Do Physicians Respond to Liability Standards?

by

Michael Frakes, Matthew Frank and Seth Seabury[*]

In this paper, we explore the sensitivity in the clinical decisions of physicians to the standards of care expected of them under the law, drawing on the abandonment by states over time of rules holding physicians to standards determined by local customs and the contemporaneous adoption of national-standard rules. Using data on broad rates of surgical interventions at the county-by-year level from the Area Resource File, we find that local surgery rates converge towards national surgery rates upon the adoption of national-standard rules. Moreover, we find that these effects are more pronounced among rural counties. (JEL: I10, K13, K41)

1 Introduction

The standard of care arguably constitutes the bedrock of tort law. In most tortious contexts, including medical malpractice, actors are not liable for the harms that they cause to the extent that they have taken the level of care that the law expects of them. As stated long ago by Oliver Wendell Holmes, the imposition of an objective standard of care under tort law has the potential to redirect the behavior of those individuals otherwise inclined to operate at levels of care in contradiction with such standards (Holmes 1881). Defining what this standard is and discussing how it is established is an exercise that occupies a substantial portion of a typical first-year Torts course and that has received considerable attention in tort theory (Shavell 2004). Yet, despite all of this classroom and theoretical attention, the empirical tort literature to date has paid little attention to this crucial component of the law. Instead, this literature has viewed the substance of the law in relatively abstract terms and has simply identified the effects of tort law by focusing on variations along the remedy dimension. For instance, in the medical context— where our focus will lie in this paper—many studies have attempted to isolate the influence of medical malpractice law by estimating physician responses to the adoption of caps on non-economic damage awards or to similar reforms that largely operate by curbing the expected amount of damages that plaintiffs may recover in the event of liability (Kessler and McClellan 1996, Currie and

MacLeod 2008). Such reforms do not directly alter what is expected of physicians; they simply diminish the harms associated with any deviation from those legal expectations. In this paper, we begin to fill these gaps in the empirical literature by asking instead how physician decisionmaking may change when we alter the standards to which physicians are held in the first place.

To be sure, remedies are a critical component to the deterrent channel of the law and our analysis does not undercut the merits of the remedy-focused approaches taken by others to date. Such studies indeed provide useful information. We simply note that this traditional approach cannot answer all of the questions that we may have regarding the effects of malpractice law. For instance, investigations into the effects of damage caps may do little to illuminate the projected impacts of next-generation malpractice reforms that are often discussed in the policy sphere and that focus more on altering the way in which physicians are substantively evaluated—e.g., extending definitive safe harbors to physicians who comply with established clinical practice guidelines. In order to shed light on the possible impacts of reforms that operate along this more substantive dimension to the law, it is important to conduct independent investigations into the sensitivity of physicians to the standards of care to which they are held.

To demonstrate the limitations that may come from simply relying on investigations into the penalty / remedial dimension of the law, let us turn briefly to an analogy in criminal law. That is, consider the case of legally imposed speed limits. Remedies are clearly critical to the effectiveness of such rules. After all, for speed limits to have any hope of regulating driving behavior, one must receive a penalty to the extent the limit is violated. With this in mind, policymakers may modify penalty levels in order to alter the behavior of drivers. Any attempt to understand fully the link between driving regulations and driving behavior, however, would be challenged to the extent one only observes what happens when penalties are changed in this manner. Consider a hypothetical road on which drivers may only desire to drive 35 miles per hour (criminal forces aside), given the particular characteristics of the road and given drivers' tolerance for risk. Imposing a 35 mile per hour speed limit on this road may not alter behavior significantly beyond this baseline. After all, drivers would continue to drive at the preferred 35 miles per hour. In this event, any increase or reduction in fines at this speed limit may not alter behavior as well. Any observed insensitivity to speed limit penalties in this case does not mean, however, that such drivers are universally unresponsive to speed-limit laws. If this jurisdiction were to instead lower the speed-limit on this road to 25 miles per hour, it is indeed quite possible that some drivers would adjust their inherent desire to drive 35 miles per hour and instead drive closer to the newly imposed limit. But we would never know the full picture here without studying both what happens when jurisdictions modify speed limit penalties and what happens when they modify speed limits themselves.

This speed limit hypothetical is indeed quite analogous to the case of medical malpractice law in the United States. In the above hypothetical, the speed limit was set at a point at which drivers already preferred to operate, leaving little independent room for the law to push drivers any further (at least on average). In

the case of malpractice law, such an outcome would not simply happen by happenstance; rather, standards of care in U.S. tort law are generally structured with this outcome precisely in mind. That is, courts do not generally determine on their own the clinical practices that they expect of physicians. Rather, they have deferred this exercise to the medical industry itself and have simply sought to define liability standards by asking what physicians customarily do in similar situations. Thus, by imposing that physicians should maintain the practice styles that they otherwise already wish to follow on their own, the present malpractice system may be doing little independently to redirect physician behavior, at least as a first order matter. Under such a structure, therefore, practices may not change substantially upon the adoption of damage caps or related reforms, consistent with much of the empirical evidence put forth to date (Frakes 2015). Nothing in this discussion, however, suggests that physicians would be unresponsive to a change in the law that held them to a different standard altogether (analogous to the reduced speed limit above). We can only begin to understand the full potential of tort law to influence physician behavior once we start paying independent attention to the way in which physicians respond to the clinical standards expected of them under the law.

The challenge, of course, comes in capturing variations in the law along the standard-of-care dimension, allowing us to glean such an understanding in the first place. Fortunately, as is discussed in most first-year Torts casebooks (for example, Farnsworth and Grady 2009), the vast majority of states have experienced a reform of this nature. While states have indeed set medical liability standards by deferring to customary industry practices, there has been some variation in the law concerning what sets of physicians to look to in identifying the operable customs. Historically, jurisdictions applied what is known as the "locality rule," whereby physicians would be held to a standard of care determined by the customary practices of those physicians practicing in the same locality, where locality generally entailed smaller regions within states—e.g., the nearby town or county. Early reforms to these rules were procedural in nature and facilitated the ability of local plaintiffs to find experts from outside of the community to testify as to local customs (to alleviate concerns that physician experts would be unwilling to testify against their local peers). In the latter part of the 20th Century, most states (via state supreme court decisions) took a more substantive turn with their reforms and abandoned the locality rule entirely in favor of rules that simply held physicians to the standards of care determined by the customary practices of physicians nationally.[1]

These distinctions in liability rules are of particular relevance in light of the rampant geographical variations that exist in clinical practice patterns, a phenomenon that has been the subject of a massive literature in medicine and health economics (Skinner 2011). In other words, in light of such substantial regional variations in practices, this move from a locality rule to a national-standard rule constitutes a meaningful and substantial change in the clinical

---

[1] Frakes (2013) reviewed both case and statutory law and documented the evolution of each state's malpractice standard-of-care laws from the mid-1970's to the present. Table 1 documents this legal evolution. Specifically, the table identifies states that amended their standard-of-care laws— along the national vs. non-national dimension—during the study period.

expectations placed upon physicians. Different theories have been proffered as to why clinical practices in different regions of the country developed along distinct pathways. It could be that variations results from physicians operating off of a geographically-limited set of information (Wennberg and Gittelsohn 1973). There could also be more nuanced, dynamic forces in play in which local conditions in one region are such that specialization in one practice style pushes towards heavy utilization of that practice style (because of, for example, productivity spillovers), even though local conditions in other regions may push towards specialization in a different type of practice (Chandra and Staiger 2007). Regardless of the precise cause, medical liability "locality rules" may operate to cement these regionally distinct practice styles, either by discouraging physicians from deviating from those local customs that have developed or by providing comfort to physicians wishing to maintain such customs.

Table 1. Variations in National-Standard Rules (1977-2005)

State	Year Adopted	Year Dropped	State	Year Adopted	Year Dropped
Alabama	1980		Nevada	1979	
Colorado	1983		New Mexico	1978	
Connecticut	1984		Oklahoma	1984	
Delaware	1999		Rhode Island	1998	
D.C.	1980		S. Carolina	1981	
Indiana	1992		South Dakota	1988	
Maryland		1994	West Virginia	1986	
Mississippi	1983		Wyoming	1981	
Montana	1985				

Notes: years of adoption and repeal (if applicable) of laws requiring that physicians follow national (as opposed to local) standards of care in malpractice actions. States are only included if their relevant malpractice laws varied within the 1977 – 2005 period. *Source*: Frakes (2013).

Frakes (2013) showed that the abandonment of such locality rules and the adoption of national-standard rules may have broken those isolating forces and led to substantial regional convergence in various obstetric, cardiac, and diagnostic procedures—i.e., local practices did indeed begin to conform more strongly with national practices upon the relevant change in the law. Frakes (2013) analyzed the effects of national-standard adoptions using rich inpatient discharge data, which, among other things, facilitated the inclusion of control variables for various comorbidities. Limited to inpatient records, however, Frakes (2013) was forced to focus on those types of clinical scenarios—i.e., obstetrics and intensive cardiac intervention—in which care is almost exclusively inpatient in nature. As such, it is unclear whether his findings generalize to a broader range of clinical settings. Another limitation with Frakes (2013) is that the geographical scope of the analysis is at the state level (given data limitations). The analysis may have validly captured the effects of national-standard adoptions on across-state variations in practices; however, given that locality rules operated along sub-state dimensions, Frakes (2013) sacrificed the ability to estimate more geographically nuanced (and geographically accurate) specifications.

In this paper, we attempt to fill some of these gaps left by Frakes (2013). Using data on overall rates of surgical intervention (inpatient and outpatient)

derived from the Area Resource File (ARF), we likewise find that local surgery rates converge (both upwards and downwards, as appropriate) in the direction of the national mean upon the abandonment of locality rules and contemporaneous adoption of national standard rules. While the ARF must sacrifice some of the clinical traction afforded by the specifications estimated in Frakes (2013), the results nonetheless provide some evidence to suggest that the convergence in utilization patterns documented in this prior research generalize to a broader range of surgical intervention settings, beyond just obstetrics and cardiac care. Accordingly, we build on this prior research in providing even further evidence of the empirical relevance of liability standards themselves.

Moreover, we confront this analysis using data at the county-year level, thereby demonstrating that the findings in Frakes (2013) generalize to specifications that perhaps better approximate the geographical expansion in liability rules stemming from national-standard adoptions. Taking advantage of the county-level data, we also stratify the analysis by the population density of the relevant region. Discussions surrounding the distinctions between locality rules and national-standard rules have historically suggested that the divergence in outcomes between these rules may be most pronounced in rural communities. Perhaps consistent with this general sentiment, we find evidence suggesting that the convergence in local surgery rates in the direction of national mean rates is strongest when focusing on those regions in the bottom quartile of population density.

The papers proceeds as follows. In Section 2, we discuss the data used and the methodology employed to explore the impacts of liability-standard reforms on treatment utilization patterns. In Section 3, we present the results of this analysis. Finally, in Section 4, we conclude.

2   Data and Methodology

For our analysis, we use data on characteristics of local health care markets from the Area Resource File (ARF).[2] At the county-by-year level, the ARF assimilates information from the Census Bureau, the American Medical Association, and the American Hospital Association (AHA), among other sources. Critical for our purposes is information stemming from the AHA annual surveys on the number of inpatient and outpatient surgeries performed in the relevant county during the given year, along with information from the AMA records on the number of physicians practicing in the relevant regions.[3] With this information, we impute the propensity of physicians within a relevant area to follow an intensive surgical practice style, as opposed to a non-intensive style (analogous to Chandra and Staiger, 2007). This intensive-versus-non-intensive

---

[2] The ARF is now called the Area Health Resource File (AHRF), but we use data form earlier years when it was called the ARF. For practical purposes the distinction is not important because both use(d) the AMA Masterfile for source data.

[3] Surgeries included in this count were performed either in inpatient facilities in the given county or in outpatient ambulatory surgery centers in that county that are affiliated with the hospitals in the AHA records.

decision captures the physician choice inherent in our investigation into the effects of liability standards on physician decisionmaking.

The data spans the 1977-2005 period explored in Frakes (2013). We use linear projections to impute data for missing years in the ARF files. The ARF also includes data on several sources of demographic information relevant for our study, including the age and race distribution of states over time (collected from the US Census) and average per-capita income, as well as county population density and educational attainment rates across various education categories. In addition to the ARF, we use as covariates data on other tort reforms from Ronen Avraham's Database of State Tort Law Reforms (DSTLR), 4th Edition.[4] Though using data on surgery rates at the county level, we merge in information on malpractice standards rules and tort reforms at the state-year level.

With these data, we estimate a difference-in-differences model of the impact of switching from local to national standards in medical malpractice cases on the rate by which physicians in the affected regions performs surgeries in both inpatient and outpatient settings. An important factor to note is that the effect of changing the local standard depends on whether those standards are initially high or low. That is, changing to a national standard calls for more intensive treatment only to the extent that the existing local custom is to provide less intensive treatment. In such instances, we effectively ask whether the shift to national standards is associated with an increase in practice sensitivity. This allows us to test whether practice intensity rises when the law changes to seek such an outcome. Similarly, the move towards a national standard will represent a decrease in the practice intensity expected of physicians in regions whose existing local custom is to provide an above-average level of practice intensity. With respect to these regions, we effectively ask whether the adoption of a national standard rule is associated with a decrease in practice intensity. Thus, following Frakes (2013), we estimate the extent to which the adoption of national standard rules promotes *convergence* in outcomes—reducing variation across areas so that counties become more similar in their surgical proclivities.

Using a general measure of practice intensity based on the number of total surgeries performed in a county extends the analysis of the impact of standards on practice style beyond the cardiac and obstetric focus of Frakes (2013). However, this extension does come with some cost. Focusing on the more general clinical setting complicates the choice of denominator to use. That is, this analysis does not capture the simplicity of the obstetric analysis, whereby we may simply model the number of cesareans performed out of the subsample of deliveries observed. In this obstetric analysis, it is straightforward to model the choice between the intensive and the non-intensive-alternative—i.e., between cesarean and vaginal delivery—since both can be observed and since the rate of cesarean deliveries out of total deliveries necessarily describes both rates. Finding a metric that fully describes the propensity of physicians in a region to choose the intensive surgical option over the non-intensive option in a more general setting is more challenging.

---

[4] The specific reforms we use include noneconomic damage caps, punitive damage caps, collateral source rule reform, joint and several liability reform and caps on attorney contingency fees.

One choice for a denominator in this broader setting would be simply to examine the number of surgeries per capita. However, this choice imposes two notable concerns. First, for much of the sample, the county population data is not available with precision. Population counts are linearly interpolated to cover the intercensal years. Second, using the total area population as the denominator in the surgery rate calculation is problematic because the change in the liability rule may be spuriously correlated with a release of certain access barriers that increase the number of physicians servicing the relevant area. This access relief might itself lead to an increase in the number of surgeries and thus an increase in the number of surgeries per population. Since physician access may have improved under this hypothesized incident, however, it is possible that the relevant region may have experienced an even bigger increase in utilization of the non-intensive practice alternative. In this event, the average utilization intensity for that region may have actually declined. By just normalizing surgery counts by population, we may not be able to capture changes in the non-intensive alternative that we want in order to form the right utilization rate.

To address these concerns, we normalize the total surgery counts by the number of total physicians practicing in an area. If we assume, as we do, that physicians take either a non-intensive or an intensive—i.e., surgical—practice style, then a given increase in the number of physicians in the region should increase both intensive and non-intensive treatment in a proportional fashion (roughly speaking). Normalizing the number of surgeries by the total physician count reflects this and allows us to focus on changes in clinical practice in an area independent of total aggregate treatment. Stated differently, normalizing the number of observed surgeries by the number of physicians in the regions allows us to capture the expected practice style of a given physician operating in that region, which arguably facilitates a more direct test of the link between liability standards and physician practice behavior.[5] Moreover, to the extent that the rate of surgeries to physicians increases, it implies roughly that surgeries may be rising relative to their non-intensive alternative, which is the physician choice we are attempting to model with this analysis. Over the sample period, the data reflect roughly 50 surgeries per physician across counties (weighted by county size).

We implement our approach using the following empirical specification (weighted by the size of the county-by-year cell):

$$(1) \qquad S_{c,t} = \alpha + \lambda_t + \delta_c + \beta_1 NS_{c,t} + \beta_2 HI_c + \beta_3 \big(NS_{c,t} \times HI_c \big) + \beta_4 X_{c,t} + \varphi_{s,t} + \varepsilon_{c,t}$$

where c indexes county and t indexes year; $NS_{c,t}$ represents an indicator for a national-standard law (which actually evolves on a state-year basis); HI_c indicates that a county is located in what is initially a high treatment-intensity county[6]; $X_{c,t}$

---

[5] Regardless the findings do not differ substantially when normalizing by population.

[6] Note that the variable *HI* would drop out of the models with county fixed effects. Initially low-intensity (high-intensity) treatment counties are identified by whether the rates of inpatient and outpatient surgeries in those counties were below (above) the national mean surgery rate during the five year period leading up to the national-standard adoption. The results are robust to alternative classification schemes—e.g., focusing on the first three years of the sample instead.

contains additional county-year or state-year factors;[7] and year fixed effects, λ_t, control for fixed differences across years. The term $\left(NS_{c,t} \times HI_c\right)$ represents the interaction between initially high-intensity areas and the national standard laws. State-specific linear time trends are captured by $\varphi_{s,t}$, which control for slowly-moving correlations between physician behavior in a state and the adoption of a national-standard law. Note that when we implement (1), we include county fixed effects as well (δ_c) so the HI_c term drops out. Surgery rates (as defined above) are captured by S_{ct}.

As above, the predicted effects of national-standard adoptions on the proclivities of physicians to treat patients intensively depends on the baseline tendencies of the relevant county during the locality-rule regime. That is, national-standard rules are not predicted to increase or decrease surgery rates across the board. Rather, they are predicted to induce convergence in surgery rates towards the national norm—i.e., an increase in rates for those areas that start below the norm and a decrease in rates for those areas that start above the norm. This interaction specification is designed to model this convergence effect. There are two primary coefficients of interest in this specification. The term β_1 identifies the extent to which the adoption of a national-standard law is associated with a change in surgery rates for those areas that began the sample with initially low-intensity practice styles—i.e., those with $HI = 0$. The term β_3 identifies how the effect of the standard differs in areas that are initially high intensity. The sum $\beta_1 + \beta_3$ represents the combined effect of national standards in the initially high intensity areas. In other words, this interaction specification allows us to explore the effects of national standard adoptions separately on: (1) those areas with initially low practice intensities and with respect to which a national standard adoption arguably entails a heightening of the practice intensity expected of medical practices and (2) those areas with initially high practice intensities and with respect to which a national standard adoption arguably entails a reduction in the practice intensity expected of medical practices.

We test several versions of this model to check for sensitivity and highlight different results. We vary the inclusion of covariates, fixed effects and time trends to determine whether other confounders are possibly driving the results. We stratify the results according to the population density of the county to determine whether the effects vary in urban or rural areas. We also estimate dynamic versions of this model to test for the presence of leads (i.e., pre-existing trends) or lags in the effect. Primarily, we log the dependent variable, so the effect of the standard rules (which are dummy variables) can be interpreted in percentage terms (among other reasons for the log transformation). Coefficients in the Tables below are multiplied by 100 to facilitate this interpretation in the case of the log-transformed specifications. Also, for all regressions, we produce variance estimates that are adjusted to allow for clustering at the state level.

---

[7] These county-by-year characteristics include the percent of the population over age 65, percent non-white, percent Hispanic origin, average per-capital household income, the number of hospital beds per population, the number of physicians per capita (general and across different specialty categories) and the overall population density, as well as dummy indicators for the tort reform variables described above and attainment rates for various education categories.

On a final note, we focus our primary specifications on those counties that on average over the sample period have at least 50 hospital beds. There are a large number of counties within the data that do not have a hospital at all (not surprisingly concentrated largely in rural areas). Many counties have a small hospital presence early in the sample period and none later in the sample, likely as a result of industry consolidation over the 1980s and 1990s. Since our goal is to evaluate the intensive-versus-non-intensive nature of medical decisionmaking within regions and how the practice style followed in an area depends on the prevailing malpractice-standard rules—as opposed to how overall surgery levels depends on malpractice forces—we choose to condition this analysis on some baseline level of supply. Given some hospital presence itself, the intensive versus non-intensive choice within regions becomes a meaningful one that we can capture continuously with the data. To be sure, the decision to abandon a hospital presence in a county could be a reflection of a decision among physicians in that county to focus more on non-intensive measures. However, a decision of this nature represents a rather discrete jump that is likely to be more driven by fundamental market factors than malpractice considerations, complicating the empirical exercise. Nonetheless, we also present results from unrestricted specifications that include data from all counties in all years, demonstrating the general insensitivity of the results to this sample selection criterion.

3 Results

3.1 Convergence Results

In Table 2, we present estimates of specification (1) above, capturing the relationship between the adoption of national-standard rules and regional convergence in the utilization rates of surgical procedures. The table allows us to determine the association between national-standard rules and surgery rates separately for counties that were initially low intensity in nature and counties that were initially high intensity in nature. The different columns report results from separate regressions with different combinations of covariates and state-specific time trends, as noted in the table. In the full covariate specification, we include controls for prevailing county-year-specific rates of physician concentration rates (that is physicians per capita, both total and across various specialties) and hospital bed concentration rates (bed counts per capita). However, since such supply outcomes may also be impacted by national-standard reforms, which may affect the interpretation of the results, we present results from specifications that do not include these supply-focused covariates (we likewise breakout the results in this manner in Table 4 discussed below).

Frakes (2013) showed that in the settings of cardiac care and obstetrics, national standard adoptions were associated with regional convergence in practice intensities. In other words, the low intensity areas became more intensive and the high intensity areas became less intensive. In Table 2, we find that this relationship holds in the more general setting of all surgical procedures. As demonstrated by the first row in Panel A of Table 2, in the base model with no

covariates, a national standards law is associated with an 8.82 percent increase in the surgery rate for the initially low intensity areas (capturing β_1 from specification (1)). There is an opposite effect in the initially high intensity areas, which can be gathered by summing the base effect (i.e., the coefficient from row 1) and the coefficient of the interaction term set forth in row 2 (capturing β_3). This combined effect suggests that there is a 4.52 percent decrease in surgical procedures in the initially high intensity areas. While this pattern of coefficients suggests convergence from both the bottom and the top, neither effect is statistically significant at conventional levels in the log-transformed specifications. Both effects, however, are significant at the 1 percent level in the non-log-transformed models, as demonstrated by Panel B. The size and significance of these effects generally increase as we add covariates, with the full log-transformed model indicating a 16.59 percent increase in the surgery rate in the initially low intensity areas and a 9.88 percent decrease in the initially high intensity areas (both significant at the 5% level or better).[8] Consistent with past work, this confirms that national standards promote convergence in clinical practices across geographic areas.

Malpractice scholars have suggested that the sensitivity of practices to malpractice pressure differs across geographic areas of varying population density. Matsa (2007) found evidence of greater sensitivity in rural regions and argued that it is due to rural physicians facing more uninsured malpractice costs and more elastic demand for health care. There are other reasons beyond Matsa's general claims to consider the effect of standard-of-care laws on rural and urban areas separately. Much of the historical discussion surrounding the need for the locality rule emphasized its rationale within rural areas – suggesting, for example, that it might be impractical to expect rural physicians to follow practice styles elsewhere. Inherent in such discussions is the notion that retreats from locality rules may be most impactful in rural regions. Moreover, one might presume that practices in urban areas are heavily shaped by so many influences - competitive forces, information networks, etc. – that arguably may be more attenuated in rural environments. If this is the case, physicians in rural regions may place greater weight on whatever signals are provided via the malpractice system regarding proper clinical practices (liability fears themselves aside). Altogether, these different factors suggest that we would expect the impact of a switch to national standards to be more pronounced in rural areas.

---

[8] In separate, unreported regressions, we estimate an alternative interaction specification in which the coefficient of the national-standard dummy can be interpreted as the effect of national standard laws for the initially high-intensity areas. This alternative specification allows for a more direct test of the statistical significance of the combined effects from Table 2. This alternative approach confirms the statistical significance of this combined effect.

Table 2. Relationship between National-Standard Laws and the Rate of Inpatient
and Outpatient Surgical Procedures

	(1)	(2)	(3)	(4)	(5)
Panel A. Log transformation of dependent variable (setting surgery counts = 1 in zero-surgery counties)					
National-Standard (NS) Law Dummy	8.82 (7.05)	11.02* (5.82)	11.27 (6.82)	12.65 (7.97)	16.59** (6.56)
NS Law * Above-Average Intensity County	-13.34 (8.05)	-19.53*** (7.22)	-19.27** (7.79)	-19.51** (8.92)	-26.47*** (9.77)
Panel B. No log transformation of dependent variable					
National-Standard (NS) Law Dummy	8.87*** (3.01)	9.71*** (2.93)	10.54*** (3.60)	11.03*** (4.10)	10.17*** (2.97)
NS Law * Above-Average Intensity County	-18.46*** (5.13)	-19.70*** (4.79)	-15.68*** (4.20)	-15.90*** (4.64)	-20.67*** (4.77)
N	49,645	49,645	49,645	49,645	49,645
County-Year Covariates	NO	YES	YES	YES	YES
Covariate Tort Reforms	NO	NO	YES	YES	YES
County-Year Physician Concentration Rates and Bed Count Rates	NO	NO	NO	YES	YES
State-Specific Linear Time Trends	NO	NO	NO	NO	YES

Note: Table reports the difference-in-differences estimates of the impact of national standards laws on the use of inpatient and outpatient surgical procedures per capita, in logs (Panel A) and levels (Panel B). Coefficients are multiplied by 100 in Panel A. Regressions include county and year fixed effects. Data are at the county-year level. An above-average county is defined as a county that initially had higher than average utilization of inpatient and outpatient surgical procedures (coefficient dropped with inclusion of county fixed effects). County-year covariates include the percent of the population over age 65, percent non-white, percent Hispanic origin, average per-capital household income, and the population density. Regressions are weighted by the size of the relevant county-year cell. Robust standard errors are reported in parentheses, adjusted to allow for correlation (clustering) at the state level. A *, ** or ***, indicates statistical significance at the 10%, 5% or 1% level, respectively.

To test this, we follow Matsa and estimate the treatment intensity convergence regressions with the full set of covariates (corresponding to Column 5 in Table 2) for counties in each quartile of population density. We report these results in Table 3. Each column reports the results of a separate regression of counties by density quartile.[9] Generally, the findings appear to support the notion that standards are more binding in rural areas. The results across population density quartiles are generally similar to the overall results, with two-sided convergence evident in most (but not all) cases. However, the relationship is most pronounced in rural counties, with national standards being associated with a 29.08 percent increase in the surgery rate in initially low intensity counties and a 25.27 percent

---

[9] Note that the imbalance in the number of observations across the four quartiles is due to the fact that a substantial number of counties in rural regions either have no hospitals consistently over the entire duration of the sample or a small number of hospital beds on average over the sample period and are thus excluded according to the sample selection criteria set forth above (with many of such zero-hospital observations simply being dropped by the natural logging of the outcome variables anyway).

decrease in the initially high intensity counties. These effects are more than twice in magnitude than in any other population density category. Interestingly, the effect does not appear to be monotonic in population size, as there is no clear relationship between population size and the effect of standards across the other three population density quartiles.

Table 3. Relationship between National-Standard Laws and the Rate of Inpatient and Outpatient Surgical Procedures (Logged), By Population Density

	(1)	(2)	(3)	(4)
	1ST POPULATION DENSITY QUARTILE (RURAL)	2ND POPULATION DENSITY QUARTILE	3RD POPULATION DENSITY QUARTILE	4TH POPULATION DENSITY QUARTILE (URBAN)
National-Standard (NS) Law Dummy	29.08 (22.31)	5.75 (30.51)	-2.25 (22.79)	13.45 (10.19)
NS Law * Above-Average Intensity County	-54.35* (31.57)	-28.85 (21.18)	-5.79 (25.66)	-19.88* (10.49)
N	5,619	9,050	14,052	20,924

Note: Table reports the difference-in-differences estimates of the impact of national standards laws on the use of inpatient and outpatient surgical procedures according to population density. Coefficients are multiplied by 100. Data are at the county-year level. An above-average county is defined as a county that initially had higher than average utilization of inpatient and outpatient surgical procedures (coefficient dropped with inclusion of county fixed effects). Population density is defined according to the county-level distribution across all sample years. Regressions also include county and year fixed effects, as well as county-year controls for the percent of the population over age 65, percent non-white, percent Hispanic origin, average per-capital household income, the number of hospital beds and physicians per population and the population density. Regressions also include a set of state-specific linear time trends. Regressions are weighted by the size of the relevant county-year cell. Robust standard errors are reported in parentheses, adjusted to allow for correlation (clustering) at the state level. A *, ** or ***, indicates statistical significance at the 10%, 5% or 1% level, respectively.

With any difference in differences analysis, one is concerned that the differential in the relevant utilization rates between treatment and control groups is actually due to some pre-existing trend which emerges before the reform. To explore for such differential pre-period trends, we re-estimate the model including several lead variables for the national standard law. If there is a differential pre-existing trend in states that adopt a national standard rule, we would expect the coefficient on these lead variables to be significant. In addition, to test for any dynamic changes in the relationship between standards and surgery rates we include lags terms as well. Given the interactions between standards and the initial level of intensity we employ in our main analysis, we simplify this dynamic exercise by breaking the analysis into two samples: initially below-average surgical intensity counties and initially above-average surgical intensity counties.

We present the results of this dynamic specification in both tabular and graphical forms. Across the columns in Table 4, we estimate the leads and lags specifications sequentially adding additional covariates, with Panel A capturing the effect of national-standard adoptions for initially below average counties and

Table 4. Relationship between National-Standard Laws and the Rate of Inpatient and Outpatient Surgical Procedures (Logged), Dynamic Specifications Separately for Initially Above- and Below-Average Intensity Areas

|	(1)	(2)	(3)	(4)
Panel A: Initially Below-Average Surgical Intensity Counties (operationalized by dropping treatment counties that are of initially above-average surgical intensity)				
Coefficient of National-Standard Law Dummy				
4-Year Lead Dummy	-4.11 (5.05)	-3.84 (4.65)	-3.83 (4.67)	-0.41 (4.96)
2-Year Lead Dummy	-0.03 (1.20)	0.59 (1.32)	0.34 (1.28)	0.20 (1.35)
Contemporaneous Dummy	9.03 (5.84)	11.17** (5.28)	10.47* (5.90)	10.97 (6.57)
2-Year Lag Dummy	0.83 (1.32)	0.72 (1.12)	0.44 (1.31)	-0.15 (1.42)
4-Year Lag Dummy	0.60 (3.02)	-0.17 (3.03)	0.01 (3.40)	0.64 (3.53)
N	40,703	40,703	40,703	40,703
Panel B: Initially Above-Average Surgical Intensity Counties (operationalized by dropping treatment counties that are of initially below-average surgical intensity)				
4-Year Lead Dummy	2.34* (1.40)	1.93 (1.70)	0.28 (2.44)	1.53 (2.59)
2-Year Lead Dummy	0.60 (1.36)	0.21 (1.45)	0.80 (1.55)	0.77 (1.66)
Contemporaneous Dummy	-3.22 (2.18)	-4.51* (2.45)	-3.95 (2.61)	-3.82 (2.81)
2-Year Lag Dummy	-1.59 (1.11)	-2.32** (0.97)	-2.06** (0.96)	-2.21** (0.88)
4-Year Lag Dummy	-1.72 (2.35)	-3.41 (2.32)	-2.98 (2.42)	-2.44 (2.51)
N	47,042	47,042	47,042	47,042
County-Year Covariates	NO	YES	YES	YES
Covariate Tort Reforms	NO	NO	YES	YES
County-Year Physician Concentration Rates and Bed Count Rates	NO	NO	NO	YES

Note: Table reports the difference-in-differences estimates of the impact of national standards laws on the use of inpatient and outpatient surgical procedures, including lag and lead terms on the standards to test for the presence of pre-existing trends in surgery utilization that predict the adoption of national standards laws. Coefficients are multiplied by 100. Data are at the county-year level. Regressions include county and year fixed effects. An above-average county is defined as a county that initially had higher than average utilization of inpatient and outpatient surgical procedures (coefficients dropped with inclusion of county fixed effects). County-year covariates include the percent of the population over age 65, percent non-white, percent Hispanic origin, average per-capital household income, and the population density. Given no evidence of pre-treatment trends, we elect not to estimate specifications that impose state-specific linear trends in these dynamic results. Regressions also include a set of state-specific linear time trends. Regressions are weighted by the size of the relevant county-year cell. Robust standard errors are reported in parentheses, adjusted to allow for correlation (clustering) at the state level. A *, ** or ***, indicates statistical significance at the 10%, 5% or 1% level, respectively.

Panel B for initially above-average counties. Note that we specify the four-year lead coefficient as turning from 0 to 1 four years prior to the national-standard-law adoption in the relevant state and staying at 1 thereafter. The other lead and lag variables are specified accordingly. With this specification, the coefficient of the 4-year lead coefficient captures the differential in surgery rates between treatment and control counties in the period of time between the two- and four-year period prior to a national-standard adoption and the years prior to that period. The coefficient of the two-year lead variable then captures the *subsequent* change in this differential as we move into the period marked by the two years prior to the national-standard adoption—i.e., the differential surgery rate between treatment and control states in the two-year-prior period relative to the two-to-four-year-prior period—and so on. To capture the cumulative time trend in the differential surgery rate between treatment and control states—with time entailing years prior to and subsequent to national-standard adoptions—one naturally adds up these subsequent coefficient levels. In Figures 1 and 2, we plot this cumulative time-trend differential, focusing on those specifications that include the full set of controls, with Figure 1 focusing on the initially-below-average surgical intensity counties and Figure 2 focusing on the initially-above-average intensity counties. Effectively, these figures allow us to trace out the effects of national standard adoptions on surgery rates over time.
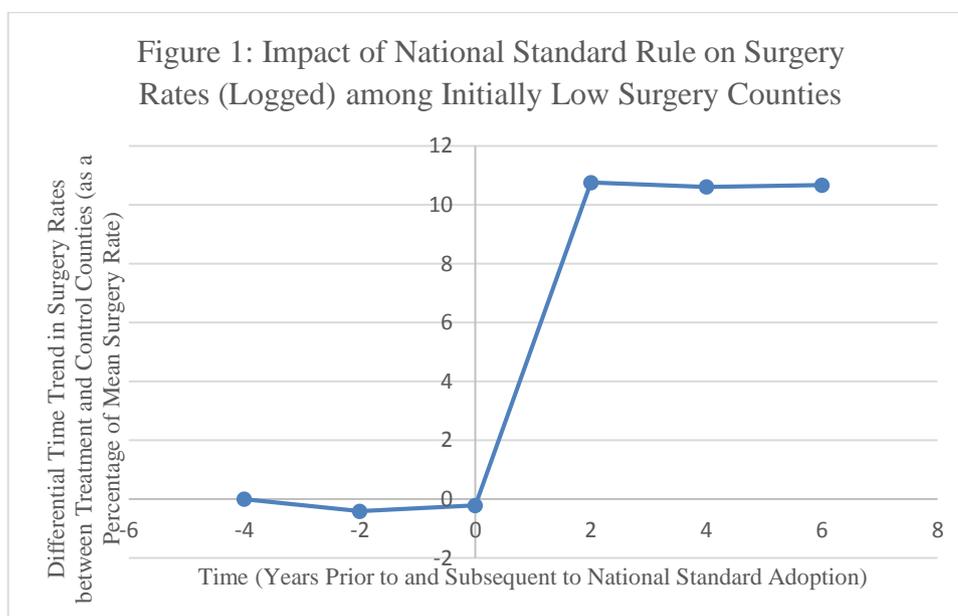
Figure 1: Impact of National Standard Rule on Surgery Rates (Logged) among Initially Low Surgery Counties

Notes: this figure plots the cumulative effects of national-standard adoptions over time in initially low intensity counties and is derived from the coefficients reported in Column 4 , Panel A of Table 4. The time marked by -4 in the figure represents the reference period for this trend and signifies the period of time that ends at the point that is four years prior to the adoption of a national standard law—i.e., capturing all earlier years. The level of this reported time trend at the time mark of -2 represents the subsequent change (beyond this reference point) in the differential surgery rate between treatment and control counties as one moves into the period that is between two and four years prior to the adoption of a national standard. Likewise, the level of the trend at the time mark of 0 represents the degree to which this differential surgery rate in the period of time between 0-2 years prior to adoption differs from the differential surgery rate during the reference period, and so on.

Our results are encouraging and support the validity of our approach. Importantly, there is no evidence of a pre-existing trend that could generate our main results. Only one of the 16 different lead terms in Table 4 is statistically significant at even the 10% level, and it has the opposite sign as predicted (it indicates a positive trend in the high intensity areas). There appears to be some evidence of a lag in the effect in high intensity counties, though the contemporaneous dummy consistently has the largest effect.[10]
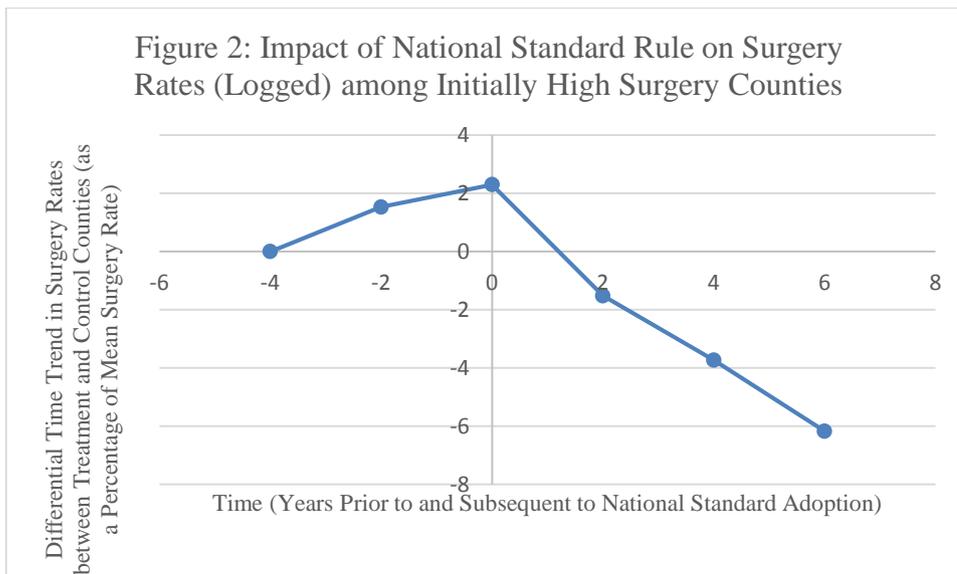
Figure 2: Impact of National Standard Rule on Surgery Rates (Logged) among Initially High Surgery Counties

Notes: this figure plots the cumulative effects of national-standard adoptions over time in initially high intensity counties and is derived from the coefficients reported in Column 4 , Panel B of Table 4. The figure is otherwise structured in accordance with the approach taken in Figure 1.

3.2 Robustness Exercises

In Table 5, we replicate the full-covariates specification from Table 2 (Column 5), while subjecting the analysis to varying sample selection criteria. For the reasons set forth above, in our primary specifications, we confine the sample to those counties that have a mean number of hospital beds over the sample period of at least 50. As stated before, a notable number of county-year cells in the sample do not have hospitals at all, with counties sometimes having dropped hospitals over the sample period, leading to the potential for discrete jumps towards 0 in surgery rates for some counties over time. As demonstrated by Table 5, this sample-selection choice appears to be of little significance, as the basic pattern of results is largely insensitive to the choice of other cutoffs (e.g., 25, 75 and 100 beds) or to the imposition of any criterion at all (Column 1). The results also do not change substantially when we drop zero-surgery cells via the log

---

[10] In Table 4, we do not actually fully separate all counties into initially high and initially low counties. Rather, we do this only for treatment counties, effectively keeping all control counties as viable controls for both panels. The results, however, are robust to an alternative approach that fully separates counties based on initial intensity classification. We elect primary specification given the choice (discussed above) to define initially high intensity and initially low intensity with reference to the years leading up to the adoptions of national standard laws themselves.

transformation of the surgery rate (Panel A), as opposed to including such cells and replacing zero-surgery counts with a surgery count of 1 before log transforming (Panel B, the primary approach utilized above). In Panel C, we show the results across the various sample-selection cutoffs with specifications that do not log-transform the dependent variable. We find comparable results in all such instances.[11]

Table 5. Robustness Checks: Alternative Sample-Selection Criteria and Treatment of Zero Cells

	(1)	(2)	(3)	(4)	(5)
	NO SAMPLE RESTRICTIONS	RESTRICTED TO MEAN BED COUNT OF AT LEAST 25	RESTRICTED TO MEAN BED COUNT OF AT LEAST 50	RESTRICTED TO MEAN BED COUNT OF AT LEAST 75	RESTRICTED TO MEAN BED COUNT OF AT LEAST 100
Panel A: Log transformation of surgery rates, thus dropping cells with zero surgeries					
National-Standard (NS) Law Dummy	9.28*	10.26**	10.73**	12.21**	-11.57**
	(5.03)	(5.09)	(5.12)	(5.02)	(5.06)
NS Law * Above-Average Intensity County	-18.14**	-19.00**	-18.55**	-19.24***	-17.40***
	(7.46)	(7.54)	(7.38)	(6.4)	(5.88)
N	61333	59437	48801	39515	33250
Panel B: Replace number of surgeries equal to 1 for zero-surgery cells before log transforming surgery rates (Primary Approach)					
National-Standard (NS) Law Dummy	6.46	13.87*	16.59**	17.34**	18.07**
	(5.76)	(7.30)	(6.56)	(6.67)	(7.89)
NS Law * Above-Average Intensity County	-15.21	-25.27**	-26.47***	-27.26***	-26.13**
	(10.12)	(12.30)	(9.77)	(9.08)	(9.75)
N	72807	61335	49645	39963	33552
Panel C: No log transformation of surgery rate					
National-Standard (NS) Law Dummy	7.59***	9.33***	10.18***	10.20***	8.82***
	(1.92)	(2.55)	(2.97)	(3.42)	(3.18)
NS Law * Above-Average Intensity County	-18.64***	-20.47***	-20.67***	-20.44***	-17.48***
	(4.69)	(4.58)	(4.78)	(4.84)	(4.42)
N	72807	61335	49645	39963	33552

Note: Table replicates the specification estimated in Column 5 of Table 2, but subject to the indicated modification. Coefficients are multiplied by 100 in Panels A and B. Columns 2-5 restrict the sample to those counties that have a mean number of inpatient beds over the sample period of the indicated amount. Column 1 presents results from specifications with no such restrictions. A *, ** or ***, indicates statistical significance at the 10%, 5% or 1% level, respectively.

Finally, we address one concern stemming from Medicare's implementation of the Prospective Payment System (PPS) in October, 1983, which broadly shifted

---

[11] Only in the case of Column 1 of Panel B (which involves no sample selection criteria and which log transforms the dependent variable while imposing a surgery count of 1 in the zero-surgery counties) do we find statistically insignificant results. While the point estimates in this specification are slightly lower, we nonetheless find the same pattern of convergence.

hospital reimbursement from a full-cost system to a casemix-adjusted flat-rate system. Acemoglu and Finkelstein (2008) demonstrate that this implementation increased hospital capital-to-labor ratios. Counties may have felt varied effects of the PPS implementation system given different pre-PPS Medicare exposures. While there is no reason to believe that any such differential Medicare exposure across counties is related to state national-standard rule adoptions, one may be concerned that a spurious correlation could confound the analysis, at least with respect to the estimated impacts of national-standard adoptions in initially low intensity counties, which may also be expected to experience increased capital-labor ratios upon national-standard adoptions (one may be less concerned that PPS implementation could explain the reductions in intensity associated with the initially high intensity regions). To first address this concern, we estimate the full-covariates specification from Table 2 (Column 5) but exclude the 1982-1985 period—i.e., the window around the implementation of PPS. In this specification, the estimated coefficient of the national standard dummy, capturing the effect of national-standard rules in initially low-intensity regions, increases slightly from 16.59 to 19.34 (still significant at 5 percent). This leaves us with greater confidence that the main findings above are not simply attributable to the confounding influence of PPS. Likewise, the estimated coefficient of the interaction term remains nearly unchanged at -29.62 (relative to a baseline of -26.47, likewise still significant at 1 percent).

4   Conclusion

Using data on surgery counts at the county-by-year level, we present evidence of regional convergence in broadly-specified surgical intervention rates—that is, initially low-intensity counties increasing their surgery rates and initially high-intensity counties decreasing their surgery rates—upon the abandonments of rules holding physicians to the standards of care applied in their locality and the contemporaneous adoption of rules requiring that physicians follow national-standards of care. Supplementing Frakes (2013), these results suggest that greater geographic standardization of medical liability rules are associated with greater geographical standardization in practices, which in turn is broadly suggestive of the empirical relevance of medical liability standards in the first instance.

References

Acemoglu, D., and A. Finkelstein (2008), "Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector," *Journal of Political Economy*, 116(5), 837-80.

Chandra, A., and D. Staiger (2007), "Productivity spillovers in healthcare: evidence from the treatment of heart attacks," *The Journal of Political Economy*, 115(1), 103-40.

Currie, J., and W. B. MacLeod (2008), "First Do No Harm? Tort Reform and Birth Outcomes," *Quarterly Journal of Economics*, 123(2), 795-830.

Farnsworth, W., and M. Grady (2009), Torts: Cases and Questions, Second Edition, Aspen Publishers.

Frakes, M. (2013), "The Impact of Medical Liability Standards on Regional Variations in Physician Behavior: Evidence from the Adoption of National-Standard Rules," *The American Economic Review*, 103(1), 257-76.

Frakes, M. (2015), "The Surprising Relevance of Medical Malpractice Law," forthcoming, The University of Chicago Law Review, 82, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2410426, retrieved July 31, 2014.

Holmes. O. W. (1881), The Common Law, Little, Brown & Company.

Kessler, D., and M. McClellan (1996), "Do Doctors Practice Defensive Medicine?," *Quarterly Journal of Economics*, 111(2), 353-90.

Matsa, D. (2007), "Does Malpractice Liability Keep the Doctor Away? Evidence from Tort Reform Damage Caps," *Journal of Legal Studies*, 36(2), S143-82.

Shavell, S. (2004), Foundations of Economic Analysis of Law. Harvard University Press.

Skinner, J. (2011), "Causes and consequences of regional variations in health care," in McGuire T, M. Pauly, and P. Barros (eds), Handbook of Health Economics, Vol 2, Elsevier, 45-93.

Wennberg, J., and A. Gittelsohn (1973), "Small Area Variations in Health Care Delivery: A population-based health information system can guide planning and regulatory decision-making," *Science*, 182.4117, 1102-1108.

M. Frakes
Law School
Northwestern University
375 East Chicago Ave., Chicago, IL 60611
U.S.A.
E-mail:
michael.frakes@law.northwestern.edu

M. Frank
Law School
Harvard University
U.S.A.
E-mail:
mfrank4@gmail.com

S. Seabury
School of Medicine, Department of
Emergency Medicine
University of Southern California
3335 S. Figueroa Street, Unit A
Los Angeles, CA 90089-7273
U.S.A.
Email:
seabury@usc.edu