Geographic Variation In The Appropriate Use Of Cesarean Delivery

Do higher usage rates reflect medically inappropriate use of this procedure?

by Katherine Baicker, Kasey S. Buckles, and Amitabh Chandra

ABSTRACT: There is enormous geographic variation in the use of cesarean delivery: For births over 2,500 grams, adjusted cesarean rates vary fourfold between low- and high-use areas. Even for births under 2,500 grams, high-use counties have rates that are double those of low-use ones. Higher cesarean rates are only partially explained by patient characteristics but are greatly influenced by nonmedical factors such as provider density, the capacity of the local health care system, and malpractice pressure. Areas with higher usage rates perform the intervention in medically less appropriate populations—that is, relatively healthier births—and do not see improvements in maternal or neonatal mortality. [*Health Affairs* 25 (2006): w355–w367; 10.1377/hlthaff.25.w355]

THE CESAREAN SECTION RATE IN THE UNITED STATES is much higher than that in other countries.[1] Even within the United States, taking patients' risk factors into account, some areas use cesareans at much higher rates than others do: In 1996–2000, in Phoenix, Arizona, there were only fifteen cesareans per hundred births over 2,500 grams, while in Long Island, New York, there were twenty-six per hundred births. This is a source of national concern: The Centers for Disease Control and Prevention's (CDC's) Healthy People 2010 initiative has the explicit goal of reducing the cesarean birth rate.[2] This objective is predicated on the belief that high rates of cesarean delivery reflect procedure use in mothers and infants who obtain little benefit from the procedure. In the extreme, higher procedure rates might even be associated with iatrogenic harm, stemming from surgical complications that are not offset by therapeutic benefit.[3]

Uncovering the relationship between areawide intensity of use and the medical

Katherine Baicker is an associate professor in the Department of Public Policy, School of Public Affairs, at the University of California, Los Angeles. Kasey Buckles is an assistant professor of economics at the University of Notre Dame, in Notre Dame, Indiana. Amitabh Chandra (Amitabh_Chandra@harvard.edu) is an assistant professor of public policy in the John F. Kennedy School of Government, Harvard University, in Cambridge, Massachusetts. Baicker and Chandra are also affiliated with Dartmouth Medical School in Hanover, New Hampshire, and are faculty research fellows at the National Bureau of Economic Research in Cambridge, Massachusetts.

DOI 10.1377/hlthaff.25.w355

appropriateness of care requires a reliable way to measure appropriateness. There are major methodological challenges in using ex-post evaluation to determine the pervasiveness of inappropriate care. Typically, a systematic review of the literature is combined with the opinions of an expert clinical panel to score patients on a scale that measures appropriateness for a given procedure.[4] Several studies have found that patients with the highest appropriateness scores benefit most from the intervention, and one notes that the views of practicing physicians are similar to those of an expert panel.[5] Critics of this approach, however, note that often the variables used to define *appropriateness* have not been validated.[6] In addition, criteria developed by different expert panels have been shown to exhibit enormous variability, particularly for procedure use classified as inappropriate, and to be greatly influenced by the composition of the panel.[7]

We introduce a new methodology to determine whether higher cesarean rates reflect less medically appropriate use of the procedure. We use the correlation between patient characteristics and whether or not that patient receives a cesarean section to construct a predicted probability of cesarean birth (PPC). For each birth, this is the probability that the typical obstetrician would perform a cesarean delivery, based on the patient's prebirth characteristics, and removing the effect of area characteristics that are unchanging over time. This measure has strengths and weaknesses: It can tell us which births are collectively viewed by doctors as being better candidates for a cesarean, but it cannot be used to infer the cutoff for a medically "appropriate" versus an "inappropriate" cesarean. Like all measures of medical appropriateness, it cannot inform us about whether the choice of cesarean delivery is associated with nonmedical factors such as patients' preferences; *appropriateness* in our paper refers only to medical appropriateness. Our index is not, however, subject to the arbitrary nature of measures calculated by panels or to the biases inherent in retrospective classification based on procedure outcomes. Nor, as we demonstrate below, is it confounded by characteristics of the area that are fixed over time. It is easily implemented and can be applied to any medical procedure.

Our analysis contains three parts. First, we demonstrate geographic variation in the use of cesarean delivery across different U.S. cities. We examined the correlates of higher use by studying the relationship between cesarean rates and birth characteristics, county socioeconomic characteristics, local provider capacity, and medical malpractice liability.[8] The larger the role played by nonclinical factors such as provider supply and malpractice liability, the more likely it is that the marginal cesarean birth occurs for less medically driven reasons.

Second, we test the hypothesis that areas with higher cesarean rates are performing the intervention in births that are less medically appropriate for the use of cesarean delivery. If our hypothesis is correct, then the typical cesarean birth in more aggressive areas will have a lower PPC than the marginal cesarean birth in less aggressive areas, because physicians would have worked their way down the distribution of patient appropriateness. Hence, this hypothesis implies a negative relationship between area cesarean rates and the average appropriateness of use of the procedure.

Third, we explore the hypothesis that even though areas with high cesarean rates might perform the procedure in less appropriate births, they might be better at the use of this procedure and achieve improved outcomes. To examine this theory, we studied the relationship between area cesarean rates and neonatal and maternal mortality ratios.

Study Data And Methods

■ **Data.** We used the National Center for Health Statistics (NCHS) linked birth and infant death data, pooled across the years 1995–1998, to calculate risk-adjusted county cesarean rates and the PPC (N = 15,592,980). We used these years of the birth data because we were able to obtain concurrent data on maternal mortality, medical malpractice, and other county characteristics. We selected the 10.2 million births that occurred during this period in the 198 U.S. counties with populations greater than 250,000 in the 1990 census because county of birth is not identified on birth certificates in smaller counties. Even if it were, sampling errors would make it difficult to calculate maternal and neonatal mortality for smaller geographic locales. Washington, D.C., was dropped from the analysis because medical malpractice data were not available for this area.

We classified infants as low birthweight (LBW) if their birthweight was less than 2,500 grams (n = 800,109) and as normal birthweight (NBW) otherwise (n = 9,361,844). Even with the large sample sizes at our disposal, we were unable to reliably estimate risk-adjusted cesarean rates separately for very-low-birthweight (VLBW) infants (those less than 1,500 grams, n = 151,480) in smaller cities. We therefore combined this group with LBW infants and use the term "LBW" in the text to designate all babies born under 2,500 grams. We defined *cesarean deliveries* to include both primary (n = 1,337,130) and repeat cesarean (n = 757,657). Using this definition, 20.6 percent of our sample had a cesarean delivery. This reflects a rate of 37.3 percent for LBW births and a rate of 19.2 percent for NBW births.[9] In two unreported secondary analyses, we repeated our analysis using (1) a subsample of our data where repeat cesarean and vaginal birth after cesarean (VBAC) deliveries were excluded from the analysis, and (2) a sample of first births (where the decision to perform a repeat cesarean is not possible). Both secondary analyses yielded results that were virtually identical to the full-sample results reported below.

Data on physician and hospital resources, including the total number of physicians, pediatricians, and obstetrician-gynecologists (OB-GYNs) per birth; neonatal intensive care unit (NICU) beds per birth; and Medicaid share of inpatient days, are from the Area Resource Files (ARF). We also obtained data on county characteristics such as population, per capita income, urban classification, and demographic composition from the ARF.

We measured malpractice liability pressures in two different ways. First, we included the number and size of malpractice payments (arising from judgments and settlements and measured separately for surgery, obstetrics, and internal medicine) per physician in each state, ascertained from the National Practitioner Data Bank (NPDB). Our choice of these measures was motivated by research finding that physicians respond to both the number of claims and the average size of malpractice awards: Being sued imposes costs on physicians, including lost time at work and psychic costs.[10] Ideally, we would have used claims per physician, but there is no nationally representative source for these data. This limitation would cause us to understate the potential role of malpractice liability. Second, we constructed a measure of malpractice liability pressure based on average physician malpractice premiums, as reported in the *Medical Liability Monitor,* a national survey of insurers. This measure addresses concerns that some payments may be missed by the NPDB and the fact that payments reported to the NPDB reflect claims filed a few years ago. A further advantage of using malpractice premiums as a measure of malpractice liability is that it reflects insurers' estimates of open and future claims—a factor that will be missed by the NPDB but might still affect physicians' practice style. Our measures are not mechanically correlated; indeed, other research has argued that a number of factors, including the interaction of state regulatory oversight and the insurance underwriting cycle, determine premiums.[11]

County-level maternal mortality was calculated from the NCHS's multiple-cause-of-death mortality data from 1995 to 1998.[12] We counted any woman in an area ages 10–54 for whom "complications of pregnancy, childbirth, and the puerperium" is listed as the primary cause of death. We calculated maternal mortality ratios by dividing the count of maternal deaths by the number of live births that occurred in each county during the same time period.

■ **Analyses.** *Correlates of area-level cesarean rates.* We calculated unadjusted county-level cesarean rates and evaluated the correlates of this county-level intensity. We report analysis of variance (ANOVA) results, which explain the variance of these rates using four sets of covariates in a prespecified order that allows birth and socioeconomic status (SES) characteristics to have maximum explanatory power; doing so minimizes the role of nonmedical factors such as provider capacity and malpractice pressure: (1) an index of patient-level characteristics (calculated using predictions from a regression model of cesarean delivery on variables in the birth certificate data); (2) county-level measures of socioeconomic factors (including average income, unemployment rate, percentage living in poverty, percentage urban, percentage white, percentage of the population with less than a high school degree, high school and college degrees, percentage of hospital patient days eligible for Medicaid, and size of the population); (3) county-level provider characteristics (including total physicians, surgeons, pediatricians, OB-GYNs, internists, and other specialists per birth, as well as neonatal intensive care beds per birth); and (4) state-level medical malpractice liability (including the number and size of judgments and settlements by medical specialty and malpractice premiums by specialty).[13]

Area usage rates and the predicted probability of cesarean birth. The second step in our analysis was to correlate risk-adjusted area-level variations in cesarean usage rates with our measure of the average appropriateness of patients who received the procedure. We hypothesized that in areas with higher cesarean rates, the intervention is being used for births that are less medically appropriate for it. If this were true, we should observe a negative relationship between the PPC for all births by cesarean and area cesarean rates. Note that this is not a mechanical relationship—both a positive relationship and no relationship are also possible. The former would occur if physicians are first performing cesareans based on some unobservable characteristic (for example, patients' income or insurance coverage), regardless of medical appropriateness, before moving on to more appropriate patients. The latter would occur if some areas are uniformly more aggressive and perform cesareans without regard to clinical appropriateness. No relationship would also be observed if half of physicians performed cesareans by the first effect and half by the second.

We computed risk-adjusted county cesarean rates by estimating a regression model of the probability that an individual infant is delivered via cesarean on patient-level covariates and indicator variables for each of the identified counties in our data. The coefficients on the county indicator variables estimated the risk-adjusted probability of receiving a cesarean delivery in each particular county.[14]

The PPC was also obtained from this regression, but it relies on only the portion of the prediction that relies on birth certificate data; it excludes the county fixed effect.[15] Thus, the PPC measures which births are more likely to involve a cesarean, based on clinical characteristics but independent of the local practice style. The regressions that generated these rates yielded an R^2 of 0.39 for NBW babies and 0.32 for LBW babies. Thus, we are able to explain at least as much of the variation in cesareans as previous research, which finds an R^2 of 0.37 using both birth certificate data and discharge data to construct risk-adjusted cesarean rates.[16] We report least squares regressions (WLS) to examine the relationship between the risk-adjusted cesarean rate for each birthweight category and the PPC at the county level.

Procedure intensity and mortality. The third step in our analysis was to evaluate the relationship between (1) variations in the intensity of the use of cesareans and (2) maternal and infant mortality across areas. Even if physicians in areas that use cesareans more intensively are performing the procedure on less and less appropriate patients, their patients could still have better outcomes if the physicians are more skilled at the procedure. We tested this hypothesis by regressing both risk-adjusted infant mortality and maternal mortality on risk-adjusted cesarean rates. We performed this analysis using both WLS and negative binomial regression and obtained quantitatively similar estimates. To preserve transparency, we report the results from the former technique.

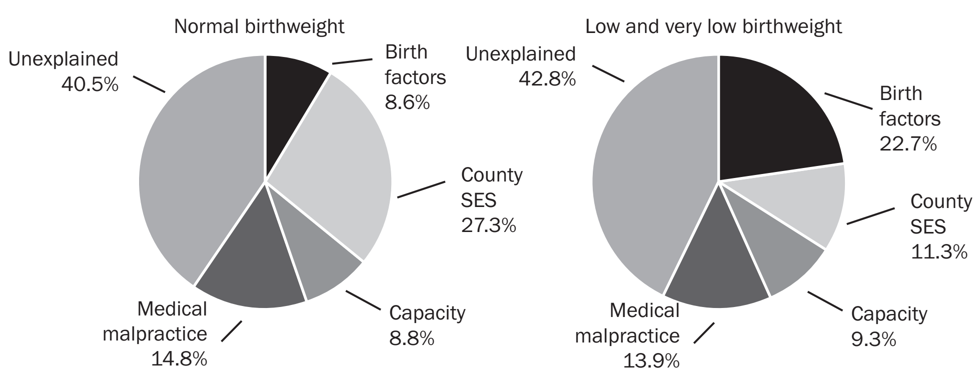
Study Results

■ **Correlates of area-level variation in the use of cesareans.** Exhibit 1 illustrates the variation in unadjusted cesarean rates that can be explained by patient and area characteristics for births over 2,500 grams and those under 2,500 grams. Characteristics of the birth explain much more of the variation in cesarean rates for LBW infants than they do for NBW infants. For both NBW and LBW births, a large part of the variation remains unexplained.

Exhibit 2 reports the range of unadjusted and risk-adjusted cesarean rates for the United States and the fifty-one counties with the most births between 1995 and 1998.[17] With few exceptions, risk adjustment does not greatly alter the ranking of the counties, and the correlation coefficients between area cesarean rates obtained from unadjusted and risk-adjusted models are 0.93 ($p < .001$) for NBW births and 0.84 ($p < .001$) for LBW births.[18] Counties whose risk-adjusted rates are higher than the unadjusted rates have cesarean usage that is more intensive than what would be predicted using observable characteristics. The variation in the rates is large: Even in the largest U.S. counties, risk-adjusted rates ranged from 6.7 percent to 28.9 percent for NWB births and from 25 percent to 50 percent for LBW births.

■ **Cesarean delivery and medical appropriateness.** In Exhibit 3 the average PPC of patients receiving a cesarean in each area is plotted against risk-adjusted area-level cesarean rates. LBW babies are more likely to be born via cesarean than

EXHIBIT 1
Correlates Of County Cesarean Rates For Births With Normal (2,500 Grams Or More) And Low And Very Low (Less Than 2,500 Grams) Birthweight, 1995–1998



SOURCE: Authors' calculations from U.S. natality data, 1995–1998.
NOTES: Fraction of variation in use of cesareans across 198 U.S. counties in 1995–1998 attributable to different correlates, including patient-level characteristics (mother's age, race, education, and marital status; adequacy of prenatal care; medical risk factors; congenital anomalies; and complications of labor and delivery), county-level measures of socioseconomic status (SES) factors (average per capita income; unemployment rate; percentage living in poverty; percentage urban; percentage minority; percentage with less than high school, high school diploma, and college degree; percentage of hospital patient days eligible for Medicaid; and population), county-level provider supply (total physicians, surgeons, pediatricians, obstetrician-gynecologists, internists, and other specialists per birth, as well as the number of neonatal intensive care unit beds per birth), and state-level malpractice liability (the number and size of judgments and settlements by specialty and malpractice premiums by specialty).

EXHIBIT 2
Cesarean Rates For Normal-Birthweight And Low/Very-Low-Birthweight Births For The Fifty-One U.S. Counties With The Largest Number Of Births, 1995–1998

			Unadjusted rates (%)		Adjusted rates (%)	
County	State	Major city	Normal birthweight	Low/very low birthweight	Normal birthweight	Low/very low birthweight
Hennepin	MN	Minneapolis	13.4	37.0	12.5	35.4
Milwaukee	WI	Milwaukee	13.5	26.7	13.5	29.8
Denver	CO	Denver	13.7	35.0	15.8	32.0
Salt Lake	UT	Salt Lake City	14.4	35.6	15.8	32.2
Maricopa	AZ	Phoenix	14.7	34.8	17.6	35.0
Franklin	OH	Columbus	15.6	36.1	18.0	35.9
Alameda	CA	Oakland	15.9	33.2	19.2	38.5
Travis	TX	Austin	16.4	35.8	17.3	36.7
Santa Clara	CA	Santa Clara	16.6	37.3	18.0	36.1
Wayne	MI	Detroit	16.7	31.1	18.5	35.2
King	WA	Seattle	16.8	36.3	14.0	31.5
Sacramento	CA	Sacramento	17.0	38.1	18.0	37.9
Cook	IL	Chicago	17.0	32.3	19.2	35.7
Baltimore City	MD	Baltimore	17.5	29.4	19.5	35.2
Hamilton	OH	Cincinnati	17.6	38.3	17.2	35.0
Bronx	NY	Bronx	17.9	31.5	19.8	36.1
Allegheny	PA	Pittsburgh	17.9	34.0	16.2	31.2
Fulton	GA	Atlanta	18.1	34.2	19.4	38.9
Orange	FL	Orlando	18.2	37.1	20.1	37.4
Cuyahoga	OH	Cleveland	18.2	34.1	17.8	34.5
Philadelphia	PA	Philadelphia	18.4	31.6	16.3	35.2
Marion	IN	Indianapolis	18.5	36.3	19.3	36.5
Clark	NV	Las Vegas	18.7	38.8	20.9	41.3
Orange	CA	Long Beach/Santa Ana	18.9	40.7	21.5	39.8
Riverside	CA	Riverside	18.9	39.1	23.7	44.2
El Paso	TX	El Paso	19.0	36.8	19.7	37.3
Middlesex	MA	Cambridge/Boston	19.0	33.7	15.0	33.6
San Diego	CA	San Diego	19.3	38.9	21.9	42.0
Suffolk	MA	Boston	19.4	42.5	15.5	32.9
Tarrant	TX	Fort Worth	19.6	40.8	20.3	40.9
Montgomery	MD	Greater DC area	19.7	38.8	17.5	35.8
San Bernardino	CA	San Bernardino	19.8	41.4	23.3	43.5
Dallas	TX	Dallas	20.0	36.2	20.2	36.3
Bexar	TX	San Antonio	20.2	39.6	20.2	39.6
St. Louis	MO	St. Louis	20.3	39.0	17.0	33.8
Kings	NY	Brooklyn	20.3	34.1	21.1	38.1
Oakland	MI	Greater Detroit	20.4	39.6	17.7	35.6
Fairfax	VA	Fairfax	20.5	42.5	19.6	40.3
Shelby	TN	Memphis	21.1	33.8	19.9	37.0
Palm Beach	FL	Palm Beach	21.1	37.1	21.8	38.2
Harris	TX	Houston	21.1	38.5	21.1	38.8
Fresno	CA	Fresno	21.1	37.5	26.4	42.5
New York	NY	Manhattan	21.3	41.7	18.4	39.4
Queens	NY	Queens	21.9	36.7	21.6	39.5
Los Angeles	CA	Los Angeles	22.0	42.5	25.0	45.9
Essex	NJ	Newark	23.6	39.9	20.5	38.2
Hillsborough	FL	Tampa	23.7	43.9	24.3	44.1
Nassau	NY	Greater New York City	23.9	45.3	21.5	40.9
Miami-Dade	FL	Miami	24.5	36.8	25.6	41.8
Broward	FL	Ft. Lauderdale	24.5	43.8	24.4	44.8
Suffolk	NY	Long Island	26.0	45.1	21.3	39.9

EXHIBIT 2
Cesarean Rates For Normal-Birthweight And Low/Very-Low-Birthweight Births For The Fifty-One U.S. Counties With The Largest Number Of Births, 1995–1998 (cont.)

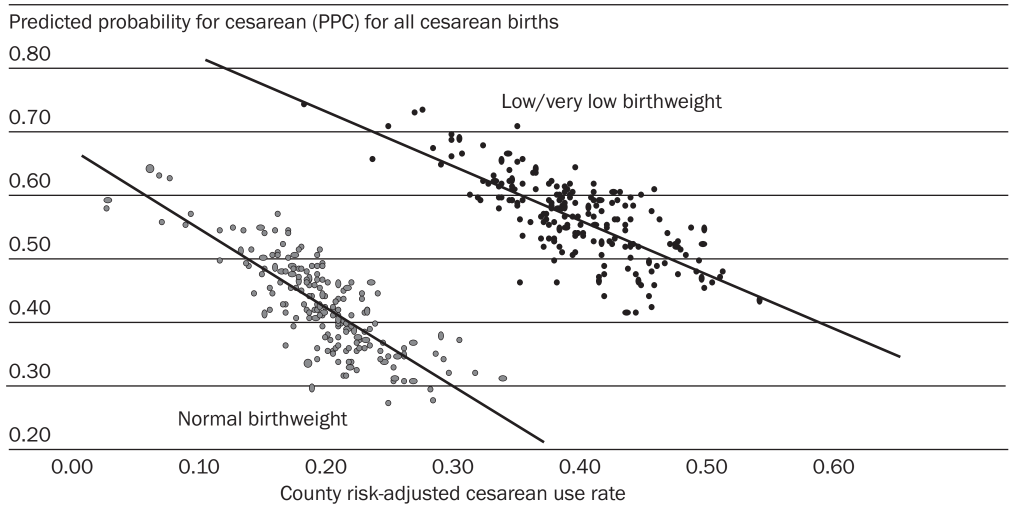
			Unadjusted rates (%)		Adjusted rates (%)	
County	State	Major city	Normal birthweight	Low/very low birthweight	Normal birthweight	Low/very low birthweight
U.S. rate[a]			19.2	37.3	19.2	37.3
Range: 1st–99th percentile			11.9–28.7	24.4–49.9	6.7–28.9	25.0–50.0

SOURCE: Authors' calculations from U.S. natality data, 1995–1998.
NOTES: Risk-adjusted cesarean rates for normal-birthweight (NBW) and low-birthweight (LBW)/very-low-birthweight (VLBW) births in the fifty-one U.S. counties with the most births in 1995–1998 (n = 10,161,953). Counties are listed in descending order of the total number of births. The correlation between NBW and LBW/VLBW cesarean rates is 0.87 ($p < .001$), and the correlation between adjusted and unadjusted rates is 0.71 ($p < .001$).
[a] The 198 largest U.S. counties.

NBW babies, but even for these births, the variation in risk-adjusted rates across large counties is striking. For both NBW and LBW babies, the average birth through cesarean has a systematically lower PPC in areas that do more cesarean sections. Increases in the cesarean rate are associated with major declines in the average medical appropriateness of births.

■ **Cesarean use and patient mortality.** Our third analytic step was to examine

EXHIBIT 3
Relationship Between Predicted Probability Of Cesarean Birth And Cesarean Rates, Normal Birthweight And Low/Very Low Birthweight, 1995–1998



SOURCE: Authors' calculations from U.S. natality data, 1995–1998.
NOTES: Relationship between each area's cesarean rate and the average probability of cesarean delivery among cesarean births in the area (a measure of appropriateness). A ten-percentage-point increase in the cesarean rate for normal-birthweight births (2,500 grams or higher) leads to a fourteen-percentage-point reduction in appropriateness for cesarean births ($p < .001$) and a nine-percentage-point decline in appropriateness for low-birthweight and very-low-birthweight births (less than 2,500 grams) ($p < .001$). Observations are 198 U.S. counties (1995–1998).

the effect of risk-adjusted area cesarean rates on risk-adjusted infant and maternal mortality rates (Exhibit 4). For NBW births, the standard deviation of the cesarean rate was 3.2 percentage points, and a change in the cesarean rate of this magnitude (from the average rate of 19.2 percent) would not alter neonatal mortality. Similarly, for LBW births, the standard deviation of the cesarean rate was 4.8 percentage points. Altering the cesarean rate from 37.5 percent upward by this amount would also not result in a change in neonatal mortality. The exhibit also shows the effect of changes in use of cesareans on maternal mortality. There is a small but statistically insignificant relationship: The average maternal mortality ratio is 7.3 deaths per 100,000 births, and a 3.2-percentage-point decrease in the cesarean rate is associated with an insignificant change in the maternal mortality rate ($p = .06$).

Discussion

We have demonstrated that there is large geographic variation in the use of cesarean delivery, only some of which is explained by patients' characteristics and county SES measures. A substantial portion of this variation remains unexplained. This unexplained variation could be labeled as the "practice style" of an area, if it is unrelated to patient and area characteristics. In fact, if the number of physicians or NICU beds in an area is also a consequence of underlying variations in practice style, then we might have understated the role of local practice style. On the other hand, if county cesarean rates are partially determined by patients' preferences, then we have overstated the case for physician practice style. To the extent that patients' preferences are explained by the demographic variables that are controlled for in our analysis (age; race; maternal schooling; and county SES measures such as income, poverty, and population), our labeling of the unexplained varia-

EXHIBIT 4
Association Between Higher Cesarean Use Rates And Neonatal And Maternal Mortality, 1995–1998

	Neonatal mortality		Maternal mortality
	Normal birthweight	Low and very low birthweight	All birthweights
Cesarean use rate	19.2%	37.5%	20.5%
Standard deviation (SD) of cesarean use rate[a]	3.2	4.8	3.2
Mortality rate per 10,000 births	9.7	330.4	0.73
Effect of decreasing cesarean rate by 1 SD on mortality per 10,000 births	0.2 ($p < .97$)	0.2 ($p < .28$)	0.096 ($p < .10$)

SOURCE: Authors' calculations from U.S. natality data, 1995–1998.

NOTES: Estimated effect of differences in cesarean rates on infant and maternal mortality. Relationships were estimated using births in 198 U.S. counties, 1995–1998. Decreasing the use of cesareans by 4.8 percentage points (one standard deviation) among low-birthweight (LBW)/very-low-birthweight (VLBW) births would decrease neonatal mortality by 0.2 per 10,000 births, but this change is not statistically significant.

[a] Percentage points.

tion as "physician practice style" is justified.

We demonstrate that more-aggressive areas perform the procedure for births that are less medically appropriate for the procedure. This finding is a prerequisite for demonstrating that the higher rates are symptomatic of "flat-of-the-curve" medicine, where physicians work into less appropriate populations. Indeed, our analysis challenges previous work that found no relationship between the intensity of diagnostic testing and clinical indications for the use of that test.[19] We demonstrate that some physicians are not systematically more aggressive than others (that is, do not have a disposition to be more aggressive on all births regardless of medical appropriateness); rather, physicians rank patients on a distribution of clinical appropriateness and work their way down that distribution. The point at which they stop in that distribution is affected by nonmedical factors such as provider capacity, malpractice liability, and local physicians' opinion.

■ **Study limitations.** Our analysis is not without limitations. We relied on birth certificate data; important information that is available to the physician, such as maternal drug use and detailed medical histories, was not available to us. We note that our flexible birth-level risk-adjustment models have explanatory power that is identical to other studies that used both birth certificate and hospital discharge data. Although this is reassuring, it does not mean that omitted factors are not an issue for our study. For omitted factors to bias our analysis (by overstating the role of physician practice style), they would have to be more prevalent in areas with higher cesarean rates and completely uncorrelated with the maternal characteristics and county socioeconomic factors that we controlled for. To explore this possibility, we restricted our sample of births to states that report information on tobacco use, alcohol consumption, and weight gain during pregnancy, and we included these variables as additional covariates in our analysis. The correlation between risk-adjusted area cesarean rates using these additional covariates and those reported in Exhibit 2 was 0.98 ($p < .001$) for LBW births and 0.99 ($p < .001$) for NBW births. These findings suggest that certain types of omitted clinical variables might not be of first-order concern.

■ **Policy implications.** Our finding that physicians in areas with higher cesarean rates are performing procedures that are of decreasing medical value to patients has important policy implications. Cesareans are an expensive intervention, with an average cost in 2003 of $12,468—twice the cost of the average vaginal birth ($6,240). There is also evidence that women undergoing a cesarean delivery are at much higher risk for rehospitalization for uterine infection and obstetrical surgical wound complications.[20] The real costs of a cesarean delivery might therefore be much higher than we have stated. Our analysis demonstrates that reductions in the cesarean rate in high-use counties (of the magnitude of three to five percentage points) will not affect mortality among newborns and mothers. Reductions in the cesarean rate have been demonstrated to be achievable in the clinical literature: A hospital was able to reduce its rate from 17.5 percent to 11 percent over two years without any adverse health outcomes. The reduction was achieved by requiring a second opinion, by providing objective criteria for when a cesarean delivery is indicated, and by a review of all cesarean deliveries. If the cesarean rate is to be reduced, we would also argue that reductions should be targeted toward primary cesareans and not repeat cesareans. The latter are believed to be safer than the VBAC alternative, and each primary cesarean that is averted also eliminates the need for a repeat cesarean.[21]

Reductions in the cesarean rate could deleteriously affect other health outcomes that were not examined in our study.[22] Furthermore, if county cesarean rates reflect patients' preferences for elective cesareans, then reducing the rate will reduce patient satisfaction. Although there has probably been an increase in patient demand for elective cesarean delivery, it is not known if this explains the geographic variation in cesarean rates, and it is difficult to rationalize preferences as being the explanation for the ten-percentage-point difference in cesarean rates between Minneapolis and Miami. The importance of preferences can be partially assessed by examining the role of county socioeconomic factors that ought to be correlated with patients' preferences (such as income, educational attainment, percentage minority, and percentage metropolitan). Relative to the unexplained variation and variation explained by medical malpractice and capacity factors, the role of county factors is small, and it suggests that preferences cannot be the principal driver of geographic variation in cesarean rates. Likewise, patient satisfaction has been shown to be unaffected by the presence of intensive health care.[23]

Variation in local medical opinion about the right cutoff for initiating a cesarean delivery, strengthened by available capacity and malpractice pressure, continues to be the best explanation for the facts in our analysis. In an era of soaring health care costs, where already strained public programs reimburse for cesarean delivery, it seems particularly important to consider the ramifications of intensive treatments whose medical benefits are uncertain when performed in less medically appropriate populations.

Katherine Baicker and Amtabh Chandra acknowledge funding from the National Institute on Aging, Grant no. NIA P01 AG19783-02; and Chandra, from the National Institute of Child Health and Human Development, Grant no. NICHD R01 HD44003-01. The authors thank three astute reviewers, Jonathan Skinner, Douglas Staiger, and Jack Wennberg, for influencing their thinking on this topic. The opinions in this paper are those of the authors and do not reflect those of any institution they are affiliated with or have received support from.

NOTES

1. J.M. Belizan et al., "Rates and Implications of Caesarean Sections in Latin America: Ecological Study," *British Medical Journal* 319, no. 7222 (1999): 1397–1400; and B.L. Flamm, "Cesarean Section: A Worldwide Epidemic?" *Birth* 27, no. 2 (2000): 139–140.
2. U.S. Department of Health and Human Services, *Healthy People 2010* (Washington: U.S. Government Printing Office, November 2000).
3. E.S. Fisher and H.G. Welch, "Avoiding the Unintended Consequences of Growth in Medical Care: How Might More Be Worse?" *Journal of the American Medical Association* 281, no. 5 (1999): 446–453.
4. M.R. Chassin et al., "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services? A Study of Three Procedures," *Journal of the American Medical Association* 258, no. 18 (1987): 2533–2537; L.L. Leape et al., "Does Inappropriate Use Explain Small-Area Variations in the Use of Health Care Services?" *Journal of the American Medical Association* 263, no. 5 (1990): 669–672; C.M. Winslow et al., "The Appropriateness of Carotid Endarterectomy," *New England Journal of Medicine* 318, no. 12 (1988): 721–727; and C.M. Winslow et al., "The Appropriateness of Performing Coronary Artery Bypass Surgery," *Journal of the American Medical Association* 260, no. 4 (1988): 505–509.
5. See, for example, Chassin et al., "Does Inappropriate Use Explain Geographic Variations?"; Winslow, "The Appropriateness of Carotid Endarterectomy"; H. Hemingway et al., "Underuse of Coronary Revascularization Procedures in Patients Considered Appropriate Candidates for Revascularization," *New England Journal of Medicine* 344, no. 9 (2001): 645–654; and J.Z. Ayanian et al., "Rating the Appropriateness of Coronary Angiography—Do Practicing Physicians Agree with an Expert Panel and with Each Other?" *New England Journal of Medicine* 338, no. 26 (1998): 1896–1904.
6. J.P. Kassirer, "The Quality of Care and the Quality of Measuring It," *New England Journal of Medicine* 329, no. 17 (1993): 1263–1265; N.R. Hicks, "Some Observations on Attempts to Measure Appropriateness of Care," *British Medical Journal* 309, no. 6956 (1994): 730–733; and C.E. Phelps "The Methodologic Foundations of Studies of the Appropriateness of Medical Care," *New England Journal of Medicine* 329, no. 17 (1993): 1241–1245.
7. P.G. Shekelle et al., "The Reproducibility of a Method to Identify the Overuse and Underuse of Medical Procedures," *New England Journal of Medicine* 338, no. 26 (1998): 1888–1895.
8. We selected these factors based on variables suggested by previous research. See, for example, A.A. Kabir et al., "Unnecessary Cesarean Delivery in Louisiana: An Analysis of Birth Certificate Data," *American Journal of Obstetrics and Gynecology* 190, no. 1 (2004): 10–19; D.C. Aron et al., "Variations in Risk-Adjusted Cesarean Delivery Rates According to Race and Health Insurance," *Medical Care* 38, no. 1 (2000): 35–44; A.A. Kabir et al., "Racial Differences in Cesareans: An Analysis of US. 2001 National Inpatient Sample Data," *Obstetrics and Gynecology* 105, no. 4 (2005): 710–718; W.J. Hueston and S. Lewis-Stevenson, "Provider Distribution and Variations in Statewide Cesarean Section Rates," *Journal of Community Health* 26, no. 1 (2001): 1–10; L. Dubay, R. Kaestner, and T. Waidmann, "The Impact of Malpractice Fears on Cesarean Section Rates," *Journal of Health Economics* 18, no. 4 (1999): 491–522; and L.R. Burns, S.E. Geller, and D.R. Wholey, "The Effect of Physician Factors on the Cesarean Section Decision," *Medical Care* 33, no. 4 (1995): 365–382.
9. Descriptive statistics for our sample are reported in Appendix Exhibit 1, available online at http://content.healthaffairs.org/cgi/content/full/25/w355/DC2.
10. D.P. Kessler and M.B. McClellan, "How Liability Law Affects Medical Productivity," *Journal of Health Economics* 21, no. 6 (2002): 931–955.
11. Congressional Budget Office, *Limiting Tort Liability for Medical Practice,* 8 January 2004, http://www.cbo.gov/showdoc.cfm?index=4968&sequence=0 (accessed 5 July 2006); U.S. Government Accountability Office, *Medical Malpractice Insurance: Multiple Factors Have Contributed to Increased Premium Rates* (Washington: GAO, 2003); and K. Baicker and A. Chandra, "The Effect of Malpractice Liability on the Delivery of Health Care," *Forum for Health Economics and Policy,* Forum: Frontiers in Health Policy Research, vol. 8, article 4 (2005), http://www.bepress.com/fhep/8/4 (accessed 21 May 2006).
12. Centers for Disease Control and Prevention, Compressed Mortality Database, 1995–1998, http://wonder.cdc.gov/mortICD9J.html (accessed 23 August 2005).

13. The patient-level controls include birthweight in grams, gestation in weeks, mother's age, mother's race (black, white, or other), mother's education (fewer than 9 years, 9–11 years, 12 years, 13–15 years, and 16 or more years), mother's marital status (married, not married), birth order (first, second, third, or fourth or higher), newborn's sex, prenatal care use based on a Kessner index (adequate, intermediate, or inadequate), and indicators for the presence of each medical risk factor, congenital anomaly, or complications of labor and delivery.

14. We also estimated more complicated models that are, in principle, more efficient. We estimated (1) a conditional logit model, (2) a random-intercept logistic regression, and (3) a generalized estimating equation (GEE) model with a logit link function and exchangeable correlation structure. The correlation coefficients between the county effects obtained from these models and our preferred model are 0.981 ($p < .001$), 0.95 ($p < .001$), and 0.962 ($p < .001$), respectively. Given the similarity of area rankings across these different procedures, we favored the more transparent model.
15. Formally, for birth *i* (with characteristics X_i) in county *j* (denoted by C_j), the propensity score for a cesarean section (CS) is: Probability $CS_{ij} = 1 = F(X_i b + C_j)$. F(.) is the cumulative distribution for the logistic density, and we estimate this model using fixed-effects regression (thereby relaxing the assumption of GEE or random effects models that C is not correlated with X). C_j represents the county risk-adjusted cesarean rate. The PPC is the predicted probability of receiving a birth based only on birth characteristics, $Pr(CS_{ij} = 1) = F(X_i b)$ averaged across all births in a county.
16. E.B. Keeler et al., "Adjusting Cesarean Delivery Rates for Case Mix," *Health Services Research* 32, no. 4 (1997): 511–528.
17. Appendix Exhibit 3 provides this information for all 198 counties in our study. See Note 9.
18. Rates for NBW and LBW births are highly correlated within counties (correlation coefficient = 0.87, $p < .001$), which allays concern that the variation in cesarean rates among LBW babies is driven by random "noise."
19. Chassin et al., "Does Inappropriate Use Explain Geographic Variations?"; and Leape et al., "Does Inappropriate Use Explain Small-Area Variations?"
20. M. Lydon-Rochelle et al., "Association between Method of Delivery and Maternal Rehospitalization," *Journal of the American Medical Association* 283, no. 18 (2000): 2411–2416.
21. M.B. Landon et al., "Maternal and Perinatal Outcomes Associated with a Trial of Labor after Prior Cesarean Delivery," *New England Journal of Medicine* 351, no. 25 (2004): 2581–2589.
22. In work not reported in this paper, we examined other outcomes such as birth injuries and complications of delivery. We did not find that higher (risk-adjusted) cesarean rates were associated with improvements in these outcomes but did not pursue this line of inquiry after noting the poor data quality of these fields in birth-certificate data. See, for example, S. Northam and T. Knapp, "The Reliability and Validity of Birth Certificates," *Journal of Obstetric, Gynecologic, and Neonatal Nursing* 35, no. 1 (2006): 3–12; and D.L. DiGiuseppe et al., "Reliability of Birth Certificate Data: A Multi-Hospital Comparison to Medical Records Information," *Maternal and Child Health Journal* 6, no. 3 (2002): 169–179.
23. E.S. Fisher et al., "The Implications of Regional Variations in Medicare Spending, Part 2: Health Outcomes and Satisfaction with Care," *Annals of Internal Medicine* 138, no. 4 (2003): 288–298; and K. Baicker and A. Chandra, "Medicare Spending, the Physician Workforce, and Beneficiaries' Quality of Care," *Health Affairs* 23 (2004): w184–w197 (published online 7 April 2004; 10.1377/hlthaff.23.w4.184).