Physician Learning and Best Practice Adoption: An Application To Cesarean Sections

November 2003

Sean Nicholson
Andrew Epstein

University of Pennsylvania
The Wharton School

Abstract

Small-area-variation studies have demonstrated that people receive a substantially different amount of medical care depending on where they live, controlling for differences in prices, income, and health. We examine why physicians have such divergent views about the appropriateness and efficacy of medical treatments, and whether physicians' treatment styles change over time. Using a data set that contains the universe of deliveries in Florida over a 9-year period and consistent physician identifiers, we find that obstetricians have distinct and widely varying treatment styles. The variation in risk-adjusted c-section rates across physicians within a region is five times greater than the variation between regions. Surprisingly, residency programs explain only two percent of the variation between physicians in their risk-adjusted c-section rates. Although treatment styles are quite stable over time, we find evidence that physicians change their treatment style according to how their peers are treating patients and, to a lesser extent, according to their patients' health outcomes.

This paper has benefited from comments by Steven Lehrer, George Macones, Jim Burgess, Will Manning, and members of the Wharton Applied Economics workshop, the University of Pennsylvania Population Studies Center Colloquium, the Brigham Young University economics workshop, the 14th annual health economics conference at RAND, and the NBER Summer Institute. We thank the University of Pennsylvania Research Foundation and the Leonard Davis Institute for Health Economics for providing funding to acquire the data.

There is ample evidence that physicians have divergent views about the efficacy and appropriateness of medical treatments. In a recent study that is representative of the "quality of care" literature, McGlynn *et al.* (2003) scrutinized the medical records of 6,700 patients and concluded that only 55 percent received the medical care recommended by 439 different clinical guidelines covering 30 diseases. For example, only 45 percent of patients were prescribed a beta-blocker following a heart attack even though published studies demonstrate that these drugs reduce the probability of dying by 13 percent in the first week of treatment.

There is a separate and equally extensive literature demonstrating that people in the United States receive a substantially different amount and type of medical care depending on where they live (e.g., Wennberg and Gittelsohn, 1973; Wennberg, Freeman, and Culp, 1987; Phelps and Parente, 1990). In these small-area-variation studies, analysts calculate the use rate of a particular treatment (e.g., back surgeries per capita) in a county or state. The amount of practice variation is characterized by the coefficient of variation in use rates across markets at a point in time.[1] If there exists a single best method for treating patients with a particular condition, there will be welfare losses when regions and physicians diverge from the medically appropriate benchmark.[2] Phelps and Parente (1990) estimate there is a $33 billion annual welfare loss (in 1987 dollars) due to variations in hospital use rates. As the authors point out, this will be an underestimate of the true welfare loss if the mean observed use rate across regions is not the optimal rate or if there is also variation in the use rate *within* a region (e.g., variation across physicians in their likelihood of admitting patients to a hospital).

In this paper we explore two related questions raised by the quality of care and small-area-variation studies. Why does the quantity and type of care differ so substantially between

---

[1] The coefficient of variation is the standard deviation divided by the mean.

[2] Use rates could vary across regions due to differences in prices, income, patients' health status, patients' preferences, or physicians' ability or willingness to induce demand for their services. Phelps and Mooney (1993) conclude that these factors collectively explain very little of the differences in the amount of

regions, and what types of interventions and policies are likely to change how physicians treat patients? The more fundamental question we address is how physicians learn: how does a physician form his treatment style, how much does that style change over time, and why does it change? We focus on deliveries, and define an obstetrician's treatment style as the proportion of deliveries he performs by cesarean section (c-section), controlling for patient characteristics such as health, the status of the pregnancy, and the patient's type of health insurance. We examine whether obstetricians have distinct treatment styles, whether residency programs influence those treatment styles, and whether and to what extent physicians change their treatment styles over time based on their own patients' health outcomes and the treatment style of their peer group. Our findings can help health insurers and the government design policies to promote effective medical practices, and can help medical device, biotech, and pharmaceutical firms predict whether and when physicians will adopt their devices and drugs.

There are a number of theoretical models in the economics literature that show how physicians in different markets might rationally adopt divergent views regarding the efficacy and appropriateness of medical care. However, there has been very little empirical work testing these predictions. Banerjee (1992), for example, describes a model where individuals share a common prior probability that a product is preferred, and each person receives a private signal regarding the product's quality. If the second person in a market observes the first person's choice and regards all signals to be of equal quality, he may infer that the first person received a positive signal regarding the quality of the chosen product, discard his own negative signal, and buy the same product. The third person in the market will then choose the same product regardless of his private signal.

In Banerjee's (1992) model each person exerts a negative externality on the rest of the population when he discards his private signal. The resulting welfare loss in health care could be

---

medical care received. Several other authors concur with this assessment, including Chasin (1987), Bikhchandani *et al.* (2001), and Newhouse (2002).

particularly large if new physicians discard information that is more technically advanced than that possessed by experienced physicians (e.g., teaching hospitals train residents with a technology that has not diffused widely to practicing physicians). Scharfstein and Stein (1990), Ellison and Fudenberg (1993, 1995), and Bikhchandani, Hirschleifer, and Welch (1992) have also presented general models where individuals observe the decisions of their peers, update their priors, and decide rationally to herd on the choices of their peers.

Phelps and Mooney (1993) propose a similar theoretical model to explain the variation in medical use rates. They argue that physicians form beliefs about the appropriateness and effectiveness of medical technologies during medical school and residency training. When a physician begins practicing medicine, she modifies her prior belief about the appropriateness of a technology or treatment method after observing her colleagues' treatment decisions. Physicians may imitate their colleagues if their colleagues' treatment decisions and outcomes provide useful and inexpensive information. Phelps and Mooney theorize that physicians will converge over time to a community standard, but community standards can differ because residency programs have unique medical styles and physicians may locate differently across markets. According to the Banerjee (1992) and Phelps and Mooney (1993) models, in the long run there would be inter-regional variation in practice patterns but little intra-regional variation across physicians. In our sample of Florida obstetricians, however, we find that the variation in risk-adjusted c-section rates across physicians within a region is five times greater than the inter-regional variation.

Fournier, Prasad, and Burke (2002) develop a theoretical model where physicians may herd on a particular treatment method due to positive network externalities: the expected outcome for a patient improves with the number of patients in that city who are treated with that particular method. As with the above models, we would expect little intra-regional variation in use rates physicians within a market should treat patients in a similar fashion. They test their model using data on Florida cardiologists' treatment decisions of heart attack patients and find that cardiologists are more likely to perform or recommend that patients receive a coronary angiography, bypass surgery, and angioplasty if their peer cardiologists are also likely to make such treatment decisions. This study uses a single cross-section, so it is difficult to interpret whether the positive peer effects are due to learning or unobserved patient characteristics that are correlated for a physician and his peer group. Our panel data set, on the other hand, allows us to control for unobserved characteristics at the level of the physician practice that are time invariant.

Bikhchandani, Chandra, Goldman, and Welch (2001) assert that when a treatment decision is discrete (e.g., either perform surgery or treat a patient medically) rather than continuous (e.g., select the dosage of a medication), there are fewer opportunities for physicians to learn by experimenting. In the former settings, physicians may be more likely to discard their private information regarding the efficacy of a treatment and herd on a single treatment method. Thus, learning may cease altogether. They test this hypothesis and show that there is indeed more regional variation in the proportion of heart attack patients who receive bypass surgery or angioplasty (discrete decisions) than the proportion of hypertensive patients who are administered specific drugs (examples of continuous, or dosage, treatment decisions that allow physicians to experiment). Although the evidence is persuasive, the sample size in this paper is essentially the number of diseases evaluated (two), so the findings may not be representative of all discrete and continuous treatment decisions.

Our approach is innovative in a number of ways. We choose the physician as the unit of observation rather than a county or a state. From a patient's perspective, it is irrelevant if physicians in her region provide the appropriate rate of treatment, on average; all that matters is whether the physician she chooses provides the appropriate treatment. Our data set contains the universe of hospital admissions in Florida over a 9-year period and consistent physician identifiers. Panel data allow us to explore why treatment styles change over time, rather than merely documenting that styles differ at a point in time. The data allow us to examine physicians' treatment decisions and their patients' health outcomes over a long duration, to characterize a physician's peer group, and to observe the treatment decisions of a physician's peer group. Finally, we have detailed information on the health status of a physician's patients, which allows us to control for a potentially important source of differences in physician treatment styles.

We find that the variation in c-section rates across physicians within a region is five times larger than the between-region variation, controlling for observed patient characteristics. This implies that existing estimates of the welfare loss stemming from practice variation are too low, and private health insurers and the government may be underestimating the value of standardizing treatment methods. Although treatment styles are associated with where physicians trained, residency programs explain only two percent of the variation between physicians in c-section rates. Almost 40 percent of the variation in risk-adjusted c-section rates across physicians and years is due to time-invariant, physician-specific factors other than experience, gender, race, and where a physician received residency training. Since we have detailed information on the characteristics of a physician's patients, our interpretation is that a considerable amount of practice variation is due to idiosyncratic physician styles.

Physicians' treatment styles are quite stable; a one percent increase in a physician's c-section rate in year t-1 is associated with a 0.1 percent increase in their rate in year t. We do, however, find evidence that physicians respond to the practice style of their peer physicians; a one-standard deviation increase in the lagged c-section rate of a physician's peer group is predicted to increase his c-section rate by 2.3 percentage points (or a 12.4 percent increase in the sample average c-section rate) the following year. We find some evidence that physicians alter their practice style according to their patients' health outcomes, but the magnitude of this effect is small.

In the next section we present some descriptive data on cesarean section rates in Florida between 1992 and 2000, by region and physician. We describe the data in Section II and present our empirical methodology in Section III. Section IV contains our estimation results and we conclude in Section V.

I. Physician Learning and Cesarean Sections

In 1998 over 900,000 cesarean sections were performed in the United States, making it the second most common surgical procedure. The proportion of deliveries performed by c-section (the c-section rate) increased markedly during the 1970s and 1980s before stabilizing in the last decade, as displayed in Figure 1. Only 5.5 percent of babies were delivered via c-section in 1970, versus 22.0 percent in 1999. The primary c-section rate, which excludes women who have had a prior c-section and thus are more likely to have a c-section on subsequent deliveries, grew at roughly the same rate as the overall rate. In Figure 1 we also display the c-section rate in Florida during the 1990s because we will examine obstetricians' decisions in Florida in the empirical section. The overall c-section rate in Florida tracks the national average.

There is a perception among health insurers, public health organizations, and obstetrician associations that too many c-sections are performed. Women who received a c-section in Florida between 1992 and 2000 remained in the hospital 3.5 days, on average, versus 1.9 days for women who have vaginal deliveries.[3] The average hospital charge for a c-section in Florida in the 1990s was $8,500, almost twice as high as the charge for a vaginal delivery, while the average physician charge for a c-section is about $500 higher than for a vaginal delivery (Gruber, Kim, and Mayzlin, 1999). Furthermore, a number of medical studies conclude that the health outcomes for children and mothers are better with vaginal deliveries than c-sections. Since there apparently are cost and health advantages of vaginal deliveries, the Public Health Service established a goal in 1990 to try to reduce the national c-section rate to 15 percent. As can be seen from Figure 1, this goal has not been achieved.

Cesarean sections are not a new technology, so one might expect information regarding the medically appropriate use of this treatment to have diffused widely, resulting in near uniformity of the c-section rate across regions. However, as with most medical treatments, there is considerable regional variation in the proportion of deliveries performed by c-section. In Figure 2 we plot the mean c-section rate in the 11 Florida health districts between 1992 and 2000, adjusted for patient health characteristics, health insurance status, and day of delivery.[4] In 2000, the risk-adjusted c-section rate ranged from 22.2 percent in the Jacksonville metropolitan area to 27.6 in the Miami area. The coefficient of variation in the risk-adjusted c-section rate across these regions is 0.06, which is actually relatively low according to small-area-variation studies of other types of medical treatments.[5]

Most of the small-area-variation studies focus on variations in regional (e.g., county or state) average use rates based on aggregate hospital data, as opposed to data on physician-specific rates. Phelps and Parente (1990) acknowledge that if the use rates vary across physicians within a region, their estimates of welfare loss due to medical practice variations will be too low. That is, even if the mean use rate of a region conforms to best medical practices, there will still be patients receiving too much or too little of the treatment if there is variation across physicians within the region.

This appears to be the case with deliveries. In Panel A of Figure 3, we depict the distribution of c-section rates of Florida physicians for the 1992 to 2000 time period. Each observation is a physician-year, and here the c-section rates are not adjusted for the patients' risk of needing a c-section, as they will be in most of our analysis. The mean unadjusted c-section rate across all physicians is 0.247, with a standard deviation of 0.096, and a coefficient of variation of 0.39. The large coefficient of variation indicates that there are considerable differences across physicians in their likelihood of performing a c-section.

---

[3] Based on data from Florida hospitals between 1992-1999.

[4] The risk adjustment method is described more fully in Section III. For Figure 2, we use ordinary least squares and regress whether or not a woman received a c-section on her age, age squared, race, type of health insurance, day of admission, indicator variables for a woman's health and the status of her pregnancy (e.g., a malpositioned fetus, or severe hypertension), and a full set of regional indicator variables. No constant is included.

[5] The inter-regional coefficient of variation in Florida is 0.11 if we do not adjust for patient characteristics.

Some of this variation is probably due to differences between physicians in the health of their patients and the status of the women's pregnancies. In panel B and panel C of Figure 3 we present the distribution of physicians' c-section rates for women whose fetus is malpositioned (e.g., breech, or feet first) and for women who have had a prior c-section, respectively. The mean c-section rates for these patients are much higher (0.79 for malpositioned fetus and 0.69 for women who have had a prior c-section), but there is still considerable variation across physicians. About two percent of physicians never perform a c-section on women who have had a prior c-section, whereas six percent of physicians always perform a c-section in such a situation. The coefficient of variation is 0.25 and 0.28 in panel B and panel C, respectively. The main purpose of this paper is to understand why physicians practice differently, whether treatment styles are stable, and what causes styles to change over time.

II. Data

We construct our sample from the Florida discharge data sets for the 1992 to 2000 time period. The data sets contain information on 1.7 million deliveries that occurred at non-federal, short-term acute care hospitals in Florida. We observe the mother's demographic information (age, race, ethnicity), her insurance coverage (e.g., HMO), codes for the primary diagnosis and secondary diagnoses, procedure codes that allow us to determine whether the baby was delivered vaginally or via c-section, a unique and consistent (across hospitals and years) physician identifier, a unique and consistent hospital identifier, and the quarter and year the patient was admitted. Sample means and standard deviations for the patient-level data set are reported in Table 1. The diagnoses codes allow us to control for objective health conditions that affect the probability a physician will perform a c-section (e.g., whether a woman has had a c-section prior to this delivery, whether the fetus was malpositioned during the delivery such as in the breech position, or whether the labor occurred before the fetus was full-term). Although none of these health conditions is common, women are much more likely to receive a c-section in these situations.[6]

Since the physician identifiers are consistent and the data include all hospital discharges, we are able to examine a physician's entire inpatient practice over time. We link the physician license numbers to data from the American Medical Association's (AMA) Masterfile to collect information on each physician's gender, race, the residency program(s) where he received training, and the year he completed residency training. We create a variable for years of post-residency experience, and sometimes include an indicator variable for physicians with fewer than four years of experience. We also have information on race for a subset of physicians from the Florida State Medical Board.

We define physician j's peer group in year t as all physicians other than physician j who delivered a baby at the same hospital or hospitals as physician j in year t. Forty-nine percent of the physicians in our data set delivered all their babies at a single hospital, 35 percent divided their deliveries between two hospitals, and 16 percent at three or more hospitals. A physician's peer group consists of an average of 51 physicians, so most obstetricians can potentially interact with and learn from a large number of their peers. We include all physicians when constructing peer group averages but omit from the regression analysis any physician who delivered fewer than 10 babies in a year.

We aggregate the patients for each physician in each year and present the means and standard deviations of the physician-level data set in Table 2. The physician-level data set contains 1,928 physicians representing 10,528 physician-years. Twenty-one percent of the physicians are women and the mean age is 44 years. Physicians in the sample delivered an average of 139 babies per year, whereas the mean quantity of deliveries for a physician's entire peer group was about 4,000. Twenty percent of the deliveries resulted in an adverse patient

---

[6] The other two diagnoses frequently associated with a c-section are fetal distress and abnormal labor. Since these assessments are fairly subjective, we do not include these as control variables.

outcome, defined as fetal-maternal hemorrhage, intrauterine death of the fetus, fetal and placental problems, abnormality in fetal heart rate or rhythm, failed forceps or vacuum extractor, prolapse of the umbilical cord, rupture of the uterus, 4th degree perineal tear, or postpartum hemorrhage.[7] One hypothesis is that physicians who experience a relatively high adverse outcome rate will subsequently perform more c-sections in an effort to improve patient outcomes. In the next section we describe the method of computing each physician's c-section rate, adjusted for patient characteristics that affect the likelihood of needing such care, as well as the adjusted c-section rate of a physician's peer group.

III. Methodology

We want our measure of a physician's treatment style to pertain to the same type of patient for all physicians. That is, ideally we would observe how each physician in Florida treated the same large sample of patients. Absent this research design, we want to control for differences in treatment decisions that are due to differences in patient characteristics, rather than physician perceptions of the appropriateness and efficacy of a c-section. The existing literature on cesarean sections indicates that patients' health conditions and demographic characteristics affect the likelihood they will receive a c-section (e.g., Burns, Geller, and Wholey, 1995). For example, women who are relatively old, have health insurance, have had a prior c-section, and whose fetus is malpositioned are more likely to receive a c-section. Since patient characteristics can differ substantially across physician practices and, to a lesser extent, within practices over time, we focus our analysis on physicians' risk-adjusted c-section rates – the proportion of a physician's deliveries that are performed by c-section, controlling for observed patient characteristics.

---

[7] We received assistance from Dr. George Macones, the Director of Fetal and Maternal Medicine at the University of Pennsylvania, with defining adverse patient outcomes

To derive the risk-adjusted c-section rates we estimate the following linear probability model, separately for each year between 1992 and 2000:

$$(2) \qquad C_{ij} = \boldsymbol{\alpha}_1\mathbf{X}_i + \mathbf{Y}_j\mathbf{J} + \varepsilon_{ij}$$

C_{ij} equals one if patient i received a c-section by physician j and is zero otherwise. We include patient characteristics **X**, such as the patient's age, type of health insurance, medical conditions (e.g., whether the woman had a c-section in a prior delivery, multiple gestation, premature labor, antepartum bleeding), and day of admission, to control for factors that might affect the likelihood of receiving a c-section. We also include a full set of physician indicator variables **J** and do not include a constant. The coefficients on the physician indicators (**Y**) measure the probability that a particular physician will perform a c-section, controlling for patient characteristics. One of the attractive features of using an ordinary least squares regression is that the distribution of the **Y** coefficients is independent of **X**, the patient characteristics. If we instead estimated equation (2) with a probit or a logit model, we could get biased coefficients (Chamberlain, 1980).

The estimated coefficient for each physician j, $\hat{Y}_j$, will incorporate both differences between physicians in unobserved patient characteristics and differences in physicians' perceptions of the appropriateness of c-sections. We are interested only in the latter component and will address below how we attempt to isolate this component. Women with strong preferences for having a c-section and women who know they are at risk of needing a c-section (e.g., a woman whose fetus is breech) may select physicians who are more willing to perform c-sections and/or are skilled at performing them. In general, most obstetricians are members of a group practice, and it is common for a single obstetrician in a group to perform all of the practice's deliveries in a given 24-hour period. Many women, therefore, choose their preferred obstetrics practice but not the actual physician who delivers her child. In the subsequent analysis, therefore, we assume that the component of Y_j that is due to unobserved patient characteristics, including patient preferences, is time invariant and will drop out of a fixed effect and a first-difference estimation method.

Phelps and Mooney (1993) propose that a physician's c-section rate for a particular type of patient will be a function of his prior belief regarding the appropriateness of a c-section, and the observed c-section rate of his peer group. We interpret a physician's adjusted c-section rate in period t-1, $Y_{j,t-1}$, as the physician's prior regarding the appropriateness of a c-section. To derive the information physician j receives from his colleagues, we calculate the weighted average adjusted c-section rate of each member of physician j's peer group in period t-1, where the weight is the proportion of the total quantity of deliveries performed by the group (other than physician j) accounted for by each physician (other than physician j).

Using ordinary least squares, we could regress physician j's risk-adjusted c-section rate in year t on his lagged rate, his peer group's lagged rate, and the proportion of deliveries performed by physician j in year t-1 that resulted in an adverse patient outcome (A_{t-1}):

$$(3) \qquad Y_{jt} = \exists_0 + \beta_1 Y_{j,t-1} + \gamma Y^{pg}{}_{j,t-1} + \vartheta A_{j,t-1} + \eta_{jt}$$

In our context, Y_{jt} is the c-section rate for physician j controlling for observed patient characteristics. If a physician's risk-adjusted c-section rate increases in year t-1, β_1 indicates the proportion of the increase that persists in year t. A value for β_1 that is close to zero would indicate that physicians' treatment patterns are persistent – little of the change is permanent. A positive value for γ would indicate that an increase in the risk-adjusted c-section rate of a physician's peer group in t-1 is associated with an increase in that physician's c-section rate in year t. Presumably this relationship exists because the physician is acquiring information from his peer group regarding the best way to treat patients. We include the adverse outcome variable, A, in equation (3) to see if physicians adjust their treatment decisions based on how their patients fared in the past. Our hypothesis is that if a physician's patients are experiencing poor health outcomes, he will respond by performing more c-sections since this method is believed to improve health outcomes for high-risk patients. All three coefficients (β_1, γ, and ϑ) can capture physician learning.

There are three potential biases in the specification of equation (3). We assume that the η_{jt} follow a one-way error component model:

(4) $\eta_{jt} = \mu_j + v_{j,t}$

where the μ_j are independent and identically distributed and the $v_{j,t}$ are independent of each other and among themselves (no serial correlation). μ_j is a time-invariant physician fixed effect based on the unobserved health and preferences of a physician's patients. The lagged dependent variable $Y_{j,t-1}$ that is included as a regressor in equation (3) is correlated with μ_j , and therefore $Y_{j,t-1}$ will also be correlated with η_{jt}, the error term. As a result, the ordinary least squares (OLS) estimate of β_1 will be biased upward.[8] If our data on patient characteristics capture most of the true differences in patients' health and if patients do not choose physicians to a great extent based on their preferences, μ_j will be close to zero and the OLS bias will be small.

A second potential bias in equation (3) may occur if physicians choose markets based, in part, on whether the physicians in that market practice a similar style of medicine. If new physicians systematically locate among established physicians who share their views, the OLS estimate of γ will be biased upward because $Y^{pg}_{j,t-1}$ will be positively correlated with $Y_{j,t-1}$, and therefore correlated with η_{jt}. Manski (1993) highlights the difficulty of separating the influence of the peer group on a person's behavior from the influence of unobserved characteristics shared by a person and his peer group. In this paper we assume that physicians are randomly assigned to peer groups and thus that $Y^{pg}_{j,t-1}$ is uncorrelated with μ_j.

The third potential bias in the OLS specification is that γ may be biased upward if a common shock in t-1, such as increased malpractice pressure, causes a physician's c-section rate and his peer group's rate to move together for reasons other than information the physician receives from his peer group.

---

[8] This part of the paper follows the discussion in Baltagi (2001) regarding the econometric problems associated the dynamic panel estimators.

One way to address the potential biases of the OLS specification is to estimate a fixed effects model that eliminates μ_j:

$$(5)\qquad Y_{j,t} - \bar{y}_{j} = \beta_0 + \beta_1(Y_{j,t-1} - \bar{y}_{j.-1}) + \gamma(Y^{pg}_{j,t-1} - \bar{y}^{pg}_{j.-1}) + \vartheta(A_{j,t-1} - \bar{A}_{j.-1}) + (v_{j,t} - \underline{v}_j)\ ,$$

where $\bar{y}_{j.-1} = \Gamma^{T}_{t=2} Y_{j,t-1}/(T\text{-}1)$, and likewise for $\bar{y}^{pg}_{j.-1}$ and $\bar{A}_{j.-1}$. (The term $\underline{v}_j$ is the average error for physician j). In some specifications we interact $(Y_{j,t-1} - \bar{y}_{j.-1})$, $(Y^{pg}_{j,t-1} - \bar{y}^{pg}_{j.-1})$, and $(A_{j,t-1} - \bar{A}_{j.-1})$ with a physician's quantity of deliveries and experience to see if the effect of learning varies by physician practice characteristics.

The coefficient γ is now identified by variations over time in the c-section rate of a physician's peer group (and β_1 and ϑ are likewise identified by variations over time). If the peer group's c-section rate is high in t-1 relative to its average rate, is the physician's c-section rate also high in year t, relative to his average? There was considerable turnover of obstetricians in the Florida markets in the 1990s, which creates plausibly exogenous variation in peer group c-section rates. In Table 3 we display the most common patterns in our data set. Of the 1,928 physicians who delivered at least 10 babies in a given year between 1992 and 2000, 623 (32 percent) were present in every year. Six percent of the physicians were present from 1993 to 1996 only, presumably because they stopped practicing medicine or moved from the state thereafter. A total of 254 physicians (13.2 percent) were first observed in 1995, 1996, or 1997 and remained in the data set through 2000, presumably because they entered the state during our sample period. If physicians have different practice styles and do not sort perfectly according to their styles, we should see considerable variation over time in the peer group variable that will help identify the learning effect.

The variable $Y_{j,t-1} - \bar{y}_{j.-1}$ in equation (5) will be correlated with $v_{j,t} - \underline{v}_j$ because by construction $Y_{j,t-1}$ is correlated with $\underline{v}_j$ (Baltagi, 2001), and likewise for the other two difference terms. As a result, the coefficients will be biased downward. Anderson and Hsiao (1981) suggest taking the first difference of equation (3),

$$(6) \qquad Y_{j,t} - Y_{j,t-1} = \beta_1(Y_{j,t-1} - Y_{j,t-2}) + \gamma(Y^{pg}_{j,t-1} - Y^{pg}_{j,t-2}) + \vartheta(A_{j,t-1} - A_{j,t-2}) + (v_{j,t} - v_{j,t-1})$$

and using the second lag of the level ($Y_{j,t-2}$) as an instrument for ($Y_{j,t-1} - Y_{j,t-2}$) as suggested by Arrelano (1989). $Y_{j,t-2}$ will be correlated with ($Y_{j,t-1} - Y_{j,t-2}$) but will not be correlated with ($v_{j,t} - v_{j,t-1}$) as long as the $v_{j,t}$ are not serially correlated.[9] If there is a common shock in a market, such as increased malpractice pressure that causes all obstetricians to perform more or fewer c-sections, then the variable ($Y^{pg}_{j,t-1} - Y^{pg}_{j,t-2}$) in equation (5) may also be correlated with ($v_{j,t} - v_{j,t-1}$). A similar situation could exist with adverse patient outcomes. Therefore, one could instrument for ($Y^{pg}_{j,t-1} - Y^{pg}_{j,t-2}$) and ($A_{j,t-1} - A_{j,t-2}$) using the second lags of the levels ($Y^{pg}_{j,t-2}$ and $A_{j,t-2}$, respectively).

Arrelano and Bond (1991) developed a generalized method of moments estimator that is more efficient than the Anderson and Hsiao (1981) estimator. We use the Arrelano and Bond estimator in this paper. Specifically, rather than just using the second lags of the levels as instruments, we use a full set of a physician's lagged values. For example, we instrument for ($Y^{pg}_{j,t-1} - Y^{pg}_{j,t-2}$) using $Y^{pg}_{j,t-3}$, $Y^{pg}_{j,t-4}$, $Y^{pg}_{j,t-5}$, and so forth, based on the number of years the particular physician appears in the data set. We test for second-order serial correlation of the residuals from the first-difference equation.

IV. Results

The first step in our analysis is to adjust a physician's c-section rate for the type of patients he treated in a particular year. To do this we estimate a linear probability model (equation 2) where the unit of observation is a delivery, and the dependent variable is one if

---

[9] One could also use ($Y_{j,t-2} - Y_{j,t-3}$) as an instrument, but Arrelano (1989) recommends $Y_{j,t-2}$ because it usually has smaller variances over a significant range of parameter values.

woman received a c-section and zero otherwise.[10] This regression is estimated separately for each year between 1992-2000, and in Table 4 we present coefficient estimates from the ordinary least squares estimation for 2000.

Although the primary purpose of this regression is to recover the coefficients on the physician indicator variables, the coefficients are interesting nevertheless. These coefficients can be interpreted as the change in a patient's probability of receiving a c-section associated with a change in the independent variable. Although the health conditions we include are uncommon (none is present in more than five percent of the women), they have a substantial impact on the probability that a woman will have a c-section. For example, a woman with severe hypertension has a probability of receiving a c-section that is 32.8 percentage points higher than women without that condition. The probability that a Hispanic woman receives a c-section is 0.9 percentage points higher than for a white woman. Women who have Medicaid and uninsured women are less likely to receive a c-section. Medicaid generally reimburses physicians and hospitals less than private insurers, so our result is consistent with Gruber, Kim, and Mayzlin (1999) who find that higher fee differentials between c-sections and normal deliveries lead to higher c-section rates. Women who are admitted on the weekend have a considerably lower chance of having a c-section, presumably because few if any c-sections are scheduled for the weekend.

Once we run a separate regression like the one reported in Table 4 for each year, we recover the physician coefficients and create a panel data set. We also use the coefficients on the physicians in physician j's peer group (physicians who deliver babies at the same hospital as physician j in year t-1) to create a lagged, delivery-weighted peer group c-section rate for each physician for each year.

---

[10] As mentioned earlier, the distribution of the **Y** coefficients in equation (2) is independent of the patient characteristics (**X**) when we use ordinary least squares, but would not be if we instead used a probit or logit (Chamberlain, 1980).

One concern is that our risk-adjusted c-section rates may be noisy, imprecise measures of a physician's true treatment style. To address this issue we focus on the 421 physicians who were in the sample for the entire 9-year period (1992-2000). In the first column of the top panel of Table 5, we report the proportion of the 421 physicians who had a risk-adjusted c-section rate in the top decile among all physicians for 9 years, 8 years, 7 years, and so on.[11] In the second column, we report the proportion of physicians one would expect to appear in the top decile by year if all physicians had the same treatment style but physicians were equally likely to be in any decile due to considerable measurement error. In the third column, we report the proportion one would expect if physician styles are fixed over time and our method were able to measure a physician's true style perfectly, with no measurement error. The bottom panel of Table 5 repeats the analysis for the bottom rather than the top decile.

Seventy-one percent of the physicians were never in the top decile over the nine-year period. If, on the other hand, physician styles as we measure them are random or if styles are fixed and measured perfectly we would expect 39 percent and 90 percent of physicians, respectively, to never be in the top decile. As another example, 1.9 percent of the physicians were in the top decile five out of the nine years, whereas if style rankings were random one would expect 0.08 percent, more than twenty times less. The bottom panel of Table 5 offers more evidence that the risk-adjusted c-section rates are fairly stable over time and convey meaningful information. For example, 8.2 percent of the 421 physicians had a risk-adjusted c-section rate in the bottom decile for four or more years, whereas if the measures were random one would only expect about one percent of physicians to be in the bottom decile that frequently.

Phelps and Parente (1990) estimate there is annual welfare loss of $33 billion (in 1987 dollars) due to practice variations. As we argued earlier, the true welfare loss will be higher if physician treatment styles vary within a region. To estimate the amount of within-region

---

[11] Kane and Staiger (2001) perform a similar analysis to see how stable student test scores are as a measure of school performance.

variation in a particular year (2000), we use the physician-specific risk-adjusted c-section rates to calculate an average rate for each of the 11 Florida regions. We then calculate the difference between each physician's rate and their region's average rate. The standard deviation of this "de-meaned" treatment style variable is 0.073, which is 5 times larger than the standard deviation between regions in 2000 (0.014). This implies that two identical women who live in the same region and choose their obstetricians randomly are likely to have very different probabilities of receiving a c-section once they have selected an obstetrician. It also implies that health insurers and the government may be substantially underestimating the benefit of standardizing physicians' treatment methods.

We now examine how much of the variation in physician treatment styles can be explained by time, region, physician characteristics, and where physicians received obstetrical residency training. We separately examine all physicians and inexperienced physicians – those with fewer than 4 years of post-residency experience – to see if the impact of a residency program diminishes over time. In column 1 of Table 6a, we pool the risk-adjusted c-section rates for all physicians over the 1992-2000 time period, and regress these rates on year indicators only. Fifteen percent of the variation in treatment styles is due to time-specific factors that affected all physicians. When we include indicator variables for the region in which each physician practices, the adjusted R^2 increases to 0.22 (column 2 of Table 6a). Thus, regional factors explain an incremental seven percent of the variation in c-section rates. In column 3 we add the number of years a physician has been practicing and indicator variables for the physician's gender, race, and specialty. These physician characteristics explain only an incremental 1.5 percent of the variation in treatment styles. Older physicians are more likely to perform c-sections, and female physicians have a 0.8 percentage point lower c-section rate than male physicians, on average. Family practitioners and internists have lower rates than obstetricians (the omitted group), presumably because patients with unobserved poor health (and a high probability of having a c-section) are usually referred to obstetricians.

In column 4 of Table 6a we include indicator variables for the 84 obstetrical residency programs that trained at least five physicians in our data set to see whether programs impart distinct styles that persist once physicians start practicing. The residency programs collectively explain an additional two percent of the variation across physicians in their risk-adjusted c-section rates. Although it appears that residency programs have a minor effect on physician treatment styles, over 40 percent of the residency program coefficients are significant at the five-percent level relative to the omitted group (physicians who attended residency programs that trained fewer than 5 physicians in our sample).

Finally, in column 5 of Table 6a we include a full set of physician indicator variables and remove the region indicators, residency program indicators, and physician characteristics. The adjusted R-squared of 0.62 indicates that 37 percent of the variation in risk-adjusted c-section rates across physicians and years (0.62 – 0.25 from Table 6a) is due to time-invariant, physician-specific factors that are not associated with where physicians trained, their experience, gender, or race. Since we have detailed information on the characteristics of a physician's patients, our interpretation is that a considerable amount of practice variation is due to idiosyncratic physician styles. Grytten and Sorensen (2003) find that physician-specific effects explain slightly more than half of the variation in average medical spending per patient among primary care physicians in Norway. However, they have a much more limited set of risk adjusters (the gender mix and average age of a physician's patients), so unobserved patient heterogeneity across practices is likely to be a concern in their study.

We repeat the above analysis for physicians with fewer than four years of post-residency experience and present the results in Table 6b. The amount of variation attributable to years, regions, physician characteristics, and residency programs is remarkably similar as with the entire physician sample. It is particularly surprising that residency programs explain such a small percentage of the incremental variation in c-section rates among inexperienced physicians (2.6 percent for inexperienced physicians versus 2.0 percent for all physicians). One might expect residency programs to have a relatively strong effect on the practice styles of newly trained physicians, before they fully incorporate the practice styles of their peers and the health outcomes of their own patients into their treatment decisions.

We present coefficient estimates of the fixed effects model (equation 5) in Table 7. In the column 1 we include the lagged c-section rate of a physician's peer group, his own lagged rate, and the lagged proportion of a physician's deliveries that resulted in an adverse outcome. These coefficients are identified by variation over time in a physician's c-section rate, his peer group's rate, and his adverse outcome rate.

The peer group coefficient in the first column of Table 7 is positive, as expected. If the risk-adjusted c-section rate of a physician's peer group increases by one standard deviation (5.1 percentage points), that physician's c-section rate is predicted to increase by 1.4 percentage points in the subsequent year. This represents a 7.3 percent increase for a physician with the sample mean risk-adjusted c-section rate of 18.7 percent. Physicians do appear to learn from the treatment decisions of their peers, and the impact is quite large.

The coefficient on a physician's lagged c-section rate is insignificant, which indicates that physician treatment styles are quite stable.[12] The coefficient on the proportion of a physician's deliveries in the prior year that resulted in an adverse patient outcome is positive but small in magnitude. As expected, physicians appear to perform more c-sections when their patients have bad health outcomes. A physician experiencing a one-standard deviation increase in his adverse outcome rate (from 0.201 to 0.283) is predicted to increase his c-section rate by 0.3 percentage points, or 1.5 percent.

In the second column of Table 7 we interact a physician's lagged quantity of deliveries with his peer group's lagged rate and his own lagged c-section rate. The former interaction term is negative; physicians who perform a relatively large number of deliveries are influenced less strongly by the treatment styles of their peers. This seems intuitive. A physician who delivers a large number of babies receives more information from his own practice and less from the practice of his peers. The positive coefficient on the interaction between a physician's own lagged rate and his quantity of deliveries indicates that a change to a physician's c-section rate is more likely to persist among relatively high-volume physicians.

In the third column of Table 7, we also interact a physician's experience with the c-section rate of his peer group, the physician's own c-section rate, and the proportion of a physician's deliveries that resulted in an adverse health event in order to examine whether peer groups and health outcomes have a relatively strong influence on inexperienced physicians. None of the experience interactions is significant.

As discussed in the previous section, the coefficients on the two lagged c-section rate variables will be biased downward because of correlation with the error term by construction. In Table 8 we report coefficient estimates from the Arrelano and Bond (1991) first-difference specification described in equation (6). We instrument for ($Y_{j,t-1}$ - $Y_{j,\,t-2}$) with $Y_{j,\,t-2}$ and all available earlier lags; we instrument for ($Y^{pg}{}_{j,t-1}$ - $Y^{pg}{}_{j,t-2}$) with $Y^{pg}{}_{j,t-2}$ and all earlier lags; and we instrument for ($A_{j,t-1} – A_{j,t-2}$) with $A_{j,t-2}$ and all earlier lags. The lagged values are valid instruments in the Arrelano and Bond model as long as the errors (the $v_{j,t}$ in equation 6) are not serially correlated. We find no evidence that the residuals are serially correlated in the specifications presented in Table 8. In column 1 of Table 8, for example, the z-statistic for the test that the average autocovariance in residuals of order 2 is zero is 0.60.

In the first column of Table 8, the coefficient of 0.453 on the lagged c-section rate of a physician's peer group is larger than in the fixed-effects specification, as expected. A one-standard deviation increase in the c-section rate of a physician's peer group is predicted to increase his c-section rate by 2.3 percentage points, or 12.4 percent of his baseline risk-adjusted

---

[12] The coefficient on a physician's lagged c-section rate is considerably smaller in magnitude in the fixed-effects specification than the OLS specification (not shown), which is consistent with physician-patient

c-section rate. The treatment style of a physician's peer group has a considerable impact on how he treats patients. The coefficient on a physician's lagged c-section rate in column 1 (0.104) is also larger than in the fixed-effect specification, as expected. Since this coefficient is still close to zero, physicians' treatment styles appear to be quite stable over time; only 10 percent of a change in a physician's c-section rate in year t-1 persists in the following period.

In columns 2 and 3 of Table 8 we include interaction terms between the c-section rate and adverse outcome variables and the quantity of a physician's deliveries (column two), and interactions with a physician's experience (column three). The only interaction term that is significant is a physician's lagged rate * experience. As expected, experienced physicians have more stable treatment styles than inexperienced physicians. For a physician who has just completed residency training, 17 percent of a change in his c-section rate in t-1 will be incorporated into his c-section rate in year t. Conversely, for a physician with 22 years of experience, a change in his c-section in year t-1 is predicted to have no impact on his c-section rate in year t.

V. Conclusions

In this paper we investigate how obstetricians form their treatment style, and whether and how their styles change over time. We use a unique data set that contains the universe of hospital admissions in Florida over a 9-year period and consistent physician identifiers. These data allow us to examine physicians' treatment decisions over an extended time period, to characterize a physician's peer group, to observe the treatment decisions of a physician's peer group, and to observe patient outcomes. We examine the extent to which an obstetrician's decision regarding whether or not to perform a cesarean section is influenced by the residency program where he trained, his peer group's c-section rate, and his patients' health outcomes, controlling for patient characteristics.

---

sorting based on unobserved health characteristics or preferences.

As with most medical care treatments, differences in the mean c-section rates between the 11 regions of Florida, controlling for patients' observed characteristics, are quite large. In 2000, for example, the risk-adjusted probability a woman would have a cesarean section ranged from a low of 0.222 in Jacksonville to a high of 0.276 in Miami. We show that there are even larger differences in c-section rates between physicians within a region; the standard deviation of the c-section rate across physicians within a region is 5 times larger than the between-region variation, controlling for observed patient characteristics. This implies that existing estimates of the welfare loss stemming from practice variation are too low, and private health insurers and the government may be underestimating the value of standardizing treatment methods.

Educational institutions appear to have a small effect on regional variation in the type of medical care received; only two percent of the variation in c-section rates among physicians can be explained by where they trained as a resident. Likewise, a physician's experience, gender, and race accounts for less than two percent of the variation in c-section rates. Almost 40 percent of the variation in risk-adjusted c-section rates across physicians and years is due to time-invariant, physician-specific factors other than experience, gender, race, and where a physician received residency training. Since we have detailed information on the characteristics of a physician's patients, our interpretation is that a considerable amount of practice variation is due to idiosyncratic physician styles.

We find that physicians' clinical styles are quite stable; a one percent increase in a physician's c-section rate in year t-1 is associated with a 0.1 percent increase in their rate in year t. We do, however, find evidence that physicians respond to the practice style of their peer physicians. In the first-difference specification, a one-standard deviation increase in a peer group's lagged c-section rate is predicted to increase a physician's c-section rate by 2.3 percentage points (or a 12.4 percent increase in the sample average c-section rate). Therefore, peer groups are likely to be a more effective mechanism for changing physician behavior than formal residency training, at least as training programs are currently constructed. We find some evidence that physicians alter their practice style according to how well their patients have fared in the past, but this effect has a small magnitude.

References

Anderson, T.W., and Cheng Hsiao, 1981, “Estimation of Dynamic Models With Error Components,” *Journal of the American Statistical Association* 76(375): 598-606.

Arcidiacono, Peter, and Sean Nicholson, 2001, “Peer Effects in Medical School,” NBER Working Paper 9025.

Arrelano, Manuel, 1989, “A Note on the Anderson-Hsiao Estimator for Panel Data,” *Economics Letters* 31: 337-341.

Arrelano, Manuel, and Stephen Bond, 1991, “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations,” *Review of Economic Studies* 58: 277-297.

Baltagi, Bani H., 2001. *Econometric Analysis of Panel Data*. Chichester, England: John Wiley & Sons, LTD.

Banerjee, Abhijit V., 1992, “A Simple Model of Herd Behavior,” *Quarterly Journal of Economics* 107(3): 797-817.

Bikhchandani, Sushil, David Hirschleifer, and Ivo Welch, 1992, “A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades,” *Journal of Political Economy* 100(5): 992-1026.

Bikhchandani, Sushil, Amitabh Chandra, Dana P. Goldman, and Ivo Welch, 2001, “The Economics of Iatroepidemics and Quakeries: Physician Learning, Information Cascades, and Geographic Variation in Medical Practice,” mimeo.

Burns, Lawton R., Stacie E. Geller, and Douglas R. Wholey, 1995, “The Effect of Physician Factors on the Cesarean Section Decision,” *Medical Care* 33(4): 365-382.

Chamberlain, Gary 1980, “Analysis of Covariance With Qualitative Data,” *Review of Economic Studies* 47: 225-238.

Ellison, Glenn, and Drew Fudenberg, 1993, “Rules of Thumb for Social Learning,” *Journal of Political Economy* 101(4): 612-643.

Ellison, Glenn, and Drew Fudenberg, 1995, “Word-of-Mouth Communication and Social Learning,” *Quarterly Journal of Economics* 110(1): 93-125.

Fournier, Gary M., Kislaya Prasad, and Mary A. Burke, 2002, “Physician Social Networks and Treatment Variations in Coronary Inpatient Care,” mimeo.

Gruber, Jon, John Kim, and Dina Mayzlin, 1999, “Physician Fees and Procedure Intensity: the Case of Cesarean Delivery,” *Journal of Health Economics* 18: 473-490.

Grytten, Jostein, and Rune Sorensen, 2002, “Practice Variation and Physician-Specific Effects,” *Journal of Health Economics* 22(3): 403-418.

Kane, Thomas J., and Douglas O. Staiger, 2001, “Improving School Accountability Measures.” NBER working paper 8156.

Manski, Charles, 1993, “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economic Studies* 60: 531-542.

McClellan, Mark, and Douglas Staiger, 1999, “The Quality of Health Care Providers.” NBER working paper 7327.

McGlynn, Elizabeth A., Steven M. Asch, John Adams, Joan Keesey, Jennifer Hicks, Alison DeCristofaro, and Eve A. Kerr, 2003, “The Quality of Health Care Delivered to Adults in the United States,” *JAMA* 348(26): 2635-2645.

Newhouse, Joseph P., 2002, “Why is There a Quality Chasm?” *Health Affairs* 21(4): 13-25.

Phelps, Charles E., and S. Parente, 1990, “Priority Setting for Medical Technology and Medical Practice Assessment,” *Medical Care* 28(8): 703-723.

Phelps, Charles E., and Cathleen Mooney, 1993, “Variations in Medical Practice Use: Causes and Consequences,” in R.J. Arnould, R.F. Rich, and W. White, eds., Competitive Approaches to Health Care Reform. Washington, D.C.: Urban Institute Press.

Phelps, Charles E., 2000, “Information Diffusion and Best Practice Adoption,” in *Handbook of Health Economics*, vol. 1A, Joseph P. Newhouse and Anthony J. Cuyler, editors: 223-264. Amsterdam: Elsevier Science Publishers.

Sacerdote, Bruce, 2001, “Peer Effects With Random Assignment: Results for Dartmouth Roomates,” *Quarterly Journal of Economics* 116(2): 681-704.

Scharfstein, David S., and Jeremy C. Stein, 1990, “Herd Behavior and Investment,” *American Economic Review* 80(2): 465-479.

Wennberg, John E., and A. Gittelsohn, 1973, “Small Area Variations in Health Care Delivery,” *Science* 182: 1102-1108.

Wennberg, John E., J.L. Freeman, and W.J. Culp, 1987, “Are Hospital Services Rationed in New Haven or Over-Utilised in Boston?,” *Lancet* 1: 1185-1189.

Table 1

Sample Means and Standard Deviations in Patient-level Data Set

(n = 1,670,991)

	Mean	Standard Deviation
Age	26.8	6.20
White	0.575	0.494
Black	0.212	0.409
Hispanic	0.154	0.361
Other race	0.0589	0.235
Health insurance		
- PPO and indemnity	0.284	0.451
- HMO	0.218	0.413
- Other private insurance	0.028	0.164
- Medicaid	0.388	0.487
- Uninsured	0.0825	0.275
Woman’s health condition:		
Woman has had a previous c-section	0.044	0.206
Malpositioned fetus	0.030	0.171
Antepartum bleeding	0.0096	0.097
Severe hypertension	0.0037	0.061
Preterm gestation	0.0318	0.176
Multiple gestation	0.0054	0.073
Soft tissue disorder	0.0165	0.127
Macrosomia	0.0171	0.130
Oligohydramnios	0.0036	0.060
Polyhydramnios	0.0118	0.108
Herpes	0.0107	0.103
Day of admission:		
- Monday – Thursday	0.643	0.479
- Friday	0.148	0.355
- Saturday	0.103	0.304
- Sunday	0.106	0.307

Table 2

Sample Means and Standard Deviations in Physician-level Data Set

(n = 1,928 physicians and 10,528 physician-years)

	Mean	Standard Deviation
Physician’s risk adjusted c-section rate	0.187	0.082
Physician’s quantity of deliveries	138.6	119.2
Female	0.214	0.410
Non-white	0.100	0.300
Race missing	0.648	0.479
Post-residency experience (years)	11.7	8.66
Physician has < 4 years of experience	0.187	0.390
Age	44.4	9.22
Specialty		
- ob/gyn	0.848	0.364
- family practice/internal medicine	0.042	0.200
- maternal and fetal medicine	0.019	0.136
- other	0.091	0.288
Peer group’s adjusted c-section rate, t-1	0.182	0.051
Peer group’s annual quantity of deliveries	3,999	2,873
Proportion of physician’s deliveries resulting in an adverse outcome	0.201	0.082

Note: a physician’s risk adjusted c-section rate is the coefficient on a physician indictor in a cross-section ordinary least squares regression where the unit of observation is a delivery and dependent variable is one if a woman received a c-section, and zero otherwise.

Table 3

Number of Physicians in Data Set by Year of Entry and Exit

	1993	1994	1995	1996	1997	1998	1999	2000
623 physicians	X----------	----------	----------	----------	----------	----------	----------	X
117 physicians	X----------	----------	----------	X				
61 physicians	X----------	----------	X					
59 physicians			X----------	----------	----------	----------	----------	X
68 physicians				X----------	----------	----------	----------	X
127 physicians					X----------	----------	----------	X
71 physicians						X----------	----------	X
74 physicians							X----------	X
69 physicians								X
659 physicians	Other patterns (e.g., 1994-1996)							
1,928 physicians total								

Table 4

Coefficient Estimates From a Patient-level Regression, 2000

	Coefficient	Standard Error
Age	-0.0087**	0.0010
Age squared	0.00019**	0.00002
Black	0.0038	0.0024
Hispanic	0.0086**	0.0026
Health insurance (PPO and indemnity omitted)		
- HMO	-0.00014	0.0023
- Medicaid	-0.020**	0.0025
- Other private insurance	-0.017**	0.0063
- Uninsured	-0.047**	0.0038
Woman's health condition:		
Woman has had a previous c-section	0.550**	0.0025
Malpositioned fetus	0.501**	0.0033
Antepartum bleeding	0.316**	0.0062
Severe hypertension	0.328**	0.0073
Preterm gestation	-0.0007	0.0032
Multiple gestation	0.154**	0.0080
Soft tissue disorder	0.209**	0.0055
Macrosomia	0.253**	0.0045
Oligohydramnios	0.148**	0.0107
Polyhydramnios	0.141**	0.0053
Herpes	0.209**	0.0063
Day of admission (Monday-Thursday omitted)		
- Friday	0.0024	0.0023
- Saturday	-0.034**	0.0028
- Sunday	-0.037**	0.0028
Observations	196,077	
R^2	0.31	

Notes: We include a full set of physician indicator variables in the above regression and omit the constant. We also include indicators for whether a patient's race is missing and whether her age is missing. ** = significantly different from zero at the 5-percent level; * = significantly different from zero at the 10-percent level.

Table 5

Proportion of Physicians Ranking in the **Highest** Decile for Risk-Adjusted C-section Rates
Among the 421 Physicians in the Sample for 9 Years

Number of years in top decile, 1992-2000	Actual	Expected Proportion if Physician Styles Are… Random	Expected Proportion if Physician Styles Are… Measured with no error
0	0.710	0.387	0.90
1	0.114	0.387	0
2	0.059	0.172	0
3	0.062	0.045	0
4	0.017	0.007	0
5	0.019	0.0008	0
6	0.005	0.000061	0
7	0.012	2.9×10^{-6}	0
8	0.002	8.1×10^{-8}	0
9	0.000	1.0×10^{-9}	0.10

Proportion of Physicians Ranking in the **Lowest** Decile for Risk-Adjusted C-section Rates
Among the 421 Physicians in the Sample for 9 Years

Number of years in bottom decile, 1992-2000	Actual	Expected Proportion if Physician Styles Are… Random	Expected Proportion if Physician Styles Are… Fixed and Perfectly Measured
0	0.674	0.387	0.90
1	0.138	0.387	0
2	0.059	0.172	0
3	0.048	0.045	0
4	0.029	0.007	0
5	0.017	0.0008	0
6	0.012	0.000061	0
7	0.014	2.9×10^{-6}	0
8	0.005	8.1×10^{-8}	0
9	0.005	1.0×10^{-9}	0.10

Table 6a: Sources of Variation in Physicians' Risk-Adjusted C-section Rates

Sample: **All Physicians**, 1992-2000 (N=10,528)

	Year Indicators	Year and Region	Year, Region and Physician Characteristics	Year, Region, MD char., and Res. programs	Year and MD Indicators
Adjusted R^2	0.148	0.219	0.234	0.254	0.618
Physician characteristics					
Post-residency experience (years)			0.00059**	0.00054**	
			(0.00017)	(0.00018)	
Female			-0.0082**	-0.0079**	
			(0.0033)	(0.0033)	
Non-white			0.0025	-0.0003	
			(0.0047)	(0.0049)	
Specialty (OB/GYN is omitted)					
Family practice			-0.032**	-0.0344**	
			(0.0099)	(0.0101)	
Internal medicine			-0.038**	-0.039**	
			(0.0139)	(0.0139)	
General practice			0.027	0.0268	
			(0.029)	(0.029)	
Maternal and fetal medicine			0.0069	0.0068	
			(0.012)	(0.011)	
Percent of residency program coefficients significant at 5-percent level (relative to MDs from programs that produced fewer than five MDs in the sample)				43%	

Notes: Dependent variable is a physician's risk-adjusted c-section rate in year t. ** = significantly different from zero at the 5-percent level; * = significantly different from zero at the 10-percent level. The sample used here is larger than in subsequent tables because we include observations from 1992 in the above regressions, but exclude these observations when we include lagged variables as regressors.

Table 6b: Sources of Variation in Physicians' Risk-Adjusted C-section Rates

Sample: **Physicians With Fewer Than 4 Years of Experience**, 1992-2000 (N=1,966)

	Year Indicators	Year and Region	Year, Region and Physician Characteristics	Year, Region, MD char., and Res. Programs	Year and MD Indicators
Adjusted R^2	0.159	0.209	0.218	0.244	0.617
Physician characteristics					
Post-residency experience (years)			0.00055	0.00098	
			(0.0018)	(0.0019)	
Female			-0.0075	-0.0099**	
			(0.0048)	(0.0050)	
Non-white			-0.0034	-0.0038	
			(0.0060)	(0.0069)	
Specialty (OB/GYN is omitted)					
Family practice			-0.0097	-0.0114	
			(0.0172)	(0.0171)	
Internal medicine			-0.0378**	-0.0401**	
			(0.0139)	(0.0140)	
General practice			-0.0197	-0.0227	
			(0.0317)	(0.0317)	
Maternal and fetal medicine			0.0019	0.0009	
			(0.0285)	(0.0282)	
Percent of residency program coefficients significant at 5-percent level (relative to MDs from programs that produced fewer than five MDs in the sample)				50%	

Notes: Dependent variable is a physician's risk-adjusted c-section rate in year t. ** = significantly different from zero at the 5-percent level; * = significantly different from zero at the 10-percent level.

Table 7: Coefficient Estimates from the Fixed Effect Model

	(1)	(2)	(3)
Peer group's c-section rate, t-1	0.267**	0.322**	0.278**
	(0.0282)	(0.0321)	(0.0429)
Peer group's rate * MD's quantity of deliveries (00)		-0.0552**	-0.0565**
		(0.0140)	(0.0141)
Peer group's lagged rate * MD's experience			0.0037
			(0.0024)
Physician's c-section rate, t-1	0.0144	-0.0355**	-0.0344
	(0.0119)	(0.0163)	(0.0228)
Physician's lagged rate * MD's # of deliveries (00)		0.0551**	0.0556**
		(0.0128)	(0.0128)
Physician's lagged rate * experience			-0.00015
			(0.0013)
Proportion of deliveries in t-1 with an adverse outcome	0.0346**	0.0412**	0.0352*
	(0.0124)	(0.0140)	(0.0188)
Adverse outcome rate * MD's quantity of deliveries (00)		-0.0060	-0.0057
		(0.0064)	(0.0064)
Adverse outcome rate * MD's experience			0.0005
			(0.0010)
Constant	0.184**	0.184**	0.182**
	(0.0062)	(0.0063)	(0.0065)
Observations	8,429	8,429	8,429
R^2	0.24	0.25	0.26

Notes: The dependent variable is a physician's risk-adjusted c-section rate in year t. The estimator includes fixed effects for physician and year. ** = significantly different from zero at the 5-percent level; * = significantly different from zero at the 10-percent level.

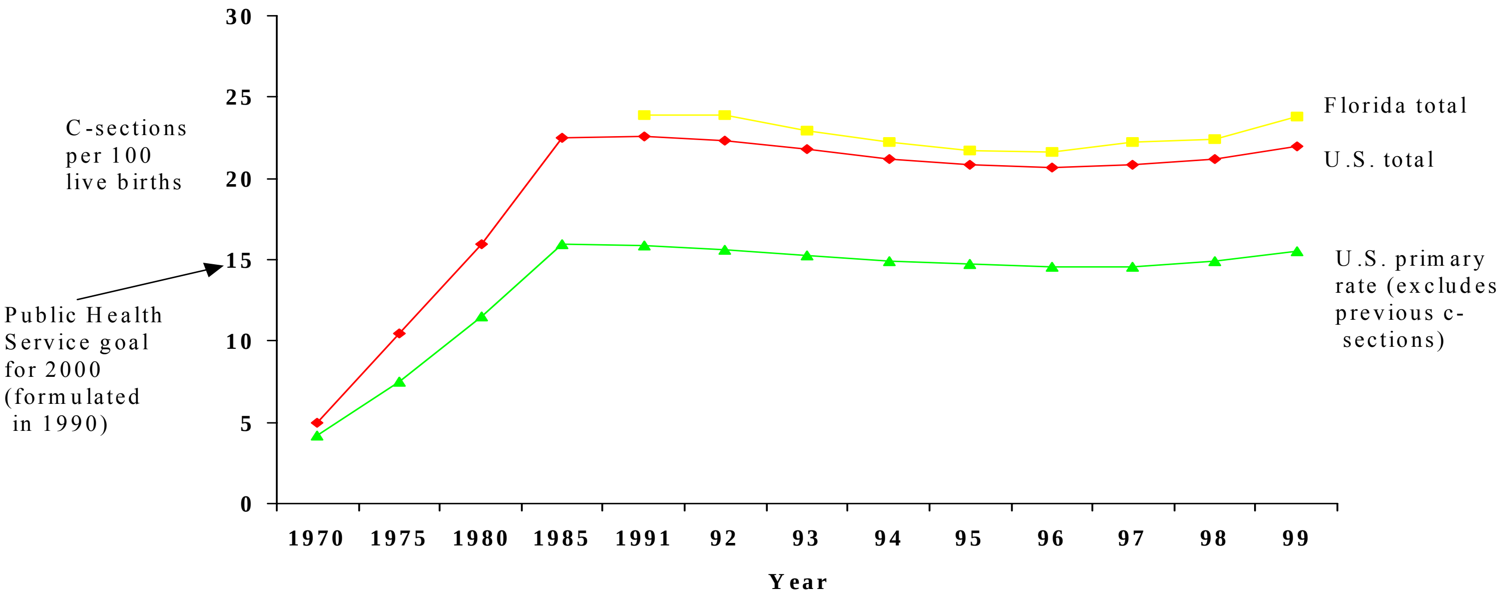
Table 8: Coefficient Estimates From the Arrelano and Bond First-Difference Model

	(1)	(2)	(3)
Peer group's c-section rate, t-1	0.453**	0.378**	0.288**
	(0.174)	(0.167)	(0.176)
Peer group's rate * MD's quantity of deliveries (00)		0.0045	-0.0393
		(0.0687)	(0.0620)
Peer group's lagged rate * MD's experience			0.0079
			(0.0058)
Physician's c-section rate, t-1	0.104**	0.125**	0.173**
	(0.0216)	(0.0549)	(0.0730)
Physician's lagged rate * MD's # of deliveries (00)		-0.0321	0.0115
		(0.0678)	(0.0610)
Physician's lagged rate * experience			-0.0078*
			(0.0043)
Proportion of deliveries in t-1 with an adverse outcome	-0.0254	-0.0494	-0.0397
	(0.0747)	(0.0655)	(0.0579)
Adverse outcome rate * MD's quantity of deliveries (00)		0.00010	0.0143
		(0.00015)	(0.0154)
Adverse outcome rate * MD's experience			0.0019
			(0.0017)
Constant	-0.0049	-0.0072	-0.0104**
	(0.0053)	(0.0047)	(0.0046)
Observations	6,669	6,669	6,669

Notes: Dependent variable is the change in a physician's risk-adjusted c-section rate between year t-1 and t. The change in a physician's c-section rate between year t-2 and t-1 is instrumented with the c-section rate in t-2 and all earlier (available) lags. Likewise for the change in the risk-adjusted c-section rate of a physician's peer group and the change in a physician's adverse outcome rate. Interaction terms are likewise instrumented. ** = significantly different from zero at the 5-percent level; * = significantly different from zero at the 10-percent level.

Figure 1

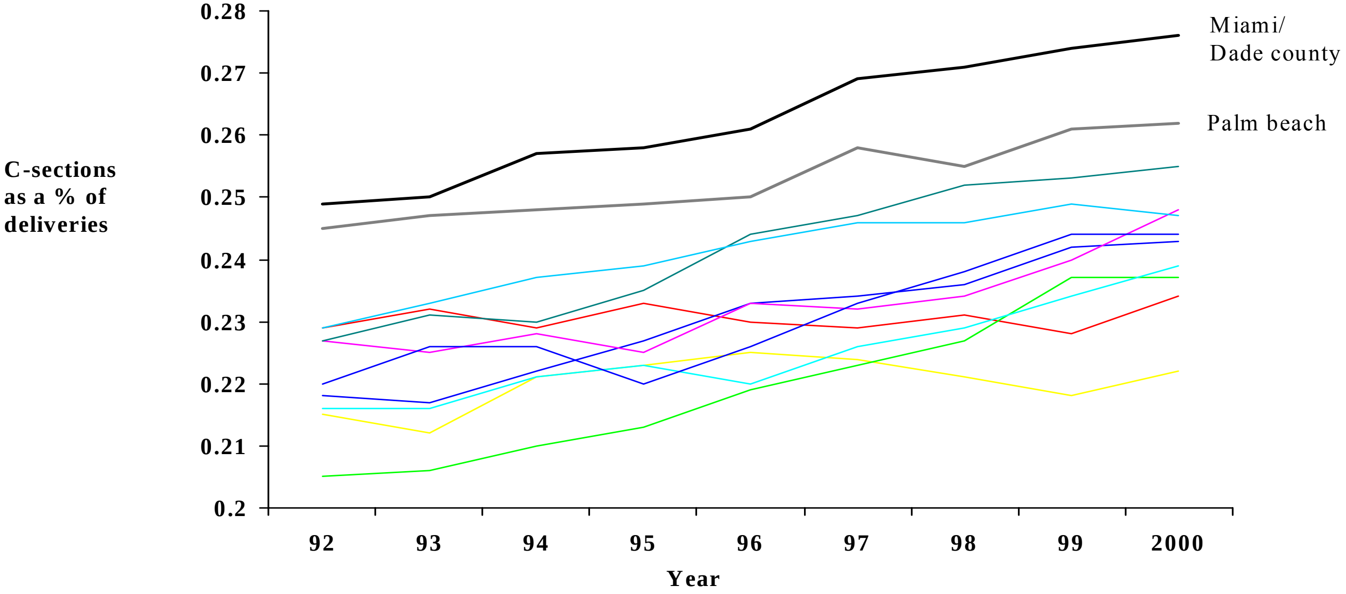
Overall and Primary Cesarean Section Rate for U.S. and Florida, 1970-1999



Source: National Center for Health Statistics.

Figure 2

Risk-Adjusted C-section Rates by Local Health District, 1992-2000



Note: this figure displays coefficients on the 11 local health district indicator variables from cross-section regressions where the dependent variable is one if a woman received a c-section and zero otherwise. The patient characteristics listed in Table 1 are included as risk-adjusters and the constant is omitted from the regressions.

Figure 3: Kernel Density Estimates of Physician C-section Rates, 1992-2000 (N = 10,528)