Is More Information Better? The E®ects of 'Report Cards' on Health Care Providers

David Dranove,[*] Daniel Kessler,[**] Mark McClellan,[**]
and Mark Satterthwaite[¤]

Preliminary Version - Please Do Not Quote or Distribute

Abstract

Health care report cards | public disclosure of patient health outcomes at the level of the individual physician and/or hospital | may address important informational asymmetries in markets for health care, but they may also give doctors and hospitals incentives to decline to treat more di±cult, severely ill patients. Whether report cards are good for patients and for society depends on whether their ¯nancial and health bene¯ts outweigh their costs in terms of the quantity, quality, and appropriateness of medical treatment that they induce. Using national data on Medicare patients at risk for cardiac surgery, we ¯nd that cardiac surgery report cards in New York and Pennsylvania led to a shift in the incidence of surgery toward healthier patients, to an increase in the overall quantity of surgeries, and to increased matching of patients with hospitals. In turn, this led to higher levels of resource use and to worse health outcomes, particularly for sicker patients. We conclude that, at least in the short run, these report cards decreased patient and social welfare.

1 Introduction

In the past few years, policy makers and researchers alike have given considerable attention to quality \report cards" in sectors such as health care and education. These report cards provide information about the performance of hospitals, physicians, and schools where performance depends both on the skill and e®ort of the producer and the characteristics of their patients/students.

[¤]Northwestern University; **Stanford University and NBER. We would like to thank David Becker for exceptional research assistance, Paul Gertler, Paul Oyer, and Patrick Romano for valuable comments, and seminar participants at the Boston University/Veterans Administration Biennial Health Economics Conference, NBER IO Workshop, Northwestern/Toulouse Joint IO Seminar, Stanford University, UCLA, University of Chicago, University of Illinois, University of Texas, and the University of Toronto for helpful suggestions. Funding from the US National Institutes on Aging and the Agency for Health Care Research and Quality through the NBER is gratefully appreciated. However, all errors are our own.

Perhaps the best known health care report card is New York State's publication of physician and hospital coronary artery bypass graft (CABG) surgery mortality rates. Other states and private consulting ¯rms also publish hospital mortality rates. Many private insurers and consortia of large employers use this information when forming physician and hospital networks and as a means of quality assurance.

The health policy community disagrees on the merits of report cards. Supporters argue that they enable patients to identify the best providers, while simultaneously giving providers powerful incentives to improve quality.[1] Skeptics counter that report cards may encourage providers to avoid di±cult cases and/or treat healthier patients so as to improve their rankings. There are two reasons why such gaming might occur. First, it is essential for the analysts who create report cards to adjust health outcomes for di®erences in patient characteristics ("risk adjustment"), for otherwise providers who treat the most serious cases necessarily appear to have low quality. But analysts can only adjust for characteristics that they can observe. Unfortunately, because of the complexity of patient care, even clinically detailed data on patient characteristics that identify illness severity and probable clinical outcomes are likely to be incomplete. This biases report cards in favor of physicians who are able to select patients on the basis of characteristics that are unobservable to the analysts but predictive of good outcomes.[2] Second, low-skilled surgeons may especially wish to select healthier patients, even if analysts have complete information about health status. When treating healthier patients, low-skilled surgeons are likely to achieve outcomes that are very nearly the same as those achieved by high-skilled surgeons, i.e., such patients are likely to survive no matter who treats them. Thus, if low-skilled surgeons restrict their practices to healthier patients, it may be di±cult for patients who use report cards to con¯dently measure skill levels. The fact that report cards are often based on small samples further aggravates this inference problem.

In this paper, we develop a comprehensive empirical framework for assessing the competing claims about report cards. We apply this framework to the adoption of mandatory CABG surgery report cards in New York and Pennsylvania in the early 1990s. We use a di®erence-in-di®erence (DD) approach. We estimate the e®ect of report cards to be the di®erence in outcome trends after the introduction of report cards in New York and Pennsylvania relative to the di®erence in outcome trends in control states. We identify the net consequences of report cards for health care costs and health outcomes. The bottom line is not favorable: report cards are associated with increases in costs and declines

[1] Dranove and Satterthwaite (1992), which examines price and quality determination in markets where consumers have noisy information about each, identi¯es su±cient conditions for report cards on quality to lead to long run improvements in welfare. While we do not study long run changes in this paper, there is anecdotal evidence that providers did take steps to boost quality after the publication of report cards in New York.

[2] For example, even if such comorbid diseases as diabetes or heart failure are measured accurately for purposes of adjusting report cards, physicians who treat patients with more severe or complex cases of diabetes or heart failure are still likely to have worse measured performance.

in outcomes. At least in the case of cardiac surgery, the concerns of report card skeptics appear to be justi¯ed.

Report cards give consumers the ability to shop for providers more carefully and give providers an incentive to select patients more guardedly. In theory, shopping is welfare-improving, but selection can be either welfare-improving or welfare-reducing, so the net e®ects of report cards on welfare are theoretically indeterminate. We identify empirically three pathways through which shopping and selection may a®ect welfare:[3]

- Matching e®ects, which may arise out of shopping or selection. Suppose, as suggested above, that treatment from a high-skill provider yields greater health bene¯ts to sicker patients. In this case, report cards can improve welfare through improved matching. Speci¯cally, report cards may increase the tendency for sicker patients to be matched with higher quality providers (and healthier patients to be matched with lower quality providers.) One reason is that, based on the quality information provided by report cards, sicker patients are more willing to incur ¯nancial, travel, and search costs to obtain treatment from a high-skilled provider. Another reason is that high-skill providers are less likely than low-skill providers to shun sicker patients because sicker patients are less likely to generate an adverse health outcome for high-skill providers.

- Incidence e®ects, which require selection, but may also arise out of shopping. Once low-skill providers begin selecting patients, then a cascade may occur in which high-skill providers ¯nd it necessary also to select healthier patients in order to avoid invidious comparisons with low-skill providers that successfully manipulate their report cards. In aggregate this can shift the incidence of CABG surgery from sicker to healthier patients, and can a®ect the incidence of other intensive treatments as well, because report cards may induce providers not only to alter their use of CABG but also to avoid treating sicker patients altogether. As clinicians have pointed out, this can be socially harmful if sicker patients derive the greatest bene¯t from bypass surgery (e.g., Topol and Cali® 1994, note 21). On the other hand, it may be socially constructive, if the equilibrium distribution of intensive treatment in the absence of report cards is too heavily weighted toward sicker patients.

- Quantity e®ects, which require selection, but may also arise out of shopping. A provider does not accept every patient who is referred to her as a candidate for surgery because, in the provider's judgement, the patient can more appropriately be treated with a di®erent therapy. Providers need not hold constant the overall quantity of procedures that they perform as they select patients in response to report cards. This too may be socially constructive or harmful, depending on whether at each level of severity

[3] These e®ects are analyzed in more detail in a companion theoretical paper that is under preparation.

the equilibrium level of CABG surgery in the absence of report cards is too high or too low.

We test for matching e®ects by (i) examining whether the release of report cards in New York and Pennsylvania led to increased sorting of patients across hospitals (in terms of their illness severity before treatment) relative to changes in sorting among hospitals in other states and (ii) examining whether the release of report cards led to an increase in the average severity of patients in teaching hospitals relative to other hospitals. To test for incidence and quantity e®ects, we study whether the release of report cards a®ected the probability that a patient at risk for CABG would actually undergo surgery, and whether that probability varies by illness severity. We assess the welfare consequences of report cards by studying how their release a®ected health outcomes and total hospital expenditures for heart attack (AMI) patients, a well-de¯ned population of patients who are at risk for CABG.

This analysis hinges on two key assumptions. First, we assume that the adoption of report cards is uncorrelated with unobserved state-level trends in the treatments, costs, and outcomes of cardiac patients. Second, we assume that AMI patients are a relevant at-risk population for CABG, but that unlike the CABG population, the composition of the AMI population is not a®ected by report cards. We explore the validity of these assumptions below.

The paper proceeds as follows. Section 2 discusses some of the institutional history behind health care quality report cards and summarizes previous research about their e®ects. Section 3 presents our empirical models. It describes in detail how we test for the presence of matching, incidence, and quantity e®ects and how we identify the consequences of report cards for treatment decisions, costs, and outcomes. Section 4 discusses our data sources. Section 5 presents our results and section 6 concludes by discussing the generalizability and implications of our ¯ndings.

2 Background and Previous Research

Basic considerations. Report cards can be categorized along two dimensions: the patient population on which the report is based and the source of information used to calculate risk-adjusted provider quality. With respect to the ¯rst dimension, some report cards publish the outcomes for speci¯c procedures, such as CABG, whereas others report the outcomes for treatment of speci¯c illnesses, such as heart attack or stroke. Because providers can a®ect the type and number of patients receiving a procedure more readily than they can a®ect the type and number of patients they treat with a given illness (see Localio et al. 1997), a procedure-based report card is more susceptible to provider selection than an illness-based report card.

Turning to the second dimension, all report cards employ measures of patient severity to risk-adjust the raw outcomes data. If this were not so, a low-skill provider treating relatively healthy patients would often score better than a high-skill provider treating relatively sick patients. Some report cards use generic hospital discharge abstract data that are collected on essentially all hospital inpatients; fewer use more detailed clinical information that is collected speci¯cally for the report card. The obvious advantages of discharge-abstract based report cards are that they use existing data, do not require any new analysis of clinical records, and are therefore far less costly.

However, predicted outcomes based on severity measures derived from discharge abstracts di®er from predicted outcomes based on severity measures derived from more detailed clinical data (e.g., Hartz and Kuhn 1994, Iezzoni et al.1998). Relative to quality assessments based on clinical data, then, quality assessments based on discharge data su®er from increased noise, which makes them less useful as a tool for distinguishing between high- and low-skill providers. But more invidiously, quality assessments based on discharge data may also suffer from greater bias (e.g., Romano and Chan 2000). First, the narrower scope of discharge data allows providers greater opportunity to engage in selection to improve measured performance. Second, discharge data have historically distinguished poorly between comorbidites present on a patient's admission and complications occurring during the course of (potentially low-quality) care (Selker et al. 1991). Third, because discharge data are self-reported by medical providers for billing and other purposes, they may provide opportunities for increased coding of adverse events, i.e., "death code creep" (Iezzoni 1997b). Fourth, if patient outcomes are not tracked beyond the initial admission, providers may in°uence measured performance by transferring patients who are likely to have adverse outcomes.

For these reasons, most experts view supplemental collection of clinical data, ideally by independent chart reviewers applying consistent standards, as a more reliable foundation for report cards (Romano et al. 1999). Still, depending on their detail and method of collection, clinical record reviews may su®er from many of these limitations as well. Even if chart abstractions are exhaustive and error-free, they are constrained by the accuracy and completeness of the medical record, and by the comprehensiveness of the abstraction instrument. Indeed, "gold standard" record abstractions generally leave unexplained the bulk of variations in patient outcomes such as mortality (e.g., Iezzoni 1997b).[4]

The ability of providers to select patients and the shortcomings of all available risk adjustment systems interact to make it di±cult to determine the net e®ects of the provision of report card information, and whether report cards improve quality of care. Providers care about the report cards they receive because the cards may in°uence the quantity of services demanded from each of them. A provider may bias his report card upward by refusing patients whose expected outcome is worse than the expected outcome that report card's risk adjustment system predicts. The provider can do this for at least some patients because (i) she has some discretion over which patients he accepts and (ii) the clinical data she observes contains much more information predictive of health

[4] For a more detailed discussion of the these and other limitations of standard measures of medical provider performance, see McClellan and Staiger (1999).

outcomes than does the data available for risk adjustment.

Brief history. Several states and the federal government have produced a variety of health care quality report cards.[5] Prior to 1994, only New York and Pennsylvania had mandatory, public report cards that utilized clinical information beyond that recorded in discharge abstracts. Both these states reported outcomes for patients receiving CABG. (Pennsylvania later developed a report card on heart attack patients' outcomes.) Several other states including California and Wisconsin, as well as the U.S. Health Care Financing Administration (HCFA), implemented discharge-abstract based reporting systems, based either on populations with speci¯c illnesses, or populations receiving one or more procedures, or both. Since the national HCFA report card preceded state-level report cards, the discharge-abstract based report cards produced by states are unlikely to have had noticeable e®ects on patient and provider behavior during our study period.[6]

For these reasons, our principal analysis treats New York and Pennsylvania as the two \treatment" states. Beginning in December of 1990 the New York Department of Health released publicly hospital-speci¯c data on raw and risk-adjusted mortality of patients receiving CABG surgery in the previous year. Beginning in 1992, New York also released surgeon-speci¯c mortality (Chassin, Hannan, and DeBuono 1996). Beginning in November of 1992, the Pennsylvania Health Care Cost Containment Council published hospital- and surgeon-speci¯c data on risk-adjusted CABG mortality (Pennsylvania Health Care Cost Containment Council 1998). This would suggest that report cards could have begun to a®ect decision-making in New York in 1991 and in Pennsylvania in 1993, though an alternative hypothesis is that a 1993 e®ective date is also appropriate for New York because the New York report card did not list individual surgeon information until then.

Previous research. The existing empirical literature on report cards can be divided into an arm that surveys the attitudes and reactions of patients and clinicians and an arm that examines treatment decisions and health outcomes. See Marshall et al. (2000) for an excellent catalogue and description of this work. In virtually all published surveys, patients or clinicians view report cards as having little e®ect on decision-making (Schneider and Epstein 1998). In part, this is due to the perceived invalidity of report cards. For example, Hartz et al. (1997) document that CABG report cards' ratings do not agree well with ratings from physician surveys. Schneider and Epstein (1996) ¯nd that cardiovascular specialists did not believe that Pennsylvania report cards provided valid assessments of physician quality. Strikingly, however, even though neither surgeons nor cardiologists considered the report cards valid, a substantial proportion of cardiac surgeons reported that, as a consequence of the report

[5] See Iezzoni (1994, 1997a) and Richards (1994) for a discussion of some of these initiatives. Mennemeyer, Morrisey, and Howard (1997) contains a detailed discussion of HCFA's reporting e®orts.

[6] We check this modeling assumption below by exploring how treatment in states with discharge-abstract based reporting di®ered from treatment in New York and Pennsylvania and from that in other states.

cards' introduction, they only accepted healthier candidates for CABG surgery. Cardiologists confirmed this, reporting that since report cards were instituted they have had more difficulty in placing severely ill candidates for CABG.

Analyses of clinical and administrative data — almost entirely from New York's report card — find important beneficial effects of report cards. This work concludes that New York's report card reduced mortality (Hannan et al. 1994; Peterson et al. 1998), in part by providing incentives for poorly-rated hospitals to improve (Dziuban et al. 1994) and in part by increasing the market shares of more highly-rated physicians and hospitals (Mukamel and Mushlin 1998).

The optimistic findings of these New York studies must be tempered by the potential presence of incidence effects due to provider selection, an issue that studies such as Schneider and Epstein (1996), Leventis (1997), and Hofer et al. (1999) suggest may be of more than academic concern. If providers perform CABG on disproportionately fewer sick patients and if sicker patients benefit more from CABG, then the mortality rate among patients who would have received CABG in the absence of report cards can increase, even as the CABG mortality rate falls. The failure of previous studies to consider the entire population at risk for CABG, rather than only those who receive it, is a potentially severe limitation. Furthermore, none of these studies assess the impact of report cards on the resources used to treat CABG patients. Even if report cards reduce mortality, they may not be socially constructive if they do so at great financial cost.

3 Empirical Models

We examine the effects of the mandatory CABG surgery report card laws adopted by New York and Pennsylvania in the early 1990s. To identify matching, incidence, and quantity effects we study two cohorts of Medicare patients: those who receive treatment for heart attack (acute myocardial infarction, or AMI) and those who receive CABG surgery. Because AMI is a medical emergency that will generally result in hospitalization if it is not immediately fatal, we assume that CABG report cards can not affect in any substantial way the population hospitalized with AMI, especially in the short run. We explore the validity of this assumption below. However, selection into the CABG cohort can be affected by report cards because CABG is an elective procedure in the vast majority of cases (Weintraub et al, 1995; Ho, 1989). That is, every patient who demands treatment for AMI will be admitted to some hospital, and almost always to the hospital to which they intially present. But this is not true of the CABG cohort. Some patients who demand a CABG (or, more realistically, whose cardiologist recommends that they see a cardiac surgeon) may not obtain the procedure because the surgeon decides they are not appropriate candidates. Indeed, as practice norms evolve, these patients will be less likely to be referred for possible CABG in the first place.

Under the assumption that the AMI population is not subject to selection (and therefore free in theory of both incidence and quantity effects), we use the

AMI cohort to test for matching effects. Matching effects may be present in the AMI cohort for two reasons. First, CABG is an important treatment for AMI, so improved matching of those patients who receive CABG will be present in a population of AMI patients. Indeed, 10-15% of elderly AMI patients will get CABG (for non-elderly AMI patients, this number is 20% or even higher); the cohorts are not mutually exclusive. AMI patients also represent a significant portion of the total CABG operations (approximately 25% in the elderly in 1994). Second, and possibly more importantly, a provider's skill at CABG is likely to be correlated with her skill at other important treatments for AMI. Thus, the quality information provided by report cards may lead sicker AMI patients to be more willing than healthier patients to incur financial or other costs to obtain treatment from a high-skill provider.

We use two types of empirical models. First, to test for matching and incidence effects, we model hospital behavior as a function of hospital characteristics and state laws that require report cards. Second, to test for quantity effects and to estimate the net consequences of report cards for health care costs and patient health outcomes, we model patient-level medical treatment decisions, health care costs, and health outcomes as a function of patient characteristics and state laws. In all the models, the question is: did utilization, expenditure and outcome trends in New York and Pennsylvania diverge significantly from trends in other states after the implementation of report cards, holding all else constant?

3.1 Hospital Level Analysis

We first estimate hospital-level models to test for incidence and matching effects. We use comprehensive individual-level Medicare claims data (described below) to calculate the average illness severity before admission or treatment of each hospital's elderly Medicare patients with AMI and elderly Medicare patients receiving CABG. We also calculate the within-hospital coefficient of variation (CV) of the illness severity before treatment of each hospital's patients. We match this information to comprehensive data on the location and characteristics of all U.S. hospitals.

In a hospital-level model, a shift in the incidence of surgery from sicker to healthier patients would imply a greater decline in the average illness severity of patients receiving CABG in hospitals in New York and Pennsylvania after versus before report cards, compared to hospitals in other states during the same time period.[7] In regression terms, this means $p < 0$, where

[7] To confirm that this is not an artifact of differential trends in medical treatment for or health of elderly cardiac patients generally in New York and Pennsylvania versus everywhere else, we compare the DD estimate of the effect of report cards on average illness severity among CABG patients to the DD estimate of the effect of report cards on illness severity of a population at risk for CABG but not subject to selection, i.e. AMI patients.

$$
\begin{aligned}
\ln(h_{lst}) \quad = \quad & A_s + B_t + g \cdot Z_{lst} \\
& + p \cdot L_{st} + q \cdot N_{st} + e_{lst}; \qquad (1)
\end{aligned}
$$

l indexes hospitals, s indexes states, and t indexes time

h_{lst} is the mean of the illness severity before admission or treatment of hospital l's elderly Medicare AMI or CABG patients

A_s is a vector of 50 state fixed effects

B_t is a vector of 8 time fixed effects

Z_{lst} is a vector of hospital characteristics

L_{st} = 1 if hospital is in NY on or after 1991, or in PA on or after 1993, 0 otherwise

N_{st} is the number of hospitals, and its square and cube, in state s at time t[8]

e_{lst} is an error term.

We estimate the models separately based on the health histories of each hospital's AMI and CABG patients. These models weight each hospital (observation) by the number of CABG or AMI patients admitted to that hospital, respectively.

One consequence of improved matching would be increased sorting by illness severity of patients to hospitals. Improved sorting would cause the average within-hospital CV of severity to decline in New York and Pennsylvania, after versus before report cards, compared to hospitals in other states during the same time period. In regression terms, increased sorting means $p < 0$, where p is from the model above and h_{lst} is the CV of the health status before admission or treatment of hospital l's patients, provided the mean of severity remains constant. We also estimate these models separately based on the health histories of each hospital's AMI and CABG patients.

Another consequence of report-card induced matching would be that high-quality hospitals would treat an increasing share of more severely ill patients. Since true quality is not observable, and indeed may or may not be measured accurately by a selection-contaminated CABG report card, we cannot test this hypothesis directly. However, we examine whether the effect of report cards varies with hospital characteristics that are likely to be correlated with true quality, such as teaching status. In regression terms, this means $r > 0$, where r is the estimated coefficient on $Z_{lst}^{TEACH} * L_{st}$ from a model analogous to the one above that also includes this interaction effect, with Z_{lst}^{TEACH} an indicator

[8] We include the number of hospitals in the state as a coarse control for provider participation. If report cards reduce the number of hospitals in a state, they would increase the measured dispersion of patients' health histories at the remaining hospitals, even in the absence of any true effect of report cards on dispersion. Our results do not change if we exclude this variable from the analysis.

variable denoting whether hospital l is a teaching hospital, and h_{lst} de¯ned as the mean of the illness severity before admission or treatment of hospital l's patients.

3.2 Patient Level Analysis

To test for a quantity e®ect, and assess the impact of report cards on medical treatment decisions, health care costs, patient health outcomes, and social welfare, we use comprehensive Medicare claims data to form a cohort of individual AMI patients. This cohort contains information on (i) illness severity in the year before treatment, (ii) the overall intensity of treatment in the year after admission, (iii) whether the individual patient received CABG surgery in the year after admission for AMI, and (iv) all-cause cardiac complications and mortality in the year after admission. Whether or not there is evidence of report-card induced patient selection among CABG patients, analysis of the AMI cohort can be used to assess the welfare implications of report cards. And, because the AMI cohort is at risk for CABG, estimates of the e®ect of report cards on the treatment decisions for AMI patients allow us to identify both a quantity e®ect and the net resource use and outcomes consequences of report cards.

In regression terms, a quantity e®ect of report cards would be measured by p, where p is from

$$C_{kst} = A_s + B_t + g \cdot Z_{kst} + p \cdot L_{st} + e_{kst}; \qquad (2)$$

k indexes patients, s indexes states, t indexes time

C_{kst} is a binary variable = 1 if patient k from state s at time t received CABG surgery

A_s is a vector of 50 state ¯xed e®ects

B_t is a vector of 8 time ¯xed e®ects

Z_{kst} is a vector of patient characteristics

L_{st} = 1 if patient k's residence is in NY on or after 1991, or in PA on or after 1993, 0 otherwise

e_{kst} is an error term.

A positive p implies that report cards increased the probability that an AMI patient receives CABG.

To assess the e®ect of report cards on social welfare, we compare estimates of the e®ect of report cards on the total resources used to treat a patient with AMI to the e®ect of report cards on AMI patients' health outcomes. In regression terms, that means comparing p estimated from two models analogous to the one above: (i) O_{kst} is the dependent variable where O_{kst} is a binary variable = 1 if patient k from state s at time t experienced an adverse health outcome and (ii) $\ln(y_{kst})$ is the dependent variable where y_{kst} is the total hospital expenditures in the year after admission with AMI. If report cards lead to greater resource use without beneficial outcome consequences, then they are welfare-reducing; if report cards lead to no change or reduced resource use with beneficial outcome consequences, then they are welfare-improving. If report cards lead to greater resource use and improved outcomes (or reduced resource use and worse outcomes), then we can calculate the "cost effectiveness" of report-card induced (or report-card restrained) treatment.

Patient-level models also allow us to assess the extent to which report cards affect differently the treatment decisions and health outcomes of sick versus healthy patients, thereby further identifying the mechanism through which report cards affect welfare. In particular, patient-level models allow us to investigate the incidence hypothesis by testing whether report cards affect differently the CABG surgery rate and the welfare of sick versus healthy AMI patients. To explore these possibilities, we estimate models that allow the effect of report cards to vary for patients with different illness severities before treatment:

$$\left\{ \begin{array}{c} \ln(y_{kst}) \\ C_{kst}; O_{kst} \end{array} \right\} = A_s + B_t + g \cdot Z_{kst} + p \cdot L_{st} + q \cdot w_{kst} + r \cdot L_{st} \cdot w_{kst} + e_{kst}; \quad (3)$$

where w_{kst} is a measure increasing in patient k's illness severity.

Finally, we estimate equations (2) and (3) based on a population of CABG patients, in order to replicate the results in the previous literature.

4 Data

We use data from two sources. First, we use comprehensive longitudinal Medicare claims data for the vast majority of individual elderly beneficiaries who were admitted to a hospital either with a new primary diagnosis of AMI or for CABG surgery from 1987-1994. The AMI sample is analogous to that used in Kessler and McClellan (2000), but extended to include rural patients. Patients with admissions for AMI in the prior year were excluded from the AMI cohort. For each individual patient, as a measure of the patient's illness severity before treatment, we calculate total inpatient hospital expenditures for the year prior to admission. We measure the intensity of treatment that the patient receives as total inpatient hospital expenditures in the year after admission. Measures of hospital expenditures were obtained by adding up all inpatient reimbursements (including copayments and deductibles not paid by Medicare) from insurance claims for all hospitalizations in the year preceding or following each patient's initial admission. We also calculate for each patient the total number of days in the hospital in the year prior to admission as an additional measure of illness severity.

We construct three measures of important cardiac health outcomes. Measures of the occurrence of cardiac complications were obtained by abstracting data on the principal diagnosis for all subsequent admissions (not counting transfers and readmissions within 30 days of the index admission) in the year following the patient's initial admission. Cardiac complications included rehospitalizations within one year of the initial event with a primary diagnosis (principal cause of hospitalization) of either subsequent AMI or heart failure (HF). Treatment of cardiac illness is intended to prevent subsequent AMIs XX, and the occurrence of HF requiring hospitalization is evidence that the damage to the patient's heart from ischemic disease has had serious functional consequences. Data on patient demographic characteristics were obtained from the Health Care Financing Administration's HISKEW enrollment ¯les, with death dates based on death reports validated by the Social Security Administration.

Our second principal data source is comprehensive information on U.S. hospital characteristics that the American Hospital Association (AHA) collects. The response rate of hospitals to the AHA survey is greater than 90 percent, with response rates above 95 percent for large hospitals (>300 beds). Because our analysis involves Medicare bene¯ciaries with serious cardiac illness, we examine only nonfederal hospitals that ever reported providing general medical or surgical services (for example, we exclude psychiatric and rehabilitation hospitals from analysis). To assess hospital size, we use total general medical/surgical beds, including intensive care, cardiac care, and emergency beds. We divide hospitals into three broad size categories (small (<100 beds), medium (100-300 beds), and large (>300 beds)) and three ownership categories (public, private for-pro¯t, and private non-pro¯t). We classify hospitals as teaching hospitals if they report at least 20 full-time residents.

Our hospital-level analysis matches the AHA survey with hospital-level statistics calculated from the Medicare cohorts. We use patient-level illness severity before admission or treatment as measured by total hospital expenditures and total number of days in the hospital in the year before admission or treatment to calculate for each hospital the within-hospital CV and mean of these two variables. We use the CV of patients' historical expenditures to measure the dispersion of severely ill patients because the CV is invariant to proportional shifts in the distribution of historical expenditures. However, the CV is not invariant to constant-level shifts in the distribution. Thus, interpretation of the estimated e®ect of report cards on the within-hospital CV of severities as a measure of the degree of sorting of patients across hospitals depends on how report cards shift the distribution of severities. This is likely to be more important in the CABG cohort than in the AMI cohort, because provider selection behavior is more likely to a®ect the distribution of illness severities of patients receiving CABG than it is to a®ect the distribution of severities of AMI patients.

5 Results

Tables 1 and 2 present descriptive statistics for hospitals and patients respectively. As reported in table 1, hospitals subject to report cards (i.e., those in New York and Pennsylvania) account for roughly 14 percent of all hospitals. The CV of patient expenditures and patient days in the year prior to admission is between 1.5 and 2.5, indicating that most hospitals treat patients with heterogeneous medical histories. As reported in table 2, AMI patients averaged between $2690 (1987) and $2977 (1994) in real hospital expenditures in the year prior to admission. These expenditures, however, were concentrated in a small subset of patients. Expenditures in the pooled 1987-94 AMI population become nonzero at the 71st percentile, and reach $9135 at the 90th percentile. CABG patients were slightly sicker in terms of prior hospital utilization (with historical expenditures averaging $3771-$4431), re°ecting the fact that they were all undergoing a procedure intended to treat serious cardiac illness.

Table 3 investigates whether report cards lead to changes in the incidence of CABG surgery, or enhanced matching of patients to hospitals, or both. The estimates in the table are the result of four sets of regressions with di®erent dependent variables: the log of the CV of severity, measured as previous year's inpatient expenditures or previous year's inpatient days, or the log of the mean severity measured in expenditures and days. The unit of analysis for the regressions is the hospital/year. Each table entry represents the coe±cient and standard error (corrected for heteroscedasticity and within state/time cell error correlation) on the dummy variable L_{st}, \Report Card Present in State" from a di®erent model. All values have been multiplied by 100 to facilitate interpretation as percentages.

The top two rows of table 3 show that report cards had an important e®ect on the illness severity of patients receiving CABG surgery, but not on the illness severity of patients with AMI. The introduction of report cards in New York and Pennsylvania is associated with a 2 to 5 percent decline in the illness severity of CABG patients from New York and Pennsylvania as compared to the severity of CABG patients from all other states. No such e®ect was present among AMI patients from New York and Pennsylvania. Indeed, the DD estimate of report cards on AMI patients health status before admission is weakly positive, although this is only statistically signi¯cant assuming the earlier New York e®ective date for report cards.

The bottom two rows of Table 3 suggest that report cards led to greater matching of patients to hospitals on the basis of patients' health status on admission. Among AMI patients, which is the cohort that providers can not shape through selection, report cards led to more homogeneous cardiac patient populations within hospitals: the CV of AMI patients' health histories declined signi¯cantly in New York and Pennsylvania versus everywhere else after report cards.[9] This pattern is not true for CABG patients' health histories. The CV

[9] We investigated the sensitivity of our test for matching to the use of the CV as a measure of dispersion of illness severity because the CV is not invariant to constant-level shifts in the distribution of severities. A constant shift up in the distribution, for example, would result

of CABG patients' historical expenditures increased signi¯cantly, and the CV of CABG patients days in the hospital was roughly unchanged. However, these coe±cients are not easily interpretable as a measure of the e®ect of report cards on matching in the CABG cohort because report cards led to a substantial decline in the ¯rst moment of the distribution of CABG patients' illness severities, which by itself increases the CV. Although our evidence of matching e®ects in the CABG cohort was inconclusive due to substantial reductions in average CABG patient severity, evidence of matching in the AMI cohort was not subject to this concern.

Table 3a explores the validity of the assumption of the exogeneity of report card adoption, i.e., whether the estimated e®ects of report cards are likely attributable to some unmeasured di®erence across states that was changing contemporaneously with the adoption of report cards. One possibility is that the observed New York/Pennsylvania di®erential may re°ect a \large state" effect. To test for this, we reestimated models in table 3 using only the ten most populous states (accounting for about half of all hospitals and patients). The estimates in the left panel of table 3a are consistent with the main results in table 3; there is no \large state" e®ect. A second possibility is that the New York/Pennsylvania di®erential may re°ect di®erent trends in the observable characteristics of patients. To test for this, we report in the right panel of table 3a the e®ects of report cards on hospital means and CVs calculated using risk adjusted patient illness severities (normalized to that of a white male resident of an urban area aged 65-69). The results are once again consistent with those in table 3.

In results not included in the tables, we explored the validity of the assumption of exogeneity of the AMI cohort, i.e., whether report cards a®ect the selection of patients with AMI. First, we investigated whether trends in AMI incidence among individuals 65 and over di®ered in New York and Pennsylvania, to provide a rough check that report cards did not a®ect selection into the AMI cohort. The point estimate of the e®ect of report cards on AMI incidence was minuscule (between two and three orders of magnitude smaller than the average AMI incidence in this period) and insigni¯cant. Second, we investigated whether the estimated e®ects in tables 3 and 3a are due to a di®erential decline in the state-level CV of AMI patients' illness severities in New York and Pennsylvania. Unreported DD estimates of the e®ect of report cards on ln(state/year average CVs of year-prior expenditures) are very small and insigni¯cant.

Table 3b investigates another predicted consequence of report-card induced matching: that an increased proportion of more severely ill patients would seek out and obtain treatment at high-quality hospitals. Since true hospital quality

---

in a decline in the CV by increasing the mean and not a®ecting the standard deviation. In this case, report cards could lead to a decline in the CV, but not a®ect patient sorting. We explored the importance of this hypothesis by estimating the e®ect of report cards on the standard deviation of historical patient expenditures and lengths-of-stay in the AMI population. We found that report cards statistically signi¯cantly decrease the log of the within-hospital standard deviation of patients' historical length of stay, although they do not signi¯cantly decrease the log of the within-hospital standard deviation of patients' historical expenditures.

is very difficult to observe and patient selection may contaminate report card rankings of quality, we use teaching status as a proxy for quality. The table's estimates show that the relative illness severity of both AMI and CABG patients admitted to teaching hospitals in New York and Pennsylvania rose dramatically, relative to the teaching/nonteaching differential in rest of the nation. These estimates confirm table 3's suggestion that report cards improve matching.

Table 4 presents our analysis of the effects of report cards on medical treatment decisions, health care costs, and health outcomes for AMI patients. We report regressions horizontally in pairs: each row of the table presents estimates from equation (2) or equation (3) with a given dependent variable, under the two alternative effective dates of report cards. Report cards increase the probability that the average AMI patient will undergo CABG surgery by .54 or .88 percentage points, depending on the assumed effective date of report cards. These quantity effects are considerable, given that the probability of CABG in an elderly AMI patient was 16.2 percent in 1994. Consistent with selection behavior, the increase in quantity is essentially accounted for by less severely ill patients — those who did not have a hospital admission in the year prior to their AMI.[10] This shift in treatment behavior led to higher levels of hospital expenditures for the average AMI patient (row 3), which is understandable, considering that the average patient is more likely to undergo costly CABG surgery. Surprisingly, however, the positive interaction term (significant only under the assumption of an effective date in New York of 1991, row 4, column 3) indicates that report cards are also driving up costs for the most severely ill patients, despite the fact that they are no more likely to receive CABG.

The bottom six rows of table 4 present estimates of the effects of report cards on patient health outcomes, for the average patient and for patients with and without a prior-year inpatient admission. It shows that report cards increase significantly the average rate of one important complication (readmission with heart failure) for all elderly AMI patients by approximately one-half of one percentage point, and also increase the average mortality rate by .41 percentage points (although this is only significant at the 10 percent level, under the assumption of a New York effective date of 1991). Much more striking, however, is the differential effect of report cards on relatively sicker AMI patients. In spite of report-card induced additional CABG surgeries, less ill elderly AMI patients experience no substantial measurable health benefits (they do have marginally lower rates of readmission with AMI in one specification). In contrast, report cards lead to substantial increases in readmission with heart failure among elderly AMI patients with a prior year's inpatient admission: approximately 2.1 percentage points on a base heart-failure readmission rate of 9.4 percent (table 2). The magnitude of the effects of report cards on the levels of the CABG surgery and the HF readmission rates suggest that their adverse impact on more severely ill patients may be due both to report-card induced reductions

[10] The effect of report cards on more severely ill patients' probability of CABG surgery is the approximately zero sum of the report cards' direct effect and the interaction effect prior year admission.

in CABG surgeries and to other changes in medical practice.[11] Report cards also lead to signi¯cant increases in recurrent AMIs among sicker AMI patients, and to signi¯cantly greater mortality among sicker AMI patients (under the assumption of a New York e®ective date of 1993).

The left panel of Table 4a reports the estimated e®ects of report cards using data from the ten most populous states. The right panel of Table 4a reports the estimated DD e®ects of report cards in models that include a separate linear time trend (1987=0) for New York and Pennsylvania, as well as the full set of state- and time-¯xed e®ects that are present in all models, in order to determine whether the estimates from Table 4 are due to an underlying di®erential trend in treatment of cardiac patients in report card versus all other states. Neither altering the control group nor including controls for a preexisting trend for report card states change the basic ¯ndings from Table 4. Although the standard errors from the models of CABG surgery rates increase enough to make those estimates statistically insigni¯cant, estimates of the e®ects of report cards on the treatment intensity and medical complications of elderly AMI patients are similar: greater intensity (particularly among sicker patients), and substantially worse health outcomes for sicker patients.

We also reestimated, but do not report results from, equations (2) and (3) including additional controls for the discharge-abstract based report cards in California (e®ective 1993) and Wisconsin (e®ective 1991). As discussed above, our principal analysis does not assess the e®ect of the state discharge-abstract based report cards, because it is unlikely that they would have had important incremental e®ects on treatment decision-making during our study period: HCFA discharge-abstract based report cards were present in every state from the start of our study period through mid-1992. The California and Wisconsin report cards di®ered from the New York and Pennsylvania report cards in that they reported mortality by illness, not by operative procedure. The estimated DD e®ects of New York/Pennsylvania report cards in a model with additional controls for California/Wisconsin report cards were virtually unchanged from the estimates in Table 4. In addition, there was no strong DD evidence of incidence or quantity e®ects from California/Wisconsin report cards, although AMI patients in California and Wisconsin showed approximately a 0.5 percentage point decline in heart failure rates after versus before report cards, relative to that in other non-report card states over the same period.

Table 5 is similar to Table 4, but reports estimates of equations (2) and (3) for the population of CABG patients rather than the population of AMI patients. Table 5 shows that applying the methods of the previous literature to our population of elderly CABG patients approximately replicates the ¯ndings of that literature. The overall health status of CABG patients appears to

[11] Estimates of q and r from analogues to equation (3), not reported in the tables, that measure w as ln(inpatient hospital expenditures for the year prior to admission) con¯rm this. These estimates show that the e®ect of report cards on the probability of CABG becomes negative at approximately the 97th percentile of the distribution of prior-year inpatient hospital expenditures, with the e®ect on the HF readmission rate becoming positive much lower in the distribution.

improve as a result of report cards, with signi¯cantly lower rates of AMI and mortality. Our DD estimate of the e®ect of report cards on 1-year mortality of about 1 percentage point is similar to the DD estimate of the e®ect of New York's report cards on 30-day mortality of 0.7 percentage points presented by Peterson et al. (1998). Table 5 further shows that there appear to be no consistent adverse di®erential e®ects of report cards by illness severity. While this is consistent with the ¯ndings in Peterson et al. (1998) and Hannan et al (1994), we o®er a di®erent explanation: observed mortality declined as a result of a shift in incidence of CABG surgeries toward healthier patients, not because CABG surgery report cards improved the outcomes of care for individuals with heart disease.

6 Conclusion

Is the publication of information on outcomes-based measures of hospital quality, which are a function of the hospital's and physicians' skill interacted with the severities of the patients served, constructive or harmful? In markets for health care, which exhibit important asymmetries of information and substantial heterogeneity of providers, patient-background adjusted hospital mortality rates would appear to enable patients to make better-informed hospital choices and to give providers the incentive to make appropriate investments in delivering quality care. On the other hand, because doctors and hospitals inevitably have more detailed information about patients' health than can the developer of a report card, any sort of mandatory reporting mechanism gives providers the incentive to decline to treat more di±cult and complicated patients, which may be either welfare reducing or enhancing depending on the appropriateness treatment decisions in the absence of report cards.

We ¯nd that in the ¯rst few years after their adoption, the New York and Pennsylvania CABG surgery report cards led to substantial selection behavior by providers. Report cards led to a decline in the illness severity of patients receiving CABG in New York and Pennsylvania relative to states without report cards, as measured by hospital utilization in the year prior to admission for surgery. This shift in the incidence of CABG surgery was accompanied by an increase in the overall quantity of surgeries, as measured by the probability that an identi¯ed at-risk patient (i.e., an AMI patient) would receive CABG. There is good evidence that report cards led to increased sorting of patients to providers on the basis of the severity of their illness, which by itself is welfare-improving. In particular, hospitals in New York and Pennsylvania experienced relative declines in the within-hospital heterogeneity of their AMI patient populations, with those two states' teaching hospitals picking up an increasing share of patients with more severe illness.

On net, however, the New York and Pennsylvania report cards reduced patient and social welfare, particularly for patients with more severe forms of cardiac illness. Report cards led on average to higher levels of Medicare hospital expenditures and greater rates of adverse health outcomes. Hospital expenditures post-treatment increased not only for healthier AMI patients (largely due to increased CABG surgery rates), but also for sicker AMI patients (despite their stable or declining surgery rate). Even as the additional CABG surgeries the healthier patients received failed to lead to substantial health benefits, more severely ill AMI patients experienced dramatically worsened health outcomes, because of both report-card induced reductions in CABG surgeries and other changes in medical practice. Among more severely ill patients, report cards led to substantial increases in the rate of heart failure, substantial increases in the rate of recurrent AMI, and, in some specifications, to greater mortality.

Caution should be exercised in interpreting our results too negatively. First, we only measure short run responses, and long run benefits to quality reporting may be positive and large (e.g., Dranove and Satterthwaite 1992). Our analysis is short run because the data we analyze is, at most, for only the first four years of the Pennsylvania and New York report card programs. This period is short enough that the population and skill distribution of providers likely remained largely fixed. In the longer run, however, some surgeons and hospitals may take self-selection to the extreme of exiting the market for CABG procedures while others invest heavily to raise their skills to a higher level.

Second, our results do not imply that report cards are harmful in general. Indeed, the fact that there is evidence of sorting in the AMI population (against which providers cannot easily select) suggests that report cards could be constructive if designed in a way to minimize the incentives and opportunities for provider selection. One potential problem with the New York and Pennsylvania report cards we analyze is that they require reporting on all patients receiving an elective operative procedure — not on a population of patients who suffer from an illness. In our companion theoretical work, we are developing some principles for optimal report card design that reflect the offsetting effects of report cards on patient welfare described here. Future empirical work should analyze recent state initiatives that use detailed clinical data to report on populations of patients with specific illnesses, in order to investigate if such design changes can address the shortcomings of procedure-based report cards. For example, if the quality of care for AMI patients is correlated with the quality of care for CABG and other types of cardiac patients, then report cards on AMI care may also be helpful for identifying high-quality CABG providers. Future work should also measure if report cards in the long run cause providers to take steps to improve quality, a behavioral response that may dominate the short-run harm that the selection response caused during the period we examine here. Finally, report cards and the incentives they create are not unique to health care. Report cards on the performance of schools raise the same issues and therefore also need careful empirical evaluation.

References

[1] Chassin, M., E. Hannan, and B. DeBuono. 1996. Benefits and hazards of reporting medical outcomes publicly. New England J. of Med. 334: 394-98.

[2] Dranove, D. and M. Satterthwaite. 1992. Monopolistic Competition when Price and Quality are not Perfectly Observable. Rand Journal of Economics 23: 518-34.

[3] Dziuban, S., J. McIldu®, S. Miller, and R. Dal Col. 1994. How a New York cardiac surgery program uses outcomes data. Annals of Thoracic Surgery 58: 1871-76.

[4] Hannan, E., H. Kilburn, Jr., M. Racz, E. Shields, M. Chassin. 1994. Improving the outcomes of coronary artery bypass surgery in New York state. J. of Amer. Med. Assoc. 271: 761-766.

[5] Hartz, A., J. Pulido, and E. Kuhn. 1997. Are the best coronary artery bypass surgeons identi¯ed by physician surveys. Amer. J. of Public Health 87: 1645-48.

[6] Hofer, T., R. Hayward, S. Green¯eld, E. Wagner, S. Kaplan, and W. Manning. 1999. The unreliability of individual physician \ report cards" for assessing the costs and quality of care of a chronic disease. J. of the Amer. Med. Assoc. 281 (9 June): 2098-105.

[7] Ho, M., M. Eisenberg, P. Litwin, S. Schae®er, and S. Damon. 1989. Delay between onset of chest pain and seeking medical care: The e®ect of public education. Annals of Emergency Medicine 187: 727-31.

[8] Iezzoni, L. I., ed. 1994. Risk-Adjustment for Measuring Health Care Outcomes. Ann Arbor, MI: Health Administration Press.

[9] Iezzoni, L. I., ed. 1997a. Risk-Adjustment for Measuring Health Care Outcomes. 2d. ed. Chicago, IL: Health Administration Press.

[10] Iezzoni, L. I. 1997b. The Risks of Risk Adjustment. J. of the Amer. Med. Assoc. 278 (19 November): 1600-07.

[11] Iezzoni, L. I., A. S. Ash, M. Shwartz, B. E. Landon, and Y. D. Mackiernan. 1998. Predicting In-Hospital Deaths from Coronary Artery Bypass Graft Surgery: Do Di®erent Severity Measures Give Di®erent Predictions? Medical Care 36: 28-39.

[12] Kessler, D. P. and M. B. McClellan. 2000. Is hospital competition socially wasteful? Quarterly J. of Economics 115: 577-615.

[13] Leventis, A. 1997. Cardiac Surgeons Under Scrutiny: A Testable Patient-Selection Model. Center for Economic Policy Stuides Working Paper 41, Princeton University.

[14] Localio, A. R., B. H. Hamory, A.C. Fisher, and T. R. TenHave. 1997. The Public Release of Hospital and Physician mortality Data in Pennsylvania: A Case Study. Medical Care 35: 272-286.

[15] McClellan, M. B. and D. Staiger. 1999. The quality of health care providers. NBER Working Paper 7327.

[16] Marshall, M. N., P. G. Shekelle, S. Leatherman, and R. H. Brook. 2000. The Public Release of Performance Data: What Do We Expect to Gain? A Review of the Evidence. J. of the Amer. Med. Assoc. 283 (12 April): 1866-74.

[17] Mennemeyer, S. M. Morrisey, and L. Howard. 1997. Death and reputation: How consumers acted upon HCFA mortality data. Inquiry 34: 117-128.

[18] Mukamel, D. and A. Mushlin. 1998. Quality of care information makes a di®erence: An analysis of market share and price changes after publication of the New York state cardiac surgery mortality reports. Medical Care 36: 945-54.

[19] Pennsylvania Health Care Cost Containment Council. 1992. Coronary artery bypass graft surgery: A technical report. Vol. I, 1990 data. Harrisburg, PA.

[20] Peterson, E. D., E. R. DeLong, J. G. Jollis, L. H. Muhlbaier, and D. B. Mark. 1998. The E®ects of New York's Bypass Surgery Provider Pro¯ling on Access to Care and Patient Outcomes in the Elderly. J. of the Am. College of Cardiology 32: 993-99.

[21] Richards, T., B. Blacketer, and D. Rittenhouse. 1994. Statewide, Metropolitan, Corporate, and National E®orts in Monitoring and Reporting Quality of Care. Sacramento, CA: Health Policy and Planning Division, O±ce of Statewide Health Planning and Development.

[22] Romano, P. S. and B. K. Chan. 2000. Risk-Adjusting Acute Myocardial Infarction Mortality: Are APR-DRGs the Right Tool? Health Services Research 34: 1469-89.

[23] Romano, P. S., J. A. Rainwater, D. Antonius. 1999. Grading the Graders: How Hospitals in California and New York Perceive and Interpret Their Report Cards. Medical Care 27: 295-305.

[24] Schneider, E. and A. Epstein. 1996. In°uence of cardiac-surgery performance reports on referral practices and access to care. New England J. of Med. 335: 251-256.

[25] Schneider, E. and A. Epstein. 1998. Use of Public Performance Reports: A Survey of Patients Undergoing Cardiac Surgery. J. of the Amer. Med. Assoc. 279 (27 May): 1638-42.

[26] Selker, H. P., J. L. Gri±th, and R. B. D'Agostino. 1991. A Time-Insensitive Predictive Instrument for Acute Myocardial Infarction Mortality: A Multicenter Study. Medical Care 29: 1196-1211.

[27] Topol, E. and R. Cali®. 1994. Scorecard cardiovascular medicine: Its impact and future directions. Annals of Internal Medicine 120: 65-70.

[28] Weintraub, W., E. Jones, J. Craver, R. Grosswald, and R. Guyton. 1995. Inhospital and long-term outcome after reoperative coronary bypass graft surgery. Circulation 92 (Supplement II): 50-57.

Table 1: Descriptive Statistics on Hospitals

	Weighted by and using health histories of AMI patients		Weighted by and using health histories of CABG patients	
	1987	1994	1987	1994
CV of patients' total hospital expenditures 1yr prior to admission	2.199 (0.445)	2.166 (0.587)	1.556 (0.351)	1.934 (0.281)
CV of patients' total days in hospital 1yr prior to admission	2.439 (0.574)	2.473 (0.751)	1.699 (0.294)	2.245 (0.418)
Number of hospitals in the state	180.3	157.7	31.58	36.52
Hospital size medium (1=yes)	49.7%	51.9%	35.8%	46.5%
Hospital size large	25.3%	20.9%	63.8%	51.0%
Teaching hospital	19.1%	20.5%	46.2%	44.1%
Public ownership	15.7%	13.3%	10.1%	8.7%
For-profit ownership	10.4%	10.1%	7.5%	8.4%
Rural location	26.4%	24.5%	2.7%	3.8%
Subject to report cards	0.00	14.2%	0.00	13.5%
Sample size	5,369 (5,077 with CV)	4,792 (4,389 with CV)	739 (714 with CV)	936 (922 with CV)

Notes: Hospital expenditures in 1995 dollars.

Table 2: Descriptive Statistics on Elderly Medicare Beneficiaries with AMI and Elderly Medicare Beneficiaries Receiving CABG Surgery

	With AMI		Receiving CABG surgery	
	1987	1994	1987	1994
Total hospital expenditures 1 year prior to admission	$2,690 (6,493)	$2,977 (7,464)	$4,431 (7,188)	$3,771 (7,586)
Total days in hospital 1 year prior to admission	4.21 (11.48)	4.22 (13.48)	4.97 (8.63)	3.39 (8.05)
Total hospital expenditures in 1 year after admission	$13,954 (12,759)	$18,078 (18,174)	$28,821 (13,213)	$32,872 (21,416)
CABG w/in 1 year of admission (1=yes)	9.2%	16.2%	XXX	XXX
Readmission w/AMI w/in 1yr of adm	5.8%	5.5%	1.1%	1.2%
Readmission with HF w/in 1yr of adm	9.0%	9.4%	6.1%	6.6%
Mortality w/in 1year of admission	40.2%	32.9%	12.2%	10.7%
Age	76.0	76.4%	71.39	72.54
Gender (1=female)	49.8%	48.7%	34.2%	34.7%
Race (1=black)	5.5%	5.9%	2.4%	3.4%
Rural residence	30.0%	30.9%	28.1%	29.0%
Sample size	218,641	229,215	88,457	146,986

Notes: Hospital expenditures in 1995 dollars. For full sample 1987-1994, sample size is 1,770,452 for AMI patients and 967,882 for CABG patients.

Table 3: Effects of Report Cards on the Within-Hospital Coefficient of Variation (CV) and Mean of Patients' Health Status Before Treatment, Medicare Beneficiaries with AMI and Medicare Beneficiaries Receiving CABG, 1987-94

Dependent Variable	Beneficiaries with AMI		Beneficiaries with CABG	
	Assumes report cards effective 1991 in NY and 1993 in PA	Assumes report cards effective 1993 in NY and PA	Assumes report cards effective 1991 in NY and 1993 in PA	Assumes report cards effective 1993 in NY and PA
1n (Mean of patients' total hospital expenditures 1 year prior to admission)	3.37* (1.77)	1.55 (1.53)	-3.92** (1.54)	-5.30** (1.28)
1n (Mean of patients' total days in hospital 1 year prior to admission)	1.11 (1.79)	1.56 (2.03)	-2.43** (0.81)	-4.51** (2.12)
1n (CV of patients' total hospital expenditures 1 year prior to admission)	-2.32** (0.81)	-2.43** (0.81)	3.00** (1.37)	3.60** (1.38)
1n (CV of patients' total days in hospital 1 year prior to admission)	-4.79** (1.41)	-4.98** (1.42)	0.94 (2.12)	2.74 (2.84)

Notes: Each table entry represents a separate model. Heteroscedasticity-consistent standard errors corrected for within state/time cell correlation in parentheses. Coefficients and standard errors multiplied by 100 to facilitate interpretation. Each observation weighted by the number of patients admitted to the hospital in the cohort in question. *-significantly different from zero at the 10 percent level. **-significantly different from zero at the 5 percent level. Sample sizes: for AMI patients, CV of expenditures = 37,672; CV of LOS = 37,681; mean expenditures = 38,066; mean of LOS = 38,084. Regressions also include controls for number of hospitals in state of residence.

Table 3a: Alternative Models of Effects of Report Cards on the Within-Hospital Coefficient of Variation (CV) and Mean of Patients' Health Status Before Treatment, Medicare Beneficiaries with AMI and Medicare Beneficiaries Receiving CABG, 1987-94

	Hospitals and patients from ten most populous states only		Hospital means and CVs calculated using risk adjusted patient histories	
Dependent Variable	Beneficiaries with AMI	Beneficiaries with CABG	Beneficiaries with AMI	Beneficiaries with CABG
1n (Mean of patients' total hospital expenditures 1 year prior to admission)	2.80 (1.95)	-5.09** (1.98)	3.35* (1.84)	-3.88** (1.53)
1n (Mean of patients' total days in hospital 1 year prior to admission)	-0.38 (1.91)	-5.65** (2.23)	1.07 (1.89)	-3.77* (2.04)
1n (CV of patients' total hospital expenditures 1 year prior to admission)	-2.23** (0.88)	3.36** (1.64)	-2.39** (0.86)	2.96** (1.37)
1n (CV of patients' total days in hospital 1 year prior to admission)	-4.39** (1.51)	1.38 (2.17)	-4.76** (1.39)	0.94 (2.16)

Notes: Each table entry represents a separate model. Models assume report cards effective 1991 in NY and 1993 in PA. Heteroscedasticity-consistent standard errors corrected for within state/time cell correlation in parentheses. Coefficients and standard errors multiplied by 100 to facilitate interpretation. Each observation weighted by the number of patients admitted to the hospital in the cohort in question. *-significantly different from zero at the 10 percent level. **-significantly different from zero at the 5 percent level. Ten most populous states are California, Florida, Illinois, Michigan, New Jersey, New York, North Carolina, Ohio, Pennsylvania and Texas. Samples sizes for regressions using hospitals and patients from ten most populous states: for AMI patients, CV of expenditures = 16,719; CV of LOS = 16,720; mean expenditures = 16,836; mean of LOS = 16,838. For CABG patients, CV of expenditures and LOS = 3,439; mean expenditures and LOS = 3,541.

Table 3b: Effects of Report Cards for Teaching and All Other Hospitals on the Mean of Patients' Health Status Before Treatment, Medicare Beneficiaries with AMI and Medicare Beneficiaries Receiving CABG, 1987-94

Dependent Variable	Beneficiaries with AMI: Report cards effective 1991 in NY and 1993 in PA	Beneficiaries with AMI: (Report card effective 1991 in NY and 1993 in PA) * teaching hospital	Beneficiaries with CABG: Report cards effective 1991 in NY and 1993 in PA	Beneficiaries with CABG: (Report card effective 1991 in NY and 1993 in PA) * teaching hospital
1n (Mean of patients' total hospital expenditures 1 year prior to admission)	-1.78 (2.66)	15.05** (5.74)	-18.63** (2.85)	19.78** (2.84)
1n (Mean of patients' total days in hospital 1 year prior to admission)	-2.06 (2.91)	9.27* (4.84)	-11.38** (2.53)	10.28** (2.41)

Notes: Each table entry represents a separate model. Heteroscedasticity-consistent standard errors corrected for within state/time cell correlation in parentheses. Coefficients and standard errors multiplied by 100 to facilitate interpretation. Each observation weighted by the number of patients admitted to the hospital in the cohort in question. *-significantly different from zero at the 10 percent level. **-significantly different from zero at the 5 percent level. Sample size: for CV of expenditures = 37,672; for CV of LOS = 37,681; for mean expenditures = 38,066; for mean of LOS = 38,084. Regressions also include controls for number of hospitals in state of residence.

Table 4: Effects of Report Cards on CABG Surgery Rates, Total Hospital Expenditures, and Health Outcomes of Individual Medicare Beneficiaries with AMI, 1987-94

	Assumes report cards effective 1991 in NY and 1993 in PA			Assumes report cards effective 1993 in NY and PA		
Dependent Variable	Effect of report cards	Inpatient admission to hospital in year before AMI	Report cards* prior year admission	Effect of report cards	Inpatient admission to hospital in year before AMI	Report cards* prior year admission
CABG surgery within 1 year of admission (1=yes)	0.54* (0.28)			0.88** (0.30)		
	0.75* (0.41)	-3.83** (0.11)	-0.63 (0.59)	1.35** (0.38)	-3.81** (0.11)	-1.50** (0.45)
ln(total hospital expends in year after admission)	4.02** (0.97)			4.05** (1.09)		
	2.99** (0.97)	7.29** (0.25)	3.37** (1.24)	3.41** (1.00)	7.40** (0.26)	1.94 (1.25)
Readmission with AMI w/in 1yr of adm (1=yes)	0.02 (0.09)			0.04 (0.10)		
	-0.15* (0.09)	1.71** (0.04)	0.55** (0.18)	-0.12 (0.11)	1.72** (0.04)	0.52** (0.12)
Readmission with HF w/in 1yr of adm (1=yes)	0.50** (0.13)			0.55** (0.15)		
	-0.20 (0.16)	4.89** (0.07)	2.27** (0.26)	-0.17 (0.20)	4.93** (0.07)	2.30** (0.34)
Mortality w/in 1year of admission (1=yes)	0.41* (0.24)			0.38 (0.30)		
	0.33 (0.29)	11.90** (0.09)	-0.02 (0.47)	0.07 (0.32)	11.88** (0.09)	0.69** (0.28)

Notes: Heteroscedasticity-consistent standard errors corrected for within state/time cell correlation in parentheses. Coefficients and standard errors multiplied by 100 to facilitate interpretation. For expenditures models N = 1,768,585; for all other models N = 1,770,452.

Table 4a: Alternative Models of Effects of Report Cards on CABG Surgery Rates, Total Hospital Expenditures, and Health Outcomes of Individual Medicare Beneficiaries with AMI, 1987-94

	Hospitals and patients from ten most populous states only			Linear time trend included for NY and PA		
Dependent Variable	Effect of report cards	Inpatient admission to hospital in year before AMI	Report cards* prior year admission	Effect of report cards	Inpatient admission to hospital in year before AMI	Report cards* prior year admission
CABG surgery within 1 year of admission (1=yes)	0.46 (0.30)			0.36 (0.40)		
	0.72* (0.43)	-3.68** (0.19)	-0.82 (0.62)	0.54 (0.51)	-3.83** (0.11)	-0.63 (0.59)
ln(total hospital expends in year after admission)	4.16** (1.14)			4.68** (2.05)		
	3.37** (1.13)	7.79** (0.41)	2.61** (1.31)	3.71* (2.11)	7.29** (0.25)	3.37** (1.24)
Readmission with AMI w/in 1yr of adm (1=yes)	-0.09 (0.10)			-0.24 (0.17)		
	-0.25** (0.10)	1.73** (0.06)	0.53** (0.19)	-0.40 (0.17)	1.70** (0.04)	0.55** (0.18)
Readmission with HF w/in 1yr of adm (1=yes)	0.50** (0.13)			0.03 (0.20)		
	-0.15 (0.14)	4.96** (0.10)	2.16** (0.27)	-0.62** (0.21)	4.89** (0.68)	2.27** (0.26)
Mortality w/in 1year of admission (1=yes)	0.12 (0.23)			-0.01 (0.30)		
	0.09 (0.28)	11.95** (0.13)	-0.02 (0.48)	0.03 (0.35)	11.90** (0.01)	-0.02 (0.47)

Notes: Models assume report cards effective 1991 in NY and 1993 in PA. Heteroscedasticity-consistent standard errors corrected for within state/time cell correlation in parentheses. Coefficients and standard errors multiplied by 100 to facilitate interpretation. Ten most populous states are California, Florida, Illinois, Michigan, New Jersey, New York, North Carolina, Ohio, Pennsylvania and Texas. For ten most populous states expenditures models N = 929,495; all most populous states models N = 930,548.

Table 5: Effects of Report Cards on Total Hospital Expenditures and Health Outcomes of Individual Medicare Beneficiaries Receiving CABG Surgery, 1987-94

	Assumes report cards effective 1991 in NY and 1993 in PA			Assumes report cards effective 1993 in NY and PA		
Dependent Variable	Effect of report cards	Inpatient admission to hospital in year before CABG	Report cards* prior year admission	Effect of report cards	Inpatient admission to hospital in year before CABG	Report cards* prior year admission
ln(total hospital expends in year after admission)	8.01** (1.90)			5.70** (0.29)		
	6.79** (1.94)	2.46** (0.22)	2.75** (0.52)	4.52** (1.79)	2.50** (0.22)	2.84** (0.57)
Readmission with AMI w/in 1yr of adm (1=yes)	-0.13** (0.06)			-0.20** (0.05)		
	-0.17** (0.06)	0.22** (0.02)	0.10 (0.08)	-0.20** (0.07)	0.23** (0.02)	0.00 (0.09)
Readmission with HF w/in 1yr of adm (1=yes)	0.12** (0.19)			0.30 (0.21)		
	-0.02 (0.19)	3.47** (0.60)	0.43 (0.32)	0.01 (0.21)	3.46** (0.06)	0.87** (0.33)
Mortality w/in 1year of admission (1=yes)	-1.18** (0.20)			-1.01** (0.27)		
	-1.03** (0.25)	2.72** (0.09)	-0.24 (0.23)	-0.86** (0.35)	2.71** (0.09)	-0.21 (0.30)

Notes: Heteroscedasticity-consistent standard errors corrected for within state/time cell correlation in parentheses. Coefficients and standard errors multiplied by 100 to facilitate interpretation. For expenditures models N = 965,942; for all other models N = 967,882.