Why Are Health Care Report Cards So Bad (Good)?: Theory and Partial Evidence *

Yijuan Chen[†]

Department of Economics, University of Pennsylvania

2008-10-12

Abstract

Issuing quality report cards for health care providers has drawn from profession and academia criticisms largely based on intuitions, conventional wisdom, and stand-alone empirical approaches. I provide a signaling-game theoretical foundation, upon which a new empirical framework is proposed, to study the effects of the report cards. I find that, when patients and health care providers are matched randomly, multidimensional measures in the existing report cards render them the optimal mechanism that reveals the provider types without causing providers to select patients. However, asymmetric distributions of patient types between providers, caused by the referring physician, may force the high-quality provider to shun patients in order to signal himself. Despite this imperfection, the existing report cards cause the *minimum* selection compared with alternative report mechanisms.

The theoretical results imply that treatment effects of the report cards vary from the signaling period to the signaled period, entailing estimating each effect with a difference-in-differences estimate in *each* period. I show that, in contrast to previous literature, a negative incidence effect is not sufficient to indicate existence of selection behavior. Moreover, the traditional single difference-in-differences estimate cannot capture the report cards' long run welfare effect. More broadly, my study can be extended and applied to other industries with experts, such as automobile service, law, and education.

Keywords: Report Cards, Signaling Game, Difference-in-differences, Experts

JEL Classification: I18, D82, C31

*I am very grateful to Jan Eeckhout and Mark Pauly for their generous guidance and support. This paper also benefits from discussions with and comments from Elena Krasnokutskaya, Andrew Postlewaite, Petra Todd, Kenneth Wolpin, Roberto Pinheiro, Philipp Kircher, Rachel Werner, Daniel Polsky, Aureo de Paula, Xun Tang and participants at Penn's Applied Micro Theory Luncheon.

[†]E-mail: yijuan@sas.upenn.edu ; Webpage: www.sas.upenn.edu/~yijuan

1 Introduction

In this paper I study the effects of quality report cards in the health care industry under both theoretical and empirical approaches, with particular focus on the well-known report cards for the coronary artery bypass graft (CABG) surgery. The report cards, planned to give outsiders better information about the quality of health care providers, have drawn doubts and criticisms since their inception. Arguably the most authoritative verdict comes from Dranove et al. (2003). They point out that, despite the risk adjustment procedures used in producing the report cards, "providers are likely to have better information on patients' conditions than even the most clinically detailed database" and may use such private information to improve their outcome by selecting patients. Using a single difference-in-differences estimate for each treatment effect, they show that report cards caused health care providers to avoid sick patients and overall decreased social welfare. In other words, a solution toward adverse selection resulted in moral hazard.

However, before being propelled to search for a new report mechanism, one should further investigate the existing report cards. The reasons are simple: First, interpretations of empirical regression results are often based on intuitions and/or conventional wisdom, lacking support of a comprehensive theoretical model. Second, even if the empirical study shows the existing report cards are broken, it does not say whether they are fixable. If the existing mechanism only requires a small fix, then it will cost much less than establishing a new mechanism, not to mention the new mechanism may call for large and time-consuming institutional changes. Fixing the existing report cards entails better understanding about them. In particular, a largely neglected fact is that the existing cards show multidimensional measures about the performance of the health care providers[1]. How do these measures collectively affect the providers' decisions? How do the patients parse the report cards, and how do they interact with the providers? Moreover, how does a provider's decision vary with his true quality? A stand-alone empirical approach fails to answer these questions in an integral way, and a theoretical signaling-game model is needed.

My studies toward the above positive questions lead to striking normative answers. First, I show that, even if the providers possess private patient information, when patients randomly choose providers, the existing report cards are actually the optimal mechanism in the sense that they fully reveal the providers' types *without* causing providers to select patients. The reason lies in the trade-off faced by each provider between two measures, volume and outcome, in the existing report cards. To improve outcome, measured by the mortality rate, a provider has to avoid sick patients, resulting in a smaller volume of patients. Due to this trade-off, when patients and providers are matched randomly, a low-quality provider has no way to mimic a

[1]For a sample of CABG report cards from New York and Pennsylvania, see Appendix 2.

high-quality one. Consequently, no providers engage in selecting patients, and the report cards reveal true quality. Though one may argue that, by shifting the distribution of illness severity toward healthy patients, the providers may improve outcome without cutting volume, the notion is self-contradictory, because it implies the providers will not be short of patient sources, and as a result they should not be concerned by the report cards, and there will be no selection behavior at the outset. My result on one hand provides a benchmark to study the effects of the existing report cards, and on the other hand calls for stopping searching for a better mechanism under the assumption of identical patient distributions between providers.

In reality, however, providers may select patients when facing report cards. The existence of the referring physician, functioning as a matching device between patients and providers, causes asymmetric patient distributions between the providers, with sicker patients matched to the high-quality provider. Sufficiently sick patients dampen the high-quality provider's outcome, forcing him to shun sick patients in order to signal his true quality. The consequent selection behavior in a separating equilibrium is characterized by three ranges of the high-quality provider's quality type: In the bottom range, the degree of selection behavior increases with the provider's type, in the middle range, the degree of selection behavior decreases with the provider's type, while in the top range, when the provider's quality is sufficiently high, there is no selection behavior.

I further show that, despite their imperfection, the existing report cards cause the *minimum* selection behavior compared with other alternative report mechanisms. The reason is, given that the existing cards have revealed the maximum information about a provider's performance, any other report mechanisms that temper the information will only make it easier for the low-quality provider to imitate the high-quality one's performance, forcing the high-quality provider to shun more sick patients to separate himself.

The theoretical results shed new light on the empirical study. In particular, they imply that when evaluating the effects of the report cards, the common difference-in-differences approach should be used with caution. Based on the two-period signaling-game model, treatment effects of the report cards, including incidence effect, quantity effect, matching effect, and welfare effect, all vary across periods after the report-card program is enacted, thus requiring a difference-in-differences estimate for each treatment effect in *each* period. Due to this, I propose a new empirical framework to re-evaluate the effects of the report cards based on the data set of Dranove et al. (2003).

My theoretical results also imply that, in contrast to that suggested by previous literature, a negative incidence effect, measured by a decrease of mean illness severity of patients undergoing surgery, is not sufficient to indicate existence of selection behavior. When providers' types are revealed in the second period and so they accept all coming patients, in one scenario all sick patients and a fraction of healthy patients will choose the high-quality provider, lowering his mean of patient illness severity, and overall also leading to a negative incidence effect. The reason of this result, which also challenges the conventional wisdom that revelation of surgeon types results in only sick patients matched with the high-quality provider, is that a patient, no matter how healthy, is still a patient and therefore is willing to seek the provider that gives her the best treatment. Hence healthy patients will continue choosing the high-quality provider up to the point where the large volume of patients brings down the high-quality provider's expected outcome.

Despite inaccessibility of the data set used by Dranove et al. (2003) due to the strict confidentiality of patient information, some specifications in Dranove et al. (2003) allow us to partially test my theory with their estimation results. In the context of my theoretical model, the single difference-in-differences estimate in the previous literature blends together both periods' data after the report cards were enacted, thus mixing a treatment effect of the first period with that of the second. Citing an augment of the New York report cards in 1993, Dranove et al. (2003) estimate each treatment effect under two assumptions: one with the New York report cards effective in 1991 and another with the effective year being 1993. Therefore a period 2 effect, if existing, should be more significant under the 1993 assumption. The regression results confirm this prediction. In particular, the quantity effect significantly increases under the 1993 assumption, indicating that revelation of providers' types leads to increased treatments.

Ultimately the report cards should be evaluated by their welfare effects. Before calculating the net effect, however, one should notice that, even though two periods suffice in the signaling-game model, in reality the welfare effect of revealing the providers' types certainly extends to more than one period. Therefore the actual effect of the report cards is the sum of the (discounted) effect of each period. Since the single difference-in-differences estimate fails to capture the period 2 effect, it cannot be used to count the long run welfare effect.

More broadly, my results imply that the existing report cards, despite the criticism about their short run cost in the health care industry, have the potential to be successfully introduced to other industries where goods and services are provided by experts while customers and providers are matched randomly. Industries of interest include automobile service, law, and education.

The rest of the paper is organized as follows: I briefly review the background of the health care report cards and the previous literature in Section 2. The theoretical model is presented in Section 3, and the empirical implications and evidence are presented in Section 4. Section 5 concludes.

2 Background and Previous Research

Brief history: Starting in 1986 the U.S. Health Care Financing Administration (HCFA) produced the first wide-scale provider report cards. Since those report cards were based on hospital discharge abstracts, making a comprehensive risk adjustment scheme impossible, HCFA stopped releasing them in 1993. In December 1990, New York Department of Health publicly released the first statewide hospital-level CABG-specific report card, and in 1992 extended it to include the surgeon-level data. Other states followed. In December 1992, the Pennsylvania Health Care Cost Containment Council released its CABG report card at both hospital and surgeon level. New Jersey started collecting data in 1994 and released its first CABG report card in November 1997. Later states establishing CABG-related assessment programs include California, Washington, Massachusetts, and Michigan. For a detailed introduction to the history of the health care report cards in the U.S., see Epstein (2004).

Previous research: Since the inception of the CABG report cards, a large amount of literature, most of which based on survey or stand-alone empirical approaches, have been focused on studying their impact on the health care industry. For a comprehensive review, see Epstein (2006). Nonetheless, as Epstein (2006) points out, in the empirical literature "prior research...has failed to distinguish the effect of public reporting from other possible confounders associated with an underlying predisposition to performance improvement". One exception is Dranove et al.(2003), who in addition point out that "the failure of previous studies to consider the entire population at severity for CABG, rather than those who received it, is a potentially severe limitation". Using longitudinal cardiac-patient data from Medicare claims from 1987 to 1994, and hospital data from American Hospital Association, they estimate treatment effects of the report cards using a difference-in-differences approach. Based on the estimated negative incidence effects, they conclude that the report cards resulted in providers avoiding sick patients. In addition, based on the estimated positive quantity effects, they suggest the report cards led the providers to shift distribution of patient illness severity toward healthy patients. Furthermore, they conclude that the report cards lead to decreased social welfare based on estimated increased post-surgical expenses and readmission rates. Using a similar approach with Florida as the control state, Epstein (2004) shows that mortality dropped sharply in New York and New Jersey around the time of the first report card publication. In addition Werner (2004) shows that racial and ethnic differences in CABG use rose significantly in New York after the state's CABG report card was released, which she attributes to physicians' belief about different clinic uncertainties in different racial and ethnic groups.

Despite the impact on surgery providers, literature that summarize survey information show that the report cards had small effect on the referring pattern of referring cardiologists. Based on a 1995 Pennsylvania survey, Schneider and Epstein (1996) show that 82% of surveyed cardiologists were aware of the report cards in 1995, but fewer than 10% discussed the guide with more than 10% of their patients needing CABG surgery. Hannan et al.(1997) show that, in New York, 85% of surveyed cardiologists received the 1995 report card, but only 22% routinely discuss the report card with patients. Nonetheless the report cards did affect the difficulty at which the referring cardiologists place patients. Schneider and Epstein (1996) show that, 59% of cardiologists reported increased difficulty since 1992 in placing their high-severity CABG patients, and 63% of cardiac surgeons reported being less willing to operate on those patients, offering circumstantial evidence for existence of selection behavior.

At the patient level, Omoigui et al. (1996) show that after the release of report cards in New York, the number of patients transferred to Ohio's Cleveland Clinic has increased by 31%, and that in general the illness severity of these transferred patients are higher than those transferred to the Cleveland Clinic from other states, offering another circumstantial evidence for selection behavior caused by report cards. In addition Gibbs et al.(1996) show that most participants considered friends and relatives as highly credible and preferred these sources to published information, indicating that word-of-mouth is another channel of information diffusion.

Conclusions drawn from most of the empirical literature stemmed from intuitions and conventional wisdom, lacking support of a comprehensive and equilibrium-based theoretical framework. Among the exceptions is Epstein (2004), who studies a single surgeon's decision problem under the report-card program. Nonetheless the model lacks necessary components of a signaling-game model. In particular, it neglects the possibility that a low-quality surgeon may pool with the high-quality counterpart by imitating his performance. In addition, the model does not specify how patients interpret the reports card and base their belief about the surgeon type on it. In other words, the necessary belief system in a signaling game is absent. Fong (2007) proposes an alternative scoring rule in a setting with one provider, whose type, drawn from a binary variable, is unknown the outsiders. The scoring rule is based on a one dimensional signal, namely the success rate. In equilibrium, however, the "good" type provider typically engages in selection behavior. Since she essentially assumes an identical distribution of patient types between both providers, the scoring rule she proposes is actually inferior to the existing report cards due to the trade-off between the measures in the existing ones. Another related paper is Lu et al. (2003). Using a Hotelling-class model, they study the effect of performance-based contracting on the providers. Though patient types are heterogeneous in their model, the types (locations) of providers are common knowledge, and so the model cannot be applied to study the effects of report cards.

3 Theoretical Model

3.1 Set-up

Time is discrete with 2 periods, indexed by $t \in \{1, 2\}$.

Each period there are a continuum of patients with measure 1, who will be active for one period. A patient, indexed by j, has a illness severity type s_j drawn from an identical and independent distribution with support $\mathbf{R}$ and a cumulative distribution function $F(.)$.

There are 3 health care institutions, indexed by $i \in \{A, B, C\}$, of whom A and B are surgery providers and C is a referring physician. A surgery provider i is characterized by his quality type $k_i \in \mathbf{R}$. One provider's quality type is k_h, the other's being $k_l < k_h$. The providers' types are known to all the physicians but unknown to the patients, whereas the patients hold a prior belief that each provider is equally likely to be type k_l or k_h.[2]

Denote by m_{it} the measure of patients that provider i performs the surgery on in period t. If $s_j = s$, $m_{it} = m$, and $k_i = k$, the minimum probability of failure of the surgery on patient j by provider i is given by $q(s, m, k)$, which satisfies the following assumption:

Assumption 1:
(i) $\frac{\partial q}{\partial s} \geq 0$, *with* $\frac{\partial q}{\partial s} = 0$ *if and only if* $q(s, m, k) = 1$;
(ii) $\frac{\partial q}{\partial k} \leq 0$, *with* $\frac{\partial q}{\partial k} = 0$ *if and only if* $q(s, m, k) = 0$;
(iii) $\frac{\partial q}{\partial m} \geq 0$, *with* $\frac{\partial q}{\partial m} = 0$ *if and only if* $q(s, m, k) = 1$. $\lim\limits_{s \to -\infty} q(s, 1, k) = 1$, $\lim\limits_{s \to +\infty} q(s, 0, k_h) = 1$ *and* $\lim\limits_{s \to -\infty} q(s, 0, k_l) \in (0, 1)$.

In Assumption 1, (i) and (ii) imply that the probability of failure is increasing with a patient's severity type and decreasing with the provider's quality. In (iii), $\frac{\partial q}{\partial m} \geq 0$ implies the capacity constraint faced by each provider, which comes from factors such as clinic facility, nursing, and logistic that affects a surgery's probability of failure[3]. $\lim\limits_{s \to -\infty} q(s, 1, k) = 1$ reflects that, as a matter of reality, any provider, no matter how skillful, will be overwhelmed when $m = 1$. The assumption $\lim\limits_{s \to +\infty} q(s, 1, k) = 1$ implies that when the severity type is sufficiently high, even the most skillful provider will fail for sure. $\lim\limits_{s \to -\infty} q(s, 0, k_l) \in (0, 1)$ reflects the minimum requirement for a provider to obtain a license and practice surgery. Moreover, the

[2] In contrast to simply assuming categorically a high type and a low type, my assumption allows continuity in the providers' types, in the sense that the difference between k_h and k_l now reflects how good a provider is relative to another. Intuitively, such a continuity captures the reality that although most of providers do differ in their quality, they may do so between "fair" and "very good", or "very good" and "excellent", instead of only between symbolically "good" and "bad".

[3] "An open-heart surgery team requires, among other things, skilled technicians, diagnostic imaging, and sophisticated laboratory support." Raffel & Barsukiewicz, "The U.S. Health System, Origins and Functions", 5th Edition, page 126

statement q being the minimum probability of failure captures the fact that a provider is able to do worse, with the worst outcome being a certain failure.[4]

Each period a patient needs to visit C before being referred to a provider i. A patient can decide which provider he would like to be referred to, and C will make the referral according to the patient's choice. Alternatively the patient can leave the referral decision to C. If a patient j is referred to provider i, his severity level is known by provider i while remains unverifiable to himself. If it is up to C to make the referral decision, C's decision will be characterized by the median severity level $s_{1/2}$ such that $F(s_{1/2}) = \frac{1}{2}$. C will refer patients with $s_j < s_{1/2}$ to the type k_l provider and those with $s_j \geq s_{1/2}$ to the type k_h provider.

My specification of C's referring pattern is worth some discussion. First, I assume that C's decision rule is not affected by report cards, which is consistent with the findings in the previous survey[5] as discussed in Section 2. Second, the specification may seem stylized at the first glance, but it captures two important facts. On the one hand, a referring physician, or more generally a primary care physician, is a generalist, whereas a surgery provider is a specialist. Hence the referring physician may only observe a category which a patient's severity level belongs to, while a surgery provider gains accurate information about the patient's type from his expertise and/or more advanced and more sophisticated examinations. On the other hand, referring physician is a characteristic feature of the now prevailing Managed Care Organizations[6]. In this sense, a referring physician is a colleague of or has close relation with the surgery providers. C's referring pattern according to my specification thus ensures that each provider receives the same measure of patients while at the same time allows patients with higher severities to be matched with the provider with higher quality. I leave discussion about a more general specification of C's referring pattern to Subsection 3.3

If patient j is referred to provider i, the provider can choose between performing the surgery or providing an alternative treatment. A provider i's payoff is $\pi_i = m_{i1} + m_{i2}$, while the referring physician C's payoff is a constant[7]. I assume all the patients are Medicare beneficiaries and so monetary charges are abstracted from patients' payoffs[8]. The payoff of patient j is 1 if the surgery succeeds and 0 if it fails. For simplicity I also assume a patient's payoff to be 0 if she undergoes the alternative treatment. Such an assumption allows a tractable analysis and also captures the worst possible scenario should the provider refuse performing surgery. While I discuss a more general assumption in Subsection 3.3, I would like

[4] Chen (2008b) shows that, as using "character evidence" is prohibited in the U.S. legal system, in the extreme case implicit collusion between providers will result in no litigation from the patient side.

[5] Schneider and Epstein (1996); Hannan et al. (1997)

[6] Phelps, C., "Health Economics", 3rd Edition, page 370, 375

[7] Note that, whether C is paid by a fixed salary or under a "fee for service" scheme, as long as the measure of patients seeking C for referral is constant, C's payoff is a constant.

[8] Raffel M. W. & Barsukiewicz C. K., "The U.S. Health System, Origins and Functions", 5th Edition, page 32-33

to stress here that I believe my results are robust to the 0 payoff assumption.

At the end of period 1, a report card of provider i will be published to the public, showing (m_{i1}, d_{i1}), where d_{i1} is the mortality rate of the surgeries performed by i in period 1.

Denote by $\mathbf{P}$ a patient's mixed action set $\{(p_A, p_B, p_C) \mid \sum_i p_i = 1 \text{ and } p_i \geq 0 \text{ for } i \in \{A, B, C\}\}$. An element of the action set, (p_A, p_B, p_C), means that for $i \in \{A, B\}$ with probability p_i the patient will ask C to refer him to provider i and with probability p_C the patient will have C make the referral decision. Denote a strategy of patient j in period t by ϕ_{jt}, where $\phi_{j1} \in \mathbf{P}$ and ϕ_{j2} is a mapping from j's information set in period 2 to $\mathbf{P}$.

An action taken by provider i in period t is summarized by a set of patient types, denoted by S_{it}. Only patients with $s_j \in S_{it}$ will receive the surgery. Denote the set of S_{it} by $\mathbf{S}$. A strategy of provider i is $\sigma_i = (\sigma_{i1}, \sigma_{i2})$, where $\sigma_{i1} : \{k_l, k_h\} \to \mathbf{S}$ and $\sigma_{i2} : \{k_l, k_h\} \times [0, 1]^4 \to \mathbf{S}$.

Denote by M_{it} the measure of patients that are referred to i in period t, and F_{it} the distribution of severity types in the patients referred to i in t.When clear in the context, below I use h in the subscript to denote the type k_h provider and l to denote the type k_l provider.

3.2 Analysis

The solution concept is a symmetric perfect Bayesian equilibrium (PBE), consisting of a strategy profile $(\sigma_A^*, \sigma_B^*, \phi_{jt}^*)$ and a belief system, such that (i) given the other players' strategies specified in the profile each player's specified strategy is sequentially rational, (ii) the belief system is consistent with the strategy profile, and (iii) $\sigma_A^* = \sigma_B^* = \sigma^*, \phi_{jt}^* = \phi_t^*$ for every patient j.

I outline some preliminary results below before studying the equilibrium outcomes:

First, in period 2 it must be that each provider performs the surgery on all coming patients.

Second, two report cards are regarded as different is they differ in either one measure or both measures. An equilibrium is a separating one if in period 1 two providers generate different report cards. The concept of perfect Bayesian equilibrium then requires the patients in period 2 correctly infer the providers' types in a separating equilibrium. Consequently, in a separating equilibrium, if in period 2 the payoff of one type of provider is less than that of another, then in period 1 the former one must perform the surgery on all the coming patients. Throughout the analysis I will focus on separating equilibrium, and I discuss pooling equilibrium in the next subsection.

Third, under the prior belief, if patient j chooses to make the referral decision on his own, he faces the probability distribution of matching between his severity type and a provider i's quality type as the table below

Prob	$k_i = k_h$	$k_i = k_l$
$s_i \leq 1/2$	1/4	1/4
$s_i > 1/2$	1/4	1/4

while if he let C make the referral decision, the matching probability table he faces is

Prob	$k_i = k_h$	$k_i = k_l$
$s_i \leq 1/2$	0	1/2
$s_i > 1/2$	1/2	0

Denote by E_{it} both a patient j's expected payoff from being referred to provider i at t when he makes the decision on his own and his expected payoff from letting C make the referral decision when $i = C$. There are, for $i \in \{A, B\}$,

$$E_{i1} = \frac{1}{2} \int_{\sigma_{i1}(k_l)} 1 - q(s, m_{l1}, k_l) dF(s) + \frac{1}{2} \int_{\sigma_{i1}(k_h)} 1 - q(s, m_{h1}, k_h) dF(s)]$$

and

$$E_{C1} = \int_{\sigma_{i1}(k_l) \cap (-\infty, s_{1/2})} 1 - q(s, m_{l1}, k_l) dF(s) + \int_{\sigma_{i1}(k_h) \cap (s_{1/2}, +\infty)} 1 - q(s, m_{h1}, k_h) dF(s)$$

with $m_{l1} = M_{i1} \cdot \int_{\sigma_{i1}(k_l)} dF_{i1}(s)$, and $m_{h1} = M_{i1} \cdot \int_{\sigma_{i1}(k_h)} dF_{i1}(s)$.

Fourth, in any equilibrium it must be $E_{A2} = E_{B2}$ in period 2, since otherwise, say $E_{A2} > E_{B2}$, all patients in period 2 will choose provider A, resulting in $M_{A2} = 1$ and so $E_{A2} = 0$ by Assumption 1, contradicting $E_{A2} > E_{B2}$. This result challenges the conventional wisdom that revelation of provider types will result in sick patients choosing the high-quality provider and healthy patients choosing the low-quality one. A healthy patient after all is still a patient, whose objective is to seek the best possible treatment. Therefore the healthy patient also wants to choose the high-quality provider as long as he provides better outcome, and this will not cease until a large volume of patients drags the high-quality provider's outcome to the same level as the low-quality one's.

Last, the period 2 patients' belief upon seeing identical report cards, i.e. $(m_{h1}, d_{h1}) = (m_{l1}, d_{l1})$, remains the same as the prior belief, since two identical cards cannot help the patients distinguish one provider from another.

Given these preliminary results, I first study a benchmark case where there is no referring physician C and so the patients have to make decisions on their own, then I study the full model with C.

3.2.1 Without C

Without the referring physician C, in period 1 the patients will choose providers on their own and so $F_{it}(.) = F(.)$. With slight abuse of notation, I restrict the patients' actions to $\mathbf{P}$ such that $p_C = 0$. Assumption 1 then implies that in equilibrium it must be $\phi_1^* = (\frac{1}{2}, \frac{1}{2}, 0)$. Intuitively, when patients have no other information about the providers' types, naturally they will randomize between the providers with equal probabilities. But when all the other patients randomize with equal probabilities, resulting in $M_{A1} = M_{B1} = \frac{1}{2}$, then a patient will be indifferent between the providers and so it is indeed his best response to also randomize with equal probabilities.

Lemma 1 below describes what will happen should the providers' types be revealed in period 2.

Lemma 1 *If the providers' types (k_i, k_{-i}) are revealed in period 2, then there exists a unique measure of patients $\hat{M}_{i2}(k_i, k_{-i})$ that will choose provider i in period 2. Moreover, $\frac{\partial \hat{M}_{i2}}{\partial k_i} > 0$ and $\frac{\partial \hat{M}_{i2}}{\partial k_{-i}} < 0$.*

Proof. In Appendix. ■

An immediate corollary of Lemma 1 is that, should the patients know the providers' types, then the provider with the higher quality type will receive more than $\frac{1}{2}$ measure of patients. For ease of notation, let $\hat{M}_h = \hat{M}_{i2}(k_h, k_l)$ and $\hat{M}_l = 1 - \hat{M}_h$.

Given M_{i1}, for every $m_{i1} \leq M_{i1}$ there exists a unique severity level s_i such that

$$m_{i1} = M_{i1} \cdot F(s_i)$$

For each $m_{i1} \leq M_{i1}$, the minimum mortality rate $\bar{d}_{i1}$ a provider i can achieve comes from avoiding treating patients with $s_j > s_i$, implying

$$\bar{d}_{i1} = \frac{\int_{-\infty}^{s_i} q(s, m_{i1}, k_i) dF(s)}{F(s_i)} = \frac{\int_{-\infty}^{s_i} q(s, M_{i1}F(s_i), k_i) dF(s)}{F(s_i)}$$

There are

$$\frac{\partial m_{i1}}{\partial s_i} = M_{i1} \cdot f(s_i)) > 0$$

and

$$\frac{\partial \bar{d}_{i1}}{\partial s_i} = \frac{\left[\begin{array}{c} q(s_i, M_{i1}F(s_i), k_i) \cdot f(s_i) \\ + \int_{-\infty}^{s_i} q_2(s, M_{i1}F(s_i), k_i) \cdot M_{i1} f(s_i) dF(s) \end{array}\right] F(s_i)}{[F(s_i)]^2} - \frac{\int_{-\infty}^{s_i} q(s, M_{i1}F(s_i), k_i) dF(s) \cdot f(s_i)}{[F(s_i)]^2}$$

$$> \frac{\int_{-\infty}^{s_i} q_3(s, M_{i1}F(s_i), k_i) \cdot M_{i1} f(s_i) dF(s)}{F(s_i)} > 0$$

where the first inequality in the $\frac{\partial \bar{d}_{i1}}{\partial s_i}$ part comes from $q(s, M_{i1} \cdot F_{i1}(s_i), k_i) < q(s_i, M_{i1} \cdot F_{i1}(s_i), k_i)$ for all $s < s_i$, as implied by Assumption 1-(i).

Since we can write s_i as $s_i = F^{-1}(\frac{m_{i1}}{M_{i1}})$, for $m_{i1} \in (0, M_{i1}]$ we can write $\bar{d}_{i1}$ as

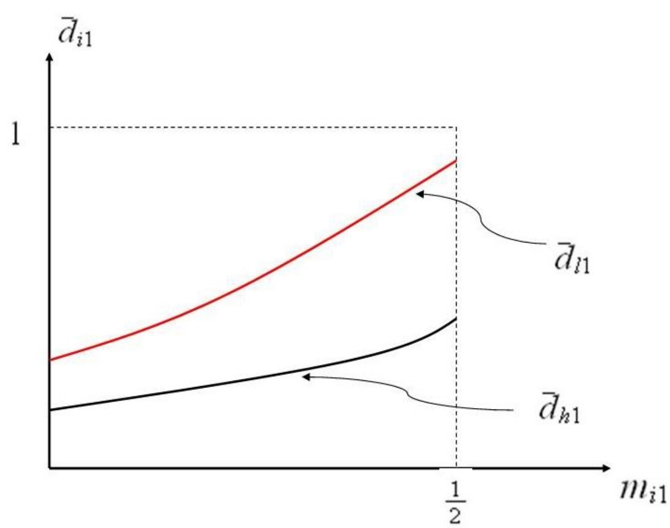
$$\bar{d}_{i1}(m_{i1}) = [\int_{-\infty}^{F^{-1}(\frac{m_{i1}}{M_{i1}})} q(s, m_{i1}, k_i) dF(s)] \cdot \frac{m_{i1}}{M_{i1}}$$

Consequently

$$\frac{\partial \bar{d}_{i1}}{\partial m_{i1}} = \frac{\partial \bar{d}_{i1}}{\partial s_i} \cdot \frac{\partial s_i}{\partial m_{i1}} > 0$$

Hence, as shown in Figure 1, given M_{i1} a provider i is facing a trade-off between $\bar{d}_{i1}$ and m_{i1}: To improve (lower) the minimum mortality rate, the provider has to shun more sicker patients, which reduces the measure of patients he treats.

Given (M_{i1}, k_i) we can define the set of provider i's possible report card results to be $\mathcal{Z}(M_{i1}, k_i) = \{(m_{i1}, d_{i1}) \,|\, m_{i1} \in (0, M_{i1}], d_{i1} \in [\bar{d}_{i1}(m_{i1}), 1]\}$and we call the set $\mathbf{Z}(M_{i1}, k_i) = \{(m_{i1}, \bar{d}_{i1}(m_{i1})) \,|\, m_{i1} \in (0, M_{i1}]\}$ the frontier of $\mathcal{Z}(M_{i1}, k_i)$.



Firgure 2-1: Trade-off between m_{i1} and $\bar{d}_{i1}$

At period 1, patients' equal randomization leads to $M_{i1} = \frac{1}{2}$ for each provider, which implies

$$\bar{d}_{i1} = \int_{-\infty}^{+\infty} q(s, \frac{1}{2}, k_i) dF(s)$$

and consequently $\bar{d}_{h1}(\frac{1}{2}) < \bar{d}_{l1}(\frac{1}{2})$, as shown in Figure 2-1. Therefore, if the type k_h provider performs the surgery on all the coming patients, then his report card will show $(\frac{1}{2}, \bar{d}_{h1}(\frac{1}{2}))$, which cannot be imitated by the type k_l provider. Hence in a perfect Bayesian equilibrium, upon seeing $(\frac{1}{2}, \bar{d}_{h1}(\frac{1}{2}))$ the patients in period 2 must form a belief that the provider's type is certainly k_h. Therefore in period 1 performing surgeries on all the coming patients gives the type k_h provider the highest possible payoff $\frac{1}{2} + \hat{M}_h$ and so dominates all the other actions. On the other hand, since the type k_l provider by no means can imitate the type k_h provider, it is the type k_l provider's best response to perform surgeries on all the coming patients in period 1 too. The consistent belief system can be easily constructed accordingly. The proposition below summarizes these results:

Proposition 1 *If there is no referring physician C, then there exists a unique separating equilibrium such that $\phi_1^* = (\frac{1}{2}, \frac{1}{2}, 0)$ and each provider performs the surgery on all coming patients. At period 2 the patients correctly infer the providers' types from the report cards and randomize between the providers in the way that they choose the type k_h provider with probability $\hat{M}_h$ and the type k_l provider with probability $\hat{M}_l$.*

Proposition 1 implies that, when patients and providers are matched randomly, thanks to the trade-off between the volume measure, m_{i1}, and the outcome measure $\bar{d}_{i1}$, the existing report cards are actually the optimal mechanism that fully reveal the providers' types without causing selection behavior. For researchers this result shows the necessity of understanding the specific features of the existing report cards and their influence on the participants of the programs. Unfortunately, the feature of multidimensional measures in the existing report cards has been largely neglected in the previous studies, with some directly calling them "mortality report cards". Normative study such as Fong (2007) has also been conducted under the assumption that only mortality rates are released, while in the mean time implicitly assuming identical distribution of patient types between providers, and as a result the proposed mechanism is inferior to the existing report cards. Paying attention to the actual features of the existing report cards thus helps to avoid conducting research under unnecessarily unrealistic assumptions. Moreover, it is worth noting that, thanks to the absence of the referring physician C here, the result in Proposition 1 is robust to the specification of C's objective and behavior. Therefore Proposition 1 provides a benchmark for further study on the report cards.

Proposition 1 also helps to clarify concerns from people that contemplate introducing the report-card program to other industries where goods and services are also provided by skilled experts. If in an industry the distributions of consumers among the providers are identical, then Proposition 1 shows the same report-card program can be successfully implemented there. If not, then one needs to exam the specific features of the industry to determine the report-card program's potential impact.

We now turn to the full model with the referring physician C.

3.2.2 With C

We first look at period 2. Suppose in period 2 the patients correctly infer the providers' types from the report cards, then in equilibrium the patients will not ask C to make the referral decision, since otherwise each provider will receive patients with measure 1/2, while $q(s, \frac{1}{2}, k_h) < q(s, \frac{1}{2}, k_l)$ implies a patient will strictly prefer choosing the type k_h provider. However, on the other hand, a patient can induce more information about his severity type from turning to C: If he asks C to make the referral decision and is referred to, say, the type k_l provider, then he will know that his severity type is below $s_{1/2}$. At this point the patient may want to reconsider the referral decision. To see how this extra information affects the patients' actions, I first show that when the patients know the providers' types, then with the extra information about severity types the period 2 patients' equilibrium actions derived in the previous section will no longer be part of an equilibrium even if

$$1-\int_{-\infty}^{+\infty} q(s,\hat{M}_h,k_h)dF(s) > 1-[\int_{-\infty}^{s_{1/2}} q(s,\hat{M}_l,k_l)dF(s)+\int_{s_{1/2}}^{+\infty} q(s,\hat{M}_h,k_h)dF(s)]$$

where the left hand side is the payoff from seeing a type k_h provider[9], and the right hand side is the payoff from letting C make the decision. The reason is that, the inequality implies

$$1-\int_{-\infty}^{s_{1/2}} q(s,\hat{M}_h,k_h)dF(s) > 1-\int_{-\infty}^{s_{1/2}} q(s,\hat{M}_l,k_l)dF(s)$$

but now if a patient induces information from C and knows that his severity type is below $s_{1/2}$, he will strictly prefer the type k_h provider to the type k_l provider, and so he would like to "renegotiate" with C and choose the type k_h provider for sure rather than randomizing between them or following C's decision. Therefore in this subsection I take into account this extra information when patients correctly infer the provider types from report cards, and allow a patient to base his action on both the providers' types and the category of his severity types characterized by $s_{1/2}$.

[9] As discussed before it is equal to the payoff from seeing a type k_l provider

If a patient knows both the provider types and the category of his severity types, he can do at least as well as letting C make the decision. Now there are three possibilities in a separating equilibrium regarding the distribution of patients' severity types facing the providers in period 2:

(i) All patients with $s_j \geq s_{1/2}$ and some with $s_j < s_{1/2}$ choose the type k_h provider

(ii) All patients with $s_j < s_{1/2}$ and some with $s_j \geq s_{1/2}$ choose the type k_h provider

(iii) Each provider is facing patients with all possible s_j's.

To analyze these possibilities, we see that, on one hand, for a patient with $s_j \geq s_{1/2}$ to be indifferent between a type k_h provider and a type k_l provider, it must be for some $\tilde{M}_h \in (\frac{1}{2}, 1]$

$$Y(\tilde{M}_h) \equiv \int_{s_{1/2}}^{+\infty} q(s, \tilde{M}_h, k_h) dF(s) - \int_{s_{1/2}}^{+\infty} q(s, 1 - \tilde{M}_h, k_l) dF(s) = 0$$

Since $Y(\frac{1}{2}) < 0$, $Y(1) > 0$, and $\frac{\partial Y}{\partial \tilde{M}_h} > 0$, there exists a unique $\tilde{M}_h^*$ such that $Y(\tilde{M}_h^*) = 0$. Note that $\tilde{M}_h^*$ is a function of (k_h, k_l), and when necessary I will use the notation $\tilde{M}_h^*(k_h, k_l)$. Moreover, there is $\frac{\partial \tilde{M}_h^*}{\partial k_h} > 0$. On the other hand, for a patient with $s_j < s_{1/2}$ to be indifferent between a type k_h provider and a type k_l provider, it must be for some $\bar{M}_h \in (\frac{1}{2}, 1]$

$$G(\bar{M}_h) \equiv \int_{s_{1/2}}^{+\infty} q(s, \bar{M}_h, k_h) - q(s, 1 - \bar{M}_h, k_l) dF(s)$$

Similarly there exists a unique $\bar{M}_h^*$ such that $G(\bar{M}_h^*) = 0$. Also $\bar{M}_h^*$ is a function of (k_h, k_l) and when necessary I will use the notation $\bar{M}_h^*(k_h, k_l)$. Moreover, there is $\frac{\partial \bar{M}_h^*}{\partial k_h} > 0$.

Now the aforementioned three possibilities correspond to the comparison between $\tilde{M}_h^*$ and $\bar{M}_h^*$:

(i) $\tilde{M}_h^* > \bar{M}_h^*$, as shown in Figure 2

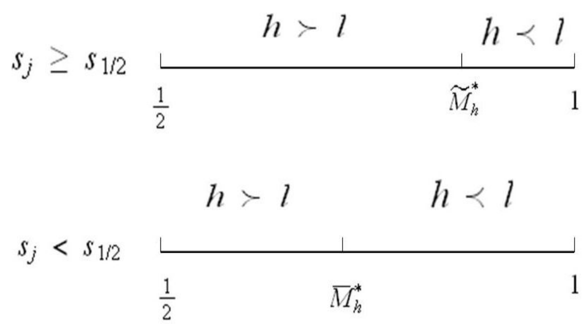


Figure 2-2: $\tilde{M}_h^* > \bar{M}_h^*$

Then in a separating equilibrium in period 2 it must be that all patients with severity types $s_j \geq s_{1/2}$ and a fraction $\frac{\bar{M}_h^* - \frac{1}{2}}{\frac{1}{2}} = 2\bar{M}_h^* - 1$ of the patients with severity types $s_j < s_{1/2}$ choose the type k_h provider, with the remaining patients choosing the type k_l provider. In total in period 2 the type k_h provider receives patients with measure $\bar{M}_h^*$ and the type k_l provider receives $1-\bar{M}_h^*$ patients. The reason of such an outcome is, as $\tilde{M}_h^* > \bar{M}_h^*$, in the specified outcome the patients with severity types $s_j \geq s_{1/2}$ strictly prefer the type k_h provider to the type k_l provider[10], whereas the patients with severity types $s_j < s_{1/2}$ are indifferent between the two providers, and so no patients have the incentive to deviate. It is then easy, though tedious, to verify that in any other outcomes at least one category of patients will have the incentive to deviate.

(ii) $\tilde{M}_h^* < \bar{M}_h^*$, as shown in Figure 2-3

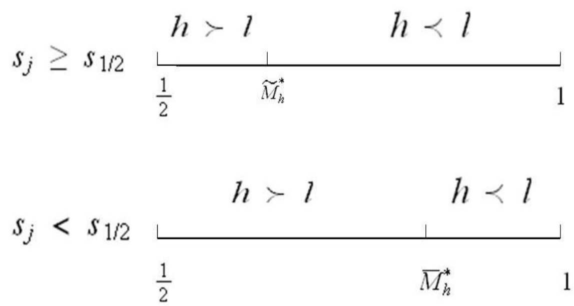


Figure 2-3: $\tilde{M}_h^* < \bar{M}_h^*$

Then in a separating equilibrium in period 2 all patients with severity types $s_j < s_{1/2}$ and a fraction $2\tilde{M}_h^* - 1$ of the patients with severity types $s_j \geq s_{1/2}$ choose the type k_h provider, with the remaining patients choosing the type k_l provider. In total in period 2 the type k_h provider receives patients with measure $\tilde{M}_h^*$ and the type k_l provider receives $1-\tilde{M}_h^*$ patients.

(iii) $\tilde{M}_h^* = \bar{M}_h^*$

Then in a separating equilibrium in period 2 the type k_h provider receives $\tilde{M}_h^*$ patients. A fraction θ of patients with $s_j \geq s_{1/2}$ and a fraction γ of patients with $s_j < s_{1/2}$ choose the type k_h provider, with $\frac{1}{2}\theta + \frac{1}{2}\gamma = \tilde{M}_h^*$.[11]

The fact that $\frac{\partial \tilde{M}_h^*}{\partial k_h} > 0$ and $\frac{\partial \bar{M}_h^*}{\partial k_h} > 0$ implies that the type k_h provider's payoff increases with k_h if the type is known by the patients. The lemma below summarizes the results obtained so far.

Lemma 2 *If in period 2 the patients know both the providers' types and the category of their severity types characterized by $s_{1/2}$, then in equilibrium the type k_h provider will receive patients with measure $M_h^*(k_h, k_l) \equiv \min\{\tilde{M}_h^*, \bar{M}_h^*\}$, and the type k_l provider will receive patients with measure $1 - M_h^*(k_h, k_l)$.*

[10] On the other hand the patients with $s_j \geq s_{1/2}$ are indifferent between choosing the type k_h surgeon on his own and letting C make the referral decision.

[11] Note that $\tilde{M}_h^* = \bar{M}_h^*$ implies that $Y(\tilde{M}_h^*) + G(\tilde{M}_h^*) = \int_{-\infty}^{+\infty} q(s, \tilde{M}_h^*, k_h)dF(s) - \int_{-\infty}^{+\infty}(s, 1-\tilde{M}_h^*, k_l)dF(s) = L(\tilde{M}_h^*) = 0$, that is, $\tilde{M}_h^* = \bar{M}_h^* = \hat{M}_h$, but on the other hand $L(\hat{M}_h) = 0$ does not necessarily mean $Y(\hat{M}_h) = 0$ and $G(\hat{M}_h) = 0$.

We now turn to period 1. There are two possible outcomes in the first period. The patients may make the referral decisions on their own, or leave the decision to the referring physician. The lemma below characterizes these possibilities.

Lemma 3 *There exists a separating equilibrium such that in period 1 the patients make referral decisions on their own if and only if*

$$\int_{s_{1/2}}^{+\infty} q(s, \frac{1}{2}, k_l) - q(s, \frac{1}{2}, k_h) dF(s) < \int_{-\infty}^{s_{1/2}} q(s, \frac{1}{2}, k_l) - q(s, \frac{1}{2}, k_h) dF(s)$$

Proof. In Appendix. ■

The intuition of Lemma 3 is: Compared with letting C make the decision, there are benefit and cost of choosing a provider on one's own. The benefit comes from choosing a type k_h provider when one's severity type is below $s_{1/2}$, while the cost comes from choosing a type k_l provider when one's severity type is above $s_{1/2}$. For a patient to prefer self-referral to letting C make the decision, it must be that the cost of self-referral, represented by the left hand side of the inequality in the lemma, is outweighed by the benefit, represented by the right hand side of the inequality.

The equilibrium outcome in the separating equilibrium with self referral in period 1 is similar as when there are no referring physician C. Equal randomization of the patients allows the type k_h provider to separate himself from the type k_l provider without shunning any patients, and consequently the type k_l provider will not shun patients either. At period 2, the report cards fully reveal the providers' types and the patients choose providers in the way implied by Lemma 2 and its preceding discussion.

When the inequality in Lemma 3 is reversed, the patients leave the referring decision to C. We now turn to this scenario.

If period 1 patients let C make the referral decision, then $M_{i1} = \frac{1}{2}$ for each provider i, but the providers face different distribution of severity types. Specifically, the type k_h provider receives only patients with $s_j \geq s_{1/2}$, so $F_{h1}(s) = \begin{cases} 2F(s) - 1 \text{ if } s_j \geq s_{1/2} \\ 0 \text{ if } s_j < s_{1/2} \end{cases}$ whereas the type k_l provider receives only patients with $s_j < s_{1/2}$, so $F_{l1}(s) = \begin{cases} 1 \text{ if } s_j \geq s_{1/2} \\ 2F(s) \text{ if } s_j < s_{1/2} \end{cases}$.

For the type k_h provider, given M_{h1}, for every $m_{h1} \leq M_{h1}$ there exists a unique $s_h \geq s_{1/2}$ such that

$$m_{h1} = M_{h1} \cdot F_{h1}(s_h)$$

For each $m_{h1} \leq M_{h1}$, the minimum mortality rate $\bar{d}_{h1}$ he can achieve comes from avoiding patients with $s_j > s_h$, implying

$$\bar{d}_{h1} = \frac{\int_{s_{1/2}}^{s_h} q(s, m_{h1}, k_h) dF_{h1}(s)}{F_{h1}(s_h)} = \frac{\int_{s_{1/2}}^{s_h} q(s, M_{h1} \cdot F_{h1}(s_h), k_h) dF_{h1}(s)}{F_{h1}(s_h)}$$

There are $\frac{\partial m_{h1}}{\partial s_h} > 0$ and $\frac{\partial \bar{d}_{h1}}{\partial s_h} > 0$. Moreover, for $s_h \geq s_{1/2}$ we can write s_h as $s_h = F_{h1}^{-1}(\frac{m_{h1}}{M_{h1}})$, and so for $m_{h1} \in (0, M_{h1}]$ we can write $\bar{d}_{h1}$ as

$$\bar{d}_{h1}(m_{h1}) = \int_{s_{1/2}}^{F_{h1}^{-1}(\frac{m_{h1}}{M_{h1}})} q(s, m_{h1}, k_h) dF_{h1}(s) \cdot \frac{m_{h1}}{M_{h1}}$$

It follows that $\frac{\partial \bar{d}_{h1}}{\partial m_{h1}} = \frac{\partial \bar{d}_{h1}}{\partial s_h} \cdot \frac{\partial s_h}{\partial m_{h1}} > 0$ and $\frac{\partial \bar{d}_{h1}}{\partial k_h} < 0$.

Similarly, for the type k_l provider, given M_{l1}, for every $m_{l1} \leq M_{l1}$ there exists a unique $s_l \leq s_{1/2}$ such that

$$m_{l1} = M_{l1} \cdot F_{l1}(s_l)$$

For each $m_{l1} \leq M_{l1}$, the minimum mortality rate $\bar{d}_{l1}$ he can achieve comes from avoiding treating patients with $s_j > s_l$, implying

$$\bar{d}_{l1} = \frac{\int_{-\infty}^{s_l} q(s, m_{l1}, k_l) dF_{l1}(s)}{F_{l1}(s_l)} = \frac{\int_{-\infty}^{s_l} q(s, M_{l1} \cdot F_{l1}(s_l), k_l) dF_{l1}(s)}{F_{l1}(s_l)}$$

There are $\frac{\partial m_{l1}}{\partial s_l} > 0$ and $\frac{\partial \bar{d}_{l1}}{\partial s_l} > 0$. Moreover, for $s_l < s_{1/2}$ we can write s_l as $s_l = F_{l1}^{-1}(\frac{m_{l1}}{M_{l1}})$, and so for $m_{l1} \in (0, M_{l1}]$ we can write $\bar{d}_{l1}$ as

$$\bar{d}_{l1}(m_{l1}) = \int_{-\infty}^{F_{l1}^{-1}(\frac{m_{l1}}{M_{l1}})} q(s, m_{l1}, k_l) dF_{l1}(s) \cdot \frac{m_{l1}}{M_{l1}}$$

It follows that $\frac{\partial \bar{d}_{l1}}{\partial m_{l1}} = \frac{\partial \bar{d}_{l1}}{\partial s_l} \cdot \frac{\partial s_l}{\partial m_{l1}} > 0$ and $\frac{\partial \bar{d}_{l1}}{\partial k_l} < 0$.

With $M_{i1} = \frac{1}{2}$, for $m \in (0, \frac{1}{2}]$ we have

$$\begin{aligned} \bar{d}_{h1}(m) &= \left[\int_{s_{1/2}}^{F_{h1}^{-1}(2m)} q(s, m, k_h) dF_{h1}(s) \right] \cdot \frac{1}{2m} = \left[\int_{s_{1/2}}^{F^{-1}(m+\frac{1}{2})} q(s, m, k_h) d(2F(s) - 1) \right] \cdot \frac{1}{2m} \\ &= \frac{1}{m} \cdot \int_{s_{1/2}}^{F^{-1}(m+\frac{1}{2})} q(s, m, k_h) dF(s) \end{aligned}$$

and

$$\bar{d}_{l1}(m) = \left[\int_{-\infty}^{F_{l1}^{-1}(2m)} q(s,m,k_l)dF_{l1}(s)\right] \cdot \frac{1}{2m} = \left[\int_{-\infty}^{F^{-1}(m)} q(s,m,k_l)d2F(s)\right] \cdot \frac{1}{2m}$$
$$= \frac{1}{m} \cdot \int_{-\infty}^{F^{-1}(m)} q(s,m,k_l)dF(s)$$

It follows that $\bar{d}_{h1}(\frac{1}{2}) = 2\int_{s_{1/2}}^{+\infty} q(s,\frac{1}{2},k_h)dF(s)$ and $\bar{d}_{l1}(\frac{1}{2}) = 2\int_{-\infty}^{s_{1/2}} q(s,\frac{1}{2},k_l)dF(s)$. Denote by $\bar{k}_h$ the type of k_h that solves $\bar{d}_{h1}(\frac{1}{2}) = \bar{d}_{l1}(\frac{1}{2})$ [12]. Also by L'Hospital's rule, there are

$$\lim_{m\to 0} \bar{d}_{h1}(m) = \lim_{m\to 0} q(F^{-1}(m+\frac{1}{2}),m,k_h) + \int_{s_{1/2}}^{F^{-1}(m+\frac{1}{2})} q_2(s,m,k_h)dF(s) = q(s_{1/2},0,k_h)$$

and

$$\lim_{m\to 0} \bar{d}_{l1}(m) = \lim_{m\to 0} q(F^{-1}(m),m,k_l) + \int_{-\infty}^{F^{-1}(m)} q_2(s,m,k_l)dF(s) = \lim_{s\to -\infty} q(s,0,k_l)$$

Denote by $\underline{k}_h$ the type k_h that solves $\lim_{m\to 0} \bar{d}_{h1}(m) = \lim_{m\to 0} \bar{d}_{l1}(m)$. The fact $\frac{\partial \bar{d}_{h1}}{\partial k_h} < 0$ implies that $k_l < \underline{k}_h < \bar{k}_h$.

$\underline{k}_h$ and $\bar{k}_h$ divide k_h into three ranges. As shown in Table 2-1, for $k_h > \bar{k}_h$, the frontier of the type k_h provider's report card results lies entirely outside the type k_l provider's set of report card results, implying that if the type k_h provider stays on the frontier, then the type k_l provider cannot imitate his report card result. For $k_h < \bar{k}_h$, the set of the type k_h provider's report card results becomes a proper subset of the type k_l provider's set of report card results, meaning that the type k_l can mimic any report card result from the type k_h provider. For $k_h \in (\underline{k}_h, \bar{k}_h)$, There is a generic subset of the type k_h provider's report card results that does not belong to the type k_l provider's set of results, implying that the type k_l provider can imitate the type k_h provider, but only to a certain extent. The remaining non-generic cases are characterized by $k_h = \bar{k}_h$ and $k_h = \underline{k}_h$, as limits of the generic cases.

[12] The existence and uniqueness of $\bar{k}_h$ result from the Intermediate Value Theorem and monotonicity of $\bar{d}_{h1}(\frac{1}{2}) - \bar{d}_{l1}(\frac{1}{2})$. The same result applies to $\underline{k}_h$ below.

	$\bar{d}_{h1}(\frac{1}{2})$ vs. $\bar{d}_{l1}(\frac{1}{2})$	$\lim_{m\to 0}\bar{d}_{h1}(m)$ vs. $\lim_{m\to 0}\bar{d}_{l1}(m)$
$k_h > \bar{k}_h$	$<$	$<$
$k_h = \bar{k}_h$	$=$	$<$
$k_h \in (\underline{k}_h, \bar{k}_h)$	$>$	$<$
$k_h = \underline{k}_h$	$>$	$=$
$k_h \in (k_l, \underline{k}_h)$	$>$	$>$

Table 2-1: Categories of k_h

For tractability in the case with $k_h \in (\underline{k}_h, \bar{k}_h)$, I impose the following assumption:

Assumption 2: *For $k_h \leq \bar{k}_h$, if $m^1 < m^2$, then $\bar{d}_{h1}(m^2) - \bar{d}_{h1}(m^1) > \bar{d}_{l1}(m^2) - \bar{d}_{l1}(m^1)$.*

Assumption 2 is essentially a single crossing condition, stating that for $k_h \leq \bar{k}_h$, the increase of the same measure of patients will result in a higher increase of the minium mortality rate for the type k_h provider than for the type k_l provider. The validity of the assumption can be better seen from $k_h = k_l$, where the assumption essentially means that a provider will face a higher increase of the minimum mortality rate when treating high-severity patients than when treating low-severity patients. Under Assumption 2, the relation between $\bar{d}_{h1}$ and $\bar{d}_{l1}$ implied by Table 2-1 can be fully shown in Figure 2-4. For $k_h \in (\underline{k}_h, \bar{k}_h)$, denote $\bar{m}(k_h, k_l)$ the solution of $\bar{d}_{h1}(\bar{m}) = \bar{d}_{l1}(\bar{m})$. The fact $\frac{\partial \bar{d}_{h1}}{\partial k_h} < 0$ implies that $\frac{\partial \bar{m}}{\partial k_h} > 0$.

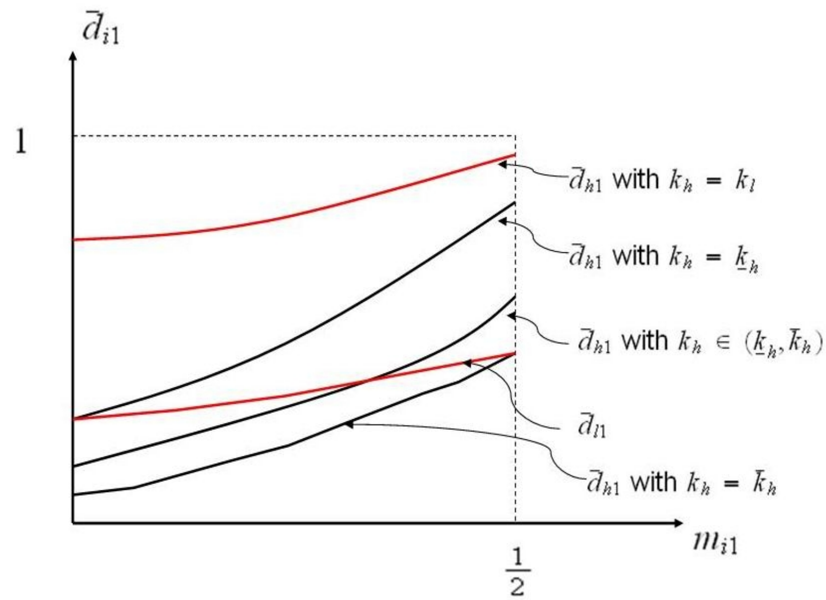


Figure 2-4: Report Card Frontiors of h and l

We can now analyze the providers' period 1 equilibrium actions according to the three ranges of k_h. For simplicity in the subsequent analysis, I further assume that, if a provider engages in selecting patients, the minimum measure of patients that he can avoid is an infinitesimal constant ε.

It is easy to see if $k_h > \bar{k}_h$ then there exists a separating equilibrium such that $m^*_{h1} = \frac{1}{2}$, because $k_h > \bar{k}_h$ implies $\bar{d}_{h1}(\frac{1}{2}) < \bar{d}_{l1}(\frac{1}{2})$, so the type k_h provider can separate himself from the type k_l provider by accepting all patients. Similarly, if $k_h = \bar{k}_h$, then $\bar{d}_{h1}(m) < \bar{d}_{l1}(m)$ for all $m < \frac{1}{2}$, so the type k_h provider can separate himself from the type k_l provider by accepting virtually all patients with shunning ε measure of patients.

Now turn to the case where $k_h < \bar{k}_h$, that is, $\bar{d}_{h1}(\frac{1}{2}) > \bar{d}_{l1}(\frac{1}{2})$. First, if $k_h < \underline{k}_h$, then a necessary condition for a separating equilibrium is that $m_{h1} \neq m_{l1}$, since otherwise the fact that $1 - M^*_{h2} < \frac{1}{2}$ implies that the type k_l provider will have the incentive to mimic the type k_h provider's signal. In other words, in a separating equilibrium it must be that the type k_h provider generates a signal that the type k_l provider does not want to mimic. Due to this, there is a separating equilibrium with

$$m_{h1} + \frac{1}{2} = \frac{1}{2} + 1 - M^*_h(k_h, k_l)$$

which implies

$$m_{h1} = 1 - M^*_h(k_h, k_l)$$

In such a separating equilibrium the type k_h provider generates a signal that the type k_l provider is indifferent to mimic. The type k_h provider receives a payoff $m_{h1} + M^*_h(k_h, k_l) = 1 > \frac{1}{2} + 1 - M^*_h(k_h, k_l)$, where the right hand side of the inequality is the payoff if type k_h provider accepts all the patients in period 1 and is regarded as a type k_l provider in period 2, and so he has no incentive to deviate given the patients' consistent belief, which can be constructed accordingly. Note that, analogous to my analysis so far, actually any $m_{h1} \in [\frac{3}{2} - 2M^*_h(k_h, k_l), 1 - M^*_h(k_h, k_l)]$ can also be supported as a part of a separating equilibrium, while I choose the one that has the maximum payoff for the providers and so implies the minimum selection behavior.

For $k_h \in (\underline{k}_h, \bar{k}_h)$, to separate himself from the type k_l provider, the type k_h provider can choose $m_{h1} = \bar{m}(k_h, k_l) - \varepsilon$, a signal that cannot be imitated by the type k_l provider, or $m_{h1} = 1 - M^*_h(k_h, k_l)$, a signal that the type k_l provider has no incentive to mimic. Since for $k_h = \underline{k}_h$, $\bar{m}(\underline{k}_h, k_l) = 0 < 1 - M^*_h(\underline{k}_h, k_l) + \varepsilon$; for $k_h = \bar{k}_h$, $\bar{m}(\bar{k}_h, k_l) = \frac{1}{2} > 1 - M^*_h(\bar{k}_h, k_l) + \varepsilon$, and $\bar{m}$ and M^*_h are increasing in k_h, there exists a unique $k_b \in (\underline{k}_h, \bar{k}_h)$ such that $\bar{m}(k_b, k_l) = 1 - M^*_h(k_b, k_l) + \varepsilon$. For $k_h < k_b$, $1 - M^*_h(k_h, k_l) > \bar{m}(k_h, k_l) - \varepsilon$, so the type k_h provider prefers the signal that the type k_l provider has no incentive to mimic. For $k_h \in (k_b, \bar{k}_h)$, $\bar{m}(k_h, k_l) - \varepsilon > 1 - M^*_h(k_h, k_l)$, so the type k_h provider prefers the signal that the type k_l provider cannot mimic.

The proposition below summarizes the analysis, with the results graphically shown in Figure 2-5. I leave the description of patients' consistent belief system in Appendix.

Proposition 2 *If in period 1 the patients let C make the referral decision, there is a separating equilibrium such that*

$$m_{h1}^{*}=\begin{cases} 1-M_h^*(k_h,k_l) \text{ if } k_h\in(k_l,k_b] \\ \bar{m}(k_h,k_l)-\varepsilon \text{ if } k_h\in(k_b,\bar{k}_h] \\ \frac{1}{2} \text{ if } k_h>\bar{k}_h \end{cases}$$

The intuition of Proposition 2 is as follows: When patients let C make the referral decision, the type k_h provider faces sicker patients than the type k_l provider does, forcing the former to shun sick patients in order to signal himself. The degree of the type k_h provider's selection behavior is characterized by k_b and $\bar{k}_h$. Intuitively, when $k_h = k_l$, there is no gain from signaling, and naturally there is no selection behavior in equilibrium. When k_h increases, the gain from separating from the type k_l provider increases as well, therefore the type k_h provider is willing to shun more patients in period 1 in exchange for more patients in period 2. Hence the degree of patient selection initially increases with k_h. But when $k_h > \underline{k}_h$, there comes the possibility that the k_h provider can generate a signal that cannot be imitated by the type k_l provider. However, initially the cost of generating such a signal is larger than generating a signal that the type k_l provider does not want to mimic, and so for $k_h < k_b$, the degree of selection behavior keeps increasing with k_h. On the other hand, the cost of generating a signal that the type k_l provider cannot mimic keeps decreasing with k_h. The turning point occurs at $k_h = k_b$, where the costs of the two kinds of signals equal. When k_h is larger than k_b, the type k_h provider switches to generating the signal that his counterpart cannot mimic, and consequently the degree of his selection behavior decreases with his type. In the end, when k_h is sufficiently high, the type k_h provider can signal himself without turning down any patients.

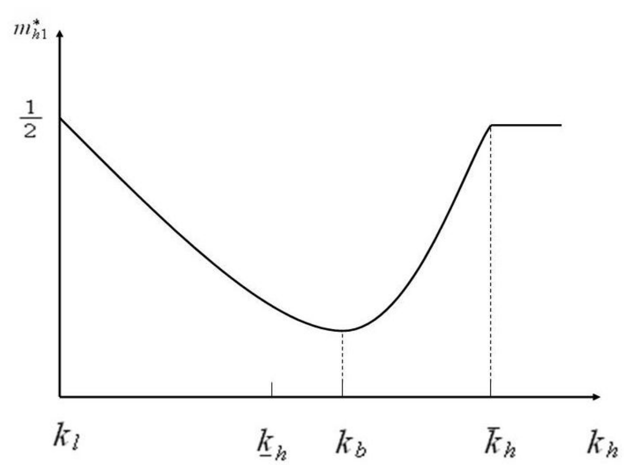


Figure 2-5: m_{h1}^* in the separating equilibrium with C making the period 1 referral decision.

The existence of selection behavior naturally makes one ponder whether there are better report mechanisms. In the current setting, however, the answer is no. In general, a report mechanism is a function $R : [0,1] \times [0,1] \to \Omega[0,1] \times \Omega[0,1]$ where $\Omega[0,1]$ is the $\sigma-$algebra on $[0,1]$. Given R, $r_i \in \Omega[0,1] \times \Omega[0,1]$ is the report card of i published to the public at the end of period 1. For example, in the existing report mechanism there is $R(m_{i1}, d_{i1}) = (m_{i1}, d_{i1})$ with $r_i = (m_{i1}, d_{i1})$. In a PBE induced by a report mechanism R, I use $\hat{x}$ to denote the equilibrium value of a variable x. Also in a PBE induced by R, if $r_i = r_{-i}$ then $M_{i2}(r_i, r_{-i}) = \frac{1}{2}$, as two identical signals yield no new information. A report mechanism R is *revealing* if $\hat{r}_l \neq \hat{r}_h$ in an induced PBE, and a revealing report mechanism R is said to cause selection behavior if in equilibrium there is a provider i such that $\hat{m}_{i1} < \hat{M}_{i1}$. Similar as the discussion before, if a report mechanism R is revealing, then it must be that in equilibrium $\hat{m}_{l1} = \frac{1}{2}$.

The following proposition shows that, despite their imperfection in term of the resulted selection behavior, the existing report cards cause the minimum selection compared with other revealing report mechanisms.

Proposition 3 *Among all the revealing report mechanisms, the existing report mechanism* $R(m_{i1}, d_{i1}) = (m_{i1}, d_{i1})$ *causes the minimum selection behavior.*

Proof. Our previous results imply that we only need to be focused on the scenario where the patients let C make the referral decision in period 1.

For $k_h > \bar{k}_h$, we already know that the existing report mechanism causes no selection behavior.

For $k_h \leq \bar{k}_h$, we only need to focus on the conditions under which the existing report mechanism $R(m_{i1}, d_{i1}) = (m_{i1}, d_{i1})$ causes selection behavior. Suppose there is a revealing mechanism such that in equilibrium $\hat{m}_{h1} > m^*_{h1}$. There are two cases:

(i) For $k_h \in (k_b, \bar{k}_h]$, $m^*_{h1} = \bar{m}(k_h, k_l) - \varepsilon$, so

$$\hat{m}_{h1} \geq \bar{m}(k_h, k_l) > 1 - M^*_h(k_h, k_l)$$

which implies

$$\hat{m}_{h1} + \frac{1}{2} > \frac{1}{2} + [1 - M^*_h(k_h, k_l)]$$

then the type k_l provider can be better off by deviating to the actions that leads to $r_l = \hat{r}_h$, which he is capable to do now thanks to the above result $\hat{m}_{h1} > 1 - M^*_h(k_h, k_l)$, contradicting the report mechanism being a revealing one.

(ii) For $k_h \in (k_l, k_b]$, $m^*_{h1} = 1 - M^*_h(k_h, k_l)$, then $\hat{m}_{h1} > m^*_{h1}$ implies that

$$\hat{m}_{h1} + \frac{1}{2} > \frac{1}{2} + [1 - M^*_h(k_h, k_l)]$$

then the type k_l provider can also be better off by deviating to the actions that leads to $r_l = \hat{r}_h$, contradicting the report mechanism being a revealing one. ■

The intuition of Proposition 3 is: In the present setting the existing report cards have shown the maximum information available, and any other revealing mechanisms necessarily temper information based on the existing cards, only to make it easier for the type k_l provider to pool with the type k_h provider, forcing the the latter to shun even more patients to signal himself. The result therefore suggests that future study aimed at a better report mechanism should be focused on eliciting information beyond the present setting, such as inducing better information about s_j and so improving the accuracy of the risk adjustment procedure.

3.3 Discussion

Pooling Equilibrium So far the analysis has been focused on separating equilibrium, where the providers' types are revealed by the report cards. However, pooling equilibria also exist, though they involve unappealing off-equilibrium-path patient belief. In a pooling equilibrium, the providers yield the same report card results, keeping the period 2 patients' belief same as the prior, and thus each provider receives measure $\frac{1}{2}$ patients in period 2. To construct a pooling equilibrium, one only needs to specify an equilibrium report card result with a sufficiently large m_{i1}, and then have the patients hold the belief that any provider unilaterally deviating from that result will be regarded as of type k_l. However, it follows that to support *any* pooling equilibrium, the period 2 patients must hold the belief that a provider is of type k_l if his report card result is slightly different[13] from the equilibrium outcome. This is because, if the period 2 patients hold the opposite belief, i.e. a provider will be of type k_h should he yield a slightly different report card result, then discontinuity between the period 2 payoff $1/2$ from pooling and $M_h^*(k_h, k_l) > 1/2$ from separating means the type k_h provider will always slightly deviate in period 1. But such a belief system implies that the patients prefer not to distinguish the high quality provider from the low quality one, which is implausible.

C's Referring Pattern More generally we can characterize C's referring pattern by a cutoff level s_c, which can be larger or smaller than $s_{1/2}$. Above s_c, C refers a patient to the type k_h provider, and otherwise refers her to the type k_l provider. For simplicity let's focus on the case where patients let C make the decision in period 1. If $s_c > s_{1/2}$ then the previous analysis carries through. The type k_h provider's report card frontier is similarly characterized by three ranges of k_h, with $\underline{k}_h$ unchanged while $\bar{k}_h$ being lower than before due to a lower measure of coming patients. Differently, if $s_c < s_{1/2}$, then in any scenario the report cards will reveal the providers' types without causing patient selection. This is because the type k_h provider is now receiving more coming patients than the type k_l provider, and so by accepting all the

[13] More exactly, "slightly different" means the report card result is ε away from the the equilibrium outcome.

patients he will generate a report card with $m_{h1} = 1 - F(s_c)$ that can not be imitated by the type k_l provider. Nonetheless, I should point out that throughout the analysis on separating equilibrium I have put no restriction on patients' off-equilibrium-path belief, as long as the belief system supports the equilibrium strategies. It will not be unrealistic to assume that the patient's belief system is characterized by "indifference curves" such that a provider with a report card result on a lower indifference curve (lower mortality rate and higher volume) will be regarded as the type k_h provider. Given any of such belief systems, the type k_h provider's incentive to signal himself from selecting patients revives.

Payoff from the Alternative Treatment For simplicity I have assumed 0 payoff for patients undergoing the alternative treatment, which captures the worst possible scenario should providers engage in selection. The set-up could be more realistic. For example, assume a patient's payoff from receiving the alternative treatment is $\alpha \in (0,1)$, then the patients' period 1 expected payoffs are, for $i \in \{A, B\}$,

$$\begin{aligned} E_{i1} &= \frac{1}{2}\int_{\sigma_{i1}(k_l)} 1 - q(s, m_{l1}, k_l)dF(s) + \frac{1}{2}\int_{\sigma_{i1}(k_h)} 1 - q(s, m_{h1}, k_h)dF(s)] \\ &\quad +\alpha[1 - \frac{1}{2}(\int_{\mathbf{R}\backslash\sigma_{i1}(k_l)} dF(s) + \int_{\mathbf{R}\backslash\sigma_{i1}(k_h)} dF(s))] \end{aligned}$$

and

$$\begin{aligned} E_{C1} &= \int_{\sigma_{i1}(k_l)\cap(-\infty,s_{1/2})} 1 - q(s, m_{l1}, k_l)dF(s) + \int_{\sigma_{i1}(k_h)\cap(s_{1/2},+\infty)} 1 - q(s, m_{h1}, k_h)dF(s) \\ &\quad +\alpha\left[1 - \int_{\mathbf{R}\backslash\sigma_{i1}(k_l)\cup\sigma_{i1}(k_h)} dF(s)\right] \end{aligned}$$

$\alpha \in (0,1)$ implies the alternative treatment yields a positive payoff lower than a successful surgery. Such a generalization captures several facts. First, a standard alternative treatment, percutaneous transluminal coronary angioplasty (PTCA), is known to suffer from recurrence of symptoms. The reported failure rates due to restenosis (recurrent narrowing) in the first 6-12 months following the PTCA procedure are 30-60%. (Rupprecht, et al. 2005; Kulick, 2005). Second, an alternative treatment may mean transferring patients to other providers that are not subject to the report cards, including those in the neighbor states (Omoigui et al. 1996) and those within-state government hospitals, including county and city hospitals[14]. However, I believe my results are robust under the original assumption, and in term of theoretical analysis the generalization will offer no more insights but computational redundancy. This is because, first, the patients' decisions will be similar as in the 0 payoff setting: In period 2, a patient

[14] "In large urban areas, these tend to be safety net hospitals (serving the poor and uninsured)." Raffel & Barsukiewicz, "The U.S. Health System, Origins and Functions", 5th Edition, page 128

will still choose the provider that gives her the highest expected payoff, and in period 1 the patients will equally randomize between the providers when there are no C, or when C is in action choose between self-referral and C in the way analogous to the 0 payoff setting. Given such similarity in patients' actions and response to report cards, providers' actions should also be close to the original setting.

Word-of-Mouth Other than resorting to published information, patients may rely on word-of-mouth from friends and relatives to gain information about the providers (Gibbs et al., 1996). Though such information is inevitably noisy, if it is positively correlated with a provider's quality, we may conjecture that existence of word-of-mouth alleviates the selection behavior, since it provides the providers with another channel to signal their types. In particular, if one assumes that patients suffering from failed surgeries (death) do not engage in spreading provider information, then the larger the patient volume, the larger the amount of successful surgeries, and consequently the larger the good word-of-mouth, which encourages providers to accept more patients. Moreover, previous studies also show that, compared with a small town will a small number of providers, it is hard for patients to gain information about providers through word-of-mouth in a big city with a large number of providers, since the chance that friends and relatives know a randomly picked provider is low (Satterthwaite, 1979; Pauly and Satterthwaite, 1981). Following this line, I conjecture that providers' selection behavior will be more severe in urban areas than in suburban and rural areas. In other words, selection will be most severe in the areas where report cards are needed the most. Future extension in this direction bears both theoretical and empirical interest.

Learning-by-Doing Existence of learning-by-doing among providers has also been documented in empirical study (Ramanarayanan, 2006). Since I assume provider types to be numerical rather than categorical, the model also has the potential to be further extended to incorporate learning-by-doing. A starting point can be assuming the increment of provider quality Δk_i to be an increasing function of m_{i1}, and so the provider's period 2 quality increases with the measure of patients he treats in period 1. Consequently, upon seeing the report cards, period 2 patients now will figure out not only k_i but also $k_i + \Delta k_i(m_{i1})$. Then we may conjecture that learning-by-doing will attenuate the selection behavior, since a provider is now facing a second layer of trade-off: Though avoiding patients can help to signal oneself or pool with the other, it also lowers Δk_i. In extreme cases, it may be that $k_l + \Delta k_l(m_{l1}) > k_h + \Delta k_h(m_{h1})$, that is, the type k_l provider grows up to be a provider with higher quality. The concern of lowered Δk_i therefore should curb a provider's incentive to avoid patients.

4 Empirical Implications, Framework, and Partial Evidence

4.1 Effects of Report Cards

The theoretical model has implications on effects of the existing report cards, including incidence effect, quantity effect, matching effect, and welfare effect. Throughout this section I will focus on the scenario where C makes the referral decision. Without loss of generality, I further restrict the attention to the generic cases of $\tilde{M}_h^* > \bar{M}_h^*$ and $\tilde{M}_h^* < \bar{M}_h^*$.

Incidence Effect Following the notion in Dranove et al. (2003), the incidence effect of report cards is measured by the average change of illness severity of the patients receiving the surgery, compared with when there are no report cards. At period 1, since the type k_l provider accepts all patients whereas the type k_h provider engages in selection behavior when $k_h < \overline{k}_h$, the report cards lead to a negative incidence effect. At period 2, however, the incidence effect depends. If $\tilde{M}_h^* > \bar{M}_h^*$, then in addition to treating all the patients with $s_j \geqslant s_{1/2}$ the type k_h provider also treats a fraction of patients with $s_j < s_{1/2}$, which lowers the mean illness severity of his patients, and thus leads to a negative incidence effect. If $\tilde{M}_h^* < \bar{M}_h^*$ then the type k_h provider treats all the patients with $s_j < s_{1/2}$ plus a fraction of patients with $s_j \geqslant s_{1/2}$, leading to a lower mean illness severity of his patients. In contrast, the type k_l provider now only treats patients with $s_j \geqslant s_{1/2}$, resulting a higher mean illness severity for him. Consequently the average incidence effect is unclear.

Quantity Effect The quantity effect is measured by the average change of the measure of the patients undergoing the surgery, compared with when there are no report cards. At period 1, the type k_h provider's selection behavior implies a negative quantity effect. At period 2, as the providers' type are revealed, the type k_h provider sees an increase of patients while the type k_l provider's patient measure decreases. Since both providers accept all the patients, on average the effects on the two cancel out, implying a zero quantity effect in period 2. In reality, however, providers' selection behavior entails a prolonged decision process, implying not only denial of surgery but also delayed treatments. Moreover, recurring symptoms from the alternative treatment such as PTCA[15] also implies an increase of patients in the 2nd period. Based on these I predict a positive quantity effect in period 2.

The theoretical model's implication about incidence effects and quantity effects allows us to visit the conclusions drawn by Dranove et al. (2003) with a second thought. Taking a difference-in-differences approach, they show that releasing CABG report cards in New York and Pennsylvania resulted in negative incidence effects and positive quantity effects. They argue that the negative incidence effect indicates "report cards have caused a shift in incidence

[15] (Rupprecht, et al. 2005; Kulick, 2005)

from sicker to healthier patients", and the positive quantity effect in addition implies that "the quantity increase was entirely accounted by surgeries on less severely ill patient". But based on my model, these results should be interpreted more carefully. First, a negative incidence effect alone can *not* be used to indicate existence of selection behavior. As aforementioned, in the theoretical model in period 2 there is no selection behavior, but when $\tilde{M}_h^* > \bar{M}_h^*$, the shift of the fraction of patients with $s_j < s_{1/2}$ from the type k_l provider to the type k_h provider also leads to a negative incidence effect. Second, a negative incidence effect together with a positive quantity effect does not mean that the quantity increase is attributed to increase of surgeries on healthier patients. More exactly, in Dranove et al. (2003), the notion that the increase of quantity came from increase of surgeries on healthier patients implicitly assume that the providers can shift the distribution of CABG patient's illness severity toward a lower mean. Such an assumption is self-contradictory, because if the providers are capable of shifting the distribution of patients' illness severity, then the providers will not be short of patient sources, which implies they will not be concerned by the report cards, and consequently they will not select patients in the first place. In my model, selection behavior exists in the first period, but it is associated with a negative incidence effect and a *negative* quality effect. Therefore, to indicate existence of selection behavior, empirically one needs to show both effects to be negative in the first period.

Matching Effect The matching effect is measured by the average change in the variation of illness severity in the surgical patients. Though Dranove et al. (2003) measures variation by the coefficient of variation, in the context of my model variance is a more accurate measure since it is exempted from the influence of mean. In the model, in period 1, the type k_h provider's avoidance of sicker patients implies a smaller variance of patient types, thus a negative matching effect. But the effect in period 2 depends. If $\tilde{M}_h^* > \bar{M}_h^*$, then the distribution of patient types confronting the type k_l provider does not change, whereas the type k_h provider faces both sick patients and healthy patients, implying a larger variance (also a higher coefficient of variation due to a lower mean), resulting in a positive matching effect. If $\tilde{M}_h^* < \bar{M}_h^*$, the average effect is uncertain, since on one hand the type k_l provider only treats patients with $s_j \geqslant s_{1/2}$, and thus the actual effect depends on comparison between $Var\left(s \mid s < s_{1/2}\right)$ and $Var\left(s \mid s \geq s_{1/2}\right)$, and on the other hand, since the type k_h provider treats all types of patients, the matching effect on the type k_h provider is positive due a higher variance (also a higher coefficient of variation due to a lower mean). Nevertheless one message is clear: As discussed in the previous section, the theoretical model challenges the conventional wisdom that revelation of provider types will result in only sick patients matched with the high-quality provider, since all patients, healthy or sick, would like to seek the best possible treatment. In line with this, a negative period 1 matching effect now is a piece of evidence supporting existence of selection behavior, instead of indication of "improved patient sorting" as suggested in Dranove et al. (2003).

Welfare Effect Social welfare in each period is measured by the sum of payoffs of providers and patients. When there are no report cards, the social welfare, denoted by W_{NR}, is the same in each period:

$$W_{NR} = 1 + \left[1 - [\int_{-\infty}^{s_{1/2}} q(s, \frac{1}{2}, k_l) dF(s) + \int_{s_{1/2}}^{+\infty} q(s, \frac{1}{2}, k_h) dF(s)] \right]$$

When there are report cards, the social welfare in the first period, denoted by W_{R1}, is

$$\begin{aligned} W_{R1} &= (\frac{1}{2} + m_{h1}^*) + \left[\int_{-\infty}^{s_{1/2}} 1 - q(s, \frac{1}{2}, k_l) dF(s) + \int_{s_{1/2}}^{F_{h1}^{-1}(m_{h1}^*)} 1 - q(s, m_{h1}^*, k_h) dF(s) \right] \\ &= \frac{1}{2} + 2m_{h1}^* - \left[\int_{-\infty}^{s_{1/2}} q(s, \frac{1}{2}, k_l) dF(s) + \int_{s_{1/2}}^{F_{h1}^{-1}(m_{h1}^*)} q(s, m_{h1}^*, k_h) dF(s) \right] \end{aligned}$$

and the the social welfare in the second period, denoted by W_{R2}, depends. If $\tilde{M}_h^* > \bar{M}_h^*$, then

$$\begin{aligned} W_{R2} &= 1 + 1 - \left\{ (\bar{M}_h^* - \frac{1}{2}) \int_{-\infty}^{s_{1/2}} q(s, \bar{M}_h^*, k_h) dF(s) + (1 - \bar{M}_h^*) \int_{-\infty}^{s_{1/2}} q(s, 1 - \bar{M}_h^*, k_l) dF(s) \right. \\ &\quad \left. + \frac{1}{2} \int_{s_{1/2}}^{+\infty} q(s, \bar{M}_h^*, k_h) dF(s) \right\} \end{aligned}$$

and if $\tilde{M}_h^* < \bar{M}_h^*$,

$$\begin{aligned} W_{R2} &= 1 + 1 - \left\{ (\tilde{M}_h^* - \frac{1}{2}) \int_{s_{1/2}}^{+\infty} q(s, \tilde{M}_h^*, k_h) dF(s) + (1 - \tilde{M}_h^*) \int_{s_{1/2}}^{+\infty} q(s, 1 - \tilde{M}_h^*, k_l) dF(s) \right. \\ &\quad \left. + \frac{1}{2} \int_{-\infty}^{s_{1/2}} q(s, 1 - \tilde{M}_h^*, k_h) dF(s) \right\} \end{aligned}$$

The welfare effect is then measured by the sum of both periods' effects, $(W_{R1} - W_{NR}) + (W_{R2} - W_{NR})$. The actual result, however, is not clear and stands for empirical tests. However, it is worth pointing out that in *each* period report cards lead to not only welfare losses but also gains. In period 1, on one hand avoidance of patients leads to decreased payoff for both the type k_h provider and the patients denied surgery; on the other hand, a decreased patient volume improves the type k_h provider's outcome, benefiting the patients that do receive surgeries from him. In period 2, if $\tilde{M}_h^* > \bar{M}_h^*$, then all the patients with $s_j < s_{1/2}$ will benefit from those switching to the type k_h provider, but the type k_h provider now faces an increased patient volume, lowering the payoffs of those with $s_j \geq s_{1/2}$. Similarly, if $\tilde{M}_h^* < \bar{M}_h^*$, all the patients with $s_j < s_{1/2}$ benefit from receiving surgeries from the type k_h provider, but the large amount of patients drives a fraction of patients with $s_j \geq s_{1/2}$ to choose the type k_l provider, and thus lowers their payoffs.

4.2 Empirical Framework

The discussion in the previous subsection shows that based on the theoretical model each effect of the existing report cards vary from period 1 to period 2. As a result, a difference-in-differences estimate should be used for each treatment effect in *each* period. Consequently, the traditional single difference-in-differences estimate only captures the average of the two periods' effects. To see this, let time period $t = 0, 1, 2$,where 0 stands for the period before the report-card program is effective, with 1 for the first period and 2 for the second. Denote the superscript T the treatment group and NT the control (non-treatment) group. Denote y a variable whose change from $t = 0$ to $t = 1, 2$ in the treatment group encompasses a treatment effect of interest. Suppose there are n members in the treatment group and m members in the control group. For simplicity, ignore covariates.

Suppose

$$\begin{array}{ll} y_{j0}^T = \alpha_0 + \varepsilon_{j0} \ , & y_{j'0}^{NT} = \alpha_0 + \varepsilon_{j'0} \\ y_{j1}^T = \alpha_1 + \beta_1 + \varepsilon_{j1} \ , & y_{j'1}^{NT} = \alpha_1 + \varepsilon_{j'1} \\ y_{j2}^T = \alpha_1 + \beta_2 + \varepsilon_{j2} \ , & y_{j'2}^{NT} = \alpha_1 + \varepsilon_{j'2} \end{array}$$

where ε_{jt} and $\varepsilon_{j't}$ are i.i.d. with $E[\varepsilon_{jt}] = 0$. Then there are

$$\begin{array}{ll} E[y_0^T] = \alpha_0 \ , & E[y_0^{NT}] = \alpha_0 \\ E[y_1^T] = \alpha_1 + \beta_1 \ , & E[y_1^{NT}] = E[y_2^{NT}] = \alpha_1 \\ E[y_2^T] = \alpha_1 + \beta_2 \ , & \end{array}$$

and identification of the period 1 effect β_1 and the period 2 effect β_2 stems from

$$\begin{aligned} \beta_1 &= \left[E[y_1^T] - E[y_0^T]\right] - [E[y_1^{NT}] - E[y_0^{NT}]] \\ \beta_2 &= \left[E[y_2^T] - E[y_0^T]\right] - [E[y_2^{NT}] - E[y_0^{NT}]] \end{aligned}$$

On the other hand the single difference-in-differences estimate $\hat{\beta}$ implies

$$\begin{aligned}
\hat{\beta} &= \left[\frac{\sum y_{j1}^T + \sum y_{j2}^T}{2n} - \frac{1}{n}\sum y_{j0}^T\right] - \left[\frac{\sum y_{j'1}^{NT} + \sum y_{j'2}^{NT}}{2m} - \frac{1}{m}\sum y_{j'0}^{NT}\right] \\
&= \left[\frac{1}{2}\left[\frac{\sum y_{j1}^T}{n} + \frac{\sum y_{j2}^T}{n}\right] - \frac{1}{n}\sum y_{j0}^T\right] - \left[\frac{\sum y_{j'1}^{NT} + \sum y_{j'2}^{NT}}{2m} - \frac{1}{m}\sum y_{j'0}^{NT}\right] \\
&\rightarrow \frac{1}{2}\left\{\left[(E[y_1^T] + E[y_2^T]) - 2E[y_0^T]\right] - 2\left[E[y_1^{NT}] - E[y_0^{NT}]\right]\right\} \\
&= \frac{1}{2}(\beta_1 + \beta_2)
\end{aligned}$$

Moreover, based on this one-estimate-for-one-period framework, we need a more careful handling of the data set in Dranove et al. (2003). The data set covers data from 1987 to 1994. However, the report-card program was introduced in New York in 1991 but was not enacted in Pennsylvania until 1993. Hence, in the context of the theoretical model, the data set contains both period 1 and period 2 data for New York, but only contains period 1 data for Pennsylvania. Therefore, assuming each period consists of two years, to estimate β_1, one needs to exclude the 1993-94 New York data, while to estimate β_2 one needs to exclude the 1991-92 New York data and the 1993-94 Pennsylvania data.

To estimate each treatment effect, the specification of the regression is the same as that in Dranove et al. (2003) except that now there are two difference-in-differences estimates for each effect. To test for incidence effects, the regression form is

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

where l indexes hospitals, s indexes states, and t indexes time, $t = 1987, ..., 1994$; h_{st} 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 state fixed effects; $\mathbf{B}_t$ is a vector of time fixed effects; $\mathbf{Z}_{lst}$ is a vector of hospital characteristics. N_{st} is the number of hospitals, and its square and cube, in state s at time t; and e_{lst} is an error term. In the first estimate, where the data set excludes the 1993-94 New York data, $L_{st} = 1$ if the hospital is in New York in 1991 or 1992, or in Pennsylvania in or after 1993. In the second estimate, where the data set excludes the 1991-92 New York data and the 1993-94 Pennsylvania data $L_{st} = 1$ if the hospital is in New York in or after 1993. The coefficient p is the difference-in-differences estimate of the effect of report cards on the severity of patients who receive CABG. To estimate the matching effect, the dependent variable will be replaced by the within-hospital coefficient of variation of illness severity.

To estimate quantity effects, the regression form is

$$C_{kst} = \mathbf{A}_s + \mathbf{B}_t + g \cdot \mathbf{Z}_{kst} + p \cdot L_{st} + e_{kst}$$

where k indexes patients, s indexes states, and t indexes time, $t = 1987, ..., 1994$; 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 acute myocardial infarction (AMI); $\mathbf{A}_s$ is a vector of sate fixed effects; $\mathbf{B}_t$ is a vector of time fixed effects; $\mathbf{Z}_{kst}$ is a vector of patient characteristics; and e_{kst} is an error term. Similar as in the previous regression form, in the first estimate, where the data set excludes the 1993-94 New York data, $L_{st} = 1$ if the hospital is in New York in 1991 or 1992, or in Pennsylvania in or after 1993. In the second estimate, where the data set excludes the 1991-92 New York data and the 1993-94 PA data $L_{st} = 1$ if the hospital is in New York in or after 1993. A positive p implies that report cards increased the probability that an AMI patient receives CABG.

My empirical framework also calls for a more careful choice of the control group states. As I discuss in Section 2, as a means of avoiding patients, the providers in New York and Pennsylvania may transfer sick patients to the neighbor states (Omoigui et al., 1996), taking resources that otherwise would be allocated to the neighbor states' own patients. Hence the report-card programs may affect those neighbor states in the first period. Moreover, in period 2, although the report cards only contain statewide provider information, their free accessibility means that their readers are not restricted to the issuing states. The argument in the theoretical model that period 2 patients prefer informed self-referral then implies that, further taking into account the sizes and closeness of the northeastern states around New York and Pennsylvania, patients in the neighbor states may travel to the issuing states for informed treatment. Hence the report-card programs may also have an impact on the neighbor states in the second period.[16] Therefore, to select the control group members, I propose to first use the states adjacent to New York and Pennsylvania as the treatment group, and use the others excepts New York and Pennsylvania as the control group. If a treatment effect on the neighbor states is significant, then the neighbor states should be excluded from the control group for the estimation on New York and Pennsylvania.

4.3 Data and Partial Evidence

Dranove et al. (2003) use data from two sources. The patient-level data are the longitudinal Medicare claim data for 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 hospital-

[16]Such an impact also lowers income for the providers in the neighbor states, which may help to explain why ensuing states releasing report cards are initially centered around New York.

level data come from the American Hospital Association. However, as one of the authors in the 2003 paper points out, "[b]ecause the data have information on particular individuals, it is subject to strict confidentiality restrictions including that we can't provide it to anyone else. You can apply to use the data if you like, although it is a long process."[17]

Despite inaccessibility of the data set at this stage, particular specifications in Dranove et al. (2003) allow us to reinterpret their estimation results, offering evidence for my theoretical predictions. Citing an augment of the New York report cards in 1992, Dranove et al. (2003) estimate each treatment effect under two assumptions: one that assumes report cards effective 1991 in New York, another assuming them effective 1993 in New York. Though the 1992 augment was to include surgeon-level information, the 2003 paper is focused on the hospital level. Hence the augment should have small effects on the hospitals. As I discussed earlier, in the single difference-in-differences estimation, blending period 1's data with that of period 2 mixes the treatment effect of one period with that of another. Following the theoretical model, the problem should be exacerbated under the assumption that report cards were effective 1993 in New York, since it excludes the period 1 treatment effect in New York and makes the period 2 treatment effect more significant. Consequently, under the two assumptions the differences between the estimation results, though not exploited in the original paper, should be significant and follow the prediction about the period 1 and period 2 effects made from the theoretical model. I investigate them below.

First, as discussed before, I predict quantity effects to be negative in period 1 and positive in period 2. Thus the difference-in-differences estimate under the 1993 assumption should be larger than that under the 1991 assumption. The result in the 2003 paper, reproduced in Table 3-1, is consistent with this prediction. The increase of probability that the average AMI patient will undergo CABG surgery within one year of admission for AMI rises from 0.60 under the 1991 assumption to 0.91 under the 1993 assumption. The result also shows that in the single difference-in-differences estimate the positive period 2 effect outweighs the negative period 1 effect, more so under the 1993 assumption.

[17] Quoted from correspondence bewteen this paper's author and the authors of the 2003 paper.

EFFECTS OF REPORT CARDS ON CABG, PTCA,AND CATHETERIZATION RATES: MEDICARE BENEFICIARIES WITH AMI, 1987-94.		
	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	Effect of Report Cards
DEPENDENT VARIABLE	(1)	(4)
CABG within one year of admission	.60**	.91**
(1 = yes)	(.21)	(.44)

** Significantly different from 0 at the 5 percent level
Reproduced from Dranove et al. (2003) Table 4

Table 3-1

The results of estimating incidence effects, reproduced in Table 3-2, indicates a negative period 2 effect, because the estimation results under the 1993 assumption are further negative than that under the 1991 assumption, regardless of whether we measure illness severity by total hospital expenditure prior to admission or by total days in hospital prior to admission. But now we can see, as I discuss before, a negative period 2 incidence effect cannot be used as evidence of providers selecting patients. Moreover, given a negative period 2 effect, it is more likely that $\tilde{M}_h^* > \bar{M}_h^*$ occurs, that is, in period 2 the type k_h provider not only treats all the patients with $s_j \geq s_{1/2}$ but also some patients with $s_j < s_{1/2}$, supporting my prediction about the matching effect, which we now turn to.

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		
	BENEFICIARIES RECEIVING CABG	
	ASSUMES REPORT CARDS EFFECTIVE 1991 IN N.Y. AND 1993 IN PA	ASSUMES REPORT CARDS EFFECTIVE 1993 IN N.Y AND PA.
DEPENDENT VARIABLE	(1)	(2)
ln(mean of patients' total hospital expenditures	-3.92**	-5.30**
one year prior to admission)	(-1.52)	(1.10)
ln(mean of patients' total days in hospital	-3.74**	-4.51**
one year prior to admission)	(1.84)	(1.54)

** Significantly different from 0 at the 5 percent level
Reproduced from Dranove et al. (2003) Table 2

Table 3-2

Based on an implicit assumption of identical patient distribution between providers, Dranove et al. (2003) interpret a report-card-induced improved matching between patients and providers as one that leads to sicker patients matched with high-quality providers and healthier patients matched with low-quality ones, suggesting a negative period 2 matching effect. In contrast, as discussed before, the theoretical model suggests that the period 2 incidence effect tends to be positive. Though the positive estimation results in the 2003 paper, reproduced in Table 3-3, do not directly support the claim made by Dranove et al., the numbers are consistent with mine. Since the estimates mix the period 1 effect with the period 2 effect, and the period 1 matching effect is negative due to the selection behavior, the positive results imply a positive period 2 matching effect, which is more significant under the 1993 assumption.

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

	BENEFICIARIES RECEIVING CABG	
	ASSUMES REPORT CARDS EFFECTIVE 1991 IN N.Y. AND 1993 IN PA	ASSUMES REPORT CARDS EFFECTIVE 1993 IN N.Y AND PA.
DEPENDENT VARIABLE	(1)	(2)
ln(CV of patients' total hospital expenditures one year prior to admission)	3.00**	3.60**
	(1.39)	(1.77)
ln(CV of patients' total days in hospital one year prior to admission)	.94	2.74
	(2.22)	(3.53)

** Significantly different from 0 at the 5 percent level
Reproduced from Dranove et al. (2003) Table 2

Table 3-3

The 2003 paper also takes a difference-in-difference-in-differences approach to estimate the quantity effect of report cards on sick patients, defined by those admitted to hospital in the year before AMI. The estimation results, reproduced in the table below, were negative under both assumptions but were significant only under the 1993 assumption. The negativity is consistent with the prediction of period 1 selection behavior, which concentrates on the sick patients. Moreover it shows that the strong negative period 1 effect outweighs the positive period 2 effect in this cohort. But the results indicate the potential problem caused by including the neighbor states in the control group. As discussed before, as a means of avoiding patients in period 1, transferring sick patients to neighbor states from New York and Pennsylvania implies a negative quantity effect on sick patients in the neighbor states. At period 2, revelation of provider types led patients in neighbor states to seek informed treatment in New York and Pennsylvania, implying a positive quantity effect in the neighbor states. As shown in Capps et al. (2001), sick patients are more willing to incur financial and travel costs to obtain treatment from high-quality providers, hence the positive period 2 quantity effect on the neighbor states is likely to be concentrated on sick patients. Via the difference-differences estimate, the negative period 1 quantity effects on the neighbor states offsets a negative quantity effect on sick patients in New York and Pennsylvania, whereas the positive period 2 quantity effects in the neighbor states amplifies the negative effect in New York and Pennsylvania. This explains why the estimate results were insignificant under the 1991 assumption but significant under the 1993 assumption.

EFFECTS OF REPORT CARDS ON CABG, PTCA, AND CATHETERIZATION RATES: MEDICARE BENEFICIARIES WITH AMI, 1987-94						
	EFFECTIVE 1991 IN N.Y. AND 1993 IN PA			EFFECTIVE 1993 IN N.Y AND PA.		
	Effect of Report Cards	Admission to Hospital in Year before AMI (2)	Report Cards x Prior Year Admission	Effect of Report Cards	Admission to Hospital in Year before AMI (2)	Report Cards x Prior Year Admission
DEPENDENT VARIABLE	(1)	(2)	(3)	(4)	(5)	(6)
CABG within one year	.81**	-3.80**	-.65	1.39**	-3.78**	-1.52**
of admission (1 = yes)	(.15)	(.15)	(.44)	(.42)	(.16)	(.19)

** Significantly different from 0 at the 5 percent level
Reproduced from Dranove et al. (2003) Table 4

Table 3-4

Dranove et al. (2003) measure the welfare effect of report cards by change of patient's post-surgery hospital expenditure and readmission rates, and they conclude that overall report cards decreased social welfare. Though these post-surgery measures are not incorporated in my theoretical model, the implication on other treatment effects suggests that these measures also vary across the periods, and thus should be estimated separately. If both periods' effects turn out to be negative, then we can conclude that the report cards resulted in lower social welfare. In other cases, however, the results need more careful interpretation. In particular, I want to stress that a single difference-in-differences estimate, despite capturing the average effect of both periods, is inappropriate for measuring the actual welfare effect. This is because, as people aging, population prone to cardiac disease is also changing. In the theoretical model, two stages suffice for the analysis, but in reality the periods after revelation of provider types will be more than one stage. To count the actual welfare effect, one needs to use the sum of the (discounted) effects in all periods. Hence even though a negative period 1 effect may outweighs a positive period 2 effect, leading to a negative single difference-in-differences estimation result, the sum of all periods' effects may well turn out to be positive, rendering the result from the single difference-in-differences estimate inconclusive.

5 Conclusion

My study of the impact of the health care report cards starts from the observation that previous literature put unbalanced emphasis on the empirical approaches, often basing interpretation of regression results on intuitions and conventional wisdom. I believe that providing a theoretical foundation to the empirical framework is necessary and will shed new light on this field. When professional and academic critics argue that the health care providers may use private patient information to "game" the system, a game-theoretical model is the best candidate to help us understand why they game and how.

Based on a two-stage signaling game, I show that, when patients and providers are matched randomly, the trade-off between the measures in the existing report cards render them the optimal mechanism that reveal provider types without causing providers avoiding patients. Asymmetric distributions of patient types between providers caused by the referring physician, however, may force the high-quality provider to shun patients in order to separate himself from the low-quality one. The results imply that, in contrast to previous literature, a negative incidence effect alone can not be used to prove existence of selection behavior. Moreover, the traditional single difference-in-differences approach cannot be used to capture the long run welfare effect. More generally, I propose a new empirical framework where a difference-in-differences estimate is used for each treatment effect in each period. In addition my empirical framework calls for more attention to selection of the control group.

My results also clarify conventional wisdom on several accounts. First, many previous studies have mistaken the existing report cards as simply "mortality report cards", neglecting the fact that they actually show multidimensional measures. Such negligence results in not only flawed understanding of the existing report cards but also searching for new mechanisms in wrong directions. Second, at the patient level I argue that a healthy patient is foremost a patient, and no matter how healthy he is he will seek the best possible treatment. Thus it is incorrect to assume that healthy patients will choose the low-quality provider without conditions. Third, at the provider level I argue that a high-quality provider is not necessarily associated with a better outcome, because a sufficiently large patient volume will lower his performance.

Though the theoretical model I employ is a standard two-stage signaling game, it has the potential to be further extended to incorporate more realistic features, including word-of-mouth among patients, learning-by-doing of providers, and overlapping generations that allow providers to enter and exit.

The existence of selection behavior documented in previous literature may daunt people planing to introduce the report-card program to other fields in the health care industry and more broadly to other industries, such as law and education, where goods and services are also provided by skilled experts. My results help to clarify certain issues of concern. First, in the health care industry, we should recognize that though from an economics perspective whether a program should be implemented hinges on weighing between its benefit and cost, from a broader perspective, the decision may depend on whether we want to bear any cost in term of denial or delay of treatments to some first period patients at all. Second, I show that providers' selection behavior is not caused by the report cards alone, rather it is caused by the combination of the report cards and the particular features of the health care industry. This implies the report-card program has the potential to be successfully introduced to other industries where good or service providers and consumers are randomly matched. Nevertheless, each industry is distinct in its own features, so the specific conclusions should be drawn upon a sufficient understanding of the industry where the report-card program is applied to.

6 Appendix 1

Proof of Lemma 1:

Proof. Suppose the providers' types (k_i, k_{-i}) are revealed at $t = 2$, then the facts $F_{it}(.) = F(.)$ and that each provider performs surgeries on all coming patients in period 2 imply

$$\begin{aligned} E_{i2} &= 1 - \int_{-\infty}^{+\infty} q(s, M_{i2}, k_i) dF(s) \\ E_{-i2} &= 1 - \int_{-\infty}^{+\infty} q(s, 1 - M_{i2}, k_{-i}) dF(s) \end{aligned}$$

Following the discussion prior to the lemma, it must be $E_{i2} = E_{-i2}$ in equilibrium, which implies

$$L(M_{i2}) \equiv \int_{-\infty}^{+\infty} q(s, M_{i2}, k_i) - q(s, 1 - M_{i2}, k_{-i}) dF(s) = 0$$

Since

$$\begin{aligned} L(1) &\equiv \int_{-\infty}^{+\infty} [1 - q(s, 0, k_{-i})] dF(s) > 0 \\ L(0) &\equiv \int_{-\infty}^{+\infty} [q(s, 0, k_i) - 1] dF(s) < 0 \end{aligned}$$

and

$$L'(M_{i2}) = \int_{-\infty}^{+\infty} \frac{\partial}{\partial M_{i2}} q(s, M_{i2}, k_i) + \frac{\partial}{\partial (1 - M_{i2})} q(s, 1 - M_{i2}, k_{-i}) dF(s) > 0$$

by the Intermediate Value Theorem, for every (k_i, k_{-i}) there exists a unique $\hat{M}_{i2}(k_i, k_{-i})$ such that $E_{ji2} = E_{j,-i2}$.

By the Implicit Function Theorem, there is

$$\begin{aligned} &\int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial k_i} q(s, \hat{M}_{i2}, k_i) + \frac{\partial}{\partial \hat{M}_{i2}} q(s, \hat{M}_{i2}, k_i) \frac{\partial \hat{M}_{i2}}{\partial k_i} + \frac{\partial}{\partial (1 - \hat{M}_{i2})} q(s, 1 - \hat{M}_{i2}, k_{-i}) \frac{\partial \hat{M}_{i2}}{\partial k_i} \right] dF(s) \\ = &\int_{-\infty}^{+\infty} \frac{\partial}{\partial k_i} q(s, \hat{M}_{i2}, k_i) dF(s) \\ &+ \frac{\partial \hat{M}_{i2}}{\partial k_i} \cdot \int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial M_{i2}} q(s, \hat{M}_{i2}, k_i) + \frac{\partial}{\partial (1 - \hat{M}_{i2})} q(s, 1 - \hat{M}_{i2}, k_{-i}) \right] dF(s) \\ = &\ 0 \end{aligned}$$

which implies $\frac{\partial \hat{M}_{i2}}{\partial k_i} > 0$ since $\frac{\partial}{\partial k_i} q(s, \hat{M}_{i2}, k_i) < 0$ and $\frac{\partial}{\partial M_{i2}} q(s, \hat{M}_{i2}, k_i) > 0$.

Similarly, there is

$$
\begin{aligned}
& \int_{-\infty}^{+\infty}\left[\begin{array}{c}\frac{\partial}{\partial M_{i2}}q(s,\hat{M}_{i2},k_i)\frac{\partial \hat{M}_{i2}}{\partial k_{-i}}-\frac{\partial}{\partial k_{-i}}q(s,1-\hat{M}_{i2},k_{-i}) \\ +\frac{\partial}{\partial(1-\hat{M}_{i2})}q(s,1-\hat{M}_{i2},k_{-i})\frac{\partial \hat{M}_{i2}}{\partial k_{-i}}\end{array}\right]dF(s) \\
= & \frac{\partial \hat{M}_{i2}}{\partial k_{-i}}\int_{-\infty}^{+\infty}\left[\frac{\partial}{\partial \hat{M}_{i2}}q(s,\hat{M}_{i2},k_i)+\frac{\partial}{\partial(1-\hat{M}_{i2})}q(s,1-\hat{M}_{i2},k_{-i})\right]dF(s) \\
& -\int_{-\infty}^{+\infty}\frac{\partial}{\partial k_{-i}}q(s,1-\hat{M}_{i2},k_{-i})dF(s) \\
= & \ 0
\end{aligned}
$$

which implies $\frac{\partial \hat{M}_{i2}}{\partial k_{-i}}<0$. ■

Proof of Lemma 3:

Proof. In the previous subsection we already know that in period 1 when all the other patients randomize between the providers with equal probabilities a patient will also be indifferent between choosing each provider on his own. Also we show that if all patients in period 1 equally randomize between the providers, then the type k_h provider will perform surgeries on all the coming patients and thus reveals his type in period 2. Consequently the type k_l provider will also perform surgeries on all the coming patients. To complete the proof, we only need to show that when all the other patients equally randomize between the providers, a patient will prefer this equal randomization to letting C make the referral decision, that is,

$$
\begin{aligned}
& \frac{1}{2}[\int_{-\infty}^{+\infty}q(s,\frac{1}{2},k_l)dF(s)+\int_{-\infty}^{+\infty}q(s,\frac{1}{2},k_h)dF(s)] \\
< & \int_{-\infty}^{s_{1/2}}q(s,\frac{1}{2},k_l)dF(s)+\int_{s_{1/2}}^{+\infty}q(s,\frac{1}{2},k_h)dF(s)
\end{aligned}
$$

which implies

$$
\begin{aligned}
& \int_{-\infty}^{+\infty}q(s,\frac{1}{2},k_l)dF(s)+\int_{-\infty}^{+\infty}q(s,\frac{1}{2},k_h)dF(s) \\
< & \ 2[\int_{-\infty}^{s_{1/2}}q(s,\frac{1}{2},k_l)dF(s)+\int_{s_{1/2}}^{+\infty}q(s,\frac{1}{2},k_h)dF(s)]
\end{aligned}
$$

which implies

$$
\begin{aligned}
& \int_{s_{1/2}}^{+\infty}q(s,\frac{1}{2},k_l)dF(s)+\int_{-\infty}^{s_{1/2}}q(s,\frac{1}{2},k_h)dF(s) \\
< & \int_{-\infty}^{s_{1/2}}q(s,\frac{1}{2},k_l)dF(s)+\int_{s_{1/2}}^{+\infty}q(s,\frac{1}{2},k_h)dF(s)
\end{aligned}
$$

which after rearrangement leads to the inequality condition stated in the lemma. ■

Consistent belief system supporting the separating equilibrium in Proposition 2:

Proof. To support the equilibrium strategies, the patients' off-equilibrium-path belief can be constructed accordingly. We focus on the two cases $k_h \in (k_b, \bar{k}_h)$ and $k_h \in (k_l, k_b]$, as the other two cases are straightforward. Note that we only need to be concerned with two categories of the patients' off-equilibrium-path belief, one with a provider i's report card being $(m_{i1}, d_{i1}) = (\frac{1}{2}, \bar{d}_{l1}(\frac{1}{2}))$ and the other with a provider i's report card being $(m_{i1}, d_{i1}) = (m^*_{h1}, \bar{d}_{h1}(m^*_{h1}))$. First, for $k_h \in (k_b, \bar{k}_h)$, for the type k_h provider not to deviate, the patients can hold the belief that if $(m_{i1}, d_{i1}) = (\frac{1}{2}, \bar{d}_{l1}(\frac{1}{2}))$ but $m_{-i1} > \bar{m}(k_h, k_l) - \varepsilon$ then $k_i = k_h$ and $k_{-i} = k_l$. For the type k_l provider not to deviate, the patients can hold the belief that if $(m_{i1}, d_{i1}) = (\bar{m}(k_h, k_l) - \varepsilon, \bar{d}_{h1}(\bar{m}(k_h, k_l) - \varepsilon))$ but $(m_{-i1}, d_{-i1}) \neq (\frac{1}{2}, \bar{d}_{l1}(\frac{1}{2}))$ then $k_i = k_h$ and $k_{-i} = k_l$. Second, for $k_h \in (k_l, k_b]$, for the type k_h provider not to deviate, the patients can hold the belief that if $(m_{i1}, d_{i1}) = (\frac{1}{2}, \bar{d}_{l1}(\frac{1}{2}))$ but $(m_{-i1}, d_{-i1}) \neq (1 - M^*_h(k_h, k_l), \bar{d}_{h1}(1 - M^*_h(k_h, k_l)))$ then $k_i = k_h$ and $k_{-i} = k_l$. For the type k_l provider not to deviate, the patients can hold the belief that (i) if $(m_{i1}, d_{i1}) = (1 - M^*_h(k_h, k_l), \bar{d}_{h1}(1 - M^*_h(k_h, k_l)))$ but $(m_{-i1}, d_{-i1}) \neq (m_{i1}, d_{i1})$ and $(m_{-i1}, d_{-i1}) \neq (\frac{1}{2}, \bar{d}_{l1}(\frac{1}{2}))$, then $k_i = k_h$, $k_{-i} = k_l$, and (ii) if $(m_{i1}, d_{i1}) = (m_{-i1}, d_{-i1}) = (1 - M^*_h(k_h, k_l), \bar{d}_{h1}(1 - M^*_h(k_h, k_l)))$, then $\Pr(k_i = k_h) = \frac{1}{2}$. ■

7 Appendix 2

Sample of New York CABG report card, hospital level, 1994

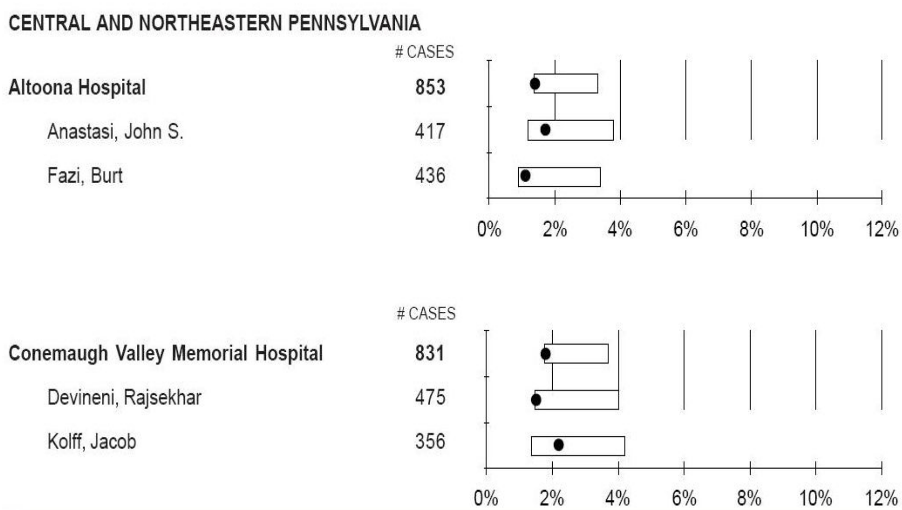
Table 2: Hospital Observed, Expected and Risk-Adjusted Mortality Rates for Coronary Artery Bypass Grafts in New York State, 1994 Discharges (Listed Alphabetically by Hospital)

Hospital	Cases	Deaths	OMR	EMR	RAMR	95% CI for RAMR
Albany Medical Center	1167	18	1.54	1.93	1.98	(1.18, 3.14)
Arnot-Ogden	236	4	1.69	2.42	1.74	(0.47, 4.45)
Bellevue	93	6	6.45	2.28	7.05 *	(2.57, 15.34)
Beth Israel	270	4	1.48	2.98	1.24	(0.33, 3.17)
Buffalo General	1173	25	2.13	1.95	2.71	(1.75, 4.00)

Sample of Pennsylvania CABG report card, hospital level and surgeon level, 1994-95

Risk-Adjusted In-hospital Mortality Outcomes

FIGURE 3B: Actual to Expected Mortality, by Cardiac Surgeon, 1994-1995



8 References

Cameron A. C., Trivedi, P. K., "Microeconometrics: Methods and Applications", ISBN 0521848059, page 768-769

Capps, C., Dranove D., Greenstein S. and Satterthwaite M., 2001. "The Silent Majority Fallacy of the Elzinga-Hogarty Criteria: A Critique and New Approach to Analyzing Hospital Mergers." NBER Working Paper 8216

Chassin, M. R.,Hannan, E. L., and DeBuono, B. A., 1996. "Benefits and hazards of reporting medical outcomes publicly". New England Journal of Medicine 334 (6):394–98.

Chen, Y., 2008, "Character Evidence and Regulation of Skilled Experts", Working Paper, University of Pennsylvania

Dranove, D., Kessler, D., McClellan M., Satterthwaite. M., 2003. "Is More Information Better? The Effects of 'Report Cards' on Health Care Providers." Journal of Political Economy 111(3):555-588.

Epstein, A. J., 2004. "Understanding incentives for responding to report cards". Ph.D. diss., University of Pennsylvania.

Epstein, A. J., 2006. "Do Cardiac Surgery Report Cards Reduce Mortality? Assessing the Evidence", Medical Care Research and Review 63(4):403-426.

Fong, K., 2007, "Evaluating Skilled Experts: Optimal Scoring Rules for Surgeons", Ph.D. Dissertation Paper, Stanford University

Fudenberg, D., Tirole, J., "Game Theory", 1991, ISBN 0262061414, page 34, 319-364

Gibbs, D.A., Sangl J. A., and Burrus B. 1996. "Consumer perspectives on information needs for health plan choice". Health Care Financ Rev, 18(1):55-74.

Hannan, E. L., C. Stone, T. L. Biddle, and B. A. DeBuono. 1997. "Public release of cardiac surgery outcomes data in New York: What do New York State cardiologists think of it?" American Heart Journal 134 (6): 1120–28.

Harlan, B. J. 2001. "Statewide reporting of coronary artery surgery results: A view from California". Journal of Thoracic and Cardiovascular Surgery 121 (3): 409–17.

Kulick D. L. (M.D.), 2005, "The New 'Coated' Stents - What's The Story?", www.medicinenet.com

Lu M., Ma C. A., Yuan L., 2003, "Risk selection and matching in performance-based contracting", Health Economics, 12: p339-354.

Mukamel, D. B., D. L. Weimer, J. Zwanziger, S. F. H. Gorthy, and A. I. Mushlin. 2004. "Quality report cards, selection of cardiac surgeons and racial disparities: A study of the publication of the New York State cardiac surgery reports." Inquiry 41 (4): 435–46.

Mukamel D.B., D.L. Weimer, J. Zwanziger, and A. I. Mushlin. 2002. "Quality of cardiac surgeons and managed care contracting practices". Health Services Research 37 (5): 1129–44.

New York State Department of Health, "Cardiovascular Disease Data and Statistics, Adult Cardiac Surgery in New York State, 1992-1994", November 1998, http://www.health.state.ny.us

Omoigui, Nowamagbe A., Dave P. Miller, Kimberly J. Brown, Kingsley Annan, Delos Cosgrove, Bruce Lytle, Floyd Loop & Eric J. Topol. 1996. "Outmigration for Coronary Bypass Surgery in An Era of Public Dissemination of Clinical Outcomes." Circulation 93:27-33.

Phelps, C., "Health Economics", 3rd Edition, ISBN 0-321-06898-X, page 370, 375

Pauly, M.V., and Satterthwaite, M.A., "The Pricing of Primary Care Physician' Services: A Test of the Role of Consumer Information," Bell Journal of Economics 1981; 12:488-506

Pennsylvania Health Care Cost Containment Council, "Pennsylvania's Guide to Coronary Artery Bypass Graft Surgery 1994-1995", May 1998, http://www.phc4.org/reports/cabg/95/default.htm

Ramanarayanan, S., 2006, "Does Practice Make Perfect? An Empirical Analysis of Learning-by-Doing in Cardiac Surgery", Job Market Paper, Northwestern University

Raffel M. W. & Barsukiewicz C. K., "The U.S. Health System, Origins and Functions", 5th Edition, ISBN 0-7668-0714-2, page 32-33, 126

Rupprecht, H., Brennecke, R., Bernhard, G., Erbel, R., Pop, T., Meyer, J., 2005, "Analysis of risk factors for restenosis after PTCA", Catheterization and Cardiovascular Diagnosis, Volume 19 Issue 3, Pages 151 - 159

Satterthwaite, M.A., 1979, "Consumer Information, Equilibrium, Industry Price, and the Number of Sellers", Bell Journal of Economics, 10(2): 483-502

Schneider, E. C., and A. M. Epstein. 1996 "Influence of cardiac-surgery performance reports on referral practices and access to care: A survey of cardiovascular specialists". New England Journal of Medicine 335 (4): 251–56.

Spence M., 1973, "Job Market Signaling", The Quarterly Journal of Economics, Vol. 87, No. 3., pp. 355-374.