The Impact of Medical Liability Standards on Regional Variations in Physician Behavior: Evidence from the Adoption of National-Standard Rules

By MICHAEL FRAKES

Web Appendix

APPENDIX A: Developments in Malpractice Standard-of-Care Laws

FOR ONLINE PUBLICATION

The traditional "locality rule" contained both a substantive and a procedural component. The substantive component simply provided that physicians must follow the practices of local physicians. The procedural restrictions of the traditional rule required that plaintiffs use local physicians to testify as to the local practices, implicating concerns regarding plaintiffs' abilities to find physicians willing to testify against their peers. Many states amended their malpractice laws by the 1970's to alleviate this concern, either by permitting the use of outside experts familiar with local practices[1] or by adopting modified locality rules that based standards on the practices of physicians in the same locality *or in a similar locality*. While relaxing certain evidentiary burdens,

[1] See, for example, *Ardoline v. Keegan*, 140 Conn. 552 (1954).

courts generally did not view these developments as significantly changing the substantive requirements that physicians comply with local practices.[2]

In the latter part of the 20th Century, many jurisdictions came to further relax the geographical limitations in malpractice law by abandoning the locality rule in favor of laws requiring physicians to comply with national standards of care, or standards of care that are not subject to geographical restrictions. Courts justified these amendments by arguing that the initial rationale for the rule had dissipated many years prior with certain early-20th-Century advances, including the standardization of medical school curriculum and board certification.[3] To the extent that the previous developments (e.g., same-or-similar-locality rule adoptions) were already effective in opening up the market for willing experts, these subsequent developments can be seen as largely substantive in nature.

Based on research of case and statutory law, I document the evolution of each state's malpractice standard-of-care laws from the mid-1970's to the present. For each state and year, I determine whether the prevailing law requires the use of a national standard of care or a local standard of care (i.e., a same-locality, same-or-similar locality, or, in rare instances, a statewide standard). The variation over this time period is largely substantive in nature and focuses on the local-national distinction. Nearly all of the procedural developments had taken place before the beginning of the sample period.

[2] As stated by the Washington Supreme Court in one of the pioneering locality-rule abdication cases, "[b]roadening the rule to include `similar localities'... alleviated, to a certain extent, the first practical difficulty of the `locality rule' – additional witnesses might be available; but it did little to remove the deficiencies springing from the second" – that is, "the possibility of a small group, who, by their laxness or carelessness, could establish a local standard of care that was below that which the law requires." *Pederson v. Dumouchel*, 431 P.2d 973 (1967).

[3] For instance, the Mississippi Supreme Court employed these arguments to justify its continuing retreatment from the locality rule in *Hall v. Hilbun* 466 So. 2d 856 (1985).

Table A1. Adoptions and Repeals of National Standard-of-Care Requirements

State	Year	Source(s)
Alabama	1980	Zellis v. Brown, 382 So. 2d 528 (March 1980). *
Colorado	1983	Green v. Thomas, 662 P.2d 491 (November 1982); Short v. Kinkade, 685 P.2d 210 (December 1983). **
Connecticut	1984	Logan v. Greenwich Hospital Association, 191 Conn. 282 (September 1983). ***
Delaware	1999	18 Del.C. § 6801 (amendment effective July 1998).
D.C.	1980	Morrison v. Macnamara, 407 A.2d 555 (October 1979).
Indiana	1992	Vergara v. Doan, 593 N.E. 2d 185 (June 1992).
Maryland	1994	Md Code Ann, [Cts&Jud Proc] §3-2A-02(c) (effective July 1993). ****
Mississippi	1983	King v. Murphy, 424 So. 2d 547 (Nov. 1982); Hall v. Hilbun, 466 So. 2d 856 (Feb. 1985). *****
Montana	1985	Aasheim v. Humberger, 215 Mont. 127 (February 1985).
Nevada	1979	Orcutt v. Miller, 595 P2d 1191 (June 1979).
New Mexico	1978	Pharmaseal Lab., Inc. v. Goffe, 90 N.M. 753 (September 1977).
Oklahoma	1984	76 O.S. Supp. 1983 § 20.1 (effective September 1983).
Rhode Island	1998	Sheeley v. Memorial Hospital, 710 A. 2d 161 (April 1998).
South Carolina	1981	King v. Williams, 276 S.C. 478 (June 1981).
South Dakota	1988	Shamburger v. Behrens, 418 N.W.2d 299 (January 1988).
West Virginia	1986	Paintiff v. City of Parkersburg, 176 W. Va. 469 (March 1986); W. Va. Code § 55-7B-3 (effective 1986). ******
Wyoming	1981	Vassos v. Roussalis, 625 P.2d 768 (March 1981). *******

* The Zellis decision was decided by a plurality and its holding did not turn directly on this geographical distinction. Nonetheless, the Supreme Court of Alabama in Zellis provided a strong, direct indication of their intention to abandon the use of a locality rule in Alabama. This stance was subsequently strengthened by the Court's decision in Bryant v. Otts, 412 So. 2d 254 (1982). The results are robust to the use of 1982 as the relevant date of adoption.

** It was not until Jordan v. Bogner, 844 P.2d 664 (January 1993) when the Supreme Court of Colorado spoke definitively on the geographical scope of the standard of care owed by a specialist physician. However, several earlier cases, including those indicated, adopted requirements that specialist physicians are to be judged by a standard commensurate with that of a reasonable physician practicing in that specialty. This approach did not limit the standard to particular geographical bounds, a fact emphasized by subsequent case law. In describing the standard to be applied to non-specialist physicians, Colorado case law continued to refer to the use of a community standard. As such, in indicating that specialists are not subject to a locality rule, the Supreme Court in the 1993 Bogner decision cited these earlier decisions as the applicable law and did not indicate that it was adopting a new approach. For these reasons, I use these earlier rulings, Green and Kinkade, as the turning point in Colorado's national-standard requirement for specialist physicians. The findings presented however are robust to the alternative use of 1993 as the year of adoption and to the exclusion of Colorado entirely.

*** This position was subsequently codified in Conn. Gen. Stat. § 52-184c.

**** Maryland had, prior to this time, adopted a national-standard requirement in Shilkret v. The Annapolis Emergency Hospital Association, 276 Md. 187 (October 1975).

***** King expanded the geographical scope of the standard-of-care requirements to include at least the entire state of Mississippi plus a reasonable distance beyond the boundaries of the state. For the purposes of this empirical analysis (structured around state-year cells) I consider this breaking of state boundaries as an abandonment of the locality rule. However, the findings are robust to the use of 1985 as the relevant national-standard adoption year, at which time the court fully embraced a national standard in Hilbun.

****** With respect to specialists, the abolition of the substantive components of the locality rule in West Virginia may arguably be seen as having begun years before with the West Virginia Supreme Court's decision in Hundley v Martinez, 151 W. Va. 977 (1967). The presented results are robust to the exclusion of West Virginia.

******* In Vassos, the Wyoming Supreme Court stated that "a physician or surgeon must exercise the skill, diligence and knowledge, and must apply the means and methods, which would reasonably be exercised and applied under similar circumstances by members of his profession in good standing and in the same line of practice… The skill, diligence, knowledge, means and methods are not those 'ordinarily' or 'generally' or 'customarily' exercised or applied, but are those that are 'reasonably' exercised or applied." Subsequent case law viewed this 1981 decision as the turning point in the abandonment of the locality rule. Wyoming subsequently codified the use of a national standard in 1986.

Table A1 documents this evolution in the law. The table provides information only for those states that amended their relevant standard-of-care laws over the sample period (along the national vs. non-national dimension). A small minority of states take varying approaches to their standard-of-care requirements depending on whether the physician in question is a general practitioner or a specialist. Given the nature of the procedures explored above and to allow for a consistent empirical approach, I specify national-standard requirements using laws that apply to specialists only or laws that take symmetrical approaches between both general practitioners and specialists. I exclude states where there is substantial uncertainty in the nature of their standard-of-care requirements over time (e.g., conflicting case law). This only affects two states: Texas and Hawaii. The results of this study are entirely robust to the assumption that Texas follows a national-standard law throughout the entire sample period, arguably the most accurate reading of its case law.

APPENDIX B: Data Description

FOR ONLINE PUBLICATION

For roughly 260,000 inpatient records per year, the National Hospital Discharge Survey (NHDS) contains information on the primary/secondary diagnosis and procedure codes associated with the discharge and on various characteristics of the relevant patient and hospital. I supplement the public NHDS files with geographic identifiers received from the Research Data Center at the National Center for Health Statistics. The resulting sample covers 1977 to 2005.

Certain characteristics of the NHDS limit its ability to tell a comprehensive story of physician behavior. First, the NHDS may only contain enough inpatient records to provide a valid indication of the treatment rates of relatively high volume surgical procedures. Second, the data contains information on inpatient practices only. A large number of procedures may be performed in both inpatient and outpatient settings, with the outpatient alternative expanding considerably over time. This is especially the case with respect to diagnostic utilization. In addition to concerns over the generalizability of the analysis, focusing only on an inpatient utilization story raises sample-selection fears (e.g., those states that adopt national standard laws may also be those that are trending more heavily towards outpatient care, leaving a selected inpatient sample behind) and also muddles the ability to evaluate a utilization rate within a tangible clinical context.

In this paper, I focus on two clinical contexts that are less susceptible to the above concerns: obstetrics and complex cardiac care. Deliveries represent the most frequent discharge type within the NHDS inpatient records, appearing an average of 600 times in a given state-year cell. Cesarean sections, in turn, represent one of the most common surgical procedures performed in the United States. Moreover, the event of childbirth itself represents a situation in which hospitalization is almost universally sought, leaving few concerns over any confounding effects from care delivered outside of an inpatient setting. Similarly, intensive cardiac interventions represent procedures that are both frequently employed throughout the NHDS records and, over the full range of the sample period considered, consistently performed in a predominantly inpatient setting.

The NHDS is especially well suited for an analysis of obstetric behavior as it contains a large and self-contained subsample of deliveries out of which one can readily calculate a local utilization rate that can be directly compared with the national rate, a necessary component of the convergence analysis. Moreover, the availability of the delivery subsample alleviates certain sampling-variability concerns[4] and facilitates the use of a denominator – i.e., the delivery count – that is itself not likely to be impacted by legal forces. More generally, an analysis of cesarean behavior presents a clean, binary decision-making context.

Notes on Cesarean Rate Calculations. In the primary convergence specifications, state-year cesarean rates are risk-adjusted for maternal age and race and for the incidence of breech presentation, multiple delivery or previous cesarean delivery. Other specifications estimated in Appendix C follow the Agency for Healthcare Research and Quality and use "primary" cesarean rates that ignore deliveries with these latter three factors.

Notes on Delivery Risk Factors and Complications. The set of risk factors used in (1) calculating the predicted probabilities of cesarean delivery, (2) standardizing regional cesarean rates for the triage specification indicated in equation (4) and in standardizing rates of vaginal births after cesarean delivery

[4] The design of the NHDS may lead to some amount of within-state variability in the set of hospitals included in the NHDS records. This sampling variability may induce variation in the observed propensity or ability of the included hospitals to accept the relevant category of patients in the first instance, which may induce variation in the state-year treatment count. These concerns are softened in the cesarean analysis given that this variability is also accounted for by the cesarean denominator. If a hospital is sampled that does not perform any (or many) deliveries, then the utilization rate is calculated using the deliveries from the remaining hospitals. Most of the cardiac procedure measures do not benefit from a logical denominator of this nature.

include the following: maternal age, breech presentation, multiple deliveries (e.g., twins), previous cesarean delivery, placenta previa, placenta abruption, dysfunctional labor, cephalopelvic disproportion, fetal distress, precipitous labor, postpartum hemorrhage, prolonged labor, premature rupture of the membranes, cord prolapse, maternal hypertension, maternal diabetes, and maternal anemia.[5]

Notes on Cardiac Treatment Utilization Calculations. Unlike the obstetric analysis, the NHDS records do not contain a self-contained denominator to form the relevant cardiac intensive treatment rate. However, it is necessary to normalize the state-year treatment counts in some fashion in order to facilitate a comparison between national physician behavior and the behavior of physicians represented in each state-year cell. I achieve the necessary normalization by dividing the state-year treatment count by the number of primary admissions for acute myocardial infarction (AMI) occurring in the given state-year cell. AMIs represent a situation in which affected patients will almost universally seek hospitalization (Wennberg and Cooper 1999; Skinner et al. 2005).[6] This denominator thus allows for a rational scaling of the cardiac treatment count using a measure that itself should not be expected to change upon the change in law.

[5] As a result of the shift from the Eighth to the Ninth Revision of the International Classification of Diseases in 1979, it is not possible to generate incidence rates for certain maternal risk factors in the 1977 – 1978 period. Accordingly, those specifications that incorporate the full set of risk factors are confined to the 1979 – 2005 period.

[6] Similar conditions include stroke, hip fracture and gastrointestinal bleeding. The results below are robust to the use of an alternative denominator equal to an index measure comprised of the sum of the occurrences of any one of these four non-discretionary medical events.

APPENDIX C: Specification Checks

FOR ONLINE PUBLICATION

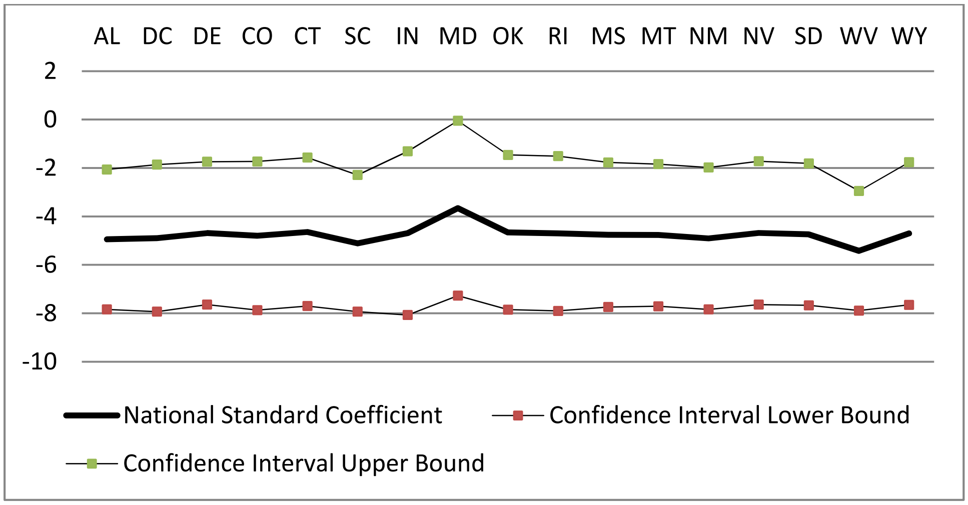
In this appendix, I discuss a number of additional specification checks to the empirical analysis set forth in the paper. For the purposes of brevity and consistent with the emphasis in the paper, much of these specification checks are focused on challenging the cesarean utilization results.

1. *Randomization inference*. Standard errors may be inaccurately estimated in difference-in-difference specifications when there are a limited number of overall analytical or treatment groups (Bertrand et al. 2004, Conley and Taber 2005). While the above specifications draw on a large number of treatment and control states, I perform hypothesis tests on the estimated national-standard coefficient using a randomization inference approach (Duflo et al. 2007) that allows for an estimation of the distribution of the treatment effect that is valid under any number of groups. Using only the set of states that did not amend their relevant laws over the sample period, I randomly generate 5,000 sets of placebo laws and estimate the primary specification on each of these simulated sets. I simulate the placebos so that the expected distribution of placebo law changes over time matches the distribution of actual law changes (Gruber and Hungerman 2008). I find that the estimated coefficient of the national-standard dummy (using actual variation in laws) falls in the 0.7th percentile of the empirical distribution of the 5000 estimated coefficients from the simulations, consistent with a p-value of 0.014.

2. *Dropping individual treatment states*. The significance and magnitude of the estimated coefficients in the primary cesarean specifications are robust to

systematic, one-by-one exclusion of each treatment state from the sample, as demonstrated by Figure C1.[7]

FIGURE C1: NATIONAL STANDARD COEFFICIENTS FOR STANDARDIZED CESAREAN SPECIFICATION: SENSITIVITY TO DROPPING INDIVIDUAL TREATMENT STATES



This graph presents regression results from 17 different specifications. Each specification modifies the specification estimated in Column 5 of Panel A of Table 2 to remove the inclusion of the indicated treatment state. Data on cesarean utilization is from the NHDS.

3. *Non-Discretionary Medical Events*. Certain medical events should not be expected to vary with national-standard adoptions due to the fact that physicians have little or no discretion over their usage or existence. As an additional check on the underlying specification, I thus test whether the abandonment of the locality rule leads to convergence between state and national rates of the following hospitalizations: (a) acute myocardial infarctions, (b)

[7] For the reasons set forth in Section 4, the specifications estimated in Table 3, which likewise explore the relationship between national-standard laws and convergence in cesarean rates, also exclude Maryland from the estimation sample and thus further demonstrates the robustness of the analysis to the exclusion of the one treatment state that modifies its national-standard laws by repealing a previous adoption.

gastrointestinal bleedings, (c) strokes, and (d) hip fractures.[8] These events represent situations in which patients will virtually always seek hospitalization (Skinner et al. 2005). As reported in Table C1, the average of the national-standard coefficients estimated in these falsification specifications – both in terms of average coefficient levels (-0.4 percentage points) and as a percentage of the associated dependent variable (roughly -2.8%) – are small in magnitude relative to the convergence estimates presented above (and not significantly different from zero).

TABLE C1. FALSIFICATION TESTS: THE RELATIONSHIP BETWEEN NATIONAL-STANDARD LAWS AND CONVERGENCE IN THE OCCURRENCE RATES OF VARIOUS NON-DISCRETIONARY MEDICAL EVENTS

	(1)	(2)	(3)	(4)
	AMI Discharge	Stroke Discharge	Gastro Bleeding Discharge	Hip Fracture Discharge
Coefficient of National Standard Law Dummy	-0.91 (1.71)	0.04 (2.12)	1.16 (3.32)	-1.99 (4.07)
Associated % Change in the Gap between the State and National Rates	-9.6%	+0.3%	+10.1%	-12.1%
N	1233	1233	1140	1233

* significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors corrected for within-state correlation in the error term are reported in parentheses. The dependent variables in each regression are the absolute deviation between state and national occurrence rates for the indicated medical event (normalized by the national rate). All regressions include state and year fixed effects and a set of state-year controls and state-specific linear time trends. Regressions are weighted by the number of discharges associated with each NHDS state-year cell. Discharge data is from the NHDS.

4. *HMO penetration rates*. While the above specifications include controls for the insurance status of the associated mother, the estimated coefficients are further robust to the inclusion of state-year controls for prevailing HMO penetration rates (from Interstudy Publications). While the coefficient of the contemporaneous national-standard indicator is -4.87 in the primary cesarean specification estimated in Column 5 of Table 2, that coefficient falls to only -4.95

[8] Appendix B provides further details on the construction of these hospitalization rates.

(still significant at 1% level) upon the addition of HMO penetration rates (and their squares). HMO-enrollment rates, however, were collected for the 1979 – 2005 period only and are unavailable for the District of Columbia over most of the sample period.

5. *Additional census variables*. While the specifications estimated above include controls for various demographic and supply characteristics provided in the NHDS records, the estimated coefficients remain nearly unchanged when I add certain additional demographic variables and other area characteristics, at the state-year level, from the decennial Censuses and the American Community Surveys, including the median household income and the percentage of the relevant state-year population that lives in an urban setting (among many others).[9] For instance, the primary coefficient of the national-standard variable changes from -4.87 to -4.42 (still significant at 1%) with the inclusion of the percent urbanization and household income variables.

6. *Marginal appropriateness analysis and low birth weight / pre-term deliveries*. The NHDS does not include data on birth weight or gestation; however, the medical literature identifies that low birth weight and pre-term delivery may be associated with a higher likelihood of cesarean delivery (see, for example, Poma 1999). As such, I consider a similar triage analysis using data from the Natality Detail files, which contain data on cesarean utilization in the post-1988 period only. As states increase their cesarean rates, I find evidence of a statistically significant reduction in the average incidence of low birth weight (less than 2500 grams) and of pre-term delivery (less than 37 weeks, or alternatively less than 27 weeks) among cesarean mothers. For instance, as the state cesarean rate increases by 10 percentage points, I find that the average

[9] These measures are obtained from the decennial Census files (1969-1999) and the American Community Surveys (2000-2006). Data for the intercensal years are linearly interpolated. With corresponding measurement-error concerns, I include these linearly-interpolated variables as an iterative specification exercise only.

incidence of low birth weight among cesarean mothers falls by 1 percentage point. These findings suggest that the marginal cesarean births are less likely to be of low birth weight and pre-term -- i.e., that marginal cesareans are less appropriate for cesarean delivery.[10] The first-stage findings associated with this reduced form analysis (using Natality Data) are consistent with the findings in Table 2 (using NHDS data) in documenting evidence of convergence in cesarean rates following national standard adoptions (estimating a roughly 3 percentage point reduction in the percentage absolute deviation between state and national cesarean rates in connection with the adoption of a national-standard law).

7. *State-year changes in clinical indications for treatment*. As discussed in Section 3, a concern arises that the estimated convergence findings may be attributable to within-state changes in the risk factors and indications for the procedures being investigated -- that is, changes in the group of patients in need of the procedure, as opposed to changes in clinical decisionmaking. The above analysis addresses this concern in two ways. For instance, in the cesarean analysis, I do the following: (1) risk-adjust state-year cesarean rates in the primary convergence specifications for the incidence of the three risk factors that very heavily indicate the usage of cesarean section (e.g., breech presentation) and (2) include right-hand-side controls for the state-year maternal (cesarean and non-

[10] I derive this finding by regressing the mean state-year incidence of low birth weight (or pre-term) among cesarean deliveries on (1) regional cesarean rates (i.e., an OLS approach), or (2) in the alternative, on the modified national standard indicator (note, however, there are only 3 treatment states over this time period), in each case with the addition of state and year fixed effects and an analogous set of state-year controls and state-specific linear time trends, though the results are robust to the exclusion of the controls and time trends. The birth weight results are also robust to the exclusion of very large (> 4000 grams) birth weight babies from the analysis (and thus simply comparing low birth weight to average birth weight).

cesarean delivery) incidence of these risk factors (e.g., the alternative convergence specifications indicated in Table 3 take this approach).[11]

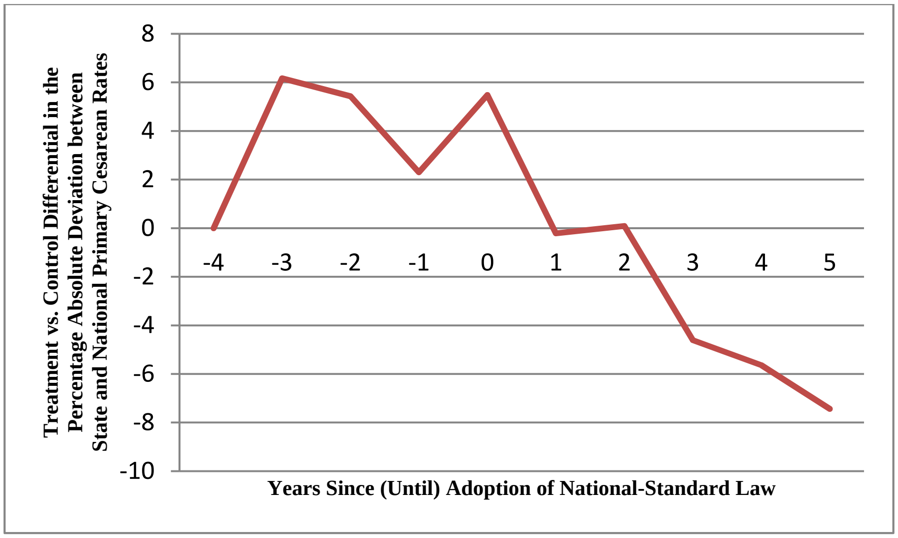
In unreported regressions, I confirm that the estimated convergence patterns found in Tables 2 and 3 persist without any of the above approaches. Moreover, I also estimate no relationship between the adoption of a national-standard law and the percentage absolute deviation between state and national means of the predicted cesarean probabilities across all mothers (cesarean and non-cesarean) -- *that is, while I find convergence in cesarean rates, I do not find convergence in a metric capturing the clinical need for cesarean delivery*. If anything, the results suggest that the gap between state and national means of the predicted cesarean probabilities across all mothers expands by a little over 10% (though statistically indistinguishable from 0) upon a national-standard adoption.[12] Similarly, using Natality Detail Data from 1977-2003, I find that the adoption of a national standard law is associated with a statistically-insignificant increase of only 0.60 percentage points in the percentage absolute deviation between state and national incidence of low birth weight (and a similarly sized, insignificant coefficient of -0.74 in an analogous regression based on pre-term deliveries).

In addition, to address these concerns, I estimate specifications that use the Agency for Health Care Research and Quality's specification of a "primary" cesarean rate, which excludes from the cesarean-rate calculation deliveries with any of the following risk factors: breech presentation, multiple delivery or previous cesarean delivery – i.e., the three factors forming the basis of the risk adjustment in the primary specifications. I find that the adoption of a national-

[11] The estimates likewise remain virtually unchanged when I include right-hand-side controls for each of the risk factors and complications included in the predicted-cesarean probability calculations, though these full set of risk factors are included in the post-1978 period only.
[12] Even at the bottom end of the 95% confidence interval, the results suggest only a 7% reduction in the gap between state and national predicted-cesarean-probability means.

standard law is associated with a roughly 8 percentage-point decrease in the percentage absolute deviation between state and national primary cesarean rates (significant at 1%), which is likewise consistent with a roughly 40% reduction in the gap between the state and national rate. Moreover, the estimated lead coefficients (both individually and jointly) are indistinguishable from 0. Figure C2 replicates Figure 1 using primary cesarean rates, demonstrating the results from a dynamic regression that includes 1-, 2-, 3- and 4-year leads and lags of the national-standard coefficient.

FIGURE C2: DYNAMIC PRIMARY CESAREAN SPECIFICATION RESULTS



This graph presents regression results from a dynamic specification that includes a 4-, 3-, 2-, and 1-year lagged dummies and 4-, 3-, 2-, and 1- year lead dummies for national-standard rule adoptions. Each point draws on the estimated lagged, lead and contemporaneous national-standard dummy coefficients and reflects the change over time in the percentage absolute deviation between state and national cesarean rates, where time is defined with reference to the years prior to and subsequent to national-standard law adoptions. The period of time prior to the 4th year before national-standard law adoptions represents the reference group (with a value set at 0). The specifications estimated in Figure 1 include the full set of state fixed effects, year fixed effects, state-specific linear time trends and state-year controls. Data on cesarean utilization is from the NHDS.

Finally, in unreported regressions, I likewise confirm that the set of cardiac results presented in Tables 4 and 5 remain virtually unchanged when I take various approaches to accounting for state-year changes in the incidence of clinical indications for intensive cardiac treatment and/or cardiac testing (e.g., angina, hypertension, and congestive heart failure, among others), including the use of right-hand-side controls for the state-year incidence, among the full adult sample, of such factors and for the standardization of cardiac utilization rates for the incidence of such factors.

As such, it does not appear that the convergence findings are being driven by shifts in clinical indications for treatment.

TABLE C2. OLS ESTIMATES OF THE RELATIONSHIP BETWEEN STATE CESAREAN RATES AND MEAN STATE NEONATAL OUTCOMES

	(1) 5-Minute Apgar Scores	(2) "Good" 5-Minute Apgar Score	(3) 28-Day Neonatal Mort. Rate
Panel A: Sample of Deliveries from All States			
Coefficient of Risk-Adjusted Area Cesarean Rate (logged)	0.40	-0.32	-5.88
	(0.55)	(0.26)	(14.60)
N	748	748	766
Panel B: Sample of Deliveries from States with Initially-Below-Average Cesarean Rates			
Coefficient of Risk-Adjusted Area Cesarean Rate (logged)	-0.28	-0.22	-10.60
	(0.44)	(0.35)	(24.43)
N	372	372	372
Panel C: Sample of Deliveries from States with Initially-Above-Average Cesarean Rates			
Coefficient of Risk-Adjusted Area Cesarean Rate (logged)	0.91	-0.29	-6.30
	(0.75)	(0.39)	(16.13)
N	376	376	394
Dependent Variable Logged?	YES	NO	YES

* significant at 10%; ** significant at 5%; ***significant at 1%. Robust standard errors are reported in parentheses (corrected for within-state correlation in the error term). Area cesarean rates are risk-adjusted for the incidence of each of the maternal risk factors identified in the Natality records. All regressions include state and year fixed effects and a set of state-year controls and state-specific linear time trends. Regressions are weighted by the number of deliveries associated with each state-year cell. Data on Apgar scores and cesarean rates is from the 1989 – 2004 Natality Detail Files. Data on neonatal deaths is from the 1989 – 2004 Vital Statistics Mortality records.

8. *Ordinary least squares estimates of the relationship between state cesarean rates and mean state neonatal health outcomes*. To supplement the primary triage and health outcomes analyses presented in Tables 7 - 9, I also present results in Table C2 from ordinary least squares regressions of state-year means of various neonatal outcomes on state-year risk-adjusted cesarean rates.

9. *Dynamic difference-in-difference regressions*. The specifications estimated in most of tables throughout the paper include a 2-year lead indicator for the adoption of a national standard law to appease concerns that the estimated impact of such laws began in the period of time prior to their adoptions. Taking an even more dynamic approach, Table 6 presents results from difference-in-difference regressions that include 1-, 2-, and 3-year lead indicators for those specifications that explore the relationship between national standard laws and the percentage absolute deviation between state and national (1) cesarean rates and (2) intensive cardiac treatment rates (CABG or PTCA). Table 6 also includes F-statistics testing whether the lead coefficients are jointly equal to zero. The results of this richer dynamic exercise largely generalize to the remaining specifications. For instance, in those dynamic specifications that estimate the relationship between the incidence of cesarean delivery and the modified national-standard indicator (i.e., an extension of Table 3), the coefficient for the contemporaneous indicator is 1.04 (significant at 10%), while the coefficients of the 1-, 2-, and 3-year leads are 0.4, -0.7, and 0.7, respectively. Each of the lead coefficients is individually insignificant, and, with a p-value of the relevant F-test of 0.5, the coefficients are likewise jointly indistinguishable from 0.

Only in the case of the VBAC specification does the relevant F-test reject a hypothesis of jointly 0 lead coefficients. In that instance, the gap between state and national VBAC rates appears to fall significantly 3 years prior to the adoption of a national standard law only to rise by an equivalent amount in the following year. As such, even in this instance, it does not appear that a consistent convergent trend in utilization rates materialized in the pre-adoption period.

Additional References

Conley, Timothy G., and Christopher R. Taber. Forthcoming. "Inference with 'Difference-in-Differences' with a Small Number of Policy Changes." Review of Economics and Statistics.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using Randomization in Development Economics: A Toolkit." In *Handbook of Development Economics*, ed. T. Paul Schults, and John Strauss, 3895-62. North Holland: Elsevier Science Ltd.

Poma, Pedro A. 1999. "Correlation of birth weights with cesarean rates." *International Journal of Gynaecology and Obstetrics*, 65(2): 117-23.