NBER WORKING PAPER SERIES

DO MARKETS RESPOND TO QUALITY INFORMATION? THE CASE OF FERTILITY CLINICS

M. Kate Bundorf
Natalie Chun
Gopi Shah Goda
Daniel P. Kessler

Working Paper 13888
http://www.nber.org/papers/w13888

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2008

Do Markets Respond to Quality Information? The Case of Fertility Clinics
M. Kate Bundorf, Natalie Chun, Gopi Shah Goda, and Daniel P. Kessler
NBER Working Paper No. 13888
March 2008
JEL No. I11,I18,L15

ABSTRACT

Although policymakers have increasingly turned to provider report cards as a tool to improve health care quality, existing studies provide mixed evidence that they influence consumer choices. We examine the effects of providing consumers with quality information in the context of fertility clinics providing Assisted Reproductive Therapies (ART). We report three main findings. First, clinics with higher birthrates had larger market shares after relative to before the adoption of report cards. Second, clinics with a disproportionate share of young, relatively easy-to-treat patients had lower market shares after adoption versus before. This suggests that consumers take into account information on patient mix when evaluating clinic outcomes. Third, report cards had larger effects on consumers and clinics from states with ART insurance coverage mandates. We conclude that quality report cards have potential to influence provider behavior in this setting.

M. Kate Bundorf
Health Research and Policy
Stanford University
HRP T108
Stanford, CA 94305-5405
and NBER
bundorf@stanford.edu

Natalie Chun
Department of Economics
Stanford University
Stanford, CA 94305
natc@stanford.edu

Gopi Shah Goda
Robert Wood Johnson Scholar
in Health Policy Research Program
Harvard University
1730 Cambridge Street
Cambridge, MA 02138
gopi.shah.goda@gmail.com

Daniel P. Kessler
Graduate School of Business
Stanford University
518 Memorial Way, Room L241
Stanford, CA 94305
and NBER
fkessler@stanford.edu

I. Introduction

Health policy makers and researchers have long debated the extent to which "report cards" – public disclosure of comparative information on the performance of doctors, hospitals, and insurers – affect the allocation of consumers to health care providers. On one hand are studies that find that more highly rated hospitals have higher market shares (Mukamel and Mushlin 1998; Cutler, Huckman et al. 2004) and that report cards have an effect, albeit a small one, on health plan enrollment (Beaulieu 2002; Scanlon, Chernew et al. 2002; Wedig and Tai-Seale 2002; Dafny and Dranove 2005; Jin and Sorensen 2005; Chernew, Gowrisankaran et al. 2007). On the other hand are studies that find that report cards have a more marginal impact on markets for health services (Schneider and Epstein 1998; Mukamel, Mushlin et al. 2000; Romano and Zhou 2004). Report cards may be ineffective in this context because they are difficult to understand, they provide information that consumers already know, or they provide information that is irrelevant to consumers (Marshall, Shekelle et al. 2000).

From the perspective of public policy, determining whether report cards affect provider market share is important. Although the absence of a link between ratings and market share does not rule out an effect of report cards on quality, one of the main mechanisms by which report cards may improve quality is by reallocating demand from low- to high-ranked providers. Yet, despite this importance, estimation of the causal effects of report cards on market share remains an open empirical issue. Not only do existing studies produce mixed findings on the effect of report cards on provider market share, but, due to data limitations, virtually all of the studies in the literature rely on untestable assumptions to identify the effect of interest.

In this paper, we examine the effects of report cards on the market shares of fertility clinics. We estimate the effect of public disclosure of a clinic’s three-year lagged birth rate, a measure of quality, on clinic market share after relative to before birth rates were disclosed to consumers. Because we observe the three-year lagged birth rate for the vast majority of clinics in the U.S. both before and after public reporting occurred, we are able to identify the effect of disclosure with a transparent "difference-in-difference" strategy that requires fewer restrictive assumptions than other existing analyses. We also observe an additional, more recent measure of quality -- the one-year lagged birth rate -- that was never publicly disclosed to consumers through report cards. This allows us to control for the possibility that the effect of other sources of information for consumers changed contemporaneously with the adoption of the report card, and assess the impact of disclosure *per se*.

We also offer two novel extensions to the literature on report cards. Because clinics had to disclose information on both their birth rates and the age distribution of their patients, we are able to test whether consumers take into account patient mix when evaluating a clinic's outcomes. The birth rate of younger women undergoing fertility treatment is widely known to be higher than the birth rate of older women; for this reason, the age distribution of a clinic's patient population captures the extent to which its birth rate may be driven by underlying patient characteristics rather than the skill or effort of its workers. Finally, the nationwide scope of our data, combined with the existence of differences across states in regulations affecting insurance coverage for fertility treatment, enables us to test whether the effects of public reporting differ based on the regulatory environment.

II. Previous literature

A key issue in identifying the effects of report cards on provider market share is distinguishing the effect of information provided by the report card from that of information that consumers have from other sources. In particular, consumers are likely to have information on provider quality from a variety of sources other than the report card, but information from these other sources is generally not observable to the econometrician. Depending on the correlation between unobserved information on quality and the information disclosed by the report card, estimates of the impact of the report card may either overstate or understate its true causal effect.

Papers in the existing literature impose a variety of assumptions to address this concern. One set of papers examines the relationship between reported quality and market share after the implementation of a report card (Cutler, Huckman et al. 2004; Howard and Kaplan 2006). These studies define the effect of the report card as the correlation between market share and reported quality over time, holding constant provider characteristics and/or provider fixed effects. Although this identification strategy allows for correlation between reported and *time-invariant,* unobserved quality, it requires the assumption that *changes* in reported quality are uncorrelated with *changes* in unobserved quality information that consumers had from other sources.

A second set of studies analyzes the market shares of providers after relative to before the adoption of a report card (Mennemeyer, Morrisey et al. 1997; Beaulieu 2002; Scanlon, Chernew et al. 2002; Wedig and Tai-Seale 2002; Dafny and Dranove 2005; Chernew, Gowrisankaran et al. 2007). These studies define the effect of the report card as the difference in the change in market shares, after versus before adoption, between providers who are highly-rated and those who are poorly-rated by the report card. Like the studies in the first set, this identification strategy allows for time-invariant correlation between reported and unobserved quality. However, because none of these studies have information on quality from *before* the adoption of the report card, they all either assume that the unobserved quality of providers did not change contemporaneously with the adoption of the report card, or at the least, assume a specific functional form for how it would have changed in the absence of the report card (e.g., Dafny and Dranove (2005) and Chernew, Gowrisankaran et. al (2007)).

In a unique approach, Jin and Sorensen (2005) examine how health plan market share varies with quality under a voluntary reporting system. They observe a published quality measure for reporting plans as well as an identical (unpublished) measure for non-reporters. They define the effect of the report card as the difference between reporters and non-reporters in the correlation between quality and market share. Although this identification strategy allows for correlation between changes in reported and unobserved quality, it requires the assumption that this correlation is the same for reporters and non-reporters. Because health plans may be more likely to report when reporting leads to an increase in market share, and less likely to report when doing so would lead to a decrease in market share, this assumption may be incorrect.

In our study, we observe quality measures and market share both before and after a mandatory report card disclosed this information. Only Wedig and Tai-Seale (2002), who compare the relationship between consumer satisfaction with health plans and enrollment rates in the FEHBP, after relative to before the FEHBP widely distributed satisfaction ratings, have a similar research design. They find that the distribution of satisfaction information affects plan choices, all else held constant, particularly for new employees. Comparing the relationship between market share and quality measures before and after quality measures were disclosed allows for changes over time in provider quality that are correlated with the information provided by the report card, assuming only that the correlation between the quality measure and unobserved sources of information did not change contemporaneously with the adoption of the report card.

While we use a research design similar to that of Wedig and Tai-Seale, an unusual feature of the setting we examine allows us to relax the assumption that the relationship between quality and market share is constant over time. We not only observe reported quality both before and after its disclosure, but we also observe an additional measure of quality that was never disclosed to consumers. Because fertility clinic reporting only informed consumers about the birth rate with a three-year lag, we use the one-year lagged birthrate as a control for unobserved quality. Since there is no reason for the less current reported information to affect consumer decision-making other than the reporting mechanism, hypothesis tests of the post-pre change in the estimated impact of the 3-year lagged birth rate, holding constant the 1-year lagged birth rate, is a strong test of the link between reporting and market share.

Our study also contributes to the existing literature in a number of other ways. First, we examine the effect of reporting birth rates, an objective measure of health outcomes, on the market share of a provider. Second, the report card includes both the unadjusted clinic birthrate and the age distribution of each clinic's patients, allowing us to test whether consumers consider information on patient mix when evaluating health care providers. In addition, participation in quality reporting was mandatory and implemented nationwide, reducing concerns regarding unobserved characteristics of those who choose to participate in voluntary settings and concerns regarding differences between states that adopt and do not adopt quality reporting in analyses of state-level mandatory initiatives. Finally, differences across states in regulations affecting insurance coverage for fertility treatment enable us to test whether the effects of public reporting differ based on the regulatory environment.

III. Fertility Clinics and Quality Report Cards

We study the effects of quality report cards in the context of fertility clinics which provide Assisted Reproductive Therapies (ART) for the treatment of infertility. Over the last 30 years, the development of drugs to stimulate egg production in women and the development of procedures to promote insemination both in utero (artificial insemination) and in the laboratory have led to dramatic improvements in the treatment of infertility. Our analysis focuses on ART in which a physician surgically removes a woman's eggs, combines the eggs with sperm in a laboratory, and returns one or more developing embryos to a women's uterus. Since the first ART cycle was performed in the U.S. in 1981, utilization has grown dramatically. In 2004, 411 fertility clinics performed 127,977 cycles resulting in 36,760 live births of 49,458 infants (CDC 2005).

The Society for Assisted Reproductive Technology (SART), an organization of ART providers, began collecting data from its members on the utilization and outcomes for ART procedures in 1989. Initially, participation by clinics was voluntary, although the vast majority of ART clinics both were members of SART and reported their information. During this period, distribution of the information was limited. Consumers were able to access the reports only by requesting a hard copy directly from SART. However, The Fertility Clinic Success Rate and Certification Act of 1992, which required the Centers for Disease Control and Prevention (CDC) to publish information about success rates for all U.S. clinics, made the collection and dissemination of this information mandatory. The first year of data collection for the mandatory reporting system was 1995, and the first report was publicly released by the CDC in December of 1997.[1] In 1998, clinic-level information on ART utilization and outcomes became publicly available and easily accessible on the CDC website. Every year since then, the CDC has updated their website with a new set of three-year lagged data.

A primary concern motivating the passage of the law was that patients were uninformed about the relatively low probability of conceiving with ART. In 1990, 16% of cycles resulted in a live birth (Medical Research International and SART 1992). Concerns about the low likelihood of success with ART were compounded by the high cost of treatment. A single cycle costs approximately $10,000 (Neumann, Gharib et al. 1994; Collins 2001), and treatment is rarely covered by insurance. Because many cycles do not result in a live birth, the average cost per delivery is over $50,000 (Collins 2001). In 1987, a group of leading endocrinologists published an article in the primary clinical journal in the field questioning whether existing credentialing mechanisms of ART providers were adequate, whether providers were performing ART on clinically appropriate patients, and whether providers were truthfully informing patients about their likely success with ART (Blackwell, Carr et al. 1987).

The report card was intended to address these issues by providing patients with accurate information on their likelihood of conceiving with ART. The initial reports

---

[1] http://www.hhs.gov/news/press/1997pres/971218c.html

published clinic-level birth rates by age and type of embryo (i.e., ART using fresh embryos from non-donor eggs, frozen embryos from non-donor eggs, or embryos from donor eggs). The report also provided information on the distribution of treatment technologies used at the clinic (i.e., IVF/GIFT/ZIFT and ICSI)[2], the proportion of patients with different diagnoses, utilization by age and cycle type, the average number of embryos transferred per cycle, and rates of multiple births per transfer (see Figure 1 for an example).

The high cost of treatment and the lack of insurance coverage also generated support for laws mandating that health insurance cover fertility treatments. A number of states enacted mandates during the late 1980s and early 1990s, and Louisiana, New Jersey, and New York passed legislation as recently as 2001. By reducing the out-of-pocket price of treatment, the most comprehensive versions of these laws increased utilization of ART (Hamilton and McManus 2005; Bundorf, Henne et al. 2007).

IV. Methods

Our basic empirical strategy is to compare the relationship between clinic market share and clinic birth rates before and after their public release. The relationship between clinic market share and birth rate prior to public reporting serves as a control for the correlation between the information on the report card and information on quality consumers obtain from other sources. The difference in this relationship between the

---

[2] IVF (in vitro fertilization), GIFT (gamete intrafallopian transfer), and ZIFT (zygote intrafallopian transfer) are types of ART. The vast majority of cycles are IVF. For example, in 1999, 97% of fresh, non-donor cycles were IVF. ICSI (intracytoplasmic sperm injection) is a procedure used in combination with an ART cycle in which a single sperm is injected directly into an egg.

reporting and non-reporting periods represents the effect of public dissemination of this information on consumer choices.

A. Data

The primary source of data is information reported by clinics to SART. In 1989, SART, an affiliate of the American Society for Reproductive Medicine (ASRM), began a voluntary reporting system to collect clinic-level information about the utilization and outcomes of ART. SART produced an annual report that provided clinic-level information for the vast majority of clinics operating in the U.S. We obtained hard copies of these reports for the years 1989 to 1994 from SART. After the implementation of mandatory reporting, this information became publicly available on the CDC website (http://www.cdc.gov/art/), where we obtained data from 1995 to 2003. Using this dataset, we linked clinics over time based on their name and address. This linkage allowed us to calculate three-year lagged performance measures for all clinics that operated from 1996 to 2003.[3]

We obtained information on annual area characteristics including the number of physicians per capita and the size of the population from the Area Resource File (U.S. Department of Health and Human Services 2005).

Although 14 states had mandates in place related to the coverage of infertility treatments in 2003, the conditions of the mandates vary significantly based on the types of plans affected, the number and types of treatments covered, the cost-sharing associated with treatment, and the population to which the mandate applies. Also, firms that self-

---

[3] We dropped market share data from 1992 to 1995 due to inconsistencies in the reporting of key variables during 1989 to 1992 which prevented us from obtaining consistent measures of lagged performance for these years.

insure are exempt from mandates in all states. Because the effects of the mandates on utilization and outcomes of fertility treatments were concentrated among states adopting the most comprehensive versions of the laws (Bundorf, Henne et al. 2007), we considered only states with comprehensive versions of a mandate as having a law. We defined a "comprehensive" mandate as a requirement that insurance companies, including health maintenance organizations (HMOs), provide coverage for the cost of diagnosis and treatment of infertility, including at least three cycles of ART, with few exclusions on the population covered by the mandate. In practice, this definition differentiates between mandates that require generous coverage of ART from those that do not. Five states had comprehensive mandates in place during our study including Illinois (1991), Maryland (1985), Massachusetts (1987), New Jersey (2001), and Rhode Island (1989). We consider the mandates to be in effect in the year following their adoption.[4]

B. Empirical Models

Our basic models specify the market share of clinic i in metropolitan or micropolitan statistical area (MSA)[5] j in year t = 1996, …, 2003, S_{ijt}, as a function of information that was ultimately disclosed by the report card, Z_{ijt}; information that was never disclosed to consumers by the report card, W_{ijt}; clinic characteristics V_{ijt}; and market characteristics X_{jt}. The models include year fixed effects and allow the effects of

---

[4] Although Maryland's law limited coverage to couples who had experienced a 5-year history of infertility when it was initially adopted, it is not clear how strictly this was enforced. In addition, this restriction was relaxed in 1994.

[5] A metropolitan statistical area has at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core as measure by commuting ties, and a micropolitan statistical area is based on a similar definition but has at least one urban cluster of at least 10,000 but less than 50,000 population (U.S. Department of Health and Human Services, 2005).

Z_{ijt}, W_{ijt}, V_{ijt}, and X_{jt} to vary after the publication of the report card (t = 1998, …, 2003) relative to before (t = 1996, 1997):

$$\ln\left(S_{ijt}\right) = \alpha_t + Z_{ijt}\beta + W_{ijt}\delta + V_{ijt}\phi + X_{jt}\gamma + I(t > 1997)\cdot(Z_{ijt}\beta^{post} + W_{ijt}\delta^{post} + V_{ijt}\phi^{post} + X_{jt}\gamma^{post}) + \varepsilon_{ijt} \quad (1)$$

The vector Z_{ijt} includes two elements. The first is the three-year lagged birth rate (live births per cycle), measured relative to the competitors of clinic *i* using the within-MSA z-score:

$$Z_{ijt}^{1} = \frac{\left(birth\ rate_{i,j,t-3} - \overline{birth\ rate}_{j,t-3}\right)}{\overline{birth\ rate}_{j,t-3}}.$$

As discussed above, quality is reported to consumers with a three-year lag; ART report cards were released at the very end of the calendar year and were based on data from cycles started during the year ending two years prior to the release date. For example, the first report card was released in December of 1997 and included data from cycles started during 1995. In our empirical work, we assume this information affected clinic market share during 1998.

The second element of Z_{ijt} is the three-year lagged share of cycles received by women 40 years of age and under:[6]

$$Z_{ijt}^{2} = \frac{\left(young\ patients_{i,j,t-3} - \overline{young\ patients}_{j,t-3}\right)}{\overline{young\ patients}_{j,t-3}}.$$

The ART report cards disclose information on the share of cycles received by relatively young patients because it is a strong predictor of ART success. For example, in the 1995 report, rates of live births per cycle are approximately 25% until age 34 when they begin to steadily decline to less than 5% for women 43 and over (CDC 1997). The report card

---

[6] The reports do not present data on the distribution of cycles by patient age consistently over time. We chose the definition of this variable to generate the most consistent definition over the years. For 1995 and 1996, this represents the portion of patients 39 and under.

explicitly provides information on this dimension of patient case mix in the form of both pregnancy rates by age and a clinic-level age standardized rate. The three-year lagged birth rate in the empirical model, in contrast, represents the clinic's aggregate performance, not adjusted for patient age. Including both the raw birthrate and the proportion of relatively easy-to-treat patients allows us to independently assess the effects of disclosing patient case mix on consumer response.[7]

The vector W_{ijt} includes two elements, the one-year lagged analogues to Z_{ijt}^{1} and Z_{ijt}^{2}:

$$W_{ijt}^{1} = \frac{\left(\mathit{birth\ rate}_{i,j,t-1} - \overline{\mathit{birth\ rate}}_{j,t-1}\right)}{\overline{\mathit{birth\ rate}}_{j,t-1}}.$$

and

$$W_{ijt}^{2} = \frac{\left(\mathit{young\ \ patients}_{i,j,t-1} - \overline{\mathit{young\ \ patients}}_{j,t-1}\right)}{\overline{\mathit{young\ \ patients}}_{j,t-1}}.$$

Because the one-year lagged quality measures were never disseminated through the report card, their correlation with market share serves as a control for changes over time in the relationship between quality and market share.

The vector V_{ijt} includes six elements. The first, V_{ijt}^{1}, is an indicator of whether the clinic entered the market within the prior two years. Because three-year lagged performance is not available for these clinics, entrants are not included in the published report. The second, V_{ijt}^{2}, captures whether the clinic is an incumbent (i.e., had entered at

---

[7] While an alternative to testing the effect of reported performance on market share would have been to base the birth rate z-score on the age adjusted rate, this specification is less flexible than the one we used because it would not have allowed us to separately identify the effect of patient mix on consumer choices.

least three years ago) who declined to report.[8] We set $Z_{ijt}^{1} = Z_{ijt}^{2} = 0$ if $V_{ijt}^{1} = 1$ or $V_{ijt}^{2} = 1$. The other elements of V_{ijt} are indicator variables capturing whether the clinic is affiliated with a teaching hospital; is a member of the SART trade association; accepts single women; and accepts egg donors.[9]

Market characteristics X_{jt} include controls for the logarithm of the number of entrants, incumbents, and non-reporting incumbents in the MSA. When there are no competitors of a given type, we set the value of the logarithm to zero and include a corresponding indicator variable. We also include an indicator of whether a non-reporting clinic exists in the market. X_{jt} also includes the number of physicians per capita and the MSA population to control for differences across areas in demand for fertility treatment as well as year fixed effects to control for time trends common to all clinics.

We estimate extended models which allow the effects of the report card to vary depending on whether the clinic's state mandated that health insurance policies cover fertility treatments:

$$\begin{aligned}
\ln\left(S_{ijt}\right) = {} & \alpha_t + Z_{ijt}\beta + W_{ijt}\delta + V_{ijt}\phi + X_{jt}\gamma + \theta M_{jt} + \\
& M_{jt}\cdot(Z_{ijt}\beta^{mandate} + W_{ijt}\delta^{mandate} + V_{ijt}\phi^{mandate} + X_{jt}\gamma^{mandate}) + \\
& I(t>1997)\cdot(Z_{ijt}\beta^{post} + W_{ijt}\delta^{post} + V_{ijt}\phi^{post} + X_{jt}\gamma^{post} + \theta^{post}M_{jt}) + \\
& I(t>1997)\cdot M_{jt}\cdot(Z_{ijt}\beta^{post*mandate} + W_{ijt}\delta^{post*mandate} + V_{ijt}\phi^{post*mandate} + X_{jt}\gamma^{post*mandate}) + \varepsilon_{ijt}
\end{aligned} \tag{2}$$

where $M_{jt} = 1$ if market j in year t had a comprehensive mandate in effect.

Finally, we estimate a variant of equations (1) and (2) that specify a clinic's number of cycles, C_{ijt}, as a function of the number of cycles in the clinic's MSA, C_{jt}:

---

[8] We identify these clinics in two ways. First, the report card identifies clinics which chose not to participate in a given year, although this only applies to the years in which the report card was in effect. We also identify non-reporters based on the existence of a gap in their reporting history.

[9] For years in which clinics did not report characteristics due to circumstances such as mergers, we used clinic data from the closest available year.

$$\ln(C_{ijt}) = \alpha_t + Z_{ijt}\beta + W_{ijt}\delta + V_{ijt}\phi + X_{jt}\gamma + I(t > 1997)\cdot(Z_{ijt}\beta^{post} + W_{ijt}\delta^{post} + V_{ijt}\phi^{post} + X_{jt}\gamma^{post}) + \lambda\ln(C_{jt}) + \varepsilon_{ijt} \quad (1a)$$

and

$$\begin{aligned}\ln(S_{ijt}) = {} & \alpha_t + Z_{ijt}\beta + W_{ijt}\delta + V_{ijt}\phi + X_{jt}\gamma + \lambda\ln(C_{jt}) + \theta M_{jt} + \\ & M_{jt}\cdot(Z_{ijt}\beta^{mandate} + W_{ijt}\delta^{mandate} + V_{ijt}\phi^{mandate} + X_{jt}\gamma^{mandate}) + \\ & I(t > 1997)\cdot(Z_{ijt}\beta^{post} + W_{ijt}\delta^{post} + V_{ijt}\phi^{post} + X_{jt}\gamma^{post} + \theta^{post}M_{jt}) + \\ & I(t > 1997)\cdot M_{jt}\cdot(Z_{ijt}\beta^{post*mandate} + W_{ijt}\delta^{post*mandate} + V_{ijt}\phi^{post*mandate} + X_{jt}\gamma^{post*mandate}) + \varepsilon_{ijt}\end{aligned} \quad (2a)$$

We report standard errors that allow for clustering at the MSA level. By construction, MSA-level market share is not independent across clinics within an MSA; in addition, it is unlikely to be independent within MSAs over time. MSA-clustered errors account for both of these concerns.

We restrict the analysis to MSAs with more than one clinic. Table 1 reports descriptive statistics on the market for ART during the period we analyze. From 1996 to 2003, the number of reporting clinics grew from 300 to 398 and the number of cycles initiated grew from 49,600 to 86,753. The average number of clinics per MSA increased from 2.73 to 3.31. Although we restrict the sample to include only clinics in MSAs with at least two clinics, we include over 87% of cycles performed in both 1996 and 2003. Table 1 also demonstrates that our data include most clinics operating in the U.S. Approximately 12% of clinics did not report in 1996 and 2003. Table 2 presents descriptive statistics for the study sample, which includes only clinics in markets with at least two clinics.

V. Results

Table 3 reports selected estimates from a restricted version of equation (1) that excludes W_{ijt}, the information that was never disclosed to consumers by the report card. According to this model (Column 1), the effect of birth rate on market share before the adoption of the report card is positive but not statistically significant. The differential effect of birthrate post- versus pre-report cards, however, is positive and statistically significant, indicating that measured performance had a larger, positive effect on market share when the information was publicly disseminated to consumers.

This difference is economically important. For example, a clinic that improves its birthrate z-score from the 25th to the 75th percentile, a 0.24 increase in our sample (not in any table), would experience a 12.6% percent greater increase in its market share after public reporting versus before.[10]

We also find that information on patient mix affected clinic market share differentially before and after mandatory public reporting. Clinics that treat a disproportionate number of young patients relative to their competitors have statistically significantly lower market shares, after public reporting relative to before. We interpret this as evidence that public reporting allowed consumers to adjust for patient characteristics when evaluating clinic quality.

Unsurprisingly, entrants have significantly lower market shares than incumbents; the coefficient on W_{ijt}^{1}, the indicator variable that captures whether the clinic entered within the prior two years, is negative and strongly significant. However, we find no evidence that the negative effect of being an entrant differed after the implementation of public reporting. We interpret this as evidence that the ability to identify potential providers of ART was not a primary benefit of public reporting. In addition, the results

---

[10] The percent change is 0.126 = (exp(0.506*0.24) - 1).

provide no evidence that consumers interpreted the expected quality of entrants negatively relative to reporting incumbents due to the provision of information on reporting incumbent quality.

Results not presented in the table show that an incumbent's choice to be a non-reporter is negatively associated with its market share; the coefficient estimate on V_{ijt}^2 is -0.19, although not statistically significant (standard error 0.151). However, this effect is smaller than the effect of being an entrant (-0.908). As above, we find no evidence that the introduction of public reporting changed the effect of being a non-reporter. The variables measuring market structure have the expected effects – clinics with more competitors, either incumbents or entrants, have lower market share.

Column (2) presents results from a similarly-restricted version of equation (1a) that uses the log of the number of cycles in a clinic as the dependent variable, controlling for the number of cycles in the MSA. The results of this specification are nearly identical to those using the log of market share as the dependent variable.

Although the results in Table 3 are consistent with the hypothesis that consumers responded to the information on the quality report card, they do not preclude the possibility that the change in responsiveness of market share to quality was due to a general increase over time in consumers’ use of quality information. Our "report card" or treatment period is later in time than our control period, and consumers may have become more sensitive to ART quality as the technology changed and became more widely used, even in the absence of the report card.

In Table 4, we explore this issue by examining the effects of both one- and three-year lagged performance on clinic market share, i.e., estimating unrestricted versions of equations (1) and (1a) that include W_{ijt}. Because the one-year lagged birth rate was never disseminated through the report card, its correlation with market share should represent only the effects of non-report card sources of information on clinic quality. The one-year lagged birth rate has a positive and statistically significant effect on market share in the pre-report card period, and this effect remained roughly constant in the post-report-card period (Column (1)). This result supports the hypothesis that consumers have other sources of quality information that are correlated with measured performance that they use when choosing among clinics.

However, even after accounting for these other sources of information, the public dissemination of the report card influences market share. Controlling for the one-year lagged birthrate and its interaction with the post-report-card period indicator, the three-year lagged birth rate has a positive and statistically significant incremental effect on market share -- but only in the post-report-card period. While consumers continued to use alternative sources of information when choosing among clinics, this result is evidence that information on the report card has an independent effect.

We find no evidence of a significant effect of either one- or three-year lagged patient age in the pre-report card period. However, three-year lagged patient age has a statistically significant negative effect on market share after report cards relative to before. Even after controlling for an up-to-date measure of a clinic's patient mix, the outdated measure matters -- but only after it was publicly disclosed to consumers.

Table 5 investigates whether the effects of the report card varied between states with and without mandates requiring insurers to include comprehensive coverage of the treatment of infertility in health insurance policies. The coefficient on the triple interaction lagged birthrate*mandate*after 1997 allows us to test whether the effect of public reporting differed between states with and without comprehensive insurance mandates. Our results indicate that this is indeed the case. The effect of the three-year lagged birth rate in the post-reporting period is larger in states with mandates than in states without mandates. While we continue to find a strong negative effect of a disproportionate number of relatively easy-to-treat patients in the post-reporting period, the results do not indicate that the effect differs between clinics in states with and without mandates.

Three possible explanations exist for a stronger response to quality information in states with mandated insurance coverage. First, insurer rather than consumer choice may be the primary mechanism by which patients are steered to higher quality providers. In states with mandated insurance coverage, more consumers are insured and insurers may direct their enrollees to clinics based on the information in the report card. Second, the incremental patients who receive ART in a mandate state may be those who would have been less informed through informal channels in the absence of a report card. By extending insurance coverage to a broader population, comprehensive mandates induce those for whom the expected benefits of infertility treatment are relatively low to seek treatment (Bundorf, Henne et al. 2007). The reduction in search costs created by the report card may have a larger effect on their decision making than those who seek treatment in the absence of generous insurance coverage. Third, the incremental patients who receive ART in a mandate state may be more responsive to quality. Because they are insured, they are less likely to trade off quality against price when choosing a provider.

The other coefficients in this model are more consistent with the latter two interpretations. Notably, the three-year lagged birth rate has a positive and statistically significant pre-report card effect in states without mandates; the pre-report card effect in states with mandates is statistically significantly *smaller*, which is inconsistent with the hypothesis that insurers serve as the vehicle through which quality information affects the allocation of patients to providers. Instead, the fact that the three-year lagged birthrate has a statistically significant smaller pre-report-card effect in mandate states suggests that patients in mandate states are either less informed, or less willing to trade off quality for price, or both.

VI. Conclusion

In this paper, we estimate the responsiveness of the market share of fertility clinics to the public disclosure of information on their birth rate. We find that public reporting of quality affects clinic market share in an economically important way in the market for ART. The implementation of mandatory quality reporting caused fertility clinics reporting better outcomes than their competitors to gain market share relative to their competitors. Although the lack of evidence that existing report cards have had a dramatic effect on consumer decision-making has led some health services researchers to argue that quality report cards are of marginal importance (Schauffler and Mordavsky 2001), our results suggest otherwise.

The divergence between the findings of our study and those of others may be due to particular characteristics of this market. First, the performance measures for fertility clinics may either be more informative or easier to understand than those used in other settings. For example, in the case of health plan report cards, some have suggested that consumer response has been muted due to limitations in the way the information is presented (Hibbard, Harris-Kojetin et al. 2000; Vaiana and McGlynn 2002), a lack of understanding on the part of consumers of the meaning of the reported measures (Hibbard and Hewett 1997), and the possibility that the measures provide relatively little information on the dimensions of plan quality that are both unobservable and important to consumers (Bundorf and Baker 2007). In the case of fertility clinics, in contrast, the main performance measure (births per cycle) is easy to understand and highly relevant to patients.

Second, consumers may place less value on or have less access to sources of information other than the report card in this context than in others. In the case of report cards for cardiovascular care, for example, some have proposed that they have little effect on patient choices because patients place greater weight on physician referrals than outcomes from a report card when making decisions (Schneider and Epstein 1998). Patients pursuing ART, in contrast, are likely to have less established relationships with either referring physician or fertility clinics, increasing the relative importance of the information presented in the report card.

Third, public reporting of information was mandatory on the part of providers in this market. In the market for commercial HMOs, a setting examined by many of the existing studies of consumer response to quality information, participation by health plans in quality reporting and dissemination of the information by employers were both voluntary. Quality information is only available for a subset of potential health insurance products, and only a subset of employers disseminate this type of information to workers.

In the setting we examine, in contrast, quality information was available for the vast majority of providers and consumers were not restricted to a subset of providers when making their choice.

Finally, the consumers in this context differ from those in other studies. Patients seeking treatment for infertility are disproportionately young, wealthy, and more highly educated than the population as a whole (Bitler and Schmidt 2006). Thus, they may be more receptive and better able to understand the information available in the report card. In addition, ART is generally an elective, rather than an emergent procedure, providing patients with the opportunity to seek out information on provider quality.

We also find that the benefits of report cards are particularly large to "marginal" consumers of health services. The effects of report cards on the relationship between quality and market share are statistically significantly larger in places where ART coverage is mandated than in places where it is not. This is likely because the consumers who would not have obtained ART in the absence of the mandate were either less informed, or less price sensitive, than their inframarginal counterparts. Although we can not definitively identify the mechanism through which this effect occurs, it has important policy implications. If report cards are more important to consumers pulled in to health services markets by mandates, then current health policy reforms that seek to expand insurance coverage should also be sure to provide an appropriate source of information for the newly-insured the reforms create.

Overall, our findings indicate that report cards in this market have the potential to influence provider behavior. Because consumers responded to information on clinic quality, they created strong incentives for clinics to improve their scores. Clinics, however, may improve measured performance either by increasing the quality of their services or by selecting good prognosis patients. While our findings create concern over the possibility that quality reporting may create incentives for providers to engage in patient selection, they also demonstrate the potential for this type of difficulty to be overcome. Even very simple risk adjustment -- such as publication of raw shares of younger patients – mutes these incentives on the part of providers. Whether fertility clinics were able to improve their scores, without a corresponding reduction in market share, by selecting patients based on characteristics that were not publicly reported remains an important question. Also, because the effect of age on probability of conception is well-known in this case, the ability of patients to effectively incorporate this information into their decision-making may not be generalizable to more complex medical treatments.

Alternatively, clinics may have improved their scores by increasing the quality of care they provide to patients. While birth rates improved over the period of our study (CDC 1997; CDC 2006), it is clearly not possible, in the absence of more detailed analyses, to attribute this to quality reporting since the public reporting mandate occurred contemporaneously with improvements in ART technology as well as a dramatic increase in the size of the population pursuing treatment. The effect of report cards on quality, through their effects on market share, remains an important topic for future research.

References

Beaulieu, N. D. (2002). "Quality Information and Consumer Health Plan Choices." Journal of Health Economics 21: 43-63.

Bitler, M. and L. Schmidt (2006). "Health Disparities and Infertility: Impacts of State-Level Insurance Mandates." Fertility and Sterility 85: 858-865.

Blackwell, R. E., B. R. Carr, et al. (1987). "Are We Exploiting the Infertile Couple?" Fertility and Sterility 48(5): 735-739.

Bundorf, M. K., M. Henne, et al. (2007). Mandated Health Insurance Benefits and the Utilization and Outcomes of Infertility Treatments.

CDC (1997). 1995 Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports. Atlanta, GA, Centers for Disease Control and Prevention.

CDC (2005). Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention: 1-84.

CDC (2006). 2004 Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports. Atlanta, GA, Centers for Disease Control and Prevention.

Chernew, M., G. Gowrisankaran, et al. (2007). Learning and the Value of Information: Evidence from Health Plan Report Cards.

Collins, J. (2001). "Cost-effectiveness of in vitro fertilization." Semin Reprod Med 19(3): 279-89.

Cutler, D. M., R. S. Huckman, et al. (2004). "The Role of Information in Medical Markets: An Analysis of Publicly Reported Outcomes in Cardiac Surgery." American Economic Review 94(2): 342-346.

Dafny, L. S. and D. Dranove (2005). Do Report Cards Tell Consumers Anything They Don't Already Know? The Case of Medicare HMOs? Cambrdge, MA, National Bureau of Economic Research: 1-38.

Hamilton, B. H. and B. McManus (2005). Infertility Treatment Markets: The Effects of Competition and Policy: 42.

Hibbard, J. H., L. Harris-Kojetin, et al. (2000). "Increasing the Impact of Health Plan Report Cards by Addressing Consumers' Concerns." Health Affairs 19(5): 138-143.

Hibbard, J. H. and J. J. Hewett (1997). "Will Quality Report Cards Help Consumers?" Health Affairs 16(3): 218-228.

Howard, D. H. and B. Kaplan (2006). "Do Report Cards Influence Hospital Choice? The Case of Kidney Transplantation." Inquiry 43: 150-159.

Jin, G. Z. and A. Sorensen (2005). "Information and consumer choice: The value of publicized health plan ratings." Journal of Health Economics 25(2): 248-275.

Marshall, M. N., P. G. Shekelle, et al. (2000). "The Public Release of Performance Data: What Do We Expect to Gain? A Review of the Evidence." Journal of the American Medical Association 283(14): 1866-1874.

Medical Research International and SART (1992). "In vitro fertilization-embryo transfer (IVF-ET) in the United States: 1990 results from the IVF-ET Registry." Fertility and Sterility 57(1).

Mennemeyer, S. T., M. A. Morrisey, et al. (1997). "Death and Reputation: How Consumers Acted Upon HCFA Mortality Information." Inquiry 34(Summer): 117-128.

Mukamel, D. B. and A. I. Mushlin (1998). "Quality of Care Information Makes a Difference: An Analysis of Market Share and Price Changes after Publication of the New York State Cardiac Surgery Mortality Reports." Medical Care 36(7): 945-954.

Mukamel, D. B., A. I. Mushlin, et al. (2000). "Do Quality Report Cards Play a Role in HMOs Contracting Practices? Evidence from New York State." Health Services Research 35(1): 319-332.

Neumann, P. J., S. D. Gharib, et al. (1994). "The Cost of a Successful Delivery with In Vitro Fertilization." The New England Journal of Medicine 331(4): 239-243.

Romano, P. S. and H. Zhou (2004). "Do Well-Publicized Risk-Adjusted Outcomes Reports Affect Hospital Volume?" Medical Care 42(4): 367-377.

Scanlon, D. P., M. Chernew, et al. (2002). "The Impact of Health Plan Report Cards on Managed Care Enrollment." Journal of Health Economics 21: 19-41.

Schauffler, H. H. and J. K. Mordavsky (2001). "Consumer reports in health care: do they make a difference?" Annual Review of Public Health 22: 69-89.

Schneider, E. C. and A. M. Epstein (1998). "Use of Public Performance Reports: A Survey of Patients Undergoing Cardiac Surgery." JAMA 279(20): 1638-1642.

U.S. Department of Health and Human Services (2005). Area Resource File [CD ROM], Quality Resource Systems, Inc.

Vaiana, M. E. and E. A. McGlynn (2002). "What Cognitive Science Tells Us About the Design of Reports for Consumers." Medical Care Research and Review 59(1): 3-35.

Wedig, G. J. and M. Tai-Seale (2002). "The effect of report cards on consumer choice in the health insurance market." Journal of Health Economics 21: 1031-1048.

Figure 1: Sample clinic report from 1995

SAMPLE CLINIC

1995 PROGRAM PROFILE

1 Program Characteristics

SART member	Yes
Single women	Yes
Surrogates	Yes
Donor eggs shared	10%

2 Type of ART Used[a]

IVF	97%
GIFT	3%
ZIFT	0%
with ICSI	24%

3 ART Patient Diagnosis[a]

Tubal factor	23%
Endometriosis	18%
Uterine factor	2%
Male factor	32%
Other factors	16%
Unexplained	9%

1995 ART PREGNANCY SUCCESS RATES

	4 Age of Woman			**5 Age-Standardized Rate**[b]
	<35	**35-39**	**>39**	**6 (95% Confidence Interval)**
7 Cycles Using Fresh Embryos From Nondonor Eggs				
Number of cycles	194	230	187	
Pregnancies per cycle (%)	32.5	22.2	10.7	24.7 (21.1 - 28.4)
Live births per cycle[c] (%)	27.3	14.8	7.0	19.1 (15.7 - 22.5)
Live births per retrieval[c] (%)	29.9	18.5	8.8	21.9 (18.1 - 25.7)
Live births per transfer[c] (%)	31.6	20.5	10.0	23.6 (19.5 - 27.6)
Cancellations (%)	6.7	17.4	16.0	
Avg. number embryos transferred	4.3	4.5	4.0	
Multiple birth rate per transfer				
Twins	13.1	6.6	0.8	
Triplets or greater	0.6	1.2	0.0	
8 Cycles Using Frozen Embryos From Nondonor Eggs				
Number of transfers	22	25	11	
Live births per transfer[c] (%)	22.7	28.0	0.0	
Avg. number embryos transferred	4.4	3.4	3.7	
9 Cycles Using Donor Eggs				
Number of fresh transfers	**10** 5	11	53	
Live births per transfer[c] (%)	3/5	2/11	30.2	
Avg. number embryos transferred	7.0	4.7	4.7	

[a] Includes only cycles using fresh embryos from nondonor eggs.

[b] No data given if there were too few cycles to permit age-standardized calculations.

[c] Pregnancies resulting in one or more children born alive; therefore, multiple births are counted as one.

Table 1: Growth in the Market for Assisted Reproductive Therapies, 1996-2003

	U.S.		MSAs with Two or More Clinics	
Variable	**1996**	**2003**	**1996**	**2003**
Total Number of Reporting Clinics	300	398	242	337
Total Number of Non-Reporting Clinics	28	45	20	35
Identified in Report Card	14	35	8	28
Not Identified in Report Card	14	10	12	7
Total Cycles Initiated	49,600	86,753	43,476	76,886
Number of MSAs	120	164	60	67
Average Number of Clinics per MSA	2.73	3.31	4.37	5.55

Table 2: Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Market share	0.214	0.236	0.000	0.974
Birth rate	0.252	0.098	0.000	0.800
% patients under age 40	0.903	0.072	0.500	1.000
Entrant	0.210	0.407	0.000	1.000
Non-reporter	0.058	0.234	0.000	1.000
Physicians per capita in MSA (000s)	3.316	0.104	1.500	11.268
MSA population (000s)	1,038.879	1,610.441	51.323	6,414.636
Number of ART clinics in MSA	12.089	11.929	2.000	45.000
Mandate state	0.132	0.338	0.000	1.000
Teaching hospital	0.182	0.386	0.000	1.000
Member of SART	0.948	0.222	0.000	1.000
Accepts single women	0.833	0.373	0.000	1.000
Accepts gestational carriers	0.542	0.498	0.000	1.000
Accepts egg donors	0.587	0.492	0.000	1.000

Notes: Number of clinic-years=2,428; Number of unique clinics=424. Sample includes ART providers in MSAs with at least two clinics during the period 1996 to 2003.

Table 3: Effect of Birth Rate, Patient Age, and Other Characteristics, Before Versus After Report Cards Were Introduced in 1998

Independent variables	(1) ln(Share of cycles in MSA)		(2) ln(# cycles)	
3-year lagged birth rate Z-score*After 1997	0.506	**	0.508	**
	(0.236)		(0.235)	
3-year lagged birthrate Z-score	0.150		0.146	
	(0.149)		(0.148)	
3-year lagged patient age Z-score*After 1997	-2.028	***	-2.029	***
	(0.760)		(0.766)	
3-year lagged patient age Z-score	0.908		0.926	
	(0.624)		(0.633)	
Entrant*After 1997	0.045		0.038	
	(0.132)		(0.132)	
Entrant	-0.826	***	-0.810	***
	(0.089)		(0.085)	
ln(# incumbents)	-0.962	***	-0.745	***
	(0.062)		(0.068)	
ln(#entrants)	-0.229	***	-0.152	***
	(0.050)		(0.055)	
No incumbents	-0.012		-0.092	
	(0.121)		(0.104)	
No entrants	0.326	***	0.238	***
	(0.076)		(0.066)	
ln(# cycles in MSA)			0.787	***
			(0.052)	
Number of observations	2,428		2,428	
R-squared	0.64		0.36	

*Note: ***,**, and * denote significance at the 1%, 5%, and 10% levels, respectively. Standard errors calculated allowing for correlation within MSA over time. All models control for additional clinic characteristics including whether the clinic is affiliated with a teaching hospital, SART membership, whether the clinic accepts single women, whether the clinic accepts egg donors, whether the clinic accepts gestational carriers, whether a clinic is a non-reporting incumbent and the interaction of these characteristics with an (After 1997) indicator. All models also control for additional market characteristics including the ln(number of non-reporting incumbents), an indicator of no non-reporting incumbents, an indicator of whether the MSA includes a non-reporting clinic, ln(number of physicians in the MSA), ln(MSA population) and the interaction of each of these with an (After 1997) indicator. Z-scores are defined as (clinic mean-MSA mean)/MSA mean. Patient age Z-score is the clinic's proportion of patients aged 40 or younger (aged 39 or younger for 1993-1996) relative to the MSA average proportion.*

Table 4: Effect of 3-Year versus 1-Year Lagged Birthrate and Patient Age, Before versus After Report Cards Were Introduced in 1998

Independent variables	(1) ln(Share of cycles in MSA)		(2) ln(# cycles)	
3-year lagged birth rate Z-score*after 1997	0.465	**	0.465	**
	(0.200)		(0.196)	
3-year lagged birth rate Z-score	-0.025		-0.027	
	(0.154)		(0.154)	
1-year lagged birth rate*After 1997	-0.086		-0.083	
	(0.181)		(0.180)	
1-year lagged birth rate	0.648	***	0.643	***
	(0.179)		(0.175)	
3-year lagged patient age*After 1997	-1.595	**	-1.607	**
	(0.639)		(0.628)	
3-year lagged patient age	0.813		0.837	
	(0.531)		(0.523)	
1-year lagged patient age*After 1997	-0.907		-0.881	
	(0.755)		(0.754)	
1-year lagged patient age	-0.497		-0.507	
	(0.641)		(0.636)	
Entrant*After 1997	0.078		0.071	
	(0.141)		(0.139)	
Entrant	-0.840	***	-0.823	***
	(0.121)		(0.119)	
ln(# incumbents)	-0.961	***	-0.747	***
	(0.081)		(0.105)	
ln(# entrants)	-0.229	***	-0.154	
	(0.086)		(0.097)	
No incumbents	-0.021		-0.101	
	(0.603)		(0.602)	
No entrants	0.311	***	0.224	*
	(0.118)		(0.118)	
ln(# cycles in MSA)			0.790	***
			(0.091)	
Number of observations	2,428		2,428	
R-squared	0.66		0.39	

Note: see table 2.

Table 5: Effect of Birth Rate and Patient Age, Before versus After Report Cards Were Introduced in 1998, States with Mandates versus States Without

Independent variables	(1) ln(Share of cycles in MSA)		(2) ln(# cycles)	
3-year lagged birth rate*Mandate state*After 1997	1.291	*	1.284	*
	(0.675)		(0.679)	
3-year lagged birth rate*Mandate state	-0.541	**	-0.532	**
	(0.266)		(0.265)	
3-year lagged birth rate*After 1997	0.224		0.232	
	(0.184)		(0.183)	
3-year lagged birth rate	0.299	*	0.291	*
	(0.157)		(0.156)	
3-year lagged age*Mandate state*After 1997	-3.102		-3.114	
	(2.553)		(2.557)	
3-year lagged age*Mandate state	0.453		0.455	
	(0.945)		(0.947)	
3-year lagged age*After 1997	-1.349	***	-1.344	***
	(0.424)		(0.427)	
3-year lagged age	0.626		0.629	
	(0.632)		(0.635)	
Mandate state*After 1997	7.149		16.560	
	(63.298)		(66.505)	
Mandate state	-0.406		-10.230	
	(64.394)		(67.645)	
ln(# cycles in MSA)			0.820	***
			(0.043)	
Number of observations	2,428		2,428	
R-squared	0.65		0.37	

*Note: see table 2. Models in this table also control for entrant (open <3 years), after 1997*entrant, ln(# incumbents in MSA), ln(# entrants in MSA), no incumbents in MSA indicator, and no entrants in MSA indicator. In addition, the models include the interaction of all clinic and market characteristics with both the mandate indicator and the mandate and (After 1997) interaction.*