

Is More Information Better? The Effects of “Report Cards” on Health Care Providers

David Dranove

Northwestern University

Daniel Kessler

Stanford University, Hoover Institution, and National Bureau of Economic Research

Mark McClellan

Council of Economic Advisers, Stanford University, and National Bureau of Economic Research

Mark Satterthwaite

Northwestern University

Health care report cards—public disclosure of patient health outcomes at the level of the individual physician or hospital or both—may address important informational asymmetries in markets for health care, but they may also give doctors and hospitals incentives to decline to treat more difficult, severely ill patients. Whether report cards are good for patients and for society depends on whether their financial and health benefits outweigh their costs in terms of the quantity,

We would like to thank David Becker for exceptional research assistance; Paul Gertler, Paul Oyer, Patrick Romano, and two referees for valuable comments; and seminar participants at the Boston University/Veterans Administration Biennial Health Economics Conference, NBER Industrial Organization Workshop, Northwestern/Toulouse Joint Industrial Organization Seminar, Stanford University, University of California at Los Angeles, University of Chicago, University of Illinois, University of Texas, University of Toronto, and Yale University for helpful suggestions. Funding from the U.S. National Institutes on Aging and the Agency for Health Care Research and Quality through the NBER is gratefully appreciated. The views expressed in this paper do not necessarily represent those of the U.S. Government or any other of the authors' institutions.

[*Journal of Political Economy*, 2003, vol. 111, no. 3]
© 2003 by The University of Chicago. All rights reserved. 0022-3808/2003/11103-0002\$10.00

quality, and appropriateness of medical treatment that they induce. Using national data on Medicare patients at risk for cardiac surgery, we find that cardiac surgery report cards in New York and Pennsylvania led both to selection behavior by providers and to improved matching of patients with hospitals. On net, 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.

I. 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 on both the skill and effort of the producer and the characteristics of its patients/students. 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 firms 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 physicians and hospitals, while simultaneously giving providers powerful incentives to improve quality.¹ Skeptics counter that there are at least three reasons why report cards may encourage providers to “game” the system by avoiding sick patients or seeking healthy patients or both. First, it is essential for the analysts who create report cards to adjust health outcomes for differences in patient characteristics (“risk adjustment”), for otherwise providers who treat the most serious cases necessarily appear to have low quality. But analysts can adjust for only characteristics that they can observe. Unfortunately, because of the complexity of patient care, providers are likely to have better information on patients’ conditions than even the most clinically detailed database. For this reason, providers may be able to improve their ranking by selecting patients on

¹ Dranove and Satterthwaite (1992), which examines price and quality determination in markets in which consumers have noisy information about each, identifies sufficient 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.

the basis of characteristics that are unobservable to the analysts but predictive of good outcomes.²

Even if providers do not have superior information on patients' conditions, they may still have two other reasons to engage in selection. Suppose that the difference in outcomes achieved by low- and high-quality providers is greater for sick patients. Considerable circumstantial evidence supports this assumption. For example, Capps et al. (2001) find that sick patients are more willing to incur financial and travel costs to obtain treatment from high-quality providers, suggesting that sick patients have more to gain from doing so. In this case, low-quality providers have strong incentives to avoid the sick and seek the healthy. By shifting their practice toward healthier patients, inferior providers make it difficult for report cards to confidently distinguish them from their high-quality counterparts, because on relatively healthy patients they have almost as good outcomes. In other words, low-quality providers pool with their high-quality counterparts.

Finally, even if risk adjustment were correct in expectation terms but incomplete—that is, risk adjustment produces noisy estimates of true quality—it may not compensate risk-averse providers sufficiently for the downside of treating sick patients. The cost in utility terms to a risk-averse provider of accepting a sick patient would be greater than the cost of accepting a healthy patient, as long as the variance in the unexplained portion of outcomes is greater for the sick than for the healthy. In practical terms, the utility loss from a few bad (risk-adjusted) outcomes that drove a provider to the bottom of the rankings, generated bad publicity, and catastrophically harmed his or her reputation exceeds the utility gain from a corresponding random positive shock.³ The fact that report cards are often based on small samples further aggravates both of these incentive problems.

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 begin by testing for three potential effects of report cards on the treatment of cardiac illness:

² For example, even if such comorbid diseases as diabetes and 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.

³ Dziuban et al. (1994) present a case study focusing on physicians' concerns about the incentives for selection generated by prediction errors in the New York CABG report card.

1. *The matching of patients to providers.*—If sick patients have more to gain by receiving treatment from high-quality providers, then report cards can improve welfare through improved matching of patients to providers. Sick patients disproportionately have an incentive to seek out the best providers. In addition, the best providers have less incentive to shun the sickest patients.
2. *The incidence and quantity of CABG surgeries.*—Provider selection can shift the incidence of CABG surgery from sicker to healthier patients. At the same time, the total number of surgeries may go up or down. As clinicians have pointed out, incidence effects can be socially harmful if sicker patients derive the greatest benefit from bypass surgery (e.g., Topol and Califf 1994, n. 21). On the other hand, they may be socially constructive if the equilibrium distribution of intensive treatment in the absence of report cards is too heavily weighted toward sicker patients.
3. *The incidence and quantity of complementary and substitute treatments.*—For example, a report card-induced decrease in CABG surgeries for sick patients could lead to a shift toward other substitute treatments, such as percutaneous transluminal coronary angioplasty (PTCA). However, if doctors and hospitals institute processes to avoid sicker patients generally, then a report card-induced decrease in CABG could be accompanied by a decrease in substitute treatments. In this case, report card-induced decreases in CABG could be accompanied by decreases in both PTCA and complementary diagnostic procedures such as cardiac catheterization. This too could be welfare-improving or welfare-reducing, depending on the consequences of the changing mix of treatment for health care costs and patient health outcomes.

Then we measure the net consequences of report cards for health care expenditures and patients' health outcomes.

We use a difference-in-difference approach to estimate the short-run effects of report cards in the population of all U.S. elderly heart attack (acute myocardial infarction [AMI]) patients and all elderly patients receiving CABG from 1987 through 1994. We estimate the effect of report cards to be the difference in trends after the introduction of report cards in New York and Pennsylvania relative to the difference in trends in control states. We find that report cards improved matching of patients with hospitals, increased the quantity of CABG surgery, and changed its incidence from sicker patients toward healthier patients. Overall this led to higher costs and a deterioration of outcomes, especially among more ill patients. We therefore conclude that the report cards were welfare-reducing.

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 in contrast to the population of patients who actually receive CABG, the composition of the AMI population is not affected by report cards. We explore the validity of these assumptions below.

The paper proceeds as follows. Section II discusses some of the institutional history behind health care quality report cards and summarizes previous research about their effects. Section III presents our empirical models. It describes in detail how we test for the presence of matching, incidence, and quantity effects and how we identify the consequences of report cards for treatment decisions, costs, and outcomes. Section IV discusses our data sources. Section V presents our results, and Section VI concludes by discussing the generalizability and implications of our findings.

II. Background and Previous Research

Brief history.—Prior to 1994, the federal government and several states produced a variety of health care quality report cards.⁴ Of these, only New York and Pennsylvania had mandatory public report cards that utilized clinical information beyond that recorded in generic hospital discharge abstracts. Both these states reported outcomes for patients receiving CABG. (Pennsylvania later developed a report card on AMI patients' outcomes.) In 1986, the HCFA, followed by several other states including California and Wisconsin, implemented discharge abstract-based reporting systems based either on populations with specific illnesses or on populations receiving one or more procedures, or on both. Since the national HCFA report card preceded state-level report cards and since discharge abstract-based report cards are more likely to suffer from noise and bias problems (e.g., Romano, Rainwater, and Antonius 1999; Romano and Chan 2000), the discharge abstract-based report cards states produced are unlikely to have had noticeable effects on patient and provider behavior during our study period.⁵

For these reasons, our principal analysis treats New York and Pennsylvania as the two "treatment" states. Beginning in December of 1990, the N.Y. Department of Health publicly released hospital-specific data

⁴ See Iezzoni (1994, 1997*a*) and Richards, Blacketer, and Rittenhouse (1994) for a discussion of some of these initiatives. Mennemeyer, Morrisey, and Howard (1997) contains a detailed discussion of the U.S. Health Care Financing Administration's (HCFA's) reporting efforts.

⁵ We check this modeling assumption below by exploring how treatment in states with discharge abstract-based reporting differed from treatment in New York and Pennsylvania and from that in other states.

on raw and risk-adjusted mortality of patients receiving CABG surgery in the previous year. Beginning in 1992, New York also released surgeon-specific mortality (Chassin, Hannan, and DeBuono 1996). Beginning in November 1992, the Pennsylvania Health Care Cost Containment Council published hospital- and surgeon-specific data on risk-adjusted CABG mortality (Pennsylvania Health Care Cost Containment Council 1992). This would suggest that report cards could have begun to affect decision making in New York in 1991 and in Pennsylvania in 1993, though an alternative hypothesis is that a 1993 effective 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 provides mixed evidence on the consequences of report cards. One arm of the literature uses surveys of patients and clinicians to assess the consequences of report cards. Although some surveys suggest that report cards have little effect on decision making (e.g., Schneider and Epstein [1998]; see Marshall et al. [2000] for an excellent catalog and description of this work), other surveys reach the opposite conclusion. For example, in one survey, 63 percent of cardiac surgeons reported that, as a consequence of the report cards' introduction, they accepted only healthier candidates for CABG surgery. Cardiologists confirmed this, with 59 percent reporting that report cards made it more difficult to place severely ill candidates for CABG (Schneider and Epstein 1996).

Another arm of the literature uses analysis of clinical and administrative data, almost entirely from New York's report card, to reach a very different conclusion: it finds that report cards led to dramatic improvements in the quality of care (Hannan et al. 1994; Peterson et al. 1998). Several researchers document the mechanism through which this may have occurred, including inducing poorly rated hospitals to change patterns of care (Dziuban et al. 1994) and enabling highly rated physicians and hospitals to increase their market shares (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 Green and Wintfeld (1995), 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 observed 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.

III. 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 cohorts of AMI patients and cohorts of patients receiving CABG who may or may not have had an AMI. We make two key assumptions. First, we assume that CABG report cards do not affect the composition of the population hospitalized with AMI, especially in the short run. The reason is that AMI is a medical emergency that, unless immediately fatal, generally results in hospitalization, almost always in the hospital at which the patient initially presented. We explore the validity of this assumption below. In contrast, report cards can affect the people who receive CABG because it is an elective procedure in the vast majority of cases (Ho et al. 1989; Weintraub et al. 1995).

Second, we assume that AMI patients are a relevant at-risk population for CABG and therefore are likely to be affected by the adoption of report cards. Bypass surgery is an important treatment for AMI: in 1994, 16 percent of elderly AMI patients got CABG (for nonelderly AMI patients, this number is 20 percent or higher); AMI patients also represent a significant portion of CABG operations (approximately 25 percent for the elderly in 1994). Possibly more important, 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-quality provider.

We estimate two types of empirical models. The first type takes the hospital as the unit of analysis and assesses the effects of report cards on the incidence of CABG and the matching of patients to hospitals. To determine the effect on incidence, we estimate the extent to which the trend over time in the mean health status of CABG patients in New York and Pennsylvania hospitals differed from the trend in hospitals in comparison states. We then compare the difference-in-difference estimates with difference-in-difference estimates for all AMI patients, to investigate whether differential trends in the health status of CABG patients merely reflect differential trends in the overall population of elderly patients with cardiac illness. To determine report cards' effect on the match of patients with hospitals, we investigate whether report cards led to greater within-hospital homogeneity of patients in New York and Pennsylvania. A reduction in the within-hospital variation in pa-

tients' health status on admission in New York and Pennsylvania hospitals relative to hospitals in comparison states is consistent with improved matching.

The second type of empirical model takes the patient as the unit of analysis and assesses the effect of report cards on both (i) the quantity and incidence of CABG and other intensive cardiac treatments and (ii) the resource use and health outcomes that determine the net consequences of report cards for social welfare. In these models, report cards affect the quantity of CABG surgery (or PTCA or cath) if they affect the probability that an AMI patient receives CABG (or PTCA or cath). These models also provide an alternative assessment of incidence effects. Report cards affect the incidence of CABG (or PTCA or cath) if, within the population of AMI patients, report cards have a differential effect on the probability of CABG (or PTCA or cath) for sick versus healthy patients. Finally, these patient-level models allow an assessment of report cards' effects on cost and outcomes.

A. Hospital-Level Analysis

To test for incidence and matching effects at the hospital level, we use comprehensive individual-level Medicare claims data (described below) to calculate the average illness severity of patients who are admitted to each hospital for CABG surgery. To test for incidence effects, we estimate regressions of the form

$$\ln(h_{lst}) = \mathbf{A}_s + \mathbf{B}_t + g \cdot \mathbf{Z}_{lst} + p \cdot L_{st} + q \cdot N_{st} + e_{lst} \quad (1)$$

where l indexes hospitals, s indexes states, and t indexes time, $t = 1987, \dots, 1994$; h_{lst} is the mean of the illness severity before admission or treatment of hospital l 's elderly Medicare CABG patients; \mathbf{A}_s is a vector of 50 state fixed effects; \mathbf{B}_t is a vector of eight time fixed effects; \mathbf{Z}_{lst} is a vector of hospital characteristics, including indicator variables for rural location, medium (100–300 beds) and large size (>300 beds) (omitted category is small size), two ownership categories (public and private for-profit; omitted category is private nonprofit), and teaching status; $L_{st} = 1$ if the hospital is in New York in or after 1991, or in Pennsylvania in or after 1993, zero otherwise; N_{st} is the number of hospitals, and its square and cube, in state s at time t ,⁶ and e_{lst} is an error term. We weight each hospital (observation) by the number of CABG patients admitted to it. The coefficient p is the difference-in-difference estimate of the

⁶ 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.

effect of report cards on the severity of patients who receive CABG. If $p < 0$, then report cards have caused a shift in incidence from sicker to healthier patients.

To confirm that this is not an artifact of differential trends in the health or care of those elderly cardiac patients who reside in New York and Pennsylvania, we also examine the trends for AMI patients. Though at risk for CABG, these patients are not subject to selection. We reestimate equation (1) using the mean illness severity of AMI patients as the dependent variable and compare this difference-in-difference estimate to the difference-in-difference estimate for CABG patients.

We also calculate the within-hospital coefficient of variation of the illness severity before treatment of each hospital's CABG and AMI patients. Improved sorting of patients among hospitals would cause the average within-hospital coefficient of variation of severity to decline in New York and Pennsylvania relative to other states (provided that the mean of severity does not increase). We therefore reestimate (1) using the within-hospital coefficient of variations as dependent variables; an estimated $p < 0$ is then consistent with improved patient sorting.

Report card-induced matching should also lead high-quality hospitals to treat an increasing share of more severely ill patients. Since true quality is not observable, and indeed may not be measured accurately by a selection-contaminated CABG report card, we cannot test this hypothesis directly. However, we can examine whether the effect of report cards varies with hospital characteristics that are likely to be correlated with true quality, such as teaching status. As in equation (1), let h_{lst} again be the mean of the illness severity of hospital l 's CABG and AMI patients, and define Z_{lst}^{TEACH} as an indicator variable denoting whether hospital l is a teaching hospital. Estimate (1) with the interaction $Z_{lst}^{\text{TEACH}} \times L_{st}$ included. If $r > 0$, where r is the estimated coefficient on the interaction, then report cards lead more severely ill patients to be treated at teaching hospitals.

B. Patient-Level Analysis

We also use 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, PTCA, or cath in the year after admission for AMI; and (iv) all-cause mortality and cardiac complications such as readmission for heart failure in the year after admission. To test for a quantity effect on CABG surgery, we estimate the regression

$$C_{kst} = \mathbf{A}_s + \mathbf{B}_l + g \cdot \mathbf{Z}_{kst} + p \cdot L_{st} + e_{kst} \quad (2)$$

where k indexes patients, s indexes states, and t indexes time, $t = 1987, \dots, 1994$; C_{kst} is a binary variable equal to one if patient k from state s at time t received CABG surgery within one year of admission to the hospital for AMI; \mathbf{A}_s is a vector of 50 state fixed effects; \mathbf{B}_t is a vector of eight time fixed effects; \mathbf{Z}_{kst} is a vector of patient characteristics, including indicator variables for rural residency, gender, race (black or nonblack), age (70–74, 75–79, 80–89, and 90–99; omitted group is 65–69), and interactions between gender, race, and age; $L_{st} = 1$ if patient k 's residence is in New York in or after 1991, or in Pennsylvania in or after 1993, zero otherwise; and e_{kst} is an error term. A positive p implies that report cards increased the probability that an AMI patient receives CABG. We measure the quantity effects of report cards on the alternative intensive treatments PTCA and cath by reestimating equation (2) for these treatments instead of CABG.

Our approach to measuring the effect of report cards on outcomes and costs follows the same line. Let O_{kst} be a binary variable equaling one if patient k from state s at time t experienced an adverse health outcome (e.g., heart failure), and let y_{kst} be his total hospital expenditures in the year after admission with AMI. Reestimate (2) with O_{kst} substituted as the dependent variable. If $p > 0$, then report cards increase the incidence of that adverse outcome. Similarly, if the model is run with $\ln(y_{kst})$ as the dependent variable, then $p > 0$ implies that report cards increase expenditures.

To assess the effect of report cards on social welfare, we compare estimates of the effect of report cards on the total resources used to treat a patient with AMI to the effect of report cards on AMI patients' health outcomes. If report cards uniformly increase adverse outcomes and increase costs, then we conclude that their effect on social welfare is negative. If report cards uniformly decrease adverse outcomes and decrease costs, then we conclude that their effect on social welfare is positive. 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 analysis also permits an alternative assessment of incidence effects. To compare the effects of report cards on sick versus healthy patients, we estimate models that include a control for patients' illness severity before treatment and its interaction with L_{st} :

$$\ln_{C_{kst}O_{kst}}(y_{kst}) = \mathbf{A}_s + \mathbf{B}_t + g \cdot \mathbf{Z}_{kst} + p \cdot L_{st} + q \cdot w_{kst} + r \cdot L_{st} \cdot w_{kst} + e_{kst} \quad (3)$$

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

model is estimated with C_{kst} as the dependent variable, then an estimate of $r \neq 0$ implies that report cards altered the incidence of CABG surgery.

In order to replicate the results in the previous literature, we also use the claims data to form a cohort of patients receiving CABG whether or not they had an AMI and estimate equations (2) and (3).

IV. 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 to 1994. The AMI sample is analogous to that used in Kessler and McClellan (2000) but is 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. Treatment of cardiac illness is intended to prevent subsequent AMIs, and the occurrence of heart failure 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 HCFA's health skeleton eligibility write-off enrollment files, 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 beneficiaries with serious cardiac illness, we examine only nonfederal hospitals that ever reported providing general medical or surgical services (e.g., we exclude psychiatric and rehabilitation hospitals from the analysis). To assess hospital size, we use total general medical/surgical beds, including intensive care, cardiac care, and emergency beds. 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 coefficient of variation and mean of these two variables. We use the coefficient of variation of patients' historical expenditures to measure the dispersion of severely ill patients because the coefficient of variation is invariant to proportional shifts in the distribution of historical expenditures. However, the coefficient of variation is not invariant to constant-level shifts in the distribution. Thus interpretation of the estimated effect of report cards on the within-hospital coefficient of variation 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 affect the distribution of illness severities of patients receiving CABG than it is to affect the distribution of severities of AMI patients.

Appendix tables A1 and A2 present descriptive statistics for hospitals and patients, respectively, for the full set of control variables and outcomes used in our analysis. As reported in Appendix table A1, hospitals subject to report cards (i.e., those in New York and Pennsylvania) account for roughly 14 percent of all hospitals. The coefficient of variation 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 Appendix table A2, AMI patients averaged between \$2,690 (1987) and \$2,977 (1994) in real 1995 dollar 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 seventy-first percentile and reach \$9,135 at the ninetieth percentile. The CABG patients were slightly sicker in terms of prior hospital utilization (with historical expenditures averaging \$3,771–\$4,431), reflecting the fact that they were all undergoing a procedure intended to treat serious cardiac illness. The relative trend in the health status of CABG versus AMI patients was strikingly different. While prior year's

TABLE 1
MEAN EXPENDITURES IN YEAR PRIOR TO ADMISSION FOR AMI OR FOR CABG SURGERY,
ELDERLY MEDICARE BENEFICIARIES, 1990 AND 1994

	1990 (1)	1994 (2)	Percentage Change (3)
A. All AMI Patients			
N.Y. and Pa.	\$3,110	\$3,373	.0846
All other states	2,660	2,910	.0940
Conn., Md., and N.J. only	3,055	3,318	.0861
B. All Patients Receiving CABG within One Year of Admission			
N.Y. and Pa.	4,850	4,511	-.0699
All other states	3,657	3,660	.0008
Conn., Md., and N.J. only	5,015	4,934	-.0162
C. AMI Patients Receiving CABG within One Year of Admission			
N.Y. and Pa.	1,867	1,702	-.0883
All other states	1,537	1,585	.0312
Conn., Md., and N.J. only	1,911	1,859	-.0272

hospital expenditures for the AMI population were rising, prior year's expenditures for the CABG population were falling, and the number of patients receiving CABG was rising dramatically as well. Over the 1980s and 1990s, CABG surgery was diffusing to an increasing number of healthier patients.

V. Results

Table 1 presents inflation-adjusted mean hospital expenditures in the year prior to entry into our study cohorts of all AMI and CABG patients from 1990 (prior to report cards) and 1994 (after report cards). Recall that mean expenditures in the year prior to admission are an indicator of that cohort's health status; that is, lower expenditures imply a healthier cohort. Table 1 previews our basic result: report cards led to a dramatic shift in the incidence of intensive cardiac treatment. The data in panel A of table 1 show that the prior year's expenditures for AMI patients in New York and Pennsylvania increased roughly 8.5 percent. Expenditures in all other states increased by 9.4 percent, and in the neighboring states of Connecticut, Maryland, and New Jersey, expenditures grew by 8.6 percent. These data reflect a nationwide increase in treatment intensity for elderly patients with cardiac illness. There is no evidence of a differential change across states in the illness severity of AMI patients, consistent with our assumption that report cards did not affect the composition of this population.

Trends in the hospitalization history of patients receiving CABG surgery looked quite different. As in Appendix table A2, the average growth in the prior year's expenditures of the average CABG patient (with or without AMI) was substantially smaller: CABG was diffusing to healthier patients. But the extent to which the incidence of CABG surgery shifted toward healthier patients differed dramatically across areas. In New York and Pennsylvania, the prior year's hospital expenditures of CABG patients (with or without AMI) fell; in all other states, the prior year's expenditures rose; in the states neighboring New York and Pennsylvania, the prior year's expenditures fell, but by a much smaller amount.

The adoption of report cards in New York and Pennsylvania coincided with a substantial decline in the relative illness severity of CABG versus AMI patients, as compared to the change in illness severity of CABG versus AMI patients in a "control" group of states. This is compelling evidence that the incidence of CABG surgery in New York and Pennsylvania shifted toward healthier patients relative to incidence trends in comparison states.

A. Hospital-Level Analysis: Testing for Incidence and Matching Effects

Table 2 confirms that report cards led to a shift in the incidence of CABG surgery toward healthier patients and provides evidence of enhanced matching of patients to hospitals. The estimates in the table are the result of four sets of regressions, each with a different dependent variable. The unit of analysis for the regressions is the hospital/year. Each table entry represents the coefficient and standard error (standard errors are based on an estimator of the variance-covariance matrix that is consistent in the presence of heteroscedasticity and of any correlation of regression errors within states over time) on the dummy variable L_{st} , report card present in state, from a different model. All values have been multiplied by 100 to facilitate interpretation as percentages. We report results for two alternative effective dates of report cards: (i) 1991 in New York and 1993 in Pennsylvania and (ii) 1993 in both states.

The top two rows of table 2 show that report cards led to a decline in the illness severity of patients receiving CABG surgery, but not in the illness severity of patients with AMI. Report cards are associated with declines of 3.74–5.30 percent (cols. 1 and 2) in the illness severity of CABG patients from New York and Pennsylvania relative to all other states. No such effect was present among AMI patients from New York and Pennsylvania (cols. 3 and 4). Indeed, the difference-in-difference estimate of report cards on AMI patients' health status before admission is weakly positive, although this is statistically significant only for the earlier New York effective date.

The bottom two rows of table 2 suggest that report cards led to greater

TABLE 2
EFFECTS OF REPORT CARDS ON THE WITHIN-HOSPITAL COEFFICIENT OF VARIATION AND MEAN OF PATIENTS' HEALTH STATUS BEFORE
TREATMENT: MEDICARE BENEFICIARIES WITH AMI AND MEDICARE BENEFICIARIES RECEIVING CABG, 1987-94

DEPENDENT VARIABLE	BENEFICIARIES RECEIVING CABG		BENEFICIARIES WITH AMI	
	Assumes Report Cards Effective 1991 in N.Y. and 1993 in Pa. (1)	Assumes Report Cards Effective 1993 in N.Y. and Pa. (2)	Assumes Report Cards Effective 1991 in N.Y. and 1993 in Pa. (3)	Assumes Report Cards Effective 1993 in N.Y. and Pa. (4)
ln(mean of patients' total hospital expenditures one year prior to admission)	-3.92** (1.52)	-5.30** (1.10)	3.37** (1.52)	1.55 (2.26)
ln(mean of patients' total days in hospital one year prior to admission)	-3.74** (1.84)	-4.51** (1.54)	1.11 (2.76)	1.56 (2.95)
ln(CV of patients' total hospital expenditures one year prior to admission)	3.00** (1.39)	3.60** (1.77)	-2.32** (.64)	-2.43** (.66)
ln(CV of patients' total days in hospital one year prior to admission)	.94 (2.22)	2.74 (3.53)	-4.79** (1.79)	-4.98** (2.01)

NOTE.—Each table entry represents a separate model. Standard errors are based on an estimator of the variance-covariance matrix that is consistent in the presence of heteroscedasticity and of any correlation of regression errors within states over time. Coefficients and standard errors are multiplied by 100 to facilitate interpretation. Each observation is weighted by the number of patients admitted to the hospital in the cohort in question. Sample sizes: for AMI patients, coefficient of variation of expenditures, 37,672; coefficient of variation of length of stay, 37,681; mean expenditures, 38,066; mean of length of stay, 38,084. Regressions also include controls for number of hospitals in state of residence.

** Significantly different from zero at the 5 percent level.

matching of patients to hospitals on the basis of patients' health status on admission. Column 3 shows that among AMI patients, which is the cohort that providers cannot shape through selection, report cards led to more homogeneous cardiac patient populations within hospitals: the coefficient of variation of AMI patients' health histories declined significantly in New York and Pennsylvania versus everywhere else. Column 1 shows a different story among CABG patients: the coefficient of variation of CABG patients' historical expenditures increased and that of CABG patients' days in the hospital was roughly unchanged. These coefficients, however, are not straightforwardly interpretable as a measure of the effect of report cards on matching in the CABG cohort because (as just discussed) report cards led to a substantial decline in the mean of the distribution of CABG patients' illness severities. This by itself increases the coefficient of variation. If we assume that the mean illness severity of CABG patients in New York and Pennsylvania would have been equal to that of AMI patients but for report card-induced changes in the incidence of CABG surgery, then the difference in trends in the coefficient of variation of CABG patients' health histories is also consistent with better matching. Depending on the particular model chosen, the difference between the difference-in-difference estimate of report cards on the mean illness severity of CABG patients and AMI patients was 3.5 to seven percentage points. Subtracting this from the difference-in-difference estimate of the effect of report cards on the coefficient of variation of CABG patients' health histories (because $\ln CV = \ln \sigma - \ln \mu$) leads in every specification to a negative net effect.⁷

Table 3 documents the presence of another predicted consequence of report card-induced matching: that an increased proportion of more severely ill patients would obtain treatment at high-quality hospitals. Since true hospital quality 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 results in columns 1 and 2 show that, in spite of the aggregate decline in the illness severity of CABG patients in New York and Pennsylvania, the illness severity of CABG patients at teaching hospitals in those states remained roughly constant. The results in column 3 show that report cards did not change the average severity of AMI patients in the nonteaching hospitals of New York and Pennsylvania. But, according to column 4, after the publication of report

⁷ As a second, direct test of the matching hypothesis, we estimated the effect 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 significantly decrease the log of the within-hospital standard deviation of patients' historical length of stay, although they do not significantly decrease the log of the within-hospital standard deviation of patients' historical expenditures.

TABLE 3
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 RECEIVING CABG		BENEFICIARIES WITH AMI	
	Report Cards Effective 1991 in N.Y. and 1993 in Pa. (1)	(Report Card Effective 1991 in N.Y. and 1993 in Pa.) × Teaching Hospital (2)	Report Cards Effective 1991 in N.Y. and 1993 in Pa. (3)	(Report Card Effective 1991 in N.Y. and 1993 in Pa.) × Teaching Hospital (4)
ln(mean of patients' total hospital expenditures one year prior to admission)	-18.63** (2.42)	19.78** (2.20)	-1.78 (3.95)	15.05** (7.46)
ln(mean of patients' total days in hospital one year prior to admission)	-11.38** (3.03)	10.28** (1.70)	-2.06 (4.89)	9.27 (6.11)

NOTE.—Each table entry represents a separate model. Standard errors are based on an estimator of the variance-covariance matrix that is consistent in the presence of heteroscedasticity and of any correlation of regression errors within states over time. Coefficients and standard errors are multiplied by 100 to facilitate interpretation. Each observation is weighted by the number of patients admitted to the hospital in the cohort in question. Sample size: for coefficient of variation of expenditures, 37,672; for coefficient of variation of length of stay, 37,681; for mean expenditures, 38,066; for mean of length of stay, 38,084. Regressions also include controls for number of hospitals in state of residence.

** Significantly different from zero at the 5 percent level.

cards began, the average severity of these patients among New York and Pennsylvania teaching hospitals increased substantially.⁸

B. Patient-Level Analysis: Testing for Quantity and Incidence Effects

Table 4 presents our analysis of the quantity and incidence effects of report cards on three important intensive treatments received by AMI patients: CABG, PTCA, and cath. We report regressions horizontally in pairs for a given dependent variable: the first row of a pair presents estimates from equation (2) and the second presents estimates from equation (3). Estimated coefficients for other covariates are not reported so as to make it easier to view the main results.

Table 4 contains three key findings. First, report cards led to an increase in the quantity of CABG surgery, and that increase was confined to healthier patients. Second, report cards led to a decrease in PTCA. Third, report cards led to increased delays in the execution of all three intensive treatments, significantly reducing the probability that an AMI patient would receive CABG, PTCA, or cath within one day of admission.

In particular, report cards increase the probability that the average AMI patient will undergo CABG surgery within one year of admission for AMI by 0.60 or 0.91 percentage point, depending on the assumed effective date of report cards. These quantity effects are considerable, given that the probability of CABG within one year for an elderly AMI patient during our sample period was 13.1 percent.⁹ Consistent with table 2's results on incidence, the quantity increase was entirely accounted for by surgeries on less severely ill patients—those who did not have a hospital admission in the year prior to their AMI.¹⁰ This increase in CABG quantity was accompanied by increased time from AMI to CABG: at least for healthier patients, the difference-in-difference estimate of the effect of report cards on the one-day CABG rate was negative and strongly significant.

Report cards also led to substantial reductions in the quantity of other intensive cardiac treatments. The use of PTCA, an alternative revascularization procedure, fell substantially in New York and Pennsylvania

⁸ We reestimated these models with controls for the competitiveness of hospital markets as calculated in Kessler and McClellan (2000, 2002), which did not change the results.

⁹ The proportion of AMI patients who had been hospitalized in the year prior to admission is .292. The first row of table 4 reports that (i) 14.76 percent of AMI patients who had not been hospitalized the previous year received CABG within one year of admission, and (ii) 9.10 percent of AMI patients who had been hospitalized the previous year received CABG within one year of admission. Therefore, the base rate is $0.708 \times 14.76 + 0.292 \times 9.10 = 13.1$ percent.

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 of prior year admission.

relative to other states, although this result is consistently statistically significant only for sicker patients. Depending on specification, the one-year angioplasty rate for all AMI patients fell by 1.69 or 1.22 percentage points, on a base of 12.43 percentage points; the one-year angioplasty rate for sicker patients fell by 1.50 or 1.72 percentage points, on a base of 8.76 percentage points.¹¹ The effect of report cards on one-day PTCA rates was significant for both sick and healthy patients. Although report cards did not affect the one-year cath rate, they led, for both sick and healthy patients, to statistically significant declines in the one-day cath rate, a measure of the rate at which patients are on a rapid track for subsequent intensive therapeutic treatment.¹²

In contrast to CABG, we found no strong pattern of how report cards changed the incidence of PTCA and cath. Except for the one-day rates, the effect of report cards on the quantities of PTCA and cath was roughly similar for sick versus healthy patients. For both PTCA and cath, there is some indication that the decline in their one-day rates was larger for healthy patients than for sick patients.

C. *Patient-Level Analysis: Testing for Outcomes and Welfare Effects*

Table 5 presents estimates of the effects of report cards on hospital expenditures, readmission with cardiac complications, and mortality in the year after initial admission. The first row shows that the shifts in treatment behavior documented in table 4 led to higher levels of hospital expenditures for the average AMI patient. This is understandable, considering that the average patient is more likely to undergo costly CABG surgery. Surprisingly, however, report cards also led to increased expenditures for the most severely ill patients (second row), despite the fact that they were no more likely to receive CABG and were less likely to receive PTCA. The bottom six rows of table 5 present estimates of the effects of report cards on patient health outcomes. They show that report cards increased significantly the average rate of readmission with heart failure by approximately 0.5 percentage point. They also provide statistically marginal evidence that the average mortality rate in New York and Pennsylvania increased by 0.45 percentage point on a base of 33 percent.

Much more striking, however, is the differential effect of report cards

¹¹ This is the sum of the col. 1 and col. 2 coefficients: $-1.50 = -1.73 + .23$; $-1.72 = -.96 - .76$. Standard errors for sick patients allowing for generalized within-state error correlation (not reported in the table) are 0.45 and 0.54 for the results in panels A and B of the table, respectively.

¹² Standard errors for sick patients' one-day cath rate allowing for generalized within-state error correlation (not reported in the table) are 0.63 and 0.62 for the results in panels A and B of the table, respectively.

TABLE 4
EFFECTS OF REPORT CARDS ON CABG, PTCA, AND CATHETERIZATION RATES: MEDICARE BENEFICIARIES WITH AMI, 1987-94

DEPENDENT VARIABLE	A. ASSUMES REPORT CARDS EFFECTIVE 1991 IN N.Y. AND 1993 IN PA.			B. ASSUMES REPORT CARDS EFFECTIVE 1993 IN N.Y. AND PA.		
	Effect of Report Cards (1)	Admission to Hospital in Year before AMI (2)	Report Cards × Prior Year Admission (3)	Effect of Report Cards (4)	Admission to Hospital in Year before AMI (5)	Report Cards × Prior Year Admission (6)
CABG within one year of admission (1=yes) [14.76, 9.10] ^a	.60** (.21)			.91** (.44)		
	.81** (.15)	-3.80** (.15)	-.65 (.44)	1.39** (.42)	-3.78** (.16)	-1.52** (.19)
CABG within one day of admission (1=yes) [5.40, 2.97]	-.78** (.29)			-.59** (.23)		
	-.97** (.40)	-1.73** (.13)	.72* (.41)	-.66** (.30)	-1.71** (.14)	.29 (.30)
PTCA within one year of admission (1=yes) [13.94, 8.76]	-1.69 (1.22)			-1.22 (1.17)		
	-1.73 (1.55)	-3.50** (.17)	.23 (1.15)	-.96 (1.46)	-3.46** (.19)	-.76 (.99)

PTCA within one day of admission (1=yes) [7.81, 4.82]	-2.21** (.85)			-2.06** (.91)		
	-2.55** (1.05)	-2.05** (.16)	1.22* (.70)	-2.22** (1.07)	-2.00** (.18)	.59 (.57)
Cath within one year of admission (1=yes) [40.65, 26.77]	-.81 (1.02)			.24 (.56)		
	-.88 (1.48)	-9.55** (.34)	.48 (1.64)	.72 (.89)	-9.47** (.38)	-1.37 (1.16)
Cath within one day of admission (1=yes) [26.81, 16.25]	-3.75** (1.51)			-2.77** (1.17)		
	-4.28** (1.90)	-7.54** (.38)	2.02 (1.40)	-2.86* (1.46)	-7.45** (.41)	.56 (1.08)

NOTE.—Standard errors are based on an estimator of the variance-covariance matrix that is consistent in the presence of heteroscedasticity and of any correlation of regression errors within states over time. Coefficients and standard errors are multiplied by 100 to facilitate interpretation. For expenditure models, $N = 1,768,585$; for all other models, $N = 1,770,452$.

^a Numbers in brackets are the means for individuals without and with a prior year hospital admission.

* Significantly different from zero at the 10 percent level.

** Significantly different from zero at the 5 percent level.

TABLE 5
EFFECTS OF REPORT CARDS ON HOSPITAL EXPENDITURES AND HEALTH OUTCOMES: MEDICARE BENEFICIARIES WITH AMI, 1987-94

DEPENDENT VARIABLE	A. ASSUMES REPORT CARDS EFFECTIVE 1991 IN N.Y. AND 1993 IN PA.			B. ASSUMES REPORT CARDS EFFECTIVE 1993 IN N.Y. AND PA.		
	Effect of Report Cards (1)	Admission to Hospital in Year before AMI (2)	Report Cards × Prior Year Admission (3)	Effect of Report Cards (4)	Admission to Hospital in Year before AMI (5)	Report Cards × Prior Year Admission (6)
ln(total hospital expenditures in year after admission)	3.92** (1.08)			3.95** (1.52)		
	2.89** (.73)	7.33** (.48)	3.35* (1.75)	3.31** (1.16)	7.44** (.53)	1.93 (1.49)
Readmission with AMI within one year of admission (1=yes)	.02 (.08)			.06 (.07)		
	-.15 (.10)	1.70** (.06)	.55** (.13)	-.11 (.09)	1.72** (.06)	.52** (.14)
Readmission with heart failure within one year of admission (1=yes)	.50** (.10)			.54** (.10)		
	-.20** (.08)	4.89** (.10)	2.27** (.26)	-.18** (.08)	4.93** (.11)	2.30** (.36)
Mortality within one year of admission (1=yes)	.45 (.32)			.45* (.26)		
	.37 (.41)	11.90** (.09)	-.02 (.44)	.13 (.27)	11.88** (.10)	.69** (.13)

NOTE.—Standard errors are based on an estimator of the variance-covariance matrix that is consistent in the presence of heteroscedasticity and of any correlation of regression errors within states over time. Coefficients and standard errors are multiplied by 100 to facilitate interpretation. For expenditure models, $N = 1,768,585$; for all other models, $N = 1,770,452$.

* Significantly different from zero at the 10 percent level.

** Significantly different from zero at the 5 percent level.

on healthy versus sick AMI patients. Owing to report card-induced additional CABG surgeries, less ill AMI patients experienced a small decline in the heart failure readmission rate. In contrast, among AMI patients with a prior year's inpatient admission, report cards led to statistically significant, quantitatively substantial increases in adverse outcomes. Relatively sicker patients experienced higher rates of readmission with heart failure (approximately 2.3 percentage points greater, on a base heart failure readmission rate of 9.4 percent) and higher rates of recurrent AMI (approximately 0.5 percentage point greater, on a base of 5.5 percent). This helps explain the expenditure increase reported above. Finally, in one specification, sicker patients experienced a 0.82-percentage-point statistically significantly higher mortality rate in the report card states; in the other specification, this effect is not significant.¹³

Taken together, our results show that report cards led to increased expenditures for both healthy and sick patients, marginal health benefits for healthy patients, and major adverse health consequences for sicker patients. Thus we conclude that report cards reduced our measure of welfare over the time period of our study.

D. Validity Checks

Table 6 presents estimates based on alternative models of the effects of report cards on key treatment decisions, expenditures, and health outcomes. Panel A of table 6 reports the estimated effects of report cards using only New Jersey, Connecticut, and Maryland (instead of all other states) as the "control" group. Although the statistical significance of some of the effects declines, the basic findings remain intact. Report cards led to a shift in the incidence of CABG from relatively sick to healthy patients. When the alternative control group was used, the quantity of CABG surgeries received by healthier patients increased by 0.98 percentage point whereas the quantity received by sick patients declined by 0.96 percentage point as a result of the introduction of report cards.¹⁴ The one-year PTCA rate for sick patients also declined, by 2.00 percentage points.¹⁵ Although the expenditure consequences of report cards are smaller in magnitude and insignificant in this alternative model, the adverse outcome consequences for sick patients remain significant and large.

Panel B of table 6 reports the estimated difference-in-difference ef-

¹³ The reported estimate equals the sum of the main effect (0.13) and the interacted effect (0.69).

¹⁴ This is calculated as the sum of the col. 1 and col. 3 coefficients: $0.96 = .98 - 1.94$. Its standard error, which is not reported in the table, is 0.48.

¹⁵ Its standard error, which is not reported in the table, is 0.37.

TABLE 6
 ALTERNATIVE MODELS OF EFFECTS OF REPORT CARDS ON CABG SURGERY RATES, HOSPITAL EXPENDITURES, AND HEALTH OUTCOMES
 OF INDIVIDUAL MEDICARE BENEFICIARIES WITH AMI, 1987-94

DEPENDENT VARIABLE	A. HOSPITALS AND PATIENTS FROM N.Y., PA., CONN., MD., AND N.J. ONLY			B. LINEAR TIME TREND INCLUDED FOR N.Y. AND PA.		
	Effect of Report Cards (1)	Admission to Hospital in Year before AMI (2)	Report Cards × Prior Year Admission (3)	Effect of Report Cards (4)	Admission to Hospital in Year before AMI (5)	Report Cards × Prior Year Admission (6)
CABG within one year of admission (1=yes)	.42 (.36)			.27** (.10)		
	.98** (.31)	-2.66** (.41)	-1.94** (.22)	.46** (.18)	-3.80** (.15)	-.65 (.44)
PTCA within one year of admission (1=yes)	-.93 (.96)			.10 (.65)		
	-.50 (1.24)	-2.11** (.44)	-1.50 (.90)	.03 (1.00)	-3.50** (.17)	.23 (1.15)
Cath within one day of admission (1=yes)	-.76 (1.82)			.70 (1.23)		
	-.22 (2.12)	-4.63** (.61)	-1.87 (1.15)	.10 (1.62)	-7.54** (.38)	2.01 (1.40)

ln(total hospital expenditures in year after admission)	1.74 (1.19)			3.49 (2.81)		
	1.96 (1.40)	10.81** (.97)	-.62 (1.83)	2.52 (3.23)	7.33** (.48)	3.36* (1.75)
Readmission with AMI within one year of admission (1=yes)	.17 (.09)			-.23** (.07)		
	.07 (.11)	1.90** (.12)	.35 (.21)	-.38** (.07)	1.70** (.06)	.55** (.13)
Readmission with heart failure within one year of admission (1=yes)	.41** (.09)			.01 (.10)		
	-.04 (.12)	5.52** (.15)	1.57** (.36)	-.64** (.12)	4.89** (.10)	2.27** (.26)
Mortality within one year of admission (1=yes)	.44* (.16)			.10 (.30)		
	.51 (.25)	12.05** (.23)	-.11 (.41)	.13 (.45)	11.90** (.09)	-.02 (.44)

NOTE.—Models assume that report cards are effective 1991 in New York and 1993 in Pennsylvania. Standard errors are based on an estimator of the variance-covariance matrix that is consistent in the presence of heteroscedasticity and of any correlation of regression errors within states over time. Coefficients and standard errors are multiplied by 100 to facilitate interpretation. For expenditure models in panel A, $N = 366,823$; for all other models in panel A, $N = 367,421$. For expenditure models in panel B, $N = 1,768,585$; for all other models in panel B, $N = 1,770,452$.

* Significantly different from zero at the 10 percent level.

** Significantly different from zero at the 5 percent level.

fects 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 fixed effects that are present in all the other models. Its purpose is to determine whether the estimates from tables 4 and 5 are due to an underlying differential trend in treatment of cardiac patients in report card versus all other states. Including controls for a preexisting trend for report card states absorbs neither the differential trends in CABG rates nor the differential trends in cardiac complication rates in report card states versus other areas. The slightly weaker results for expenditures and PTCA rates are not surprising given the correlation between the time trend and the indicator for the presence of report cards in New York and Pennsylvania.

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 (effective 1994) and Wisconsin (effective 1991). As discussed above, our principal analysis does not assess the effect of the state discharge abstract-based report cards because it is unlikely that they would have had important effects 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 differed from the New York and Pennsylvania report cards in that they reported mortality by illness, not by operative procedure. The estimated difference-in-difference effects of the New York/Pennsylvania report cards in a model with additional controls for California/Wisconsin report cards are virtually unchanged from the estimates in table 4. In addition, we did not find robust evidence of incidence or quantity effects from California/Wisconsin report cards, although AMI patients in those two states showed a statistically significant 0.6-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.

In other results not included in the tables, we explored the validity of the assumption of exogeneity of the AMI cohort to states' adoption of report cards, that is, whether report cards affected the selection of patients with AMI across states and over time. First, we investigated whether trends in AMI incidence among individuals 65 and over differed in New York and Pennsylvania in order to provide a rough check that report cards did not affect selection into the AMI cohort. The point estimate of the effect 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 insignificant. Second, we investigated whether the estimated effects in tables 2 and 3 are due to a differential decline in the state-level coefficient of variation of AMI patients' illness severities in New York and Pennsylvania. Unreported difference-in-

difference estimates of the effect of report cards on $\ln(\text{state/year average coefficient of variations of prior year expenditures})$ are very small and insignificant.

Table 7 is similar to table 5 but reports estimates of equations (2) and (3) for the population of CABG patients rather than the population of AMI patients. It shows that applying the methods of the previous literature to our population of elderly CABG patients approximately replicates the findings of that literature. The overall health status of CABG patients appears to improve as a result of report cards, with significantly lower rates of AMI and mortality. Our difference-in-difference estimate of the effect of report cards on one-year mortality of about one percentage point is similar to the difference-in-difference estimate of the effect of New York's report cards on 30-day mortality of 0.7 percentage point that Peterson et al. (1998) presented. Table 7 further shows that there appear to be no consistent adverse differential effects of report cards by illness severity. While this is consistent with the findings in Hannan et al. (1994) and Peterson et al. (1998), we offer a different explanation: observed mortality declined as a result of a shift in the incidence of CABG surgeries toward healthier patients, not because CABG report cards improved the outcomes of care for individuals with heart disease.

VI. Conclusion

Is the publication of information on health outcomes achieved by physicians and hospitals 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, mandatory reporting mechanisms inevitably give providers the incentive to decline to treat more difficult and complicated patients. Doctors and hospitals likely have more detailed information about patients' health than the developer of a report card can, allowing them to choose to treat unobservably (to the analyst) healthier patients. And even if they do not, providers' risk aversion and low-quality providers' desire to pool with their high-quality counterparts may lead them to engage in selection behavior. For these reasons, the net consequences of report cards for patient and social welfare are theoretically indeterminate. Report cards may be either welfare-reducing or welfare-enhancing, depending on the extent of provider selection and the appropriateness of treatment decisions in the absence of report cards.

We report three key findings. First, the New York and Pennsylvania

TABLE 7
EFFECTS OF REPORT CARDS ON TOTAL HOSPITAL EXPENDITURES AND HEALTH OUTCOMES OF INDIVIDUAL MEDICARE BENEFICIARIES
RECEIVING CABG SURGERY, 1987-94

DEPENDENT VARIABLE	A. ASSUMES REPORT CARDS EFFECTIVE 1991 IN N.Y. AND 1993 IN PA.			B. ASSUMES REPORT CARDS EFFECTIVE 1993 IN N.Y. AND PA.		
	Effect of Report Cards (1)	Admission to Hospital in Year before AMI (2)	Report Cards × Prior Year Admission (3)	Effect of Report Cards (4)	Admission to Hospital in Year before AMI (5)	Report Cards × Prior Year Admission (6)
ln(total hospital expenditures in year after admission)	8.28** (3.30)			5.93** (2.67)		
	7.08** (3.42)	2.48** (.39)	2.72** (.80)	4.78** (2.68)	2.52** (.37)	2.80** (.73)
Readmission with AMI within one year of admission (1=yes)	-.10* (.05)			-.17** (.04)		
	-.15** (.04)	.22** (.02)	.10 (.09)	-.17** (.06)	.23** (.02)	.00 (.09)
Readmission with heart failure within one year of admission (1=yes)	.14 (.18)			.31** (.14)		
	-.01 (.11)	3.47** (.07)	.42 (.49)	.02 (.14)	3.46** (.07)	.86* (.46)
Mortality within one year of admission (1=yes)	-1.17** (.28)			-1.02** (.38)		
	-1.02** (.32)	2.72** (.16)	-.24 (.23)	-.86* (.46)	2.72** (.16)	-.21 (.25)

NOTE.—Standard errors are based on an estimator of the variance-covariance matrix that is consistent in the presence of heteroscedasticity and of any correlation of regression errors within states over time. Coefficients and standard errors are multiplied by 100 to facilitate interpretation. For expenditure models, $N = 965,942$; for all other models, $N = 967,882$.

* Significantly different from zero at the 10 percent level.

** Significantly different from zero at the 5 percent level.

CABG surgery report cards led to substantial selection by providers. Report cards led to a decline in the illness severity of patients receiving CABG in New York and Pennsylvania relative to patients in states without report cards, as measured by hospital utilization in the year prior to admission for surgery. In addition, report cards led to significant declines in other intensive cardiac procedures for relatively sick AMI patients.

Second, report cards led to increased sorting of patients to providers on the basis of the severity of their illness. 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. The fact that report cards led to increased delays for both healthy and sick patients in the execution of the three intensive treatments we examine supports our findings of increased selection and increased sorting, because the processes of selection and sorting are likely to take time.

Third, on net, the New York and Pennsylvania report cards reduced our measure of welfare, particularly for patients with more severe forms of cardiac illness. Report cards led to higher levels of Medicare hospital expenditures (although this finding was not statistically significant in specifications using New Jersey, Connecticut, and Maryland as a control group) and greater rates of adverse health outcomes. Hospital expenditures in the year after admission increased not only for healthier AMI patients but also for sicker AMI patients. 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. Among more severely ill patients, report cards led to substantial increases in the rate of heart failure and recurrent AMI and, in some specifications, to greater mortality. The magnitude of the increase in the rate of adverse health outcomes among sick patients is large but plausible, given that it is roughly proportional to the magnitude of the total decrease in the use of the intensive cardiac treatments that we observe and that it was likely accompanied by other changes in medical practice that we do not observe.

How might we explain these seemingly disparate empirical findings? For healthier patients, the increase in CABG surgeries increased Medicare expenditures and led to a small decline in the rate of readmission with heart failure. For sicker patients, doctors and hospitals avoided performing both CABG and PTCA.¹⁶ In response to report cards, hos-

¹⁶ Although we did not find statistically significant decreases in all specifications in the quantity of CABG for AMI patients with a prior year hospital admission, we did find (in supplementary analysis not presented in the tables) other evidence of a decline in the quantity of CABG provided to sicker AMI patients. The prior year's expenditures of AMI

pitals implemented a broad range of changes in marketing, governance, and patient care (Bentley and Nash 1998) that may well have led to greater caution in the utilization of all invasive procedures in sick patients. On net, these changes were particularly harmful. The less effective medical therapies that were substituted for CABG and PTCA, combined with delays in treatment, led sicker patients to have substantially higher frequencies of heart failure and repeated AMIs and ultimately higher total costs of care.

Caution should be exercised in interpreting our results too negatively. First, we measure only 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 pertain, at most, to 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 whereas 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. 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 whether 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 whether 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

patients receiving CABG with a prior year's hospital admission rose everywhere in the United States between 1990 and 1994 but rose by approximately half as much in New York and Pennsylvania (from \$8,315 to \$8,793, or 5.8 percent) as in all other states (from \$7,365 to \$8,389, or 13.9 percent) or as in Connecticut, Maryland, and New Jersey (from \$8,457 to \$9,334, or 10.4 percent).

cards on the performance of schools raise the same issues and therefore also need careful empirical evaluation.

Appendix

TABLE A1
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 one year prior to admission	2.199 (.445)	2.166 (.587)	1.556 (.351)	1.934 (.281)
CV of patients' total days in hospital one year prior to admission	2.439 (.574)	2.473 (.751)	1.699 (.294)	2.245 (.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	.00	14.2%	.00	13.5%
Sample size	5,369	4,792	739	936
Sample size with CV	5,077	4,389	714	922

NOTE.—Hospital expenditures in 1995 dollars. Standard deviations are in parentheses.

TABLE A2
 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	\$2,690	\$2,977	\$4,431	\$3,771
one year prior to admission	(6,493)	(7,464)	(7,188)	(7,586)
Total days in hospital one year	4.21	4.22	4.97	3.39
prior to admission	(11.48)	(13.48)	(8.63)	(8.05)
Total hospital expenditures	\$14,634	\$18,959	\$30,226	\$34,474
one year after admission	(13,381)	(19,060)	(13,857)	(22,460)
CABG within one year of ad- mission (1 = yes)	9.2%	16.2%	100%	100%
Readmission with AMI within one year of admission	5.8%	5.5%	1.1%	1.2%
Readmission with heart failure within one year of admission	9.0%	9.4%	6.1%	6.6%
Mortality within one year 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

NOTE.—Hospital expenditures are in 1995 dollars. For full sample 1987–94, the sample size is 1,770,452 for AMI patients and 967,882 for CABG patients. Standard deviations are in parentheses.

References

- Bentley, J. Marvin, and Nash, David B. "How Pennsylvania Hospitals Have Responded to Publicly Released Reports on Coronary Artery Bypass Graft Surgery." *Joint Comm. J. Quality Improvement* 24 (January 1998): 40–49.
- Capps, Cory S.; Dranove, David; Greenstein, Shane; and Satterthwaite, Mark. "The Silent Majority Fallacy of the Elzinga-Hogarty Criteria: A Critique and New Approach to Analyzing Hospital Mergers." Working Paper no. 8216. Cambridge, Mass.: NBER, April 2001.
- Chassin, Mark R.; Hannan, Edward L.; and DeBuono, Barbara A. "Benefits and Hazards of Reporting Medical Outcomes Publicly." *New England J. Medicine* 334 (February 8, 1996): 394–98.
- Dranove, David, and Satterthwaite, Mark A. "Monopolistic Competition When Price and Quality Are Imperfectly Observable." *Rand J. Econ.* 23 (Winter 1992): 518–34.
- Dziuban, Stanley W., Jr.; McIlhuff, Joseph B.; Miller, Stuart J.; and Dal Col, Richard H. "How a New York Cardiac Surgery Program Uses Outcomes Data." *Ann. Thoracic Surgery* 58 (December 1994): 1871–76.
- Green, Jesse, and Wintfeld, Neil. "Report Cards on Cardiac Surgeons: Assessing New York State's Approach." *New England J. Medicine* 332 (May 4, 1995): 1229–33.
- Hannan, Edward L., et al. "Improving the Outcomes of Coronary Artery Bypass

- Surgery in New York State." *J. American Medical Assoc.* 271 (March 9, 1994): 761–66.
- Ho, Mary T., et al. "Delay between Onset of Chest Pain and Seeking Medical Care: The Effect of Public Education." *Ann. Emergency Medicine* 187 (July 1989): 727–31.
- Hofer, Timothy P., et al. "The Unreliability of Individual Physician 'Report Cards' for Assessing the Costs and Quality of Care of a Chronic Disease." *J. American Medical Assoc.* 281 (June 9, 1999): 2098–2105.
- Iezzoni, Lisa I., ed. *Risk Adjustment for Measuring Health Care Outcomes*. Ann Arbor, Mich.: Health Admin. Press, 1994.
- . ed. *Risk Adjustment for Measuring Health Care Outcomes*. 2d ed. Chicago: Health Admin. Press, 1997.
- Kessler, Daniel P., and McClellan, Mark B. "Is Hospital Competition Socially Wasteful?" *Q.J.E.* 115 (May 2000): 577–615.
- . "The Effects of Hospital Ownership on Medical Productivity." *Rand J. Econ.* 33 (Autumn 2002): 488–506.
- Leventis, A. "Cardiac Surgeons under Scrutiny: A Testable Patient-Selection Model." Working Paper no. 41. Princeton, N.J.: Princeton Univ., Center Econ. Policy Studies, 1997.
- Marshall, Martin N.; Shekelle, Paul G.; Leatherman, Sheila; and Brook, Robert H. "The Public Release of Performance Data: What Do We Expect to Gain? A Review of the Evidence." *J. American Medical Assoc.* 283 (April 12, 2000): 1866–74.
- Mennemeyer, Stephen T.; Morrissey, Michael A.; and Howard, Leslie Z. "Death and Reputation: How Consumers Acted upon HCFA Mortality Information." *Inquiry* 34 (Summer 1997): 117–28.
- Mukamel, Dana B., and Mushlin, Alvin I. "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 (July 1998): 945–54.
- Pennsylvania Health Care Cost Containment Council. *Coronary Artery Bypass Graft Surgery: A Technical Report*. Vol. 1, 1990 data. Harrisburg: Pennsylvania Health Care Cost Containment Council, 1992.
- Peterson, Eric D., et al. "The Effects of New York's Bypass Surgery Provider Profiling on Access to Care and Patient Outcomes in the Elderly." *J. American Coll. Cardiology* 32 (October 1998): 993–99.
- Richards, Toni; Blacketer, Bethany; and Rittenhouse, Diane. *Statewide, Metropolitan, Corporate, and National Efforts in Monitoring and Reporting Quality Care*. Sacramento, Calif.: Off. Statewide Health Planning and Development, Health Policy and Planning Div., 1994.
- Romano, Patrick S., and Chan, Benjamin K. "Risk-Adjusting Acute Myocardial Infarction Mortality: Are APR-DRGs the Right Tool?" *Health Services Res.* 34 (March 2000): 1469–89.
- Romano, Patrick S.; Rainwater, Julie A.; and Antonius, Deirdre. "Grading the Graders: How Hospitals in California and New York Perceive and Interpret Their Report Cards." *Medical Care* 37 (March 1999): 295–305.
- Schneider, Eric C., and Epstein, Arnold M. "Influence of Cardiac-Surgery Performance Reports on Referral Practices and Access to Care." *New England J. Medicine* 335 (July 25, 1996): 251–56.
- . "Use of Public Performance Reports: A Survey of Patients Undergoing Cardiac Surgery." *J. American Medical Assoc.* 279 (May 27, 1998): 1638–42.
- Topol, Eric J., and Califf, Robert M. "Scorecard Cardiovascular Medicine: Its

Impact and Future Directions." *Ann. Internal Medicine* 120 (January 1, 1994): 65-70.

Weintraub, William S., et al. "Inhospital and Long-Term Outcome after Reoperative Coronary Artery Bypass Graft Surgery." *Circulation* 92, no. 1 (suppl. 2; 1995): 50-57.