Journal of Health Economics 28 (2009) 1126–1140

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

The formation and evolution of physician treatment styles: An application to cesarean sections

Andrew J. Epstein[a,*], Sean Nicholson[b]

[a] Yale University, School of Public Health, Division of Health Policy and Administration, 60 College Street, 3rd Floor, New Haven, CT 06520-8034, United States
[b] Cornell University, and NBER, United States

ARTICLE INFO

Article history:
Received 3 August 2005
Received in revised form 15 April 2009
Accepted 18 August 2009
Available online 25 August 2009

JEL classification:
I11
D83

Keywords:
Physician treatment styles
Small-area variation
Peer effects

ABSTRACT

Small-area-variation studies have shown that physician treatment styles differ substantially both between and within markets, controlling for patient characteristics. Using data on the universe of deliveries in Florida and New York over a 15-year period, we examine why treatment styles differ across obstetricians at a point in time and why styles change over time. We find that variation in c-section rates across physicians within a market is about twice as large as variation between markets. Surprisingly, residency programs explain no more than four percent of the variation in physicians' risk-adjusted c-section rates, even among newly trained physicians. Although we find evidence that physicians learn from their peers, they do not substantially revise their prior beliefs regarding treatment due to the local exchange of information. Our results indicate that physicians are not likely to converge over time to a community standard; thus, within-market variation in treatment styles is likely to persist.

1. Introduction

There is an 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 et al., 1987, 2002). These studies usually compare the use rate of a particular treatment (e.g., back surgeries per capita) or medical expenditures across cities, counties, or states. If there is a single treatment method that patients prefer, there will be welfare losses when the use rate diverges from the medically appropriate standard.[1] Phelps and Parente (1990) estimated an annual welfare loss in 1987 due to variations in hospital use rates of $33 billion.

What matters to a consumer is whether the physician she chooses provides the appropriate treatment, not whether physicians in her market provide the appropriate treatment on average. As Phelps and Parente (1990) point out, their $33 billion estimate will understate the true welfare loss if there is variation in use rates *within* a market (e.g., variation across physicians in their likelihood of admitting patients to a hospital) as well as between markets. That is, even if the mean use rate of a market conforms to best medical practices, some patients may still receive too much or too little of the treatment if physicians in that market treat patients quite differently. If, however, patients have preferences for different treatment styles and choose physicians accordingly, then some component of the within-market variation will enhance welfare. Epstein et al. (2008) find that group practices facilitate the matching of patients who are clinically appropriate for c-sections to members of the practice who are skilled at performing c-sections.

A less-frequently cited set of studies show that there is indeed considerable variation across physicians within a market in how they treat patients, controlling for patients' observed health. Stano and Folland (1988), for example, report that variation in the amount of medical care patients receive, measured by relative value units (RVUs),[2] is three to four times larger across physicians within a market as across markets. Other studies find substantial variation across physicians in cesarean-section rates (Grant and McInnes,

* Corresponding author. Tel.: +1 203 785 6924; fax: +1 203 785 6287.
E-mail address: andrew.epstein@yale.edu (A.J. Epstein).

[1] 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 (2000) concludes that these factors collectively explain very little of the differences in the amount of medical care received. Several other authors concur with this assessment, including Chassin et al. (1986), Bikhchandani et al. (2001), and Newhouse (2002).

[2] Relative value units are a way to aggregate heterogeneous services into a single measure. Stano and Folland (1988) assigned a routine office visit a value of one, and every other service a value proportional to its charge relative to a routine office visit.

0167-6296/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhealeco.2009.08.003

2004; Grant, 2005; Goyert et al., 1989), RVUs per hospital admission (Welch et al., 1994), hospitalization rates, hospital days, and length of hospital stays (Roos et al., 1986), total medical expenditures (Phelps, 2000), and medical expenditures on outpatient care (Grytten and Sorensen, 2003).

One concern with the latter set of studies, however, is that within-market variations will be overstated if patients' unmeasured health differs across practices due to physician specialization, or if the number of patients per physician is small and treatment styles are measured with error (Hofer et al., 1999). Grant and McInnes (2004), Grant (2005), and Roos et al. (1986) have detailed health information, but the other studies cited above use no or few risk adjusters. Welch et al. (1994) and Roos et al. (1986) analyze all physicians with 10 or more and 15 or more admissions per year, respectively, which raises concerns about measurement error. In a sample where each physician treated an average of 16 diabetic patients, Hofer et al. (1999) find that at least 96 percent of the variation across physicians in hospitalization and outpatient visit rates is due to unmeasured patient factors or chance, rather than physician practice styles.

The first objective of this paper is to measure the amount of variation in treatment styles between obstetricians practicing in the same market. Our primary measure of treatment style is the proportion of a physician's deliveries performed by cesarean section, but we also examine elective c-section rates and c-section rates conditional on a patient's going into labor. We use information from hospital discharge abstracts to adjust physicians' treatment styles for patients' health, a potentially important source of variation. Moreover, the analysis is restricted to physicians who delivered 50 or more babies in a year in order to measure treatment styles precisely.

The choice of delivery method has implications for physicians, patients and payers. With over 1.2 million cesarean sections performed annually in the United States, c-sections are one of the most common surgical procedures (DeFrances and Hall, 2007). Based on our data, women who received a c-section in Florida between 1992 and 2006 remained in the hospital 3.5 days on average, versus 2.1 days for women who had vaginal deliveries. The average hospital charge for a c-section in Florida in the 1990s was $8500, almost twice 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 et al., 1999).

Our second objective is to examine the source and importance of physician learning. Our data set contains the universe of hospital admissions in Florida and New York over a 15-year period, and includes consistent physician identifiers and characteristics, such as information on where physicians trained. We first test whether residency programs produce physicians with distinct treatment styles, and whether those styles persist beyond the first few years of practice. If so, then residency programs would be an effective means of promoting evidence-based medicine.

The panel nature of the data allows us to explore whether a physician learns from his immediate peers once he begins practicing, as posited by Phelps and Mooney (1993), and whether learning is important relative to other market-specific forces, such as changes in reimbursement, the malpractice environment, and programs to promote adherence to clinical guidelines.[3] We construct two peer group variables—the change in the treatment style of physicians who deliver in the same hospital(s) as physician *j* (the "local" peer group), and the change in treatment style of physicians who deliver in all other hospitals in physician *j*'s market (the "regional" peer group). If members of the local and regional peer groups are exposed to the same policies and clinical programs, we can separately measure the effect of the local exchange of information between physicians (via the local peer group variable) and other market forces (via the regional peer group variable) on an individual physician's treatment style. Because some policy changes and clinical programs may be specific to a single hospital or subset of hospitals in a market, we also estimate models using only the changes in treatment styles of physicians who enter or exit a local peer group, after showing that the practice styles of the physicians who enter and exit are uncorrelated with the styles of incumbent physicians.

We find that the variation in c-section rates across physicians within a market is about twice as large as the inter-market variation, controlling for observed patient characteristics. Treatment styles are not strongly shaped by residency training programs. Residency programs explain no more than four percent of the variation between physicians in c-section rates, even among physicians who have been practicing for fewer than four years. Almost 30 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. Because we have fairly 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 perceptions regarding the appropriateness of specific treatments.

Although we find evidence that physicians learn from their peers, they do not substantially revise their prior beliefs regarding how patients should be treated due to the local exchange of information. A one-standard deviation (2.4 percentage points) increase in the portion of a physician's local peer group's c-section rate driven by the entry and exit of physicians is associated with a 0.16 percentage point (or 1.0 percent) increase in his own rate. Our results indicate that physicians are not likely to converge over time to a community standard and thereby eliminate the within-market variation in treatment styles.

In the next section we present the conceptual framework for the paper and some descriptive data on c-section rates in Florida and New York between 1992 and 2006. We describe the data and methods in Section 3. Section 4 contains our estimation results, and we conclude in Section 5.

2. Conceptual framework

Our primary measure of a physician's treatment style is the proportion of deliveries he performs by cesarean section, controlling for patients' observed demographic and risk characteristics. This proportion, Y, can be decomposed into the proportion of patients who go into labor (θ) multiplied by the obstetrician's c-section rate conditional on labor (Y_θ), plus the proportion of patients who do not go into labor but are instead scheduled to receive a c-section ($1-\theta$):

$$Y = \theta Y_\theta + (1-\theta) \tag{1}$$

For each physician we separately measure the overall c-section rate (Y), the proportion of patients receiving a c-section without laboring ($1-\theta$), and the c-section rate conditional on labor (Y_θ), and we adjust all three treatment style measures based on each patient's observed risk characteristics.

The overall, unadjusted c-section rate in Florida changed markedly during our 1992–2006 sample period. The c-section rate

[3] Banerjee (1992), Scharfstein and Stein (1990), Ellison and Fudenberg (1993, 1995), and Bikhchandani et al. (1992) have also presented general models where individuals observe the decisions of their peers and update their priors.

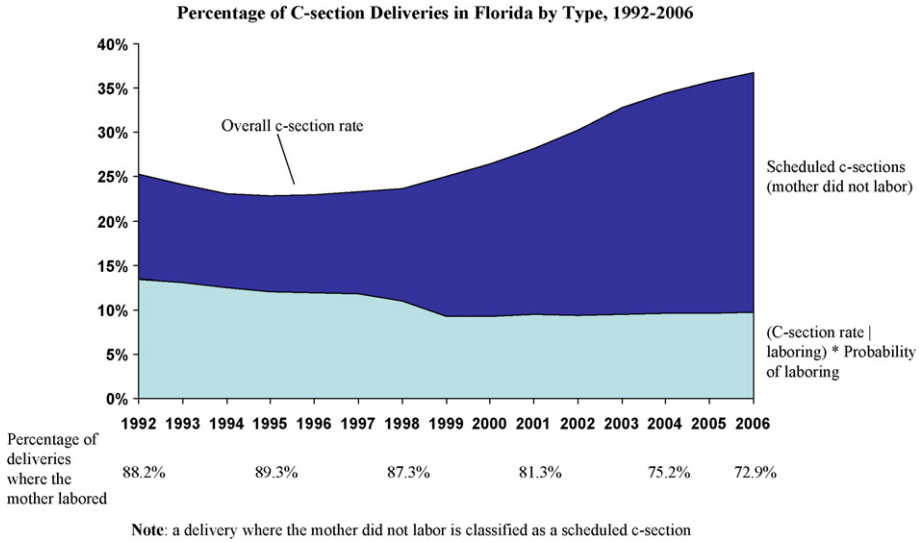

Fig. 1. Percentage of c-section Deliveries in Florida by Type, 1992–2006. *Note*: a delivery where the mother did not labor is classified as a scheduled c-section.

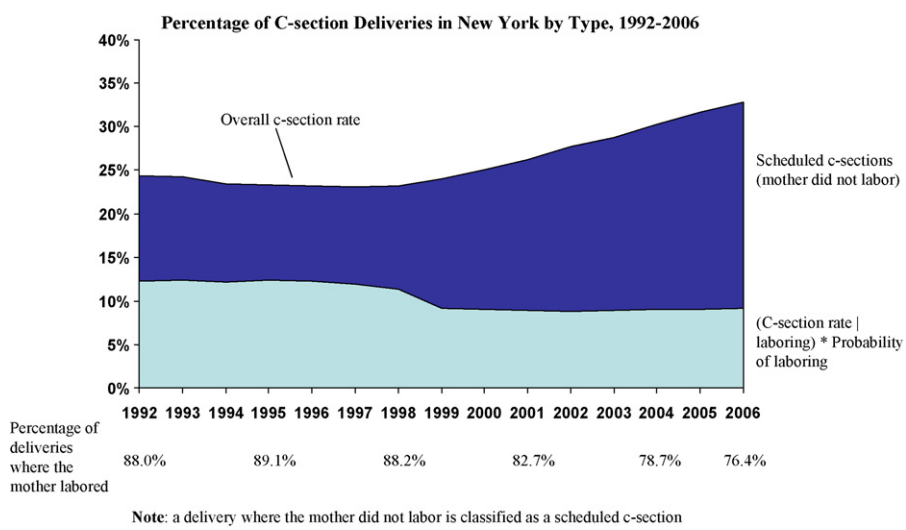

Fig. 2. Percentage of c-section Deliveries in New York by Type, 1992–2006. *Note*: a delivery where the mother did not labor is classified as a scheduled c-section.

in Florida fell from 25.4 percent in 1992 to 22.9 percent in 1996, as displayed in Fig. 1. During this period slightly fewer c-sections were scheduled (the percentage of deliveries where the mother went into labor is reported at the bottom of Fig. 1), and the c-section rate among women who went into labor was also declining. The c-section rate increased steadily after 1996, reaching 36.8 percent in 2006. The rise in the c-section rate was driven primarily by a more-than-doubling of the elective (or scheduled) c-section rate, from 10.9 percent in 1996 to 27.1 percent of all deliveries in 2006. Although the c-section rate in New York (Fig. 2) is consistently lower than in Florida, the pattern over time is very similar in the two states.

C-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 Table 1 we report the mean c-section rate in the 11 Florida local health districts and the eight New York health service areas in 2003, adjusted for patient health characteristics and health insurance status.[4] The c-section rate in Florida ranges from 18.7 percent to 27.3 percent, and regions with a high overall c-section rate tend to have both a relatively high scheduled c-section rate and a relatively high c-section rate conditional on a woman going into labor. The coefficients of variation in New York are similar to those in Florida for the overall c-sections and c-sections conditional on labor, and higher for scheduled c-sections.

Chetty (1998) develops a theoretical model of the decision to use c-section versus vaginal delivery in which some women (and/or their babies) will have better outcomes with a vaginal delivery and others with a c-section. Physicians choose a risk threshold such that women with values below this threshold receive a vaginal delivery,

[4] The 11 local health districts in Florida were formed by the state for purposes of health planning (see http://www.flhealthplanning.org/). The eight health services areas in New York were formed for a similar purpose. Each consists of one or more contiguous counties. We refer to them generically as markets or regions.

Table 1
Variation in obstetrical treatment styles between regions, 2003.

	Risk-adjusted c-section rate	Risk-adjusted scheduled c-section rate	Risk-adjusted c-section rate conditional on labor
Florida local health district			
1	0.187	0.068	0.151
4	0.192	0.072	0.149
7	0.203	0.084	0.151
3	0.205	0.080	0.159
6	0.217	0.081	0.169
2	0.217	0.090	0.166
8	0.220	0.086	0.171
5	0.229	0.090	0.174
9	0.231	0.102	0.170
10	0.257	0.114	0.190
11	0.273	0.118	0.209
Mean	0.221	0.089	0.169
Std Dev	0.026	0.016	0.018
CoV	0.12	0.18	0.11
New York health service area			
4	0.142	0.040	0.122
3	0.143	0.057	0.111
2	0.146	0.063	0.108
5	0.157	0.068	0.119
1	0.162	0.071	0.119
7	0.179	0.076	0.139
6	0.188	0.088	0.138
8	0.200	0.103	0.142
Mean	0.165	0.071	0.125
Std Dev	0.022	0.019	0.013
CoV	0.13	0.27	0.11
Florida and New York combined			
Mean	0.197	0.082	0.150
Std Dev	0.037	0.019	0.027
CoV	0.19	0.24	0.18

Note: We use variables displayed in Table 2 to adjust the probability a patient receives a c-section for her health, the status of the pregnancy, and type of insurance coverage. Std Dev is standard deviation; CoV is coefficient of variation.

and women with values above this threshold receive a c-section. Because physicians estimate a woman's true risk with error, some c-sections will be performed when a vaginal delivery would have produced a better outcome, ex post, and vice versa with some vaginal deliveries. If physicians were perfect agents, they would choose a risk threshold such that the expected cost to the patient of a c-section when, ex post, a vaginal delivery would have produced a superior outcome, is equal to the expected cost of a vaginal delivery when a c-section would have produced a superior outcome.

Phelps and Mooney (1993) propose that physicians form beliefs about the appropriateness and effectiveness of medical technologies during medical school and residency training, similar to the views expressed by Wennberg and Gittelsohn (1973), Wennberg (1985), and Stano (1993). In our context, variation in the c-section rate within a region would occur if physicians choose different risk thresholds, controlling for observed patient characteristics, either due to information they acquired during their formal medical education or due to their idiosyncratic views. Physicians who attach a high cost to performing a vaginal delivery when a c-section was merited, ex post, will choose a low risk threshold and will have a high c-section rate, while physicians who attach a high cost to performing a c-section when a vaginal delivery was merited will choose a high threshold and will have a low rate.

If residency programs produce physicians with distinctive treatment styles and graduates locate unevenly across markets, then treatment styles will also vary between regions. Phelps and Mooney (1993) hypothesize that physicians will update their prior beliefs regarding the appropriateness of a technology or treatment method based on how their colleagues treat similar patients. The treatment styles of a physician's peer group serve as an inexpensive source of clinical information for continuing education. If this form of learning has a strong effect, physicians' treatment styles would converge within a market to a single standard, but would differ between markets.

The substantial changes in the c-section rate that occurred between 1992 and 2006 make obstetrics a good case study for physician learning. There is no consensus view regarding which factors were the key drivers of the changes displayed in Figs. 1 and 2. Pressure from managed care organizations and a concerted effort by physicians and hospitals to meet the Healthy People 2000 goal of a 15 percent c-section rate may have contributed to the decrease in the c-section rate between 1992 and 1996. The subsequent increase in the c-section rate may have been driven by changes in the fee physicians received for delivering by c-section versus vaginally (Gruber et al., 1999), physician-induced demand as a response to declining income (Gruber and Owings, 1996), an increase in the probability of being sued for malpractice (Marieskind, 1979; Grant and McInnes, 2004), and changes in patients' preferences, possibly driven by studies showing that there may be long-term negative health consequences associated with vaginal delivery (Rortveit et al., 2003). Furthermore, the American College of Obstetrics and Gynecology (ACOG) regularly disseminates treatment recommendations in the form of "practice bulletins."[5] For example, all 17 of ACOG's recommended clinical guidelines related to the manage-

[5] In 2000, for example, the American College of Obstetricians and Gynecologists proposed a c-section benchmark of 15.5 percent of women at 37 weeks of gestation or greater, having their first child, with a single fetus in the vertex (non-breech) position (American College of Obstetricians and Gynecologists, 2000). The report contained a number of specific recommendations to help obstetricians meet this goal.

ment of labor and delivery posted on the Agency for Healthcare Quality and Research (AHRQ) National Guideline Clearinghouse website have been created or revised since 2003.[6]

How do physicians learn about these new developments and determine whether or not to alter their treatment decisions as a result? Physicians generally acquire information on new medical technologies from medical journals, professional meetings, conferences, and informal conversations with their colleagues (Stinson and Mueller, 1980). There is evidence that physicians obtain substantial information from physicians they interact with regularly, as well as from the hospital where they practice (Coleman et al., 1966; Nair et al., 2008; Burke et al., 2007; Chung et al., 2003). Escarce (1996), for example, finds that once a general surgeon began performing laparoscopic cholecystectomies, it promoted early adoption among other surgeons at the same hospital. Many hospitals have implemented programs to encourage obstetricians to adhere to recommended clinical guidelines. Chaillet et al. (2006) review 33 such programs, 16 of which were trying to reduce c-sections, and conclude that audit and feedback programs, computer or paper reminders, and mixed strategies are more effective than academic detailing, programs where opinion leaders counsel other physicians, and general education programs.

We begin with a general question: how closely correlated is the change in a physician's risk-adjusted c-section rate with the change in the rate of his local peers—physicians who practice at the same hospital(s)? We then try to measure separately the extent to which changes in physician treatment styles are driven by: (1) factors that should affect all physicians in a region, such as changes in reimbursement incentives and the malpractice environment; (2) local factors such as changes in patient preferences and hospital-initiated programs to change physicians' treatment decisions; and (3) plausibly exogenous changes in the treatment decisions of a physician's local peers. It is difficult to measure separately the second and third factors because physicians in a peer group usually practice in the same hospital. Our empirical method, which is described in Sections 3 and 4, uses three distinct peer groups to try to separate the causal effect of a change in the treatment decisions of a peer group from market forces and hospital programs that affect both a physician and his peer group.

3. Data and measurement of treatment styles and peer groups

3.1. Data

We construct our sample from the 1992–2006 Florida and New York hospital discharge data sets, which contain information on 6.7 million deliveries performed at all non-federal, short-term acute care hospitals. Although we do not have access to medical records, we do observe information on a patient's demographic characteristics, pre-delivery health, and services received during and after delivery, including: the mother's age, race, ethnicity, and insurance coverage (e.g., HMO), codes for her primary diagnosis and secondary diagnoses, procedure codes that specify 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 discharged. Sample means and standard deviations for the patient-level data set are reported in Table 2.

The diagnosis 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, and whether the labor occurred before the fetus was full-term).[7] We use a method developed by Henry et al. (1995) and Gregory et al. (2002) to determine whether a woman went into labor. Women who delivered vaginally or had diagnosis codes indicating fetal distress, labor abnormalities, cord prolapse, or a breech converted to vertex presentation were interpreted as having gone into labor; all other women were coded as having a scheduled c-section.

Because the data contain all hospital discharges with consistent physician identifiers, 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, whether he graduated from an international medical school, his self-reported specialty, the residency program(s) where he received training, and the year he completed obstetrics residency training. We use the latter information to create a variable for years of post-residency experience.

Of the 6.7 million delivery discharges in the original data set, 6.3 million (95.2 percent) were performed by 15,340 physicians with non-missing license numbers, and 6.0 million (90.2 percent) were performed by 6261 physicians with license numbers that matched to the AMA Masterfile. In order to include data on all deliveries when constructing peer group practice styles, deliveries performed by physicians with missing license numbers, with license numbers not matched to the AMA Masterfile, or with fewer than 50 deliveries in that year were assigned to the hospital in which they were performed. We limited subsequent analyses of the determinants of physicians' practice styles, however, to data from physicians with non-missing license numbers who performed at least 50 deliveries in a given year in order to increase the precision of the practice styles. The final analytic data set contains information on the practice styles of 6097 physicians who performed at least 50 deliveries in at least one year and collectively performed 5.9 million deliveries (88.6 percent of the original total).

3.2. Measuring physicians' treatment styles

We would like our measures of treatment style to characterize how each physician would treat the same set of patients. Absent a randomized design, we develop risk-adjusted treatment styles that control for differences across physicians in the observed characteristics of their patients. For each physician in each year of practice, we separately measure the overall c-section rate, elective c-section rate, and c-section rate conditional on labor.

To derive a physician's overall risk-adjusted c-section rate, for example, we estimate the following linear probability model, separately for each year between 1992 and 2006:

$$C_{ij} = \boldsymbol{\alpha X}_i + \boldsymbol{YJ} + \varepsilon_{ij} \tag{2}$$

C_{ij} equals one if patient i received a c-section by physician j and is zero otherwise. We include patient characteristics $\mathbf{X}_i$, such as the patient's age, type of health insurance, existing medical conditions (e.g., severe hypertension), and the status of the pregnancy (e.g., multiple gestation, preterm gestation, antepartum bleeding, whether the woman had a c-section in a prior delivery) that may affect the risks and/or benefits of a c-section. We also include a full set of physician indicator variables $\mathbf{J}$ and do not include a constant. The coefficients on the physician indicators ($\hat{\boldsymbol{Y}}$) measure the

[6] Accessed http://www.guideline.gov/browse/guideline_index.aspx on August 19, 2008.

[7] Two diagnoses that are frequently associated with a c-section, fetal distress and abnormal labor, are fairly subjective, so we do not include these as control variables.

Table 2
Sample means and standard deviations by state in delivery-level data set, 1992–2006.

Variable	Florida		New York	
	Mean	Std Dev	Mean	Std Dev
Patient demographics				
Age	27.0	6.2	28.4	6.2
Race/ethnicity				
White	0.541	0.498	0.503	0.500
Black	0.210	0.407	0.165	0.371
Hispanic	0.185	0.388	0.130	0.337
Race other/missing	0.064	0.244	0.201	0.401
Patient health insurance				
Commercial	0.486	0.500	0.557	0.497
Medicaid	0.414	0.492	0.376	0.484
Uninsured	0.075	0.263	0.039	0.194
Insurance other	0.025	0.158	0.028	0.166
Health of mother or baby				
Woman had a previous c-section	0.133	0.340	0.127	0.333
Malpositioned fetus	0.062	0.241	0.060	0.238
Antepartum bleeding	0.018	0.132	0.017	0.128
Severe hypertension	0.012	0.108	0.009	0.093
Preterm gestation	0.072	0.259	0.067	0.251
Multiple gestation	0.010	0.100	0.013	0.114
Maternal soft tissue disorder	0.024	0.154	0.027	0.162
Macrosomia	0.030	0.170	0.019	0.136
Oligohydramnios	0.006	0.077	0.006	0.075
Polyhydramnios	0.023	0.150	0.037	0.188
Herpes	0.016	0.125	0.010	0.097
Uterine scar	0.002	0.041	0.002	0.041
Uterine rupture	0.001	0.026	0.001	0.026
Unengaged fetal head	0.012	0.110	0.017	0.128
Congenital fetal CNS anomaly	0.001	0.031	0.001	0.028
Cerebral hemorrhage	0.00005	0.007	0.00005	0.007
Diabetes	0.009	0.093	0.005	0.069
Chorioamnionitis	0.020	0.141	0.018	0.131
Ruptured membrane > 24 h	0.013	0.112	0.021	0.142
Maternal hypotension	0.001	0.034	0.001	0.032
Intrauterine growth restriction	0.016	0.125	0.012	0.111
Maternal heart disease	0.012	0.109	0.015	0.123
Asthma	0.014	0.120	0.017	0.131
Maternal renal abnormality	0.001	0.035	0.001	0.033
Other maternal infection	0.013	0.113	0.008	0.091
N	2,902,064		3,764,161	

probability that a particular physician will perform a c-section, controlling for observed characteristics of the mother's health and the status of the pregnancy.

We estimate Eq. (2) with ordinary least squares rather than a probit or logit model. Nonlinear regression has some disadvantages for our application. Probit/logit estimates of $\hat{\boldsymbol{Y}}$ are potentially inconsistent when the number of observations per physician is finite (Neyman and Scott, 1948; Chamberlain, 1980).[8] In order to interpret $\hat{\boldsymbol{Y}}$ we would need to convert these estimates to marginal effects. Also, estimating probit or logit models using our sample would take a prohibitively long time with available computing resources. In contrast, OLS estimates are consistent, easily interpretable as a percentage point change in the probability a woman will receive a c-section, and quick to generate. Regardless, our results appear to be robust to using an alternative model. For the 1,142 physicians in 2003 in Florida, for example, we compared the $\hat{\boldsymbol{Y}}$ from an OLS regression with the $\hat{\boldsymbol{Y}}$ from a logit regression by calculating a Pearson correlation coefficient, which was 0.95.

In Table 3 we present selected coefficient estimates from the OLS estimation for Florida for 2003. Because the main purpose of this regression is to recover the coefficients on the physician indicator variables ($\hat{\boldsymbol{Y}}$), we do not discuss the coefficients on the patient characteristics ($\boldsymbol{\alpha}$) at great length. Each coefficient can be interpreted as the average difference in a patient's probability of receiving a c-section associated with a one-unit difference in the independent variable. Most of the displayed health conditions substantially increase the chance a woman will have a c-section. For example, a woman with severe hypertension has a probability of receiving a c-section that is 33 percentage points higher than a woman without that condition.

We perform similar regressions to derive a physician's risk-adjusted elective c-section rate and risk-adjusted c-section rate conditional on labor. The latter regression is estimated only on women who went into labor, and the dependent variable is one if she received a c-section and zero if she delivered vaginally.[9] After we run a separate regression like the one reported in Table 3 for each of the 15 years, three practice style measures and two states (i.e., 90 regressions total), we recover the physician coefficients and create a panel data set where observations are at the level of physician *j*'s practice style for year *t*.

One concern is that substantial unobserved patient heterogeneity may cause the physician risk-adjusted c-section rates to

[8] Specifically, the inconsistency of $\boldsymbol{\alpha}$ is transmitted to the fixed effect coefficients.

[9] Results from these regressions are available from the authors by request.

Table 3
Selected coefficients from a patient-level linear probability model, overall c-section rate, Florida 2003.

Variable	Coefficient	Std Err
Patient demographics		
Age	−0.008	0.001
Age squared	0.0002	0.00002
Race/ethnicity		
Black	0.004	0.002
Hispanic	0.008	0.002
Patient health insurance		
Medicaid	−0.018	0.002
HMO	−0.001	0.002
Uninsured	−0.037	0.004
Health of mother or baby		
Woman had a previous c-section	0.670	0.002
Malpositioned fetus	0.470	0.003
Antepartum bleeding	0.316	0.006
Severe hypertension	0.333	0.007
Preterm gestation	−0.004	0.003
Multiple gestation	0.205	0.008
Maternal soft tissue disorder	0.165	0.005
Macrosomia	0.323	0.004
Oligohydramnios	0.132	0.009
Polyhydramnios	0.135	0.005
Herpes	0.183	0.006
Uterine scar	0.429	0.015
Uterine rupture	0.151	0.039
Unengaged fetal head	0.678	0.006
Congenital fetal CNS anomaly	0.126	0.023
Cerebral hemorrhage	0.245	0.145
Diabetes	0.069	0.007
Chorioamnionitis	0.220	0.006
Ruptured membrane >24 h	0.045	0.007
Maternal hypotension	0.068	0.022
Intrauterine growth restriction	0.103	0.006
Maternal heart disease	0.029	0.007
Asthma	0.035	0.006
Maternal renal abnormality	0.037	0.019
Other maternal infection	0.075	0.007
Observations	202,616	
R^2	0.63	
Mean of dependent variable	0.33	

Notes: Dependent variable is an indicator for whether a woman received a c-section. 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, whether her age is missing, and various other health measures. Indicators are included for eight other health conditions (e.g., thyroid abnormality).

differ more than they would if we had a more complete set of demand-side control variables. This could occur, for example, if certain physicians specialize in treating patients with poor unobserved health, or if there are ethnic differences in preferences for c-sections and patients of the same ethnic group tend to pick the same physicians. The first-difference models we use will eliminate the component of unobserved patient-heterogeneity that is fixed at the physician level in adjacent years.

Another concern is that the practice style measures may be noisy, imprecise estimates of a physician's true style. This could occur if the component of patient health that physicians observe and we do not has a substantial effect on treatment decisions, and the mean unobserved health of a physician's patient population changes substantially from year to year due in part to small patient volume. Limiting the sample to patients of physicians who perform at least 50 deliveries per year should reduce the possibility of spurious treatment style measures. We find that the correlation between a physician's treatment style in year t and year $t-1$ is high, ranging from 0.68 for the c-section rate conditional on labor to 0.79 for the overall c-section rate.

Kane and Staiger (2002) develop a test to measure how much of the change in school test scores is due to non-persistent variation, such as high teacher-turnover in a school during a particular year, the chemistry between a teacher and her students, or test-taking conditions. If the correlation of the change in test scores in adjacent years is negative and large (i.e., a class that experienced a large increase from year 1 to year 2 tends to have a large decrease in scores from year 2 to year 3), most of the change is transitory; when the correlation is close to zero, most of the change is persistent and the scores are an accurate measure of performance.[10] Applying this test, we find that the proportion of the change in physician treatment styles that is attributable to persistent factors ranges from 0.10 for the elective c-section rate to 0.32 for the overall c-section rate. We therefore focus primarily on the overall c-section rate when examining how physicians learn over time.

3.3. Defining local and regional peer groups

We divide Florida into 11 health care markets using the Florida Department of Health's definition of local health districts and divide New York into eight markets as defined by the New York Department of Health's definition of health service areas. Each market consists of one or more contiguous counties. These markets are large enough so that 92 percent of the deliveries occur at a hospital in the same market where the mother resides. For each physician we construct a measure of the practice style of his local and regional peers. We define physician j's local peer group in year t as all physicians other than physician j who delivered a baby at the same hospital(s) as physician j in year t.[11] Physician j's regional peer group in year t includes all physicians who delivered a baby at a hospital in the same market as physician j other than the hospital(s) where physician j delivered.[12] We use the physician-specific practice style coefficients for the members of physician j's local and regional peer groups to create delivery-weighted local and regional peer group treatment styles for each physician for each year, 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).[13]

The median number of physicians in a local peer group is 33 and in a regional peer group is 244. We posit that a physician is more likely to exchange information with his local than his regional peers, but is likely to be affected by some of the same market shocks and policy changes as his regional peers. We use this assumption to try to separate the influence of the local exchange of information from policies that affect all physicians in a market. In a single year, 60 percent of the physicians in our data set delivered all their babies at a single hospital, 30 percent divided their deliveries between two hospitals, and 10 percent at three or more hospitals.

[10] Specifically, the proportion of the change in the performance measure that is due to non-persistent factors is equal to −2 multiplied by the correlation of the change in adjacent years (Kane and Staiger, 2002). This test assumes that the permanent and transitory shocks in adjacent years are independent.

[11] For each year, we pooled the experience of individual physicians who performed fewer than 50 deliveries in that year in hospital-specific residual groups and included these residuals in the peer group practice style calculations.

[12] A physician who delivers in both a hospital where physician j delivers as well as a hospital in the same region where physician j does not deliver will appear in both physician j's local and regional peer groups. Although this will not bias our empirical estimates, it will make it more difficult to measure the separate effects of regional policies and of the local exchange of information. Fortunately, 60% of physicians in our sample deliver at a single hospital.

[13] Physicians may, in fact, interact with a subset of the other physicians who deliver babies at the same hospital (e.g., members of their group practice). If so, the resulting measurement error would attenuate the coefficient on the local peer group toward zero.

Table 4
Sample means and standard deviations in physician-year-level data set.

	Florida & New York combined		Florida only		New York only	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Physician characteristics						
Female	0.298	0.458	0.237	0.425	0.334	0.472
Gender information missing	0.035	0.185	0.025	0.155	0.042	0.200
Non-white	0.198	0.399	0.206	0.405	0.193	0.395
Race information missing	0.324	0.468	0.304	0.460	0.336	0.472
International medical graduate	0.289	0.453	0.216	0.411	0.331	0.471
IMG information is missing	0.047	0.212	0.033	0.179	0.055	0.229
Post-residency experience (years)	13.0	8.8	13.0	8.5	13.1	8.9
Specialty						
Ob/Gyn	0.878	0.327	0.885	0.319	0.874	0.332
Family practice/internal medicine	0.010	0.100	0.008	0.092	0.011	0.105
Maternal and fetal medicine	0.015	0.123	0.015	0.120	0.016	0.124
Information missing	0.053	0.224	0.039	0.193	0.061	0.240
Physician's practice characteristics						
Number of deliveries	144	103	173	122	127	85
Unadjusted c-section rate	0.278	0.104	0.291	0.104	0.271	0.103
Risk-adjusted c-section rate	0.156	0.089	0.176	0.094	0.145	0.085
Unadjusted elective c-section rate	0.170	0.091	0.179	0.093	0.165	0.090
Risk-adjusted elective c-section rate	0.067	0.063	0.062	0.060	0.070	0.064
Unadjusted c-section rate | labor	0.132	0.063	0.138	0.063	0.128	0.063
Risk-adjusted c-section rate | labor	0.119	0.070	0.144	0.075	0.104	0.062
Local peer group's practice characteristics						
Risk-adjusted c-section rate	0.153	0.070	0.171	0.078	0.142	0.064
Risk-adjusted elective c-section rate	0.065	0.048	0.061	0.046	0.067	0.050
Risk-adjusted c-section rate | labor	0.116	0.051	0.140	0.059	0.102	0.040
Local peer group's first differences	(N = 33,608)		(N = 12,600)		(N = 21,008)	
Overall risk-adjusted c-section rate	0.011	0.032	0.012	0.033	0.011	0.031
Stayers' risk-adjusted c-section rate	0.010	0.037	0.011	0.040	0.010	0.035
Enterers' – Exiters' risk-adjusted c-section rate	0.001	0.024	0.001	0.026	0.001	0.022
Overall risk-adjusted elective c-section rate	0.009	0.028	0.008	0.027	0.010	0.029
Stayers' risk-adjusted elective c-section rate	0.009	0.028	0.008	0.027	0.009	0.028
Enterers' – Exiters' risk-adjusted elective c-section rate	0.001	0.012	0.001	0.011	0.001	0.013
Overall risk-adjusted c-section rate | labor	0.004	0.029	0.007	0.030	0.003	0.028
Stayers' risk-adjusted c-section rate | labor	0.004	0.032	0.006	0.035	0.003	0.030
Enterers' – Exiters' risk-adjusted c-section rate | labor	0.0003	0.018	0.0002	0.021	0.0004	0.016
Regional peer group's practice characteristics						
Risk-adjusted c-section rate	0.152	0.061	0.169	0.069	0.143	0.053
Risk-adjusted elective c-section rate	0.064	0.041	0.058	0.038	0.068	0.042
Risk-adjusted c-section rate | labor	0.116	0.043	0.139	0.053	0.103	0.029
Regional peer group's first differences	(N = 33,608)		(N = 12,600)		(N = 21,008)	
Risk-adjusted c-section rate	0.011	0.025	0.012	0.027	0.011	0.024
Risk-adjusted elective c-section rate	0.009	0.024	0.008	0.023	0.010	0.025
Risk-adjusted c-section rate | labor	0.004	0.024	0.007	0.025	0.003	0.023
N (unless specified otherwise)	41, 089		15, 017		26, 072	

Note: There are 6097 physicians in the panel data set. 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. A physician's risk-adjusted elective c-section rate and c-section rate conditional on labor are likewise coefficients on physician indicator variables where the dependent variable is one if a woman received an elective c-section and one if she received a c-section conditional upon going into labor, respectively.

In some regression specifications we decompose physician j's local peer group into physicians who remained in his peer group between year $t-1$ and t, physicians who were in the peer group in $t-1$ but exited in year t, and physicians who were not in the peer group in $t-1$ but entered in year t. We argue below that the practice styles of the latter two constituents are more likely to represent an exogenous shock to physician j's information set than the change in the practice styles of the former constituents.

We present the means and standard deviations of the variables in the physician-year-level data set in Table 4. This data set contains 6097 physicians representing 41,089 physician-years. At the physician-year level, 30 percent of the physicians are women, 20 percent are non-white, 29 percent are international medical graduates, and 88 percent are obstetricians/gynecologists. The physicians have 13 years of post-residency experience and performed 144 deliveries per year, on average.

4. Analysis and results

4.1. Amount of within-market variation in physician treatment styles

The first objective of this paper is to measure the amount of variation in treatment styles among obstetricians practicing in the same market. We argued above that a patient who is searching for a physician will care more about the variation in treatment styles among physicians within a market than across markets. In order to quantify the variation that exists between physicians within a

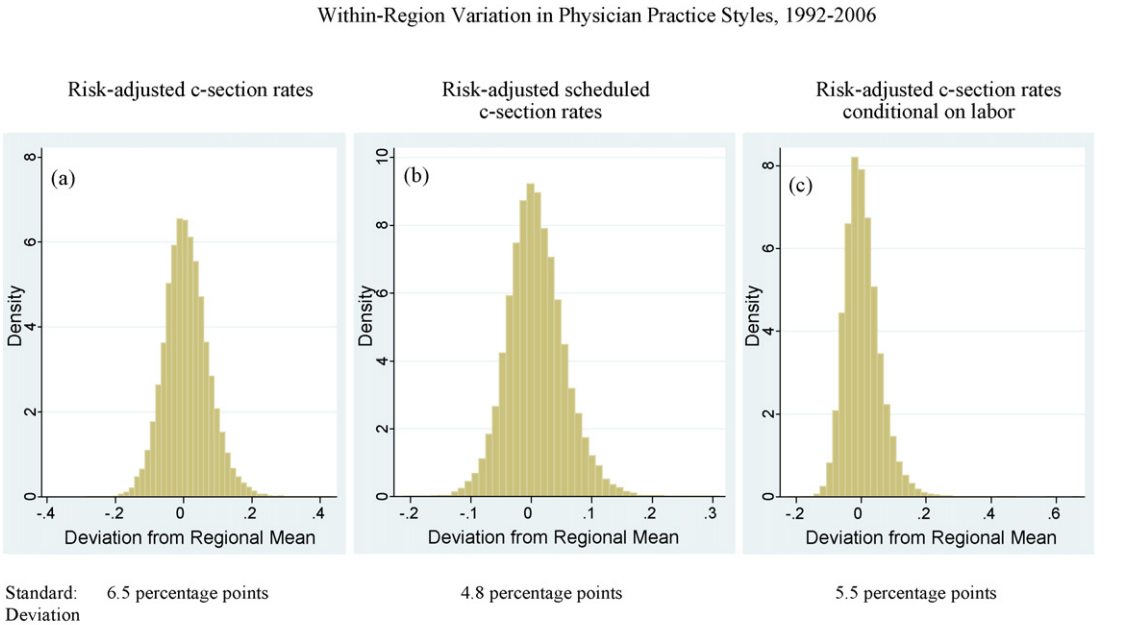

Fig. 3. Within-region variation in physician practice styles, 1992–2006; (a) risk-adjusted c-section rates; (b) risk-adjusted scheduled c-section rates; (c) risk-adjusted c-section rates conditional on labor. *Note*: an observation is the deviation between a physician's practice style and the regional average for a particular year. There are 6097 physicians and 41,089 physician-years in the above figures. Physician practice styles are risk-adjusted based on patients' observed characteristics.

market, we calculate the difference between a physician's treatment style ($\hat{Y}_{jt}$, as estimated in Eq. (2)) and the mean treatment style for the physician's market in year t.[14] We then calculate the standard deviations of the within-market practice style measures (constructed as the physician-level style minus the market-level style) for the overall risk-adjusted c-section rate, the risk-adjusted elective c-section rate, and the risk-adjusted c-section rate conditional on labor, and compare these to the between-market standard deviations.

We pool the risk-adjusted treatment style differences between physician j and the market average over the 1992–2006 time period and depict the distribution separately for the three treatment styles in Fig. 3. In 24 percent of the cases the physician's risk-adjusted c-section rate in the first panel of Fig. 3 is statistically different from the regional mean at the five percent level.[15] The standard deviation of the within-market variation in risk-adjusted c-section rates is 6.5 percentage points, about 75% larger than the variation across regions at a point in time (as reported in Table 1). The within-market variations for the elective c-section rate and c-section rate conditional on labor are also about twice as large as the between-market variation. Two identical women who live in the same market and choose their obstetricians randomly are likely to have very different ex ante probabilities of receiving a c-section.[16] If there is a single correct risk-adjusted c-section rate that all women prefer, then this within-market variation reduces welfare. We cannot, however, rule out the possibility that patients have preferences for different treatment styles and choose physicians accordingly, in which case some component of the within-market variation would enhance welfare.

One concern is that sampling error may exaggerate the amount of variation that actually exists between physicians in their treatment styles. That is, unobserved (to us but not to the physician) patient characteristics may create the appearance of differences across physicians in treatment styles when in fact the true differences are small. Unmeasured characteristics will be most problematic for physicians who perform a relatively small number of deliveries. To address this concern we use a shrinkage technique (Hofer et al., 1999; Kane and Staiger, 2002) that adjusts the treatment styles of physicians who have low signal-to-noise ratios toward the year-specific market mean.[17] After application of the shrinkage technique, the standard deviation of the within-market variation in risk-adjusted c-section rates is 4.9 percentage points, still 32% larger than the variation across markets.

[14] We derive the market risk-adjusted c-section rate by replacing the physician indicator variables in Eq. (2) with a complete set of 11 indicators for the Florida local health districts or eight indicators for the New York health service areas. The coefficient on a market indicator variable is the risk-adjusted c-section rate for that market in that year.

[15] This is similar to the result in Phelps (2000) where one-third of the primary care physicians have mean expenditures per patient that differ from the market average.

[16] Grant (2005, p. 717) likewise finds that the within-region variation in physician risk-adjusted c-section rates in Florida in 1992 was almost as large as the between-region variation.

[17] A physician's practice style is assumed to consist of a signal, Y_j, and a noise component, u_j : $\hat{Y}_j = Y_j + u_j$. The coefficients $\hat{Y}_j$ are estimates of a physician's practice style and are likely to be measured with error, particularly for physicians who perform a relatively small number of deliveries. We can measure the variance of the noise component for each physician from the covariance matrix of Eq. (2). If we assume the variance of Y_j, the true practice style of a physician, is the same for all physicians, we can estimate a unique signal to noise ratio for each physician in the data set. We derive a "filtered" estimate of each physician's practice style by taking a weighted average of each physician's coefficient $\hat{Y}_j$ and the sample mean $\hat{Y}$ for the physician's market, where the weights are a physician's signal to noise ratio (which ranges from zero to one) and one minus this ratio, respectively (McClellan and Staiger, 1999). The coefficients for physicians with a small number of deliveries will therefore be shifted toward the sample mean, whereas the coefficients for physicians who perform a large number of deliveries will not be affected much by the shrinkage method.

Table 5
Sources of variation in physicians' practice styles.

	Incremental R^2 (Florida and New York combined)					
	All physician years (n = 41,089)			Physician years with fewer than 4 years experience (n = 5572)		
	C-section rate	Elective c-section rate	C-section rate | labor	C-section rate	Elective c-section rate	C-section rate | labor
A: Total variation explained	0.777	0.716	0.676	0.869	0.833	0.808
Incremental variation due to						
B: Unmeasured, time-invariant physician-specific factors	0.283	0.275	0.305	0.396	0.384	0.428
C: Year physician is practicing	0.358	0.336	0.209	0.277	0.310	0.163
D: Market where physician practices	0.034	0.017	0.057	0.047	0.021	0.073
E: Physician characteristics: gender, race, experience, specialty and IMG status	0.018	0.015	0.013	0.020	0.009	0.020
F: Residency program indicators	0.021	0.021	0.021	0.030	0.028	0.031
G: Variation common to the above variables	0.064	0.052	0.071	0.101	0.081	0.094
R^2						
Regression 1: MD and year indicators	0.777	0.716	0.676	0.869	0.833	0.808
Regression 2: MD characteristics and year indicators	0.494	0.441	0.371	0.474	0.449	0.380
Regression 3: regression 2 without year indicators	0.136	0.105	0.162	0.197	0.139	0.217
Regression 4: regression 2 without market indicators	0.461	0.424	0.314	0.427	0.428	0.307
Regression 5: regression 2 without MD characteristics	0.477	0.426	0.358	0.454	0.440	0.360
Regression 6: regression 2 without residency indicators	0.474	0.420	0.350	0.444	0.421	0.349
N	41,089			5572		

Notes: Observations are physician-year practice style measures, pooled from 1992 to 2006. There are 6097 unique physicians in the data set, of whom 3015 had fewer than four years of experience at some point. The practice style measures are adjusted for observed patient characteristics such as patient's age and health. To account for this, models were estimated with generalized least squares using weights based on Borjas (1987). The incremental R^2 for the physician indicators is the difference between the R^2 when physician and year indicators are included in the regression and the R^2 when the physician indicators are omitted but all other variables are included. Indicators were included for the 159 residency programs in which five or more physicians received their obstetrical/gynecology residency training.

4.2. Sources of variation in physician treatment styles

Our second objective is to examine why treatment styles differ between physicians within the same market. For purposes of estimating the determinants of variation in treatment styles, an observation is a practice style measure for a physician-year ($\hat{Y}_{jt}$). We pool the observations across the 15 years and 6097 physicians and estimate a series of OLS regressions. We use the incremental R^2 method proposed by Theil (1971) to decompose the total explained variation into incremental components that can be attributed to the year, region, measured physician characteristics, location of residency training, and all other unmeasured factors specific to a physician. We separately examine all physicians and inexperienced physicians – those with fewer than 4 years of post-residency experience – to test whether the impacts of residency training and physician characteristics diminish over a physician's career.

These models will have heteroskedasticity, because the dependent variables, $\hat{Y}_{jt}$, are the physician fixed effect coefficients estimated in the first stage, patient-level regressions (Eq. (2)) and thus the practice styles of high-volume physicians will generally be estimated more precisely than for low-volume physicians (Borjas, 1987). We therefore estimate these models with generalized least squares, where the weight for each physician-year observation increases with the precision of the coefficient from Eq. (2) (Borjas, 1987).[18] There is very little difference in the results when the models are estimated with OLS.

We first regress $\hat{Y}_{jt}$ on a full set of physician and year indicator variables. The R^2 from this regression, reported at the bottom of Table 5 in the "Regression 1" row as well as in row A of Table 5, represents the total amount of variation in treatment styles that can be explained collectively by all physician-specific and year-specific factors. Between 68 (for the risk-adjusted c-section rate conditional on labor) and 78 percent (for the overall risk-adjusted c-section rate) of the variation in practice styles can be explained by these two broad sets of factors.

We then repeat the regression described above after replacing the physician indicator variables with specific (and measurable) physician characteristics: indicator variables for a physician's gender, race, specialty, attending medical school outside the United States, specific residency program attended, region/market, a continuous variable measuring years of post-residency experience, and experience squared. We include indicator variables for the 159 residency programs that trained five or more physicians in our sample. This R^2 is reported in the "Regression 2" row at the bottom of Table 5.

The difference between the R^2 from the first and second regressions, reported in row B of Table 5, is the incremental variation due to factors that are included in the first regression but not the second—i.e., time-invariant, physician-specific factors *other than* those we explicitly control for in the second regression (i.e., gender,

[18] Specifically, the weight is one divided by the sum of the squared standard error of $\hat{Y}_{jt}$ from Eq. (2) and the difference between the sum of squared errors from an OLS estimate of the second stage model and the sum of a physician's squared standard errors from Eq. (2) for all years, divided by the degrees of freedom. See footnote 21 in Borjas (1987).

race, specialty, medical school location, residency program, practice location, and experience). Almost 30 percent of the variation in treatment styles among physicians is due to these time-invariant, physician-specific factors. That is, almost one-third of the differences between physicians in how they treat patients are due to characteristics of physicians that are not easily measured, such as risk aversion, skill, and attitudes about the benefits of medical treatments. If the values in row B were instead close to zero, it would indicate that physician treatment styles are determined by demographics and their training, specialty, and practice location choices. Because we have fairly detailed information on patient characteristics, our interpretation is that a considerable amount of practice variation is due to idiosyncratic physician perceptions regarding how patients should be treated.[19] 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, because they have a much more limited set of risk adjusters (the gender mix and average age of a physician's patients only), there is likely to be more unobserved patient heterogeneity across practices in their sample than in ours.

In order to measure the amount of variation explained by time, we perform a third regression after removing the year indicator variables from the specification of the second regression, and measure the difference in the R^2 between the second and third regressions. The year indicators explain 21–36 percent of the variation in treatment styles, as reported in row C of Table 5. This highlights that practice styles changed systematically over this time period. After restoring the year indicators, we then successively omit from the second regression specification the region indicators, the physician characteristics, and the residency program indicators, and measure an incremental R^2 for each of these factors.

Regional indicators (row D of Table 5) account for only two to six percent of the variation. Likewise, little of the variation in treatment styles is explained by observed physician characteristics (gender, race, specialty, attending medical school outside the United States) or the specific residency program where a physician trained. The former set of characteristics collectively explain less than two percent of the variation and the latter about two percent.

Apparently residency programs do not produce obstetricians with distinct views regarding how patients should be treated. One alternative explanation is that the time-invariant coding of residency programs misses important variation over time in the styles that residency programs impart to their trainees. Another alternative explanation for the apparent unimportance of residency training is that the treatment style imprinted by their residency programs dissipates once physicians begin practicing. We would 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 treatment decisions. To test this, we repeat the R^2 decomposition after restricting the sample to observations where the physician has fewer than four years of post-residency experience. These results are reported in the final three columns of Table 5. Relative to the all-physician analysis, only slightly more of the variation in treatment styles among inexperienced physicians is attributed to the specific residency program where a physician trained; residency programs still explain only three percent of the variation.

[19] Our data are not complete, however. Variations in unobservable (to us) patient preferences and clinical conditions may encourage physicians to specialize (Epstein et al., 2008), which would also lead to practice style variation.

4.3. Physician learning and changes in treatment styles over time

The third objective of this paper is to determine whether physicians learn from their immediate colleagues and if this learning leads to substantial changes in treatment styles over time. Manski (1993) notes that one of the major challenges of empirically estimating social interactions is to measure separately the effect on an individual's behavior of his peer group's behavior (in our context the change in the treatment style of a physician's peers), unobserved characteristics shared by members of a peer group (e.g., technical skill or risk aversion), and the environment in which the members of a peer group operate (e.g., changes in reimbursement incentives or the probability of being sued for malpractice).

If one regresses a physician's treatment style on the style of his peer group, the coefficient on the latter variable would capture the net effect of all three factors. We address this problem by estimating first-difference models to eliminate unobserved, time-invariant characteristics that a physician may share in common with his peer group (e.g., unmeasured patient health), and including two mutually exclusive peer group variables rather than a single peer variable. We begin by performing a regression where the dependent variable is the change in physician j's treatment style between year $t-1$ and year t ($\Delta\hat{Y}_{jt}$), and the complete specification has the following form:

$$\Delta\hat{Y}_{jt} = \gamma_0 + \gamma_1\,\Delta L_{jt} + \gamma_2\,\Delta R_{jt} + \gamma_3\,\Delta L_{jt-1} + \gamma_4\,\Delta R_{jt-1} + \gamma_5 S + \boldsymbol{\gamma_6 T} + \boldsymbol{\gamma_7} S * \boldsymbol{T} + v_{jt} \quad (3)$$

S is an indicator variable for New York and $\mathbf{T}$ is a set of year indicators that capture, along with the state-year interactions, the trend in how babies were delivered in Florida and New York over the 1992–2006 time period, including the effect of state policies and changes in clinical guidelines. ΔL_{jt} is the contemporaneous change in the treatment style for physician j's local peer group (physicians who perform deliveries at the same hospital(s) as physician j), and ΔR_{jt} is the contemporaneous change in the treatment style of physician j's regional peers (physicians in j's market who deliver babies at hospitals *other than* those where physician j delivers babies). In order to examine how quickly physicians respond to information, we include measures of the lagged changes in peer group treatment styles, ΔL_{jt-1} and ΔR_{jt-1}.

The dependent variables in the second stage (Eq. (3)), $\Delta\hat{Y}_{jt}$, are constructed from the physician fixed effect coefficients estimated in the first stage, patient-level regressions (Eq. (2)).[20] This will create heteroskedasticity in Eq. (3), as discussed above (Borjas, 1987). We therefore employ heteroskedasticity-robust standard errors, and also adjust the standard errors for clustering by physician to allow the error terms to be correlated across years. We check the sensitivity of the standard errors two ways: by bootstraping the standard errors of Eq. (3) using 1000 iterations, and by estimating Eq. (3) with generalized least squares (Borjas, 1987). There is very little difference in the standard errors among these three approaches.[21]

The coefficient γ_1 measures the change in physician j's treatment style associated with a contemporaneous change in the

[20] The peer group treatment styles are also constructed from first stage physician fixed effect coefficients. As existing approaches to the generated regressors problem (e.g., Hoffman, 1987; Gawande, 1997) are computationally burdensome and we are aware of no correction that applies to the particular case here (i.e., generated regressand and generated regressors that are constructed as weighted averages of first stage coefficients), however, we focus on dealing with the generated regressand problem.

[21] A few of the peer effect coefficient p-values change at the third decimal point. The complete set of regressions is available from the authors by request.

Table 6
First difference coefficient estimates of a physician's risk-adjusted overall c-section rate.

	A	B	C
Change in local peer group's rate, $t-1$ to t	0.451*** [0.019]	0.450*** [0.019]	
Change in local peer group's rate, $t-2$ to $t-1$	0.065*** [0.015]	0.064*** [0.015]	
Change in regional peer group's rate, $t-1$ to t		0.050 [0.039]	0.070* [0.040]
Change in regional peer group's rate, $t-2$ to $t-1$		0.055 [0.038]	0.114*** [0.038]
Enterer (t) – exiter $(t-1)$ rate			0.068*** [0.015]
Enterer $(t-1)$ – exiter $(t-2)$ rate			−0.014 [0.014]
Constant	0.011*** [0.002]	0.012*** [0.002]	0.016*** [0.002]
Observations	27765	27765	27765
R^2	0.20	0.20	0.17
Mean of dependent variable	0.0146	0.0146	0.0146

Note: All models include indicators for year, state and year-state interactions. Robust standard errors adjusted for physician-level clustering in brackets.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

treatment style of his local peer group. In column A of Table 6 we report coefficient estimates from a specification of Eq. (3) where the two regional peer group terms are omitted and the dependent variable is the change in physician j's overall risk-adjusted c-section rate. The coefficient of 0.45 on ΔL_{jt}, which is identified by variation between physicians in the treatment style change of their local peers, indicates that changes in physician treatment styles at the local level are highly correlated; they tend to move together, although not perfectly. A one percentage point increase in the overall c-section rate of a physician's immediate colleagues is associated with an increase of 0.45 percentage points in his own rate. Moreover, the coefficient on ΔL_{jt-1}, is about one-seventh as large as the coefficient on ΔL_{jt}. This indicates that physicians adjust their treatments styles quickly in response to local changes.

The γ_1 coefficient in this specification captures both the causal effect of a change in the practice style of a physician's local peers on a change in his own practice style *and* the effect of policies and shocks (e.g., changes in patients' preferences or changes in private health insurance reimbursement incentives) on changes in the treatment styles of all physicians in the same market. We try to measure separately these two effects by including the contemporaneous and lagged regional peer group variables. Consider a situation where members of a physician's local and regional peer groups are exposed to a common set of policies and shocks, such as a change in the state's Medicaid reimbursement policy for c-sections or a change in the malpractice environment. The coefficient γ_2 will capture the impact of market-level policies and shocks on a physician's treatment style, and the coefficient γ_1 will measure the effects of information exchange among local physicians and local shocks/policies.[22]

In column B of Table 6 we report the regression results of a specification that includes the two regional peer group variables along with the two local peer variables. The coefficients on the contemporaneous and lagged regional peer group change variables are positive and small in magnitude, and not significant. Furthermore, including the regional peer variables has very little effect on the γ_1 and γ_3 coefficients. Local factors appear to have a stronger influence on obstetricians' treatment styles than market-wide policies and shocks.

If there are policies, programs, and shocks that affect a physician's local peers differently than his regional peers, such as a change in patient preferences by a subset of the patients in a market, a change in reimbursement incentives for a single physician group practice, or a guidelines-adherence program initiated by a single hospital, then γ_1 will be an upward-biased estimate of the causal effect of local learning due to the exchange of information between physicians who practice in the same hospital(s). This is a legitimate concern because the majority of physicians in the sample practice in a single hospital, and hospitals often initiate programs to change physicians' treatment decisions (Chaillet et al., 2006).

To address this, for each physician we decompose ΔL_{jt} into two components as follows:[23]

$$\Delta L_{jt} = \left[\left(\sum_{i=1}^{S_t} n_{ti} + \sum_{i=1}^{N_t} n_{ti}\right)^{-1}\left(\sum_{i=1}^{S_t} n_{ti}Y_{ti}\right) - \left(\sum_{i=1}^{S_{t-1}} n_{t-1,i} + \sum_{i=1}^{X_{t-1}} n_{t-1,i}\right)^{-1}\left(\sum_{i=1}^{S_{t-1}} n_{t-1,i}Y_{t-1,i}\right)\right] + \left[\left(\sum_{i=1}^{S_t} n_{ti} + \sum_{i=1}^{N_t} n_{ti}\right)^{-1}\left(\sum_{i=1}^{N_t} n_{ti}Y_{ti}\right) - \left(\sum_{i=1}^{S_{t-1}} n_{t-1,i} + \sum_{i=1}^{X_{t-1}} n_{t-1,i}\right)^{-1}\left(\sum_{i=1}^{X_{t-1}} n_{t-1,i}Y_{t-1,i}\right)\right] \quad (4)$$

The first row above is the change in the treatment style between $t-1$ and t for physicians who were in physician j's local peer group in both $t-1$ and t ("stayers"); the second row is the difference between the treatment style in t of physicians who entered physician j's peer group in t but were not in the group in $t-1$ and the treatment style of physicians in $t-1$ who were in the peer group in $t-1$ but not in t. In Eq. (4), n_{ti} is the number of deliveries performed by physician i in year t, S refers to physicians who remain in the peer group, N to physicians who enter the peer group, and X to physicians who exit the peer group. Each peer physician's practice style,

[22] Hong et al. (2005) use a similar method. They allow the stock-holding decisions of a mutual fund manager to be affected differently by the decisions of other managers in the same geographic market and managers in more remote geographic markets.

[23] To simplify exposition, the formula as presented applies to a single hospital. For physicians practicing at multiple hospitals in a given market in year t, we weight each hospital's contribution by the share of physician j's delivery volume in that year at that hospital (and we weight similarly when a physician practices in more than one market in a given year).

Table 7
First difference coefficient estimates of a physician's risk-adjusted elective c-section rate.

	A	B	C
Change in local peer group's rate, $t-1$ to t	0.511*** [0.019]	0.509*** [0.019]	
Change in local peer group's rate, $t-2$ to $t-1$	0.081*** [0.016]	0.080*** [0.016]	
Change in regional peer group's rate, $t-1$ to t		0.0002 [0.042]	0.052 [0.043]
Change in regional peer group's rate, $t-2$ to $t-1$		0.068* [0.039]	0.142*** [0.040]
Enterer (t) – exiter ($t-1$) rate			0.159*** [0.023]
Enterer ($t-1$) – exiter ($t-2$) rate			−0.067*** [0.025]
Constant	0.014*** [0.001]	0.015*** [0.002]	0.025*** [0.002]
Observations	27765	27765	27765
R^2	0.28	0.28	0.26
Mean of dependent variable	0.0107	0.0107	0.0107

Note: All models include indicators for year, state and year-state interactions. Robust standard errors adjusted for physician-level clustering in brackets.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Y_{ti}, is weighted by the proportion of the peer group's deliveries in that year performed by physician i.[24]

The turnover of obstetricians that occurred in Florida and New York over the sample period helps identify the coefficient on the entering/exiting physician practice style variable. During our 15-year study period, each year an average of 255 obstetricians exited the sample because they stopped practicing, stopped delivering babies, or moved out of state, and 237 obstetricians entered the sample for the first time. A small number of obstetricians also moved to a different market in the same state: 3.3 percent of the physicians in our sample drew a majority or plurality of their annual caseload from more than one region (local health district in Florida or health service area in New York) during their tenure in the data set.

If the practice styles of physicians who enter and exit a market are independent with respect to the practice styles of physicians who remain in the market, then the second row of Eq. (4) represents an exogenous change in the treatment decisions that physician j is exposed to that is unrelated to new local policies and programs faced by physician j. The data generally support this assumption. The correlation between the overall risk-adjusted c-section rate in year t of physicians who remain in a local peer group and physicians who enter a peer group in year t is −0.01, and the correlation in year $t-1$ between the "stayers" and the physicians who exit a peer group is −0.03. That is, local peer groups that have a relatively high (or low) c-section rate do not tend to attract or repel physicians with relatively high (or low) c-section rates.[25]

In column C of Table 6 we report coefficient estimates of a specification of Eq. (3) where we replace ΔL_{jt} and $\Delta L_{j,t-1}$ with the contemporaneous (entering – exiting) peer variable (i.e., the second row of Eq. (4)), as well as the lagged version of the variable. The coefficient on the contemporaneous (entering – exiting) physician practice style variable (0.068) is positive and significant, but much smaller than the coefficient on ΔL_{jt} in the second column. The local exchange of information between physicians has a statistically significant but rather small effect on obstetricians' treatment styles. Consider an obstetrician with the mean risk-adjusted c-section rate (0.156). A one-standard deviation (2.4 percentage points) increase in the (entering – exiting) c-section rate is associated with a 0.16 percentage point (or 1.0 percent) increase in his own rate. Apparently, most of the association between the change in a local peer group's treatment style and the change in an individual physician's style is due to policies and programs experienced by all physicians. This small effect of a change in a physician's information set is consistent with our results in Table 5 that a substantial amount of the variation between physicians in their treatment styles is due to idiosyncratic differences in their perceptions regarding treatment efficacy and appropriateness. That is, physicians are not easily persuaded to change their treatment style based on observing how other physicians are treating patients, but they are by hospital administrators or other local policy changes.

The regional peer group coefficients are slightly larger and more precisely estimated in this specification. The estimated effect of a regional policy on a physician's own c-section rate is still rather small, however. A policy that increases the change in the c-section rate of a physician's regional peers by one standard deviation (2.5 percentage points) is predicted to increase a physician's own rate by 1.1 percent in the first year and 1.8 percent in the second year. These results indicate that it takes physicians longer to respond to a regional than a local change.

In Table 7 we present a similar set of regressions where the dependent variable is the change in a physician's risk-adjusted elective c-section rate – the proportion of deliveries that are c-sections and the woman does not go into labor, and in Table 8 regressions where the dependent variable is the change in a physician's risk-adjusted c-section rate conditional on a woman going into labor. The results in are quite similar to those in Table 6: the coefficient on ΔL_{jt} is positive and large, the coefficient on $\Delta L_{j,t-1}$ is much smaller (but still positive and significant), neither of the local peer style coefficients change much when the regional peer variables are included, the regional coefficients are positive, small, and often insignificant, and the coefficient on the (entering – exiting) variable is much smaller than on the change in the local peer group variable (ΔL_{jt}). One should be cautious, however, in attributing a causal interpretation to the (entering – exiting) coefficient in column C of Table 7 because the correlation between the elective c-section rate in year t of physicians who remain in a local peer group and physicians who enter a peer group in year t is 0.17, and the correlation in year $t-1$ between the "stayers" and the physicians who exit a peer group is 0.22. That is, the elective c-section practice styles of physicians who enter and exit appear to be related to the incumbent physicians.[26]

[24] We thank an anonymous referee for suggesting this approach. Note that the total number of physicians is $S_t + N_t$ in year t and $S_{t-1} + X_{t-1}$ in year $t-1$. An algebraic development of Eq. (4) is available from the authors by request.

[25] This evidence is not conclusive, however. It is possible that in response to a local policy change some physicians "move" immediately and others later, which would be masked in comparisons of contemporaneous "movers" and "stayers." Although it is difficult to test this without explicit data on local policy changes, we compared practice styles across physicians entering a peer group in years t, $t+1$ and $t+2$ as well as across physicians exiting in the same years to test how consistently similar "enterers" and "exiters" are over time. The cross-year correlations are fairly small (between 0.1 and 0.2 everywhere) for overall c-section, elective c-section and c-section after laboring.

[26] The analogous correlations for the c-section rate conditional on labor for Table 8 are −0.03 and −0.04.

Table 8
First difference coefficient estimates of a physician's risk-adjusted c-section rate conditional on labor.

	A	B	C
Change in local peer group's rate, $t-1$ to t	0.408*** [0.021]	0.407*** [0.021]	
Change in local peer group's rate, $t-2$ to $t-1$	0.085*** [0.018]	0.085*** [0.018]	
Change in regional peer group's rate, $t-1$ to t		0.016 [0.049]	−0.014 [0.051]
Change in regional peer group's rate, $t-2$ to $t-1$		0.056 [0.047]	0.097** [0.048]
Enterer (t) − exiter ($t-1$) rate			0.073*** [0.019]
Enterer ($t-1$) − exiter ($t-2$) rate			0.023 [0.018]
Constant	−0.003 [0.002]	−0.002 [0.002]	−0.007*** [0.002]
Observations	27765	27765	27765
R^2	0.19	0.19	0.17
Mean of dependent variable	0.0074	0.0074	0.0074

Note: All models include indicators for year, state and year-state interactions. Robust standard errors adjusted for physician-level clustering in brackets.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

We performed two sensitivity checks.[27] First, we re-estimated the three model specifications in Tables 6–8 after limiting the sample to physicians who practiced in a single market during their entire tenure in the data set (thus removing the impact of physicians who experienced severe peer group changes by relocating to another market in the same state). The results are qualitatively unchanged. To test the consistency of the effects of peer groups over calendar time, we also re-estimated the models after including interactions between the peer group variables and indicator variables for five-year blocks (1997–2001 and 2002–2006, omitting 1992–1996). The positive and significant contemporaneous local peer group effects and (enterer − exiter) effects are larger at the start of the study period and get smaller over time. These findings may reflect a geographic widening in the sources of information for and influences on physician practices; with the advent of the Internet and other information-sharing technologies, local peers appear to influence physicians' practice styles less.[28]

5. Conclusions

In this paper we investigate how obstetricians form their treatment styles, and whether and how much their styles evolve over time. We assemble a comprehensive data set that contains the universe of inpatient births in Florida and New York over a 15-year period matched with detailed information on the physicians performing the deliveries. We construct annual measures of each physician's propensity to provide cesarean section that control for a range of patient health and demographic characteristics. We then examine the influence of physician-level attributes on practice styles, focusing on the role of residency training, learning from peers, and market factors. We explore the relative contributions of factors that should affect all physicians in a region, such as changes in reimbursement incentives and the malpractice environment (the "regional" peer group); local factors such as changes in patient preferences and hospital-initiated programs to change physicians' treatment decisions (the "local" peer group); and plausibly exogenous changes in the treatment decisions of a physician's local peers (the subset of the "local" peer group that recently exits or enters).

As with most medical care treatments, differences in the mean c-section rates between the 11 regions in Florida and eight regions in New York, controlling for patients' observed characteristics, are quite large. In 2003, for example, the risk-adjusted probability a woman in Florida would have a cesarean section ranged from a low of 0.188 in Northwest Florida to a high of 0.273 in Miami. We show that there are even larger differences in c-section rates among physicians within a region; the standard deviation of the c-section rate across physicians within a region is almost twice as large as the between-region variation, controlling for observed patient characteristics. If there is there is a single correct risk-adjusted c-section rate that all women prefer, within-market variation reduces welfare. If, on the other hand, patients have preferences for different treatment styles and choose physicians accordingly, part of the within-market variation enhances welfare.

Physician demographic factors and training experiences appear to have a small effect on inter-physician variation in the type of medical care received. About 2 percent of the variation in risk-adjusted c-section rates among all physicians can be explained by where they trained as a resident. Likewise, observable physician characteristics, such as experience, gender, and race, account for between 1 and 2 percent of the variation in c-section rates. Nearly 30 percent of the variation in risk-adjusted c-section rates among all physicians is due to time-invariant, physician-specific factors separate from observable physician characteristics, residency training and region. The explanatory power of these factors is somewhat higher among new physicians; in particular, residency training explains around 4 percent of the variation. Because we have fairly 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.

Although physicians' practice styles do not appear to be influenced much by shocks at the regional level, they are fairly responsive to changes in practice styles among other physicians in the same hospitals. Physicians' own c-section rates increased about half a percentage point for an increase of 1 percentage point in the local peer groups' c-section rate. However, most of this effect is attributable to common shocks at the hospital level; the influence of information exchange with local peers is fairly small. The effect of a plausibly exogenous 1 percentage point increase in the local peer group's overall c-section rate led to a predicted 0.068 percentage point change in own overall c-section rates, although the analogous effect for elective c-section rates was about double.

Physicians appear to be quite independent. Treatment styles are not strongly shaped by residency training programs and are rather impervious to market-specific shocks and changes in peer treatment styles. Although physicians learn from their peers, they do not substantially revise their prior beliefs regarding how patients should be treated due to the local exchange of information. We do find, however, that physicians respond quickly to local, hospital-level shocks and to national forces (as evidenced by our finding in Table 5 that the year variables explain a considerable amount of the variation), both of which may reflect changes in clinical guide-

[27] Results are available by request.
[28] We thank an anonymous referee for this suggestion.

lines and patient preferences. Taken together, these results imply that cross-sectional variations in regional treatment rates that are commonly observed are unlikely to dissipate substantially over time, nor is the within-market variation across physicians likely to dissipate over time due to convergence to a community standard.

Acknowledgements

This paper has benefited from comments by Jim Burgess, Jason Fletcher, Frances Frei, Michael Grossman, Steven Lehrer, George Macones, Will Manning, Jon Skinner, and many other seminar participants. We thank the University of Pennsylvania Research Foundation and the Leonard Davis Institute for Health Economics for providing funding to acquire the data.

References

Banerjee, A.V., 1992. A simple model of herd behavior. Quarterly Journal of Economics 107 (3), 797–817.
Bikhchandani, S., Hirschleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100 (5), 992–1026.
Bikhchandani, S., Chandra, A., Goldman, D., Welch, I., 2001. The economics of iatroepidemics: Bayesian learning, informational cascades and geographical variation. Working Paper.
Borjas, G.J., 1987. Self-selection and the earnings of immigrants. American Economic Review 77 (4), 531–553.
Burke, M.A., Fournier, G.M., Prasad, K., 2007. The diffusion of a medical innovation: is success in the stars? Southern Economic Journal 73 (3), 588–603.
Chaillet, N., Dube, E., Dugas, M., Audibert, F., Tourigny, C., Fraser, W.D., Dumont, A., 2006. Evidence-based strategies for implementing guidelines in obstetrics: a systematic review. Obstetrics & Gynecology 108 (5), 1234–1245.
Chamberlain, G., 1980. Analysis of covariance with qualitative data. Review of Economic Studies 48 (1), 225–238.
Chassin, M.R., Brook, R.H., Park, R.E., Keesey, J., Fink, A., Kosecoff, J., Kahn, K., Merrick, N., Solomon, D.H., 1986. variations in the use of medical and surgical services by the medicare population. The New England Journal of Medicine 314 (5), 285–290.
Chetty, V.K., 1998. Stochastic technology, production organization and costs. Journal of Health Economics 17 (2), 187–210.
Chung, P.J., Chung, J., Shah, M.N., Meltzer, D.O., 2003. How do residents learn? The development of practice styles in a residency program. Ambulatory Pediatrics 3 (4), 166–172.
Coleman, J.S., Katz, E., Menzel, H., 1966. Medical Innovation: A Diffusion Study. Bobbs-Merrill Company, New York.
DeFrances, C.J., Hall, M.J., 2007. 2005 National hospital discharge survey. Advance Data from Vital and Health Statistics, 385, No.
Ellison, G., Fudenberg, D., 1993. Rules of thumb for social learning. Journal of Political Economy 101 (4), 612–643.
Ellison, G., Fudenberg, D., 1995. Word-of-mouth communication and social learning. Quarterly Journal of Economics 110 (1), 93–125.
Epstein, A., Ketcham, J., Nicholson, S., 2008. Professional partnerships and matching in obstetrics. NBER Working Paper 14070.
Escarce, J.J., 1996. Externalities in hospitals and physician adoption of a new surgical technology: an exploratory analysis. Journal of Health Economics 15, 715–734.
Gawande, K., 1997. Generated regressors in linear and nonlinear models. Economics Letters 54, 119–126.
Goyert, G.L., Bottoms, S.F., Treadwell, M.C., Nehra, P.C., 1989. The physician factor in cesarean birth rates. The New England Journal of Medicine 320 (11), 706–709.
Grant, D., 2005. Information and sorting in the market for obstetrical services. Health Economics 14, 703–719.
Grant, D., McInnes, M.M., 2004. Malpractice experience and the incidence of cesarean delivery: a physician-level longitudinal analysis. Inquiry 41, 170–188.
Gregory, K.D., Korst, L.M., Gornbein, J.A., Platt, L.D., 2002. Using administrative data to identify indications for elective primary cesarean delivery. Health Services Research 37 (5), 1387–1401.
Gruber, J., Kim, J., Mayzlin, D., 1999. Physician fees and procedure intensity: the case of cesarean delivery. Journal of Health Economics 18, 473–490.
Gruber, J., Owings, M., 1996. Physician financial incentives and cesarean section. Rand Journal of Economics 27, 99–123.
Grytten, J., Sorensen, R., 2003. Practice variation and physician-specific effects. Journal of Health Economics 22 (3), 403–418.
Henry, O.A., Gregory, K.D., Hobel, C.J., Platt, L.D., 1995. Using ICD-9 codes to identify indications for primary and repeat cesarean sections: agreement with clinical records. American Journal of Public Health 85 (8), 1143–1145.
Hofer, T.P., Hayward, R.A., Greenfield, S., Wagner, E.H., Kaplan, S.H., Manning, W.G., 1999. The unreliability of individual physician 'report cards' for assessing the costs and quality of care of a chronic disease. The Journal Of the American Medical Association 281 (22), 2098–2105.
Hoffman, D.L., 1987. Two-step generalized least squares estimators in multi-equation generated regressor models. The Review of Economics and Statistics 69 (2), 336–346.
Hong, H., Kubik, J.D., Stein, J.C., 2005. Thy neighbor's portfolio: word-of-mouth effects in the holdings and trades of money managers. Journal of Finance 60, 2801–2824.
Kane, T.J., Staiger, D.O., 2002. In: Ravitch, D. (Ed.), Volatility in School Test Scores: Implications for Test-Based Accountability Systems, Brookings Papers on Education Policy. Brooking Institution, Washington, D.C., pp. 235–283.
Manski, C., 1993. Identification of endogenous social effects: the reflection problem. Review of Economic Studies 60, 531–542.
Marieskind, H.I., 1979. An Evaluation of Cesarean Section in the United States. Final Report Submitted to the Department of Health, Education and Welfare Office of the Assistant Secretary for Planning and Evaluation/Health. GPO, Washington, D.C.
McClellan, M., Staiger, D., 1999. The quality of health care providers. NBER Working Paper 7327.
Nair, H., Manchanda, P., Bhatia, T., 2008. Asymmetric social interactions in physician prescription behavior: the role of opinion leaders. Working Paper.
Newhouse, J.P., 2002. Why is there a quality chasm? Health Affairs 21 (4), 13–25.
Neyman, J., Scott, E.L., 1948. Consistent estimates based on partially consistent observations. Econometrica 16 (1), 1–32.
Phelps, C.E., Parente, S.T., 1990. Priority setting for medical technology and medical practice assessment. Medical Care 28 (8), 703–723.
Phelps, C.E., Mooney, C., 1993. Variations in medical practice use: causes and consequences. In: Arnould, R.J., Rich, R.F., White, W. (Eds.), Competitive Approaches to Health Care Reform. Urban Institute Press, Washington, D.C..
Phelps, C.E., 2000. Information diffusion and best practice adoption. In: Cuyler, A.J., Newhouse, J.P. (Eds.), Handbook of Health Economics. Elsevier Science, Amsterdam.
Roos, N.P., Flowerdew, G., Wajda, A., Tate, R.B., 1986. Variations in physicians' hospitalization practices: a population-based study in Manitoba, Canada. American Journal of Public Health 76 (1), 45–51.
Rortveit, G., Daltveit, A.K., Hannestad, Y.S., Hunskaar, S., 2003. Urinary incontinence after vaginal delivery or cesarean section. The New England Journal of Medicine 348 (10), 900–907.
Scharfstein, D.S., Stein, J.C., 1990. Herd behavior and investment. American Economic Review 80 (2), 465–479.
Stano, M., Folland, S., 1988. Variations in the use of physician services by medicare beneficiaries. Health Care Financing Review 9 (3), 51–58.
Stano, M., 1993. Evaluating the policy role of the small area variations and physician practice style hypotheses. Health Policy 24 (1), 9–17.
Stinson, E.R., Mueller, D.A., 1980. Survey of health professionals' information habits and needs. Conducted through personal interviews. The Journal Of the American Medical Association 243 (2), 140–143.
Theil, H., 1971. Principles of Econometrics. John Wiley & Sons, Inc., New York, NY.
Welch, H.G., Miller, M.E., Welch, W.P., 1994. Physician profiling: an analysis of inpatient practice patterns in Florida and Oregon. The New England Journal of Medicine 330 (9), 607–612.
Wennberg, J.E., Gittelsohn, A., 1973. Small area variations in health care delivery. Science 182, 1102–1108.
Wennberg, J.E., Freeman, J.L., Culp, W.J., 1987. Are hospital services rationed in new haven or over-utilised in Boston? Lancet 1, 1185–1189.
Wennberg, J.E., Fisher, E.S., Skinner, J., 2002. Geography and the Debate Over Medicare Reform, Health Affairs, Web Exclusive, February 13, 2002.
Wennberg, J.E., 1985. On patient need, equity, supplier-induced demand, and the need to assess the outcome of common medical practices. Medical Care 23 (5), 512–520.