

The following is a preview of my results. I do not find an overall effect of report cards on within-hospital sorting in New York State. I find that hospitals in Manhattan (including

Brooklyn) improved within-hospital sorting after the publication of hospital report cards relative to other hospitals. Moreover, the publication of surgeon report cards weakened the (positive) effect of hospital report cards on within-hospital sorting for hospitals in Manhattan. This Manhattan effect appears to be partly explained by the fact that hospitals in Manhattan experience greater competitive pressure than hospitals in the rest of the State. Moreover, among hospitals facing substantial competitive pressure, the impact of hospital report cards on within-hospital sorting increases with the degree of competition.

The rest of the paper is organized as follows. Section 2 presents the institutional and theoretical background and places the research in the context of the existing literature. Section 3 provides a detailed description of the data sets I use in this study as well as key measures and definitions. The effect of report cards on sorting is analyzed in section 4. Section 5 concludes.

2 Background Information and the Related Literature

2.1 CABG Surgery

CABG surgery was first invented in the 1960s and has been widely used since the 1980s. It is one of the few treatments available for coronary artery disease (CAD). Other common treatments for this disease include medical management using drugs such as beta blockers and PTCA (cardiac angioplasty), another surgical procedure. Among the three, CABG is the most invasive. During the surgery, the cardiac surgeon opens the patient's chest wall and grafts arteries or veins from the patient's body to coronary arteries to bypass blockages. The surgery usually involves using a heart and lung machine with the patient's heart temporarily stopped. After the surgery, the patient is typically transferred to the Intensive Care Unit (ICU) for recovery.

Cardiologists I interviewed pointed out that CABG is the most effective treatment for patients with a relatively severe condition. In these cases, medical treatment alone does not work well, while CABG is always weakly superior to PTCA in terms of treatment effects, after controlling for surgical risks. For certain patient characteristics, CABG is strictly superior.

The path to CABG surgery typically involves the following decision makers: the patient, the cardiologist, and the cardiac surgeon.³ The cardiologist diagnoses the patient and suggests appropriate treatment. If CABG is recommended, the cardiologist refers the patient to a

³In some instances, primary care physicians may also be involved. The main role of the primary care physician is to refer patients to cardiologists. However, primary care physicians occasionally refer patients to cardiac surgeons directly. Patients may also seek treatment from cardiac surgeons through self-referral.

cardiac surgeon.⁴ If PTCA or medical treatment is selected, the patient is treated by the cardiologist and is not referred to a cardiac surgeon.

In practice, the choice between CABG and PTCA is made jointly by the patient and the cardiologist. The decision is usually based on the patient's medical characteristics, but may also be influenced by other factors, such as the patient's preference and the cardiologist's incentive. For example, although the treatment outcome of PTCA is often not as good as that of CABG, some patients may prefer the former because it is less invasive and requires a much shorter, if any, in-hospital stay. Other patients, however, may favor CABG even when PTCA would be sufficient. Furthermore, since PTCA procedures are usually performed by cardiologists,⁵ a cardiologist who performs PTCA procedures may have the incentive to recommend PTCA to patients even when CABG is the better choice given the circumstances.⁶

As a result, patients who undergo CABG surgery vary considerably in terms of the severity of their condition. They may have no symptoms at all, relatively mild symptoms—such as chest heaviness and chest pain—or a life-threatening acute myocardial infarction (heart attack). Also, they may or may not have other diseases, such as diabetes or chronic lung disease, which tend to complicate CAD.

The outcome of a CABG patient is jointly determined by the patient's characteristics, the cardiac surgeon's skills, the quality of the supporting surgical staff,⁷ and the quality of post-surgery care. It is commonly believed that the operating surgeon's quality significantly influences the outcome of the surgery. The most severe adverse outcome is death. Examples of other major complications after surgery include heart failure, stroke, and infection.

Matching CABG patients with surgeons is a complicated process that may involve all or some of the following parties: patients, primary care physicians, cardiologists, surgeons, hospitals, and insurers.⁸ Sorting may occur at different levels as well. For example, one can investigate the sorting between different hospitals within the same health-services market. The aspect of sorting I focus on in this paper is within-hospital sorting, that is, whether riskier patients are matched with the relatively better surgeons in a hospital.

⁴The patient may also be referred to a different cardiac surgeon or by the cardiologist or the original cardiac surgeon if either the patient or the original cardiac surgeon feels that that is necessary.

⁵Cardiologists who perform PTCA procedures are sometimes referred to as interventional cardiologists. Cardiologists who do not perform such procedures are known as medical cardiologists.

⁶Afendulis and Kessler (2007) found that cardiologists who perform PTCA procedures tend to make different treatment decisions when choosing between PTCA and CABG compared to cardiologists who do not perform PTCA.

⁷Performing CABG surgery typically requires a cardiac surgeon, an anesthesiologist, a perfusionist (who operates the heart and lung machine) and surgical nurses.

⁸Insurers may be involved by ranking providers based on cost-effectiveness or performance and by charging patients less if they choose one of the "preferred" providers.

2.2 Background on New York State CABG Report Cards

New York is the first state to implement mandatory reporting of CABG surgery quality in terms of clinical outcomes.⁹ CABG surgery report cards were initiated by the New York State Department of Health and the Cardiac Advisory Committee (CAC). The CAC is a committee of cardiologists, cardiac surgeons, general physicians, and consumers set up to assist the Department of Health with the design of the report cards. The effort to reduce CABG mortality began in 1989 (Chassin, Hannan, and DeBuono., 1996) and the first public reporting of information on quality with respect to CABG surgery in New York State took place on December 5, 1990. On that day, *The New York Times* published the patient volume and risk-adjusted mortality rate (RAMR) for every hospital that performed CABG surgeries in New York State (Mukamel and Mushlin, 1998). The information released covered the second half of 1989 and the first half of 1990. On June 6, 1991, the Department of Health and CAC published hospital RAMRs covering CABG surgeries for all of 1990.

The first report cards did not release surgeon-level information. However, after a successful lawsuit against the New York State Department of Health, *Newsday*, a newspaper based in Long Island, published the surgeon-level information in December of 1991 (Chassin et al., 1996).¹⁰ In other words, the initial publication of hospital-level report cards was planned by the Department of Health while the initial release of surgeon-level data was not.

From 1992 on, hospital- and surgeon-level report cards for all isolated CABG surgeries¹¹ were released by the Department of Health and CAC on an annual basis. The report cards would cover hospital-level outcomes for the preceding year and surgeon-level outcomes for the preceding three years. For example, the report cards published in December 1992 reported the 1991 hospital-level mortality rates and the 1989-1991 surgeon-level mortality rates. In addition to the patient volume and RAMRs, the post-1992 report cards also published the actual number of deaths, actual mortality rates, expected mortality rates, 95% confidence intervals, and indicators for hospitals and surgeons with significantly higher or lower RAMRs. Surgeons who performed CABG surgeries at more than one hospital during the reported period were also identified.

Several regulatory requirements deserve to be mentioned here. First, in New York State, because of the Certificate of Need (CON) review, government approval is needed for a hospital to start a cardiac surgery program. Accordingly, the number of hospitals performing CABG surgeries has been fairly stable compared to states without a CON program, such as

⁹Several other States followed the New York model and issued CABG report cards in later years. Examples include Pennsylvania, New Jersey, California, Massachusetts and Virginia.

¹⁰*Newsday* sued the Department of Health under the Freedom of Information Law to acquire surgeon-level performance information (Chassin et al., 1996).

¹¹Isolated CABG surgeries are CABG surgeries performed with no other major cardiac procedure.

Pennsylvania (Epstein, 2006).¹² Second, not all surgeons who performed CABG surgeries in New York State were included in the surgeon report cards; only those who had performed at least 200 surgeries in the three-year reporting period were included. Lastly, hospitals and surgeons with RAMRs over 150% higher than the state average are subject to review by the CAC and may be asked to stop performing CABG surgeries temporarily (Epstein, 2006).¹³

2.3 Report-Card Incentives and Within-Hospital Sorting

One of the potential benefits of report cards is improved patient sorting. Existing literature (Dranove et al., 2003; Epstein, 2006) has highlighted several reasons why report cards may improve sorting. First, report cards provide information on quality to patients, enabling them to identify the best providers. This is particularly important for the sickest patients who have the most to gain from seeing the best providers. Second, low-quality hospitals and surgeons may voluntarily turn away the sicker patients to improve their report-card scores, while the best providers are under less pressure to do so. Finally, hospital administration may help guide the sicker patients away from low-quality surgeons.

The last argument is especially relevant for within-hospital sorting. Given that hospitals have relatively good information about the quality and performance of their medical staff, they can exercise their administrative power and take steps to improve the quality of care when presented with the right incentives. These steps may include acquiring new equipment, recruiting more nurses and providing existing ones with more training, firing incompetent physicians and hiring new ones, improving patient management, and advising top surgeons to take on more risky patients.

The question is whether report cards did actually incentivize hospitals to take the above steps. There is some anecdotal evidence showing that they did. Chassin (2002) documented some changes that took place in New York State hospitals after the introduction of CABG report cards. In one particular hospital, an internal review was conducted to study how performance could be improved. It was found that the most skilled surgeon in that hospital was heavily booked for elective surgeries and, consequently, most of the urgent and usually difficult cases had to be performed by two other surgeons who were not well trained in adult cardiac surgery. Subsequently, the hospital hired a new surgeon to take over some of the elective surgeries so that their best surgeon could devote more time to the difficult cases, while the two less-skilled surgeons were asked to stop performing CABG surgeries.

¹²Only one hospital started a cardiac surgery program during 1990-1995. Two more hospitals started their respective programs between 1995 and 2000.

¹³Chassin (2002) provided two examples of hospitals suspending CABG services. One of those hospitals reopened its service in four months, while the other resumed its service after a year.

Although I am interested in whether report cards enabled hospitals and surgeons to improve within-hospital sorting, and not in how report cards affected patient-initiated sorting, it is worthwhile to point out that report cards may not provide the right incentives for patients to match themselves with appropriate providers when information is not perfect.

If there is perfect information, patients know the severity of their condition and the treatment technology that is available. Assuming that prices do not change, risky patients will have the incentive to choose the good surgeons and hospitals, while non-risky ones may or may not choose the good surgeons depending on the technology. In this situation, report cards will encourage patients to self-sort.

However, if patients are uncertain about either their own medical condition or the health care production function, the publication of report cards provides incentives for all patients to select the good hospitals and the good surgeons. Contrary to the case of perfect information, within-hospital sorting may not improve based on the actions of the patients.

As for the hospitals and surgeons, the incentives provided by report cards are mixed. Because risk adjustment in report cards is usually not perfect (Green and Wintfield, 1995), treating risky patients may result in higher RAMRs compared to non-risky patients. Even if risk adjustment were perfect, as long as hospitals and surgeons have doubts about how the adjustment will be carried out, they will tend to prefer non-risky patients to risky ones.

Given these considerations, there are two ways for hospitals to deal with risky patients. One is simply to turn them away, while the other is to direct them to the hospital's best surgeons. Surgeons have fewer options. In fact, their only recourse is to reject risky patients whenever possible.

Because of this, the introduction of surgeon-level report cards on top of hospital-level report cards may create conflicting interests between hospitals and surgeons. When there are only hospital-level report cards, the performance of individual surgeons does not matter. What matters to the hospitals and surgeons alike is the performance of the hospital. Under this scenario, surgeons have the incentives to help hospitals achieve better scores. Therefore, good surgeons are more likely to accommodate risky patients.

However, the situation changes once surgeon-level report cards come into play. Now surgeons, high-quality and low-quality alike, have strong incentives to avoid treating risky patients, for even the best surgeon will worry about her score being ruined by several difficult cases. As a result, even though hospitals still desire to allocate risky patients to good surgeons, the surgeons may refuse to cooperate. In this respect, the addition of surgeon-level report cards may lead to unintended consequences by weakening the positive incentives that hospital-level report cards create to promote within-hospital sorting. If, on the other hand, low-quality surgeons are more likely to reject risky patients, the publication of surgeon-level report cards

may actually strengthen the effect of hospital report cards.

Moreover, the effect of report cards may not be homogeneous across all hospitals and surgeons. If hospitals and surgeons care about report cards mainly because report-card results affect demand, the degree of competition may play a role. Economic theory does not yield uniform prediction on how competition affects quality.¹⁴ The effect depends on whether the price is fixed or set by firms. In the first scenario, theory predicts that competition leads to better quality.¹⁵ In the second scenario, however, the effect of competition on quality is ambiguous. Quality will improve if competition increases the quality elasticity of demand relative to the price elasticity of demand, and vice versa.¹⁶

In the context of this paper, hospitals face fixed prices for some of the patients, such as Medicare patients, but may set prices for privately insured patients. Moreover, the presence of report cards provides patients with relatively precise information on quality compared to the absence of report cards, which could potentially increase the quality elasticity of demand. Intuitively, for hospitals facing very little competition, demand is inelastic in terms of both price and quality. Even if the publication of report cards increases the precision of the information on quality, it is unlikely that report cards will have a significant impact on quality elasticity. Thus, report cards may not provide enough incentive for the less-competitive hospitals to improve within-hospital sorting. On the other hand, for hospitals facing greater competitive pressure, report cards may cause quality elasticity to increase significantly, and if price elasticity is not affected by report cards, the more-competitive hospitals will have the incentive to improve within-hospital sorting so as to compete with respect to quality. By the same token, a surgeon in a more-competitive market will have a greater incentive to select against risky patients because a bad report-card score will have a more-negative effect on him.

2.4 Related Literature

There exists a sizeable literature on the consequences of public reporting of information regarding quality for CABG surgeries.¹⁷ Some of the early work explored the impact of CABG report cards on mortality rates, and studies such as Hannan, Kilburn, Racz, Shields, and Chassin (1994) and Peterson, DeLong, Jollis, Muhlbaier, and Mark (1998) found that CABG mortality rates in New York State declined in absolute terms as well as relative to

¹⁴See Gaynor (2006) for a recent, comprehensive review of both the theoretical and empirical aspects of this line of research.

¹⁵Examples of papers with this conclusion in the context of health care include Allen and Gertler (1991), Held and Pauly (1983), and Pope (1989).

¹⁶A number of papers discussed models with this insight; see, for example, Dorfman and Steiner (1954), Dranove and Satterthwaite (1992), and Allard, Léger, and Rochaix. (2009).

¹⁷See Epstein (2006) for a comprehensive review on CABG report cards and Dranove and Jin (2010) for a recent review on quality report cards in general.

control states after the publication of report cards.

A number of papers have examined whether and how report cards changed market shares and prices. For example, Mukamel and Mushlin (1998) found that hospitals and surgeons with lower RAMRs saw higher rates of growth in market shares and that these surgeons also benefited from higher rates of growth in CABG charges. Cutler et al. (2004) found that hospitals with a high-mortality flag experienced a decrease in patient volume, primarily driven by a drop in the number of non-risky patients seeking treatment. Similarly, Romano and Zhou (2004) detected an increase in patient volume for hospitals with low RAMRs and a decrease for hospitals with high RAMRs after the introduction of report cards in New York State. More recently, Dranove and Sfekas (2008) estimated a structural model and concluded that hospitals with report-card scores lower than patients' prior expectations subsequently lost market shares. Kolstad (2011) also adopted a structural approach and found that the demand for surgeons with good report-card scores increased slightly in Pennsylvania. On the other hand, Wang et al. (2011) found a reduction in patient volume for high-mortality and unrated surgeons in Pennsylvania.

Another strand within this literature documented the presence of a selection bias against risky patients among health providers after the introduction of report cards. In a survey that involved about half the cardiologist and cardiac-surgeon population in Pennsylvania, Schneider and Epstein (1996) reported that 63% of cardiac surgeons were less willing to operate on risky patients and 59% of cardiologists had difficulty finding surgeons for their risky patients. This finding is supported by Dranove et al. (2003), who found that CABG patients in states with report cards were healthier compared to CABG patients in other states. Moreover, Omoigui, Miller, Brown, Annan, Cosgrove, Lytle, Loop, and Topol (1996) found that CABG patients from New York State receiving treatments at Cleveland Clinic were sicker than patients from other states after the publication of report cards in New York, suggesting that some risky patients had migrated out of New York State due to report cards. Werner, Asch, and Polsky (2005) documented selection based on patients' race—racial disparity in CABG surgeries increased in New York State after the publication of report cards.

Previous work on whether report cards improved patient sorting reported mixed evidence. Dranove et al. (2003) found that patients admitted to teaching hospitals were on average more risky than patients in other hospitals after the introduction of report cards, and the coefficient of variation of the within-hospital patients-condition severity declined after the introduction of report cards. This suggests that report cards improved sorting between hospitals. Werner et al. (2011) studied sorting in the context of nursing homes report cards. They found that high-risk patients became more likely to visit better-rated facilities, indicating improved sorting across nursing home facilities.

On the other hand, Cutler et al. (2004) did not find evidence of improved between-hospital sorting. They observed that hospitals with a high mortality flag in New York State cut back on the non-risky patients rather than the risky ones. With respect to surgeons, Wang et al. (2011) found that those with low report-card scores lost patients at all levels of condition severity. Moreover, Epstein (2010) examined the referral patterns from cardiologists to cardiac surgeons and did not find that the publication of surgeon report cards in Pennsylvania changed the referral pattern for either low-quality or high-quality surgeons.

To my knowledge, the creation of potentially conflicting incentives for health providers by the publication of both hospital-level report cards and surgeon-level report cards has yet to be discussed, let alone empirically tested. Furthermore, none of these papers has studied whether and how report cards have affected within-hospital sorting. It is in these areas that this paper seeks to make some marginal contributions.

3 Data

3.1 Data Sources and Measures

This paper employs data from the following sources. First, I use patient-level discharge data from all New York hospitals from 1986 to 1993. The data contains patient demographics, diagnosis and procedure codes, treatment outcomes, and physician and hospital identifiers. I look at patients undergoing isolated CABG surgeries as their principal procedure during this period.¹⁸

To analyze the effect of report cards on within-hospital sorting, I focus on patients who did not choose their providers. To do so, I create a sub-sample that includes only patients whose principal diagnosis codes indicated a heart attack (AMI), who were admitted to the hospital through the emergency department, and whose admission was not scheduled. Moreover, this sub-sample omits patients who had previously had heart surgeries, as those patients likely knew exactly which hospital or which doctor to visit in case of emergency because they had already experienced similar conditions in the past.¹⁹ After these eliminations, the remaining sub-sample accounts for about 10% of all patients in the data set.

The data set also allows me to construct measures for the patients' condition severity and treatment outcome. The measure of severity is derived using each patient's 15 diagnosis codes. I follow the risk factors listed by the Society of Thoracic Surgeons (STS) for isolated CABG

¹⁸Patients simultaneously undergoing additional major cardiac procedures, such as valve replacement surgery, are excluded from the sample.

¹⁹I can only identify these patients to the extent that their diagnosis codes include complications due to previous heart surgeries. This is likely to be a subset of patients who had heart surgeries before.

surgery and identify from the diagnosis codes three major cardiac risk factors for patients in the sub-sample, namely congestive heart failure, cardiac arrhythmias, and valvular disease. Within the sub-sample, 31% of the patients were risky.

The second data set used here comprises CABG surgery report cards issued by the New York State Department of Health. Specifically, I use the 1989-1991 and 1990-1992 surgeon-level risk-adjusted mortality rates as a measure of surgeon quality. There are 77 surgeons in the 1989-1991 report cards and 86 surgeons in the 1990-1992 report cards. Three of the surgeons included in the 1989-1991 report cards do not appear in the 1990-1992 report cards.²⁰ In addition, 11 surgeons appear only in the 1990-1992 report cards. Surgeons with at least one report-card score were responsible for over 80% of all CABG surgeries performed in New York State during the period of study. I am able to match each surgeon's report-card score with the patient-level discharge data set by linking the name of the surgeon in the report cards with his state license number. This is done by looking up each physician on the New York State Department of Health website. I will only focus on hospitals with at least two surgeons who have received individual report-card scores each year, which accounts for two-thirds of the hospitals, and I will only look at surgeons with both 1989-1991 and 1990-1992 report-card scores.²¹

There are a total of 10,282 CABG surgeries in 31 hospitals in New York State from 1986 to 1993 in the sub-sample.²² Table 1 and Table 2 report the sample means and standard deviations for patient and hospital characteristics, respectively. The average patient age in the sample is 64 and Medicare is the primary insurer of 40% of the patients. The number of surgeons in each hospital ranges from one to nine, with an average of 3.8, while the number of surgeons in each hospital with both 1989-1991 and 1990-1992 report-card scores varies from zero to six.²³ Each surgeon performs 13 surgeries on patients in the sub-sample annually, while the number of surgeries performed by surgeons with both report-card scores on these patients is 15 per year on average.²⁴

²⁰These surgeons received the highest RAMRs in the 1989-1991 report cards in New York State. Two of them stopped performing CABG surgeries permanently shortly after 1991 and the other one transferred to a low-volume hospital. This is consistent with what previous research (Chassin, 2002) has found.

²¹I exclude the three surgeons with only 1989-1991 report cards who dropped out shortly after 1991. Moreover, for reasons that will be discussed later, I mainly rely on the 1989-1991 report-card scores; thus, I dropped the surgeons with only 1990-1992 report-card scores.

²²The total number of CABG surgeries in the full sample is 99,281 during the same period.

²³Four low-volume hospitals did not have any individual surgeons who exceeded the 200-patient threshold during the period of study. Thus, no surgeon in these hospitals had a report-card score.

²⁴In the full sample, on average, each surgeon performs 116 surgeries annually, and the number of surgeries performed by surgeons with both report-card scores is 138 per year.

3.2 Competition

In this section I describe how I quantify the degree of competition. Two methods are employed here. First, I use a hospital-specific HHI (Herfindahl-Hirschman Index) as a continuous measure of the degree of competition faced by each hospital. The measure is between zero and one, with a higher value indicating more market power. Using a hospital-specific HHI allows me to capture the degree of competition for each hospital without specifying geographic markets. Nonetheless, I use the HHI for each geographic market, which I call the market HHI, as a second measure of competition. Note that I calculate the hospital and market HHIs using the actual patient admission patterns observed in the full sample. As previous research has pointed out, doing so may result in biased estimates due to unobserved hospital quality and prices.²⁵

I calculate the hospital-specific HHI following the method developed by Zwanziger and Melnick (1988). For each five-digit zip code area, I compute the zip code level HHI, which is the sum of the squared market shares of each hospital with patients from that zip code area. The market share for a hospital is simply the number of patients from that zip code area in that hospital over the total number of patients from that zip code area. Subsequently, I weight the zip code level HHIs by the zip code's share of the hospital's total number of patients. Finally, the weighted zip code level HHIs for the hospital are added up and the resulting number is defined as the hospital-specific HHI.²⁶

Markets are defined based on the pattern of the catchment area for each hospital observed directly from the full sample and the hospital referral region (HRR) defined by the Dartmouth Atlas of Health Care. To distinguish from HRRs, I define the markets as CABG markets. The reason for not using the HRR directly is because the definition is not constructed specifically for CABG surgery markets. One feature of New York State is that it contains one distinct, densely populated metropolitan area, which is the NYC-metro area.²⁷ Although small in size compared to the rest of New York State, this area has 19 hospitals performing CABG surgeries, which account for over 60% of all hospitals performing CABG surgeries in New York State. The rest of the state, which I refer to as the Upstate region, has 12 hospitals scattered over five HRRs. In the Upstate region, because the number of hospitals performing CABG surgeries is small, patients may have to travel a longer distance for CABG surgeries than for other services. Thus, the size of each geographic market for CABG surgeries in the Upstate region is usually larger than an HRR. On the other hand, the existing market definition works

²⁵Examples of papers addressing this issue include Capps and Dranove (2004), Kessler and McClellan (2000), and Werden (1989).

²⁶I compute the hospital-specific HHI for each hospital in each year.

²⁷I refer to the NYC-metro area as the part of the New York Metropolitan Area located in New York State, which includes the five boroughs of New York City, Long Island, and three counties in the lower Hudson Valley.

better in the NYC-metro area because hospitals are densely located there.

Figure 1 shows the catchment area for Upstate hospitals during the period of study. A black dot implies that there is at least one hospital performing CABG surgeries in that zip code area.²⁸ There is a line connecting a zip code area with a black dot if at least three patients from that zip code area have visited the hospital(s) represented by that dot.²⁹ Lines associated with hospitals in the same HRR are assigned the same color.

Intuitively, if lines with two different colors overlap a lot, the hospitals in those two HRRs are competing over the same pool of patients. As such, they should be considered as being in the same market. Similarly, if lines with two different colors overlap very little, the two HRRs are relatively self-contained and, hence, should be treated as separate markets. Without showing the details here, I have checked that hospitals within the same HRR usually draw patients from the same pool. The question is whether hospitals in different HRRs are competing with one another. It is clear from Figure 1 that Buffalo, Rochester, and Albany are relatively well-defined markets, while hospitals in Syracuse, Elmira and Binghamton appear to be competing with one another for patients in the same locations. Thus, I pool these three HRRs together into what I call the Central NY market.

Hospitals in the NYC-metro area belong to four different HRRs: the Bronx, East Long Island, Manhattan (including Brooklyn), and White Plains. To avoid confusion, I label the Manhattan market—which includes Brooklyn—as Manhattan, and the island of Manhattan as Manhattan Island. Figure 2 shows the catchment area for hospitals in the NYC-metro area excluding the nine hospitals on Manhattan Island. Although there is some overlap between Bronx and White Plains as well as between Brooklyn and East Long Island, these four areas are still self-contained in the absence of Manhattan Island. Figure 3 introduces Manhattan Island hospitals into the picture. It is evident that hospitals on Manhattan Island draw patients from every neighboring area and even some from Albany. This is not surprising given the density of hospitals on Manhattan Island. Although 97% of the patients from Manhattan Island went to the Manhattan Island hospitals, this only accounts for 17% of those hospitals' total admissions. Thus the majority of the patients treated at Manhattan Island hospitals came from other areas.³⁰ Manhattan Island overlaps considerably with nearby areas, including Brooklyn, the South Bronx, and the west part of East Long Island. Thus, the HRR definition of market is generally appropriate here and I will use it as the baseline market definition for

²⁸Hospitals in same zip code areas will appear in the graph as a single dot.

²⁹I drop zip code areas with fewer than three patients visiting hospitals in one zip code area for ease of illustration. The patterns are similar otherwise. Furthermore, patients who live in the same zip code areas as the hospitals they visit are not represented in the graph.

³⁰Five percent of all patients came from outside New York State and generally sought treatment at hospitals on Manhattan Island. These patients are not included in the graph.

the NYC-metro area.

There are eight CABG markets in New York State.³¹ Table 3 shows the number of hospitals, the mean market HHI, and the mean hospital-specific HHI in each of these markets. Manhattan has the lowest market HHI, followed by East Long Island. Similarly, Manhattan has a significantly larger number of hospitals relative to the other markets. Manhattan also has the lowest average hospital-specific HHI, followed by East Long Island.

The hospital-specific HHIs show a similar pattern. The mean hospital-specific HHI across all hospitals and all years is 0.36, with a minimum of 0.23 and a maximum of 0.63. Every Upstate hospital has an average hospital-specific HHI higher than the mean, and none of them has an average hospital-specific HHI lower than any of the hospitals in the NYC-metro area. Generally speaking, hospitals in Buffalo, Rochester, and Central NY have the highest hospital-specific HHIs, while hospitals on Manhattan Island and some hospitals in Brooklyn and East Long Island have the lowest hospital-specific HHIs. In the sub-sample of hospitals I use in the report-card analysis, all but one in Manhattan are in the lowest 25% percentile, and the rest of the hospitals in the NYC-metro area are below or slightly above the median (0.35). The highest 25% includes every hospital in Buffalo, Rochester, Elmira, and Binghamton. All of these findings lead to the conclusion that hospitals on Manhattan Island and in nearby areas face the highest levels of competitive pressure within New York State.

4 The Effect of Report Cards on Within-Hospital Sorting

In this section I study the effect of the two-stage reporting of mortality rates for CABG surgeries in New York State on within-hospital sorting. The empirical models presented below share several features. First, I only use a sub-sample of patients, surgeons, and hospitals. The sub-sample of patients is chosen so as to limit the amount of patient choices of surgeons and hospitals. Surgeons in the sub-sample are those who have received both 1989-1991 and 1990-1992 report-card scores, and the sub-sample of hospitals includes only hospitals with at least two surgeons who have received report-card scores each year.

Second, although the effect of report cards on within-hospital sorting is a hospital-level phenomenon, I conduct patient-level analysis for two reasons. It allows me to control for patient-level characteristics which may be correlated with condition severity and surgeon choices. Moreover, since the number of hospitals is small, patient-level analysis expands the sample size considerably. To account for the possibly correlated standard errors for patients in the same hospital in the same year, the standard errors are clustered at the hospital-year level.

³¹I will consider alternative market definitions in the empirical analysis.

Third, I use the report-card scores for each surgeon as a proxy for true quality. Since correctly measuring true quality is often difficult, there might be some concerns about using this measure. First, the methodology used to calculate the report-card scores may fail to produce unbiased results. The New York State surgeon-level CABG report-card scores are computed in the following way. First, a logit model is employed to predict the expected probability of death for each patient, controlling for a set of risk factors. Then, for each surgeon, the actual mortality rate is divided by the expected mortality rate to obtain the O/E ratio. Finally, the risk-adjusted mortality rate for each surgeon is calculated by multiplying the O/E ratio of that surgeon with the state-wide average mortality rate. This method may generate biased estimates as surgeon quality could be correlated with observable patient characteristics. Alternative methods, such as fixed effects estimation and correlated random effects estimation, may help resolve this issue (Glance, Dick, Osler, Li, and Mukamel, 2006; Johnson, 2011).³² Yet Glance et al. (2006) found that these three methods showed a near-perfect degree of agreement in determining the quality outliers for New York CABG report cards, which supports the use of report-card scores as a quality measure.

In addition, since the set of risk factors is unlikely to be exhaustive and surgeons may have private information about the patients' condition severity, report cards may induce selection behavior against risky patients (Dranove et al., 2003; Schneider and Epstein, 1996). As a result, a bad surgeon who treats many non-risky patients may appear as good as a better surgeon who treats more risky patients. In other words, surgeon quality may also be correlated with unobservable patient characteristics that affect outcomes. To mitigate this potential bias, I chose the 1989-1991 RAMRs because they evaluate surgeons based on surgeries performed before the publication of the first surgeon report cards, which may considerably exacerbate the selection problem.

Moreover, using this measure implies that the quality of a surgeon is constant in the short run. However, in reality, new surgeons may improve their quality quickly through learning by doing (Ramanarayanan, 2009), while the quality of experienced surgeons may be declining over time because they become less energetic or less focused. One way to justify the use of report-card scores is to assume that the increase or decrease in surgeon quality is linear in time. Since the report-card scores measure the quality of surgeons around 1990, they are in essence measuring the average quality of each surgeon during the period of study.

Finally, to distinguish the effects of hospital report cards and surgeon report cards, I divide the post-report-card period in two, 1990-1991 and 1992-1993. Because I am primarily

³²Theoretically, I can replicate the report-card scores using all three methods. However, it would be difficult to obtain precise estimates since my information on the patients' condition severity is significantly less complete than the data used by the Department of Health in New York to generate the actual report-card scores.

interested in how hospitals and surgeons responded to report cards, the choice of the post-report-card periods should correspond with the time when providers learned about the report cards instead of the time when report cards were actually published. Since data collection and other preparation for the initial hospital report cards began in 1989, it is reasonable to assume that, by the beginning of 1990, hospitals and surgeons were already informed about the forthcoming quality reporting. For surgeon-level reporting, the lawsuit was adjudicated in court in 1991, and the publication of surgeon report cards took place towards the end of that year. Thus, I choose 1990-1991 as the period when only hospital report cards were available, and 1992-1993 as the period when both hospital and surgeon report cards were available.

4.1 The Manhattan Effect

The empirical question is whether risky patients were matched with relatively better surgeons within the same hospital after the publication of hospital report cards and surgeon report cards, respectively, compared to the period prior to the publication of the report cards. To answer this question I use the following simple model:

$$M_l = \beta_0 + \beta_1 R_{iljt} + \beta_2 R_{iljt} \times P_t + \beta_3 R_{iljt} \times Q_t + X_{iljt} + N_{jt} + A_{ljt} + H_j + Y_t + \varepsilon_{iljt}, \quad (4.1)$$

where M_l is the quality measure for surgeon l and R_{iljt} indicates whether patient i treated by surgeon l in hospital j and year t is risky. P_t and Q_t are the dummy variables for the period following the publication of hospital report cards, with the former representing the period preceding, and the latter the period following, the publication of surgeon report cards. X_{iljt} represents patient-level characteristics, including age, age squared, sex, race, insurance type and medical risk factors. I use a comprehensive set of medical risk factors developed specifically for using hospital discharge data sets by Elixhauser, Steiner, Harris, and Coffey. (1998). N_{jt} is the number of surgeons with report-card scores in hospital j in year t . A_{ljt} includes the number of patients treated by surgeon l at hospital j in year t and the number of years of experience of surgeon l at hospital j as of year t . H_j is the hospital fixed effect, Y_t is the year fixed effect, and ε_{iljt} is the error term.

I assume that P_t and Q_t capture the report cards' specific time trend, and patients do not choose hospitals or surgeons. Moreover, controlling for hospital fixed effects allows me to estimate the level of within-hospital sorting. Since a lower RAMR is equivalent to a better surgeon, a negative β_2 would imply that after the publication of hospital report cards, risky patients were on average treated by better surgeons relative to non-risky patients within the same hospital than before their publication. Furthermore, if $\beta_3 > \beta_2$ and the difference is statistically significant, we may conclude that surgeon report cards weakened the effect of

hospital report cards on within-hospital sorting.

I start by investigating the overall effect of report cards and the effect by region. Table 4 shows the results of the above model for all hospitals in New York State, hospitals in the Upstate region only, and hospitals in the NYC-metro area only. Neither hospital nor surgeon report cards seem to affect the level of within-hospital sorting. However, $\hat{\beta}_2$ in column (3) is much larger (in absolute value) than that in column (2). This suggests that the effect of hospital report cards may be heterogenous across hospitals or markets. In particular, the effect on hospitals in the NYC-metro area may be different than on hospitals elsewhere.

I then focus on the NYC-metro area. I repeat the same regression, dropping one CABG market in the NYC-metro area at a time in order of decreasing mean market HHIs.³³ The results are shown in Table 5. It is clear from the table that hospital report cards improved within-hospital sorting for hospitals in Manhattan. On average, the difference in the RAMR of a surgeon who treated a risky patient and that of a surgeon who treated a non-risky patient in the same hospital fell by 0.201 after the publication of hospital report cards for Manhattan hospitals. Because report cards tend to under-adjust for risky conditions, surgeons who treat many risky patients may be of higher quality than observed, which implies that $\hat{\beta}_2$ is an underestimate of the impact of hospital report cards.³⁴ However, the result of the F-test ($\beta_2 = \beta_3$) indicates that the publication of surgeon report cards decreased the level of within-hospital sorting in Manhattan compared to the period with only hospital report cards.

Table 5 suggests that Manhattan is where report cards did have an impact. I confirm this by repeating the analysis for Manhattan using different specifications. Table 6 presents the results. The direction of the effect of hospital report cards is consistent, although the magnitude varies across different specifications. In column (3) I exclude surgeons with the highest RAMRs, and the result does not change much. However, when I exclude surgeons with the lowest RAMRs in column (4), the magnitude of $\hat{\beta}_2$ drops quite a bit. This indicates that a large amount of within-hospital sorting in Manhattan hospitals after the introduction of hospital report cards came from directing the risky patients to the best surgeons. Furthermore, surgeon report cards led to a decrease in the level of within-hospital sorting compared to the period when only hospitals report cards were available. I also try using hospitals on Manhattan Island and hospitals with mean hospital-specific HHIs in the lowest 25% percentile instead of all hospitals in Manhattan, and the results are consistent with Table 6.³⁵

So far I have found that hospitals in Manhattan increased the level of within-hospital sort-

³³I also performed the analysis for each geographic market individually and the results were consistent.

³⁴This is true under the assumption that surgeons' selection behavior did not change differently across surgeons before or after the publication of hospital report cards.

³⁵I did not find a statistically significant effect of report cards on within-hospital sorting for any market in the Upstate region, and this holds for both CABG markets and HRRs.

ing after the publication of hospital report cards, but the effect vanished after the appearance of surgeon report cards. However, this does not necessarily imply that the effect of report cards for hospitals in Manhattan is statistically *different* from the effect of report cards for hospitals elsewhere.

To verify if Manhattan hospitals did indeed respond differently to report cards, consider the following triple-difference variation of the model in the previous section:

$$\begin{aligned}
M_l &= \beta_0 + \beta_1 R_{iljt} + \beta_2 \text{Manhattan} \times R_{iljt} + \beta_3 \text{Manhattan} \times P_t \\
&+ \beta_4 \text{Manhattan} \times Q_t + \beta_5 R_{iljt} \times P_t + \beta_6 R_{iljt} \times Q_t \\
&+ \beta_7 \text{Manhattan} \times R_{iljt} \times P_t + \beta_8 \text{Manhattan} \times R_{iljt} \times Q_t \\
&+ X_{iljt} + N_{jt} + A_{ljt} + H_j + Y_t + \varepsilon_{iljt},
\end{aligned} \tag{4.2}$$

where Manhattan is an indicator variable for hospitals in Manhattan. If Manhattan hospitals increased the level of within-hospital sorting after the publication of hospital report cards relative to the other hospitals, β_7 should be negative. In addition, if $\beta_7 \neq \beta_8$, relative to the other hospitals, the effect of hospital report cards was different from that of surgeon report cards for Manhattan hospitals.

Column (1) of Table 7 shows the effect of report cards on Manhattan hospitals relative to all other hospitals. After the publication of hospital report cards, the difference-in-difference in the RAMR of a surgeon who treated an average risky patient and that of a surgeon who treated an average non-risky patient within a Manhattan hospital and within a hospital elsewhere dropped by 0.209. Moreover, relative to hospitals outside Manhattan, the improvement in within-hospital sorting disappeared in Manhattan following the publication of surgeon report cards.

This suggests that hospitals in Manhattan were indeed more responsive to the publication of report cards. The question is why. Given that Manhattan has the lowest mean market HHI and Manhattan hospitals have the lowest mean hospital-specific HHIs, a natural explanation is the higher degree of competition faced by Manhattan hospitals. However, there could be other hospital characteristics that may affect hospitals' sorting ability.

One alternative explanation is the size of hospitals. It may be that bigger hospitals sort better because they have more surgeons available. It is also plausible that smaller hospitals are better in within-hospital sorting since it is easier for them to coordinate internally. However, size is unlikely to be the driving factor here because Manhattan hospitals were not uniformly larger or smaller than hospitals outside Manhattan.³⁶

³⁶There were small hospitals (with two surgeons), middle-sized hospitals (with three to four surgeons), and large hospitals (with more than four surgeons) in Manhattan.

Moreover, all Manhattan hospitals are teaching hospitals according to the definition given by the Association of American Medical Colleges (AAMC). Teaching hospitals may very well sort better because their organizational structure enables them to do so. Many surgeons in teaching hospitals are also faculty members of the associated medical schools. Unlike a typical surgeon who either has his own practice or belongs to a surgeon group, many surgeons in teaching hospitals are actually employees of medical school faculty foundations. Therefore, teaching hospitals may face fewer coordination barriers and may, thus, be able to adjust the level of within-hospital sorting more easily.

Column (2) of Table 7 shows the result confined to only teaching hospitals. If only teaching hospitals are considered, the effect of hospital report cards on within-hospital sorting for hospitals in Manhattan relative to the other teaching hospitals becomes smaller and noisier. Manhattan hospitals did not seem to improve the level of within-hospital sorting relative to teaching hospitals elsewhere after the publication of hospital report cards. However, the effect of hospital report cards on within-hospital sorting was still statistically different from the effect of surgeon report cards for hospitals in Manhattan. This suggests that while there was indeed a teaching-hospital effect, it can not fully explain why the publication of surgeon report cards weakened the effect of hospital report cards for Manhattan hospitals. Furthermore, if the Manhattan effect was just a teaching-hospital effect, we would expect teaching hospitals outside Manhattan to also respond differently than non-teaching hospitals.

To confirm this I replace the variable Manhattan in the previous equation with an indicator variable for teaching status and re-run the regressions. Table 8 shows that when Manhattan hospitals are included, the publication of hospital report cards does improve the level of within-hospital sorting for teaching hospitals relative to the other hospitals. However, the effect is not statistically significant once Manhattan hospitals are taken out of the analysis. Additionally, I replace the time trend 1990-1991 with 1990-1993 in Table 7 to estimate if surgeon report cards created any additional effect on within-hospital sorting. As shown in Table 9, the publication of surgeon report cards decreased the level of within-hospital sorting for Manhattan hospitals relative to the others even when only teaching hospitals are included. This suggests that although teaching hospitals may indeed find it easier to improve the level of within-hospital sorting relative to other hospitals, there are other factors at play that allowed teaching hospitals in Manhattan to take actions that led to changes in within-hospital sorting. As a robustness test, I replicate the results in Table 7 and Table 8 using hospitals on Manhattan Island and hospitals with hospital-specific HHIs in the lowest 25% percentile instead of all hospitals in Manhattan. The results are qualitatively consistent.

4.2 The Competition Effect

To examine the effect of competition on within-hospital sorting directly, I replace the Manhattan variable by the mean hospital-specific HHI in the following equation:

$$\begin{aligned}
 M_{it} &= \beta_0 + \beta_1 R_{iljt} + \beta_2 HHI \times R_{iljt} + \beta_3 HHI \times P_t + \beta_4 HHI \times Q_t \\
 &+ \beta_5 R_{iljt} \times P_t + \beta_6 R_{iljt} \times Q_t + \beta_7 HHI \times R_{iljt} \times P_t \\
 &+ \beta_8 HHI \times R_{iljt} \times Q_t + X_{iljt} + N_{jt} + A_{ljt} + H_j + Y_t + \varepsilon_{iljt}.
 \end{aligned} \tag{4.3}$$

Note that this is a generalization of the previous model. Here, the degree of competition is measured as a continuous variable, while in the previous model it is a binary variable—whether a hospital is in Manhattan. Each hospital’s mean hospital-specific HHI measures the average degree of competition faced by that hospital. A higher hospital-specific HHI indicates lower competition. Thus, if $\beta_7 > 0$, hospitals under higher competitive pressure improved within-hospital sorting after the publication of hospital report cards compared to hospitals facing less competition.

Table 10 presents the results. Column (1) shows that overall, hospitals with lower mean hospital-specific HHIs did not have higher levels of within-hospital sorting than their counterparts with higher mean hospital-specific HHIs. Although $\hat{\beta}_7 > 0$, it is not statistically significant. One possibility is that the effect of competition might be non-linear in the degree of competition. To account for this, I group hospitals by quartile according to their mean hospital-specific HHIs. Column (2) shows the results for hospitals in the lowest quartile. These hospitals are all in Manhattan and facing a high degree of competition. If the hospital-specific HHI decreases by one standard deviation (0.1), the difference in the RAMR of a surgeon for an average risky patient and that of a surgeon for an average non-risky patient in the same hospital will fall by 0.9 after the publication of hospital report cards. This suggests that in Manhattan, hospitals facing more-intense competition increased the level of within-hospital sorting more.

In column (3), we see qualitatively similar but quantitatively smaller results. Here, the hospitals analyzed are located in Manhattan, the Bronx, and East Long Island, and they face less but still substantial competition compared to the hospitals in column (2). For hospitals in the other two quartiles, which are mostly located in the Upstate region, the effect of competition disappeared. A marginal decrease in the hospital-specific HHI does not have a statistically significant effect on within-hospital sorting after the publication of report cards for these hospitals. This is not an unexpected result given that hospitals in the Upstate region did not seem to respond to report cards at all.

Overall, these results show that the Manhattan effect cannot be fully explained by hospital

characteristics, such as size or teaching status. They also suggest that competition may play a part in driving the Manhattan effect. Note that the competition effect is not very strong. On the one hand, when the degree of competition is treated as binary, hospitals facing more competitive pressure increased the level of within-hospital sorting after the publication of hospital report cards and decreased the level of within-hospital sorting after the publication of surgeon report cards relative to the other hospitals. On the other hand, when using a continuous measure for competition, the impact of competition on within-hospital sorting after the publication of hospital report cards increased with the degree of competition only among hospitals facing high competitive pressure.

5 Conclusion

In this paper I study the different effects of hospital-level and surgeon-level mortality report cards on within-hospital sorting in New York State during the early 1990s. I find that hospitals and surgeons in Manhattan increased the level of within-hospital sorting after the publication of hospital report cards relative to the other hospitals. However, the positive effect of hospital report cards vanished after the publication of surgeon report cards. This may have been driven by the intense competition faced by hospitals and surgeons in Manhattan. If higher level of within-hospital sorting improves patients' treatment outcomes, these results suggest that in Manhattan, the publication of hospital report cards improved patients' welfare, but the effect was weakened by the publication of surgeon report cards.

These findings may provide some insight on the design of quality report cards. In this case, reporting quality at hospital level is more desirable than at surgeon level from the perspective of within-hospital sorting. In other words, measuring and disclosing quality on a higher level—organizations versus sub-sections organizations, may have the advantage of mitigating gaming behaviors of each sub-section and improving the allocation of tasks within organizations.

These results also have implications for information disclosure more generally. We have already learned from previous research that disclose too little information can lead to undesirable outcomes. For example, when quality is multi-dimensional but only some of the dimensions are disclosed, sellers may shift resources from dimensions that are not evaluated (but are valuable to consumers) to dimensions that are evaluated (Lu, 2011). This paper finds that the opposite may also be true—disclosing too much information could also hurt consumers. This implication has two dimensions, measuring quality and reporting quality. First, precisely measuring quality is the key in any quality report cards. In this paper, the unintended consequence of surgeon report cards is driven by the imperfection in the quality

measure. If quality is precisely measured, there will be no incentive for surgeons to reject risky patients, and surgeon report cards may not discourage within-hospital sorting. However, in reality, precisely measuring quality may be very difficult. In this case, what information to disclose becomes crucial. As the theory of second best suggests (Lipsey and Lancaster, 1957), given that a biased quality measure may create room for gaming, more information is not necessarily better.

My results also shed light on the interaction between quality disclosure and competition. Consistent with exiting evidence(Chen, 2008), report cards may have little effect in non-competitive markets. However, competition does not necessarily lead to better quality—while the positive effect of quality disclosure is stronger in competitive markets, its adverse effect is also larger. To the extent that even the best-designed report cards may generate some perverse incentives for sellers, competition may encourage more gaming of the system rather than higher quality.

References

- Afendulis, C. and D. Kessler (2007). Tradeoffs from Integrating Diagnosis and Treatment in Markets for Health Care. *American Economic Review* (97), 1013–1020.
- Allard, M., P. T. Léger, and L. Rochaix. (2009). Provider Competition in a Dynamic Setting. *Journal of Economics & Management Strategy* (18), 457–486.
- Allen, R. and P. Gertler (1991). Regulation and the Provision of Quality to Heterogeneous Consumers. *Journal of Regulatory Economics* (3), 361–375.
- Bundorf, M. K., N. Chun, G. S. Goda, and D. P. Kessler (2009). Do Markets Respond to Quality Information? The Case of Fertility Clinics. *Journal of Health Economics* (28), 718–727.
- Capps, C. and D. Dranove (2004). Hospital Consolidation and Negotiated PPO Prices. *Health Affairs* (23), 175–181.
- Chassin, M. R. (2002). Achieving and Sustaining Improved Quality: Lessons from New York State and Cardiac Surgery. *Quality of Care* (21), 40–51.
- Chassin, M. R., E. L. Hannan, and B. A. DeBuono. (1996). Benefits and Hazards of Reporting Medical Outcomes Publicly. *New England Journal of Medicine* (334), 394–398.
- Chen, M. (2008). Minimum Quality Standards and Strategic Vertical Differentiation: An Empirical Study of Nursing Homes. *Working Paper*.
- Chen, Y. (2011). Why are Health Care Report Cards So Bad (Good)? *Journal of Health Economics* (30), 575–590.
- Cutler, D., R. S. Huckman, and M. B. Landrum (2004). The Role of Information in Medical Markets: An Analysis of Publicly Reported Outcomes in Cardiac Surgery. *American Economic Review* (94), 342–346.
- Dafny, L. and D. Dranove (2008). Do Report Cards Tell Consumers Anything They Don't Already Know? The Case of Medicare HMOs. *RAND Journal of Economics* (39), 790–821.
- Dorfman, R. and P. Steiner (1954). Optimal Advertising and Optimal Quality. *American Economic Review* (44), 826–836.
- Dranove, D. and G. Z. Jin (2010). Quality Disclosure and Certification: Theory and Practice. *Journal of Economic Literature* (48), 935–963.

- Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite (2003). Is More Information Better? The Effects of “Report Cards” on Health Care Providers. *Journal of Political Economy* (111), 555–588.
- Dranove, D. and M. Satterthwaite (1992). Monopolistic Competition when Price and Quality are Imperfectly Observable. *RAND Journal of Economics* (23), 518–534.
- Dranove, D. and A. Sfekas (2008). Start Spreading the News: A Structural Estimate of the Effects of New York Hospital Report Cards. *Journal of Health Economics* (27), 1201–1207.
- Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey. (1998). Comorbidity Measures for Use with Administrative Data. *Medical Care* (36), 8–27.
- Epstein, A. J. (2004). *Understanding Incentives for Responding to Report Cards*. Ph.D. Dissertation, University of Pennsylvania.
- Epstein, A. J. (2006). Do Cardiac Surgery Report Cards Reduce Mortality? Assessing the Evidence. *Medical Care Research and Review* (63), 403–426.
- Epstein, A. J. (2010). Effects of Report Cards on Referral Patterns to Cardiac Surgeons. *Journal of Health Economics* (29), 718–731.
- Figlio, D. N. and M. E. Lucas (2004). What’s in a Grade? School Report Cards and the Housing Market. *American Economic Review* (94), 591–604.
- Fong, K. (2009). Evaluating skilled experts: Optimal scoring rules for surgeons. *Working Paper*.
- Garicano, L. and T. Santos (2004). Referrals. *American Economic Review* (94), 499–525.
- Gaynor, M. (2006). What Do We Know about Competition and Quality in Health Care Markets? *NBER Working Paper No. 12301*.
- Glance, L. G., A. Dick, T. M. Osler, Y. Li, and D. B. Mukamel (2006). Impact of Changing the Statistical Methodology on Hospital and Surgeon Ranking: The Case of the New York State Cardiac Surgery Report Card. *Medical Care* (44), 311–319.
- Green, J. and N. Wintfield (1995). Report Cards on Cardiac Surgeons—Assessing New York State’s Approach. *New England Journal of Medicine* (332), 1229–1233.
- Hannan, E. L., H. Kilburn, Jr., M. Racz, E. Shields, and M. R. Chassin (1994). Improving the Outcomes of Coronary Artery Bypass Surgery in New York State. *Journal of American Medical Association* (271), 761–766.

- Held, P. and M. Pauly (1983). Competition and Efficiency in the End Stage Renal Disease Program. *Journal of Health Economics* (2), 95–118.
- Jin, G. Z. and P. Leslie (2003). The Effect of Informatoin on Product Quality: Evidence from Restaurant Hygiene Grade Cards. *Quarterly Journal of Economics* (118), 409–451.
- Johnson, E. (2011). Ability, Learning, and Career Path of Cardiac Specialists. *Working Paper*.
- Kessler, D. P. and M. B. McClellan (2000). Is Hospital Competition Socially Wasteful? *Quarterly Journal of Economics* (115), 557–615.
- Kim, H. K. and B. S. Black (2011). Does Hospital Infection Reporting Affect Actual Infection Rates, Reported Rates, or Both? A Case Study of Pennsylvania. *Northwestern Law Econ Research Paper No. 11-19*.
- Kolstad, J. (2011). Information and Quality When Motivation is Intrinsic: Evidence from Surgeon Report Cards. *Working Paper*.
- Lipsey, R. and K. Lancaster (1956-1957). The General Theory of Second Best. *Review of Economic Studies* (24), 11–32.
- Lu, S. F. (2011). Information Disclosure, Multitasking and Product Quality: Evidence from Nursing Homes. *Journal of Economics and Management Strategy*, forthcoming.
- Mukamel, D. and A. Mushlin (1998). Quality of Care Information Makes a Difference: An Analysis of Market Share and Price Changes After Publication of the New York State Cardiac Surgery Mortality Reports. *Medical Care* (36), 945–954.
- Omoigui, N., D. Miller, K. Brown, K. Annan, D. Cosgrove, B. Lytle, F. Loop, and E. Topol (1996). Outmigration for Coronary Bypass Surgery in an Era of Public Dissemination of Clinical Outcomes. *Circulation* (93), 27–33.
- Peterson, E., E. DeLong, J. Jollis, L. Muhlbaier, and D. Mark (1998). The Effects of New York's Bypass Surgery Provider Profiling on Access to Care and Patient Outcomes in the Elderly. *Journal of American College of Cardiology* (32), 993–999.
- Pope, D. G. (2009). Reacting to Rankings: Evidence from "America's Best Hospitals". *Journal of Health Economics* (28), 1154–1165.
- Pope, G. C. (1989). Hospital Nonprice Competition and Medicare Reimbursement Policy. *Journal of Health Economics* (8), 147–172.

- Ramanarayanan, S. (2009). Does Practice Make Perfect: An Empirical Analysis of Learning-by-Doing in Cardiac Surgery. *Working Paper*.
- Romano, P. and H. Zhou (2004). Do Well-Publicized Risk-Adjusted Outcomes Reports Affect Hospital Volume? *Medical Care* (42), 367–377.
- Schneider, E. and A. Epstein (1996). Influence of Cardiac-Surgery Performance Reports on Referral Practices and Access to Care. *New England Journal of Medicine* (335), 251–256.
- Wang, J., J. Hockenberry, S.-Y. Chou, and M. Yang (2011). Do Bad Report Cards Have Consequences? Impacts of Publicly Reported Provider Quality Information on the CABG Market in Pennsylvania. *Journal of Health Economics* (30), 392–407.
- Werden, G. J. (1989). The Limited Relevance of Patient Migration Data. *Journal of Health Economics* (8), 363–376.
- Werner, R. M., D. A. Asch, and D. Polsky (2005). Racial Profiling: The Unintended Consequences of Coronary Artery Bypass Graft Report Cards. *Circulation* (111), 1257–1263.
- Werner, R. M., R. T. Konetzka, E. A. Stuart, and D. Polsky (2011). Changes in Patient Sorting to Nursing Homes under Public Reporting: Improved Patient Matching or Provider Gaming? *Health Services Research* (46), 557–571.
- Zhang, Y. (2011). *Essays on Patient Sorting and Welfare*. Ph.D. Dissertation, Northwestern University.
- Zwanziger, J. and G. Melnick (1988). The Effects of Hospital Competition and the Medicare PPS Program on Hospital Cost Behavior in California. *Journal of Health Economics* (8), 457–464.

Tables and Figures

Table 1: Descriptive Statistics for 1986-1993 Patient-Level Data (Sub-Sample)

Variable	Mean	Std. Dev.	Min.	Max.
Age	64	10	23	99
Male	0.7	0.46	0	1
White	0.86	0.35	0	1
Black	0.04	0.2	0	1
Medicare	0.4	0.49	0	1
Medicaid	0.03	0.18	0	1
Selfpay	0.003	0.07	0	1
Any cardiac risk factor	0.31	0.46	0	1
Congestive heart failure	0.18	0.38	0	1
Cardiac arrhythmias	0.14	0.35	0	1
Valvular disease	0.04	0.19	0	1
Died	0.06	0.23	0	1
Post-surgery los (died=0)	11.8	15	0	980

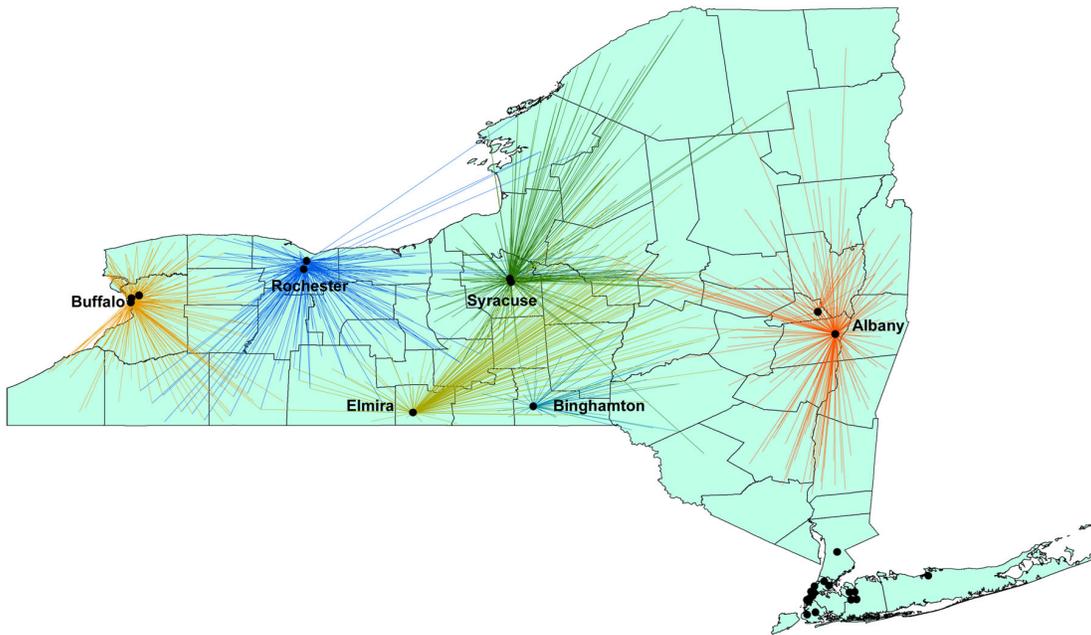
Table 2: Descriptive Statistics for 1986-1993 Hospital/Year-Level Data (Sub-Sample)

Variable	Mean	Std. Dev.	Min.	Max.
Teaching	0.74	0.44	0	1
# of surgeons each year	3.8	1.8	1	9
# of CABG surgeries each year	53	38	1	193

Table 3: CABG Markets and Competition Measure

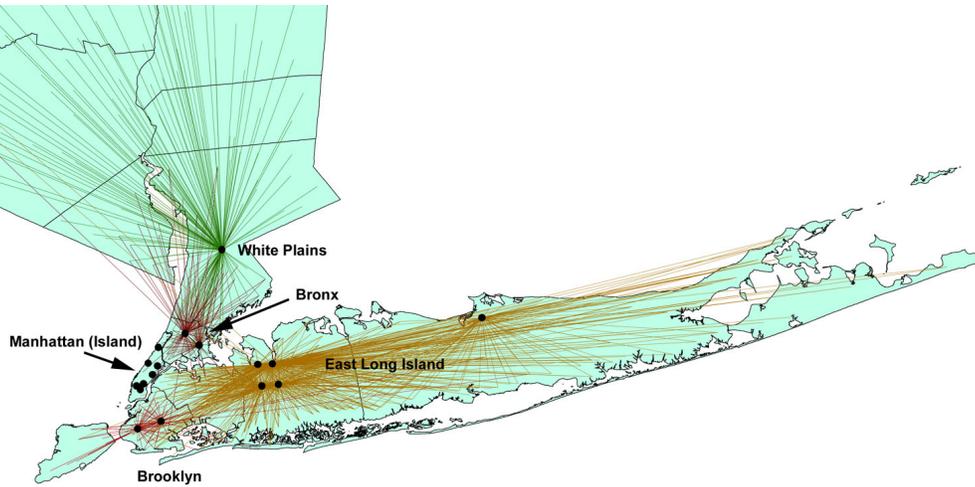
CABG Markets	Number of Hospitals	Mean Hospital Specific HHI	Mean Market HHI	HRR	Region
Albany	3	0.42	0.61	Albany	Upstate
Central NY	4	0.49	0.31	Binghamton	
				Elmira	
				Syracuse	
Rochester	2	0.53	0.39	Rochester	
Buffalo	3	0.48	0.54	Buffalo	NYC-Metro
White Plains	1	0.37	1	White Plains	
Bronx	2	0.35	0.54	Bronx	
Manhattan	11	0.26	0.13	Manhattan	
East Long Island	5	0.29	0.32	East Long Island	

Figure 1: Catchment Areas for Upstate Hospitals



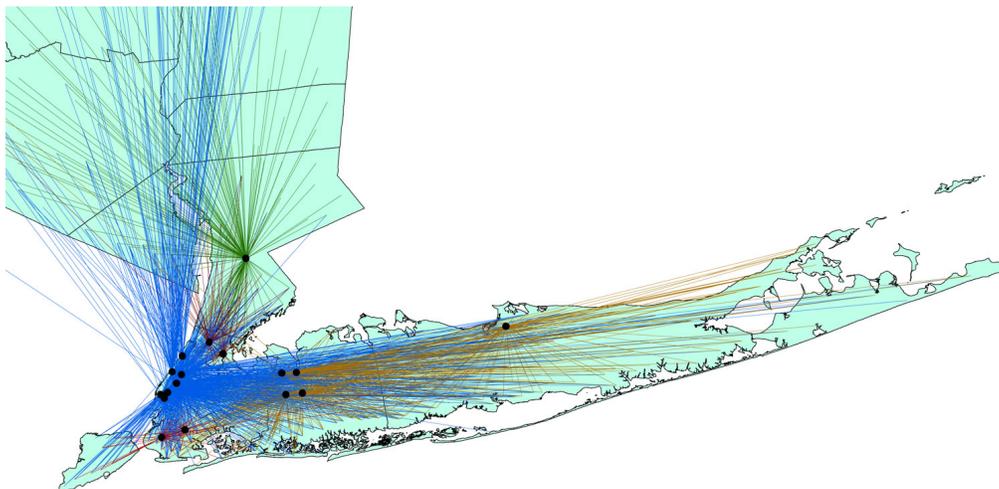
Notes: A black dot represents hospitals in a zip code area. A line connecting a zip code area with a black dot if in that hospital(s) there are at least 3 patients from that zip code area. Lines for hospitals in the same HRR have the same color.

Figure 2: Catchment Areas for NYC-metro Hospitals Excluding Manhattan Island



Notes: A black dot represents hospitals in a zip code area. A line connecting a zip code area with a black dot if in that hospital(s) there are at least 3 patients from that zip code area. Lines for hospitals in the same HRR have the same color.

Figure 3: Catchment Areas for NYC-metro Hospitals



Notes: A black dot represents hospitals in a zip code area. A line connecting a zip code area with a black dot if in that hospital(s) there are at least 3 patients from that zip code area. Lines for hospitals in the same HRR have the same color.

Table 4: Effect of Report Cards on Within-Hospital Sorting by Regions 1986-1993

Dependent Variable	RAMR 89-91	RAMR 89-91	RAMR 89-91
	(1)	(2)	(3)
Risky=1	0.046 (0.029)	0.052 (0.045)	0.021 (0.032)
Risky \times 1990-1991	-0.05 (0.051)	-0.004 (0.066)	-0.084 (0.076)
Risky \times 1992-1993	-0.06 (0.048)	-0.031 (0.063)	-0.03 (0.06)
Regions	All	Upstate	NYC-metro
N	6,052	2,879	3,173
Adjusted R^2	0.57	0.46	0.66
$F(\beta_2 = \beta_3)$	0.03	0.17	0.45

Notes: For emergency AMI patients only. Includes surgeons who have both report card scores. Patient characteristics, medical risk factors, surgeon volume per year, surgeon experience, number of surgeons in each hospital, hospital and year fixed effects included in every column. Clustered (by hospital-year) standard errors in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 5: Effect of Report Cards on Within-Hospital Sorting in the NYC-metro Area 1986-1993

Dependent Variable	RAMR 89-91	RAMR 89-91	RAMR 89-91
	(1)	(2)	(3)
Risky=1	0.015 (0.032)	0.037 (0.035)	0.007 (0.028)
Risky \times 1990-1991	-0.092 (0.08)	-0.138* (0.083)	-0.201*** (0.066)
Risky \times 1992-1993	-0.028 (0.063)	-0.03 (0.074)	0.062 (0.103)
CABG Markets	M+B+E	M+E	Manhattan
N	2,821	2,377	1,617
Adjusted R^2	0.72	0.73	0.62
$F(\beta_2 = \beta_3)$	0.54	1.26	4.57

Notes: M+B+E stands for Manhattan plus Bronx plus East Long Island. Similarly, M+E means Manhattan plus East Long Island. For emergency AMI patients only. Includes surgeons who have both report card scores. Patient characteristics, medical risk factors, surgeon volume per year, surgeon experience, number of surgeons in each hospital, hospital and year fixed effects included in every column. Clustered (by hospital-year) standard errors in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 6: Effect of Report Cards on Within-Hospital Sorting in Manhattan 1986-1993

Dependent Variable	RAMR 89-91 (1)	RAMR 90-92 (2)	RAMR 89-91 (3)	RAMR 89-91 (4)	RAMR 89-91 (5)	RAMR 89-91 (6)
Risky=1	0.007 (0.028)	-0.042 (0.031)	0.01 (0.03)	0.028 (0.031)	0.014 (0.031)	0.003 (0.03)
Risky × 1990-1991	-0.201*** (0.066)	-0.148** (0.071)	-0.219*** (0.06)	-0.09** (0.039)	-0.157*** (0.056)	-0.206*** (0.068)
Risky × 1992-1993	0.062 (0.103)	0.111 (0.105)	0.063 (0.108)	0.041 (0.078)	0.037 (0.092)	0.074 (0.109)
Patient Characteristics	Yes	Yes	Yes	Yes	Yes	No
Surgeon Characteristics	Yes	Yes	Yes	Yes	No	Yes
Surgeons	All	All	No worst 10%	No best 10%	All	All
N	1,617	1,617	1,551	1,429	1,617	1,617
Adjusted R^2	0.62	0.44	0.57	0.69	0.54	0.6
$F(\beta_2 = \beta_3)$	4.57	4.57	5.3	3.31	3.35	4.8

Notes: For emergency AMI patients only. Includes surgeons who have both report card scores. Patient characteristics, medical risk factors, surgeon volume per year, surgeon experience, number of surgeons in each hospital, hospital and year fixed effects included in every column. Clustered (by hospital-year) standard errors in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 7: Effect of Report Cards on Within-Hospital Sorting in Manhattan vs. Others 1986-1993

Dependent Variable	RAMR 89-91	
	(1)	(2)
Risky=1	0.071 (0.043)	0.125** (0.06)
Manhattan \times Risky	-0.072 (0.051)	-0.131** (0.065)
Manhattan \times 1990-1991	0.035 (0.125)	0.08 (0.132)
Manhattan \times 1992-1993	-0.273** (0.126)	-0.212* (0.116)
Risky \times 1990-1991	0.008 (0.057)	-0.047 (0.078)
Risky \times 1992-1993	-0.087 (0.061)	-0.125 (0.082)
Manhattan \times Risky \times 1990-1991	-0.209** (0.097)	-0.155 (0.12)
Manhattan \times Risky \times 1992-1993	0.128 (0.127)	0.184 (0.139)
Hospitals	All	Teaching
N	6,052	3,902
Adjusted R^2	0.57	0.56
$F(\beta_7 = \beta_8)$	5.68	4.72

Notes: For emergency AMI patients only. Includes surgeons who have both report card scores. Patient characteristics, medical risk factors, surgeon volume per year, surgeon experience, number of surgeons in each hospital, hospital and year fixed effects included in every column. Clustered (by hospital-year) standard errors in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 8: Effect of Report Cards on Within-Hospital Sorting in Teaching Hospitals vs. Others 1986-1993

Dependent Variable	RAMR 89-91	
	(1)	(2)
Risky=1	0.005 (0.05)	0.007 (0.048)
Teaching× Risky	0.049 (0.059)	0.11 (0.075)
Teaching× 1990-1991	0.034 (0.077)	0.095 (0.067)
Teaching× 1992-1993	-0.177** (0.073)	-0.077 (0.067)
Risky× 1990-1991	0.068 (0.078)	0.065 (0.078)
Risky× 1992-1993	-0.011 (0.083)	-0.022 (0.081)
Teaching× Risky× 1990-1991	-0.17* (0.097)	-0.105 (0.107)
Teaching× Risky× 1992-1993	-0.04 (0.101)	-0.1 (0.112)
Markets	All	No Manhattan
N	6,052	4,435
Adjusted R^2	0.57	0.55
$F(\beta_7 = \beta_8)$	1.37	0

Notes: For emergency AMI patients only. Includes surgeons who have both report card scores. Patient characteristics, medical risk factors, surgeon volume per year, surgeon experience, number of surgeons in each hospital, hospital and year fixed effects included in every column. Clustered (by hospital-year) standard errors in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 9: Addistional Effect of Surgeon Report Cards in Manhattan vs. Others 1986-1993

Dependent Variable	RAMR 89-91	
	(1)	(2)
Risky=1	0.071 (0.043)	0.125** (0.06)
Manhattan× Risky	-0.072 (0.051)	-0.131** (0.065)
Manhattan× 1990-1993	0.035 (0.125)	0.08 (0.132)
Manhattan× 1992-1993	-0.307** (0.14)	-0.293** (0.121)
Risky× 1990-1993	0.008 (0.057)	-0.047 (0.078)
Risky× 1992-1993	-0.094* (0.056)	-0.078 (0.075)
Manhattan× Risky× 1990-1993	-0.209** (0.097)	-0.155 (0.12)
Manhattan× Risky× 1992-1993	0.337** (0.141)	0.339** (0.156)
Hospitals	All	Teaching
N	6,052	3,902
Adjusted R^2	0.57	0.56

Notes: For emergency AMI patients only. Includes surgeons who have both report card scores. Patient characteristics, medical risk factors, surgeon volume per year, surgeon experience, number of surgeons in each hospital, hospital and year fixed effects included in every column. Clustered (by hospital-year) standard errors in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 10: Competition and the Effect of Report Cards on Within-Hospital Sorting 1986-1993

Dependent Variable	(1)	(2)	(3)	(4)	(5)
	RAMR 89-91	RAMR 89-91	RAMR 89-91	RAMR 89-91	RAMR 89-91
Risky=1	0.033 (0.09)	0.084 (0.469)	1.771 (1.051)	-0.082 (0.849)	0.743* (0.435)
Mean Hospital HHI × Risky	0.01 (0.25)	-0.404 (1.841)	-5.628* (3.296)	0.209 (2.094)	-1.291 (0.79)
Mean Hospital HHI × 1990-1991	-0.063 (0.37)	-24.993*** (3.315)	-5.268* (2.863)	3.872 (4.531)	-0.784 (0.507)
Mean Hospital HHI × 1992-1993	0.877*** (0.332)	-17.556*** (3.976)	-1.718 (2.817)	11.447** (4.981)	-0.203 (0.578)
Risky × 1990-1991	-0.256 (0.168)	-2.423*** (0.602)	-2.71** (1.287)	1.358 (1.177)	-0.783 (0.562)
Risky × 1992-1993	-0.033 (0.169)	0.236 (1.572)	-1.912 (1.269)	0.013 (1.708)	-0.242 (0.533)
Mean Hospital HHI × Risky × 1990-1991	0.542 (0.401)	9.024*** (2.359)	8.821** (4.097)	-3.076 (2.951)	1.362 (1.025)
Mean Hospital HHI × Risky × 1992-1993	-0.001 (0.42)	-0.686 (5.847)	5.956 (4.101)	-0.156 (4.057)	0.387 (0.975)
Hospitals with Mean Hospitals Specific HHI	All	Lowest 25%	25%-50%	50%-75%	Above 75%
N	6,052	1,234	1,587	684	2,547
Adjusted R^2	0.57	0.65	0.79	0.51	0.4
$F(\beta_7 = \beta_8)$	1.5	2.87	0.7	0.39	1.26

Notes: For emergency AMI patients only. Includes surgeons who have both report card scores. Patient characteristics, medical risk factors, surgeon volume per year, surgeon experience, number of surgeons in each hospital, hospital and year fixed effects included in every column. Clustered (by hospital-year) standard errors in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.