Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Public Econ. Author manuscript; available in PMC 2017 November 01.

Published in final edited form as:

J Public Econ. 2016 November ; 143: 142–158. doi:10.1016/j.jpubeco.2016.09.002.

Does Medical Malpractice Law Improve Health Care Quality?

Michael Frakes[*] and

Duke University and NBER

Anupam B. Jena

Harvard Medical School, Massachusetts General Hospital, and NBER

Abstract

We assess the potential for medical liability forces to deter medical errors and improve health care treatment quality, identifying liability's influence by drawing on variations in the manner by which states formulate the negligence standard facing physicians. Using hospital discharge records from the National Hospital Discharge Survey and clinically-validated quality metrics inspired by the Agency for Health Care Research and Quality, we find evidence suggesting that treatment quality may improve upon reforms that expect physicians to adhere to higher quality clinical standards. We do not find evidence, however, suggesting that treatment quality may deteriorate following reforms to liability standards that arguably condone the delivery of lower quality care. Similarly, we do not find evidence of deterioration in health care quality following remedy-focused liability reforms such as caps on non-economic damages awards.

The imposition of liability under tort law is sometimes said to serve a purely private function—to correct the injustice created by a wrongdoer and/or to provide compensation to those harmed by that wrongdoer. Tort law is also often viewed, especially by economists, as serving a broader public function—to deter potential wrongdoers from committing costly and harmful errors in the first place.[1] Despite the prominent role that deterrence plays in the theoretical justification for tort law, very little evidence has been put forth to date suggesting that it fulfills this promise.[2] With the substantial direct costs that society expends in administering a system of tort liability,[3] it is crucial for scholars to continue to challenge the

[*]Duke University School of Law, 210 Science Drive, Durham, NC 27708 (michael.frakes@law.duke.edu).

[31]For a recent defensive-medicine analysis, see Mello et al. (2010).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

[1]Of course, those viewing tort law through the lens of its deterrence rationale do not contend that tort law is the sole institution in place that is designed to encourage safer practices.

[2]Dewees and Trebilcock (1992) and Schwartz (1994) provide a review of early investigations into deterrence in a range of tortious contexts: automobile accidents, products liability, workers compensation and medical malpractice. Mello and Brennan (2002) likewise review the limited evidence of deterrence put forth to date in the medical malpractice context. Polinsky and Shavell (2010) provide a recent review of studies on deterrence in products liability contexts, generally finding no convincing evidence of improvements in the level of care arising from liability fears (as opposed to market forces that otherwise compel the provision of sound care).

[3]In the context of medical malpractice jurisprudence alone, the United States spends over $7 billion annually in litigation expenses for defendants and plaintiffs and for the overhead incurred by medical liability insurers (Mello et al, 2010). Surveying the relevant literature on the costs of the tort system, Polinsky and Shavell (2010) claim that for every dollar received by plaintiffs from defendants in tort cases, roughly a dollar is spent in legal or administrative expenses.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

empirical foundations of this regulatory premise behind the law. In this paper, we address this challenge in the specific context of medical malpractice law, exploring whether medical liability forces induce—or at least have the potential to induce—the provision of higher quality care.

For these purposes, we take a novel approach and confront this inquiry by exploiting variations along the bedrock of tort law—the negligence standard. Physicians are only liable under medical malpractice law when their harmful actions fail to comply with an operable standard of care. Empirical investigations into the effects of the malpractice system on physician behavior often ignore this fundamental fact. They speak about liability "pressure," without necessarily asking: "pressure to do what?" In evaluating the potential for malpractice law to improve physician practice quality, we move beyond this abstract treatment of the law and center our analysis on a consideration of the particular expectations placed upon physicians by the liability system.

In other words, we ask how the quality of physician care is affected when the clinical standards against which they are judged in court are themselves altered. We explore this question by drawing upon the one significant liability reform of this substantive variety that the majority of states have implemented over time: the abandonment of rules requiring that physicians be judged according to the customary practices of local physicians and the contemporaneous adoption of rules requiring that physicians be judged according to national (or non-geographically-limited) standards of care (Frakes 2013). In light of the rampant regional disparities in care that have persisted across regions for decades and that have been the subject of a massive literature in health economics and medicine (see, for example, Chandra and Staiger 2007), one can view the move from a local to a national-standard rule as a meaningful and large alteration of the standards clinically expected of physicians (Frakes 2013).

Focusing on liability-standard reforms holds two key advantages over the traditional approach taken by the literature to tease out the influences of medical liability on physician practices. This typical route estimates the impacts of reforms that do not affect the standards against which physicians are judged, but rather reduce the expected damages awarded when a physician is found liable—e.g., the adoption of caps on awards for non-economic damages (pain and suffering damages). If one views liability-standard reforms as akin to reducing driving speeds by lowering the speed limit, adopting / removing damages caps is akin to changing driving behaviors by altering the associated speeding fines. One important limitation of this traditional, remedy-focused approach, however—and thus one key motivation for the primary liability-standards approach of this paper—is the mixed set of findings to date bearing on the necessary first-stage to this damage-cap inquiry. That is, scholars have not consistently found a link between damage cap adoptions and the frequency of malpractice suits. Ultimately, it is not clear that damage caps have systematically represented a sizeable enough change in liability pressure (or in the perception of liability pressure) from which to tease out the law's influence on behavior.

A second limitation of using damage-cap adoptions to illuminate the law's influence on health care quality stems from the fact that caps were generally adopted in the face of a

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

system that determined liability standards by deferring to customary industry practices. One might predict that damage cap adoptions would not have substantial impacts on *average* prevailing practice patterns when malpractice standards are simply designed to reign in errant practitioners and to reinforce those practices that the industry has already decided it wishes to perform in the first place. The danger of relying solely upon damage-cap adoptions to understand liability's influence is that one might infer from such modest impacts of cap adoptions that physicians are only weakly sensitive to the parameters of the liability system. Any such inference would be misguided, however, insofar as the modest effects of caps might simply arise from the particular structure of the industry-custom-focused system in which they were enacted. These findings do not suggest that physicians would fail to respond to a more substantive alteration of that system altogether, as some modern reform proposals aim to do—e.g., retreating from a reliance on industry custom and instead setting liability standards with reference to evidence-based clinical practice guidelines (sometimes phrased by policymakers and analysts as the extension of liability "safe harbors" to those who comply with certain delineated guidelines). By assessing whether physicians respond to a meaningful shift in liability standards—i.e., by assessing the empirical relevance of liability standards in the first instance—our analysis provides insights into the possible impacts of next generation liability reforms of this nature (in a way that could arguably not be done by simply studying how physicians respond to the adoption of non-substantive reforms such as damage caps).[4]

To take full advantage of the quasi-experimental framework made possible by the adoption of national-standard rules, it is necessary to obtain data on physician practices that span all states and several decades (to capture meaningful variation in the law). For these purposes, we primarily turn to the 1979–2005 National Hospital Discharge Survey (NHDS) records. Frakes (2013) likewise uses NHDS records to estimate the impact on physician practices of abandoning the use of a locality rule and the adoption of a national-standard rule, documenting substantial convergence in treatment and procedure rates for various cardiac and obstetric practices. Though novel in its focus on the substance of liability rules, the analysis by Frakes (2013) follows much of the malpractice literature to date in its emphasis on explaining physician behavior though utilization rates of particular intensive medical procedures. The focus on intensive utilization measures and costs leaves many unanswered questions regarding the potential of tort law to deter medical errors and improve quality to the extent that the intensity of medical care provided may not track very well with the quality of care delivered. As scholars have shown in observing the care received by Medicare beneficiaries, patients receiving care in high-spending regions do not necessarily receive more efficacious care—i.e., care linked to better health outcomes via randomized

[4]Bovbjerg and Berenson (2012) provides a summary of safe-harbor proposals of this nature, indicating that they remain on the short list of next-generation tort reform alternatives to damage-caps that policymakers have been considering and noting that the Bipartison Policy Center's deficit reduction proposal in 2010 identified safe harbors as a suggested liability reform. While policymakers have been encouraging greater experimentation with safe harbor plans, little remains known about their potential impacts. Under a grant from the Agency for Health Care Research and Quality, Kachalia et al. (2014) retrospectively reviewed closed malpractice claim records in the State of Oregon and determined that in nearly one third of the cases in which a clinical practice guideline applied to the clinical circumstances at hand, the adverse outcome could have been avoided had the guideline been followed. This retrospective analysis, however, leaves open the fundamental question of whether physicians would indeed alter their behavior in this manner should the state redefine the standard of care. We confront the analysis in this paper in an attempt to shed some light on this margin of response.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

control trials, such as the receipt of aspirin upon admission for acute myocardial infarction (Fisher et al, 2003). Accordingly, we attempt to go beyond traditional cost- and utilization-based studies and obtain more direct evidence regarding the link between medical liability forces and health care quality.

Of course, a major difficulty with this exercise is data availability. While it may be straightforward with traditional claims data such as the NHDS to identify if a particular intensive and costly intervention was employed—e.g., coronary artery bypass grafting—it is not always easy to ascertain if less expensive and more nuanced clinical practices or techniques were followed that may nonetheless be quality enhancing (to the extent the incidence of such refined, process-based measures are not recorded in the administrative records). One might naturally abandon the idea altogether of evaluating health care quality by exploring the inputs used in health care and focus entirely on metrics of health outcomes. To be sure, an output based approach is feasible when considering mortality as the outcome of interest—i.e., assessing whether medical liability forces decreased the likelihood that particular medical encounters will culminate in or contribute to the death of the patient. Indeed, one of the metrics of quality that we will employ in this paper follows precisely this approach. However, we do not end the inquiry there as patients may experience effects from health care that fall short of the mortality margin but that nonetheless hold significant welfare implications. When all researchers have at their disposal are encounter-level records of unidentified patients—such as the case with the NHDS—it is challenging to attribute any observed morbidity to the quality of care received during the encounter itself.

Fortunately, these are challenges that medical scholars have been working to overcome in recent decades in their attempts to develop markers of health care quality that can be used to judge provider practices. To explore the impacts of medical liability forces on health care quality, we rely upon these efforts by medical scholars and embrace a range of clinically-validated health care quality metrics that have been designed for use with inpatient discharge records, including, among others, avoidable hospitalization rates and rates of maternal trauma during childbirth. Such measures (1) arguably serve as proxies for non-intensive, quality-enhancing decisions that physicians make[5] and (2) represent intermediate outcome measures that bear on health statuses beyond mere mortality. Despite being limited by the use of historical inpatient discharge records, the various quality indicators that we employ collectively account for four of the five domains of quality targeted by the OECD's Health Care Quality Indicator's project and for each of the three domains of quality promulgated by the Agency for Health Care Research and Quality (AHRQ), whose widely-adopted indicators served as the inspiration for the specific metrics we employ. Though alternative sets of indicators exist, including those developed, for instance, by the Joint Commission on

[5]Again, Frakes (2013) documented a strong response in the utilization rate of various intensive procedures following the move to national-standard reforms. By studying the effect of national-standard rule adoptions on health care quality, we do not necessarily intend for this analysis to merely illuminate the effects of the utilization-rate responses found in Frakes (2013) —e.g., to illuminate the health status consequences of increases (decreases) in cesarean rates found in initially low-intensity (high-intensity) regions following national-standard rule adoptions. While our analysis may, in part, be capturing effects of this nature, it may also be picking up the effects of non-intensive, process-based decisions that physicians are making (and that are believed to improve health care quality), which are again difficult to directly observe with data sources such as the NHDS. For these reasons, we hesitate to treat Frakes (2013) as establishing evidence of the first-stage underlying our inquiry into the effects of medical liability on health care quality. Rather, we intend for our AHRQ-inspired measures to shed light on this more generally unobservable first-stage while simultaneously illuminating the effects of liability forces on the health outcomes of the affected patients.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Accreditation of Healthcare Organizations (JCAHO) and the Hospital Quality Alliance (HQA), AHRQ emerged as the market leader for users wishing to construct quality indicators with hospital administrative records (Hussey et al, 2007).[6]

For each measure of treatment quality explored, we find that when states modify their standard-of-care rules to expect physicians to provide higher levels of quality—e.g., when initially low quality regions adopt national standard rules—observed levels of quality increase substantially in the direction of such new expectations. Consistent with a causal response to the altered liability environment, we find little evidence to suggest that the observed improvements in the quality indicators (for the treatment states relative to control states) pre-dated national-standard reforms. However, when states modify their rules so as to condone the provision of lower quality care, physicians do not respond by reducing the quality of their practices. Collectively, these findings suggest that medical liability forces—under the right structural framework—hold the potential to elevate the quality floor.

The paper proceeds as follows. In Section I, we provide a background on liability-standard rules and discuss their potential impacts on physician practices. Section II discusses the data and empirical methodology. Section III presents the results of the empirical deterrence analysis. Finally, Section IV concludes.

I. Malpractice Law and Physician Behavior

A. The effect of substantive malpractice reforms

The clinical decisions that physicians make, including decisions to perform particular treatments and to deliver certain levels of quality, are likely to be shaped by a number of forces. One such force may stem from fear over liability for harming a patient through actions that fail to comply with the operable standard of care expected of physicians by law. For instance, if tort law expects that physicians perform surgical technique X when administering a particular treatment, physicians may feel pressured to adhere to technique X despite an inclination to otherwise deviate. Accordingly, depending on how physicians weigh liability considerations against other determinants of clinical practices, the imposition of a liability system may compel at least some physicians to follow the legally expected standards inherent in that system. It stands to reason that if the liability system is reformed so as to now expect that physicians follow a different set of practices—e.g., to perform technique Y in the hypothesized context—a new equilibrium will be reached in which at least some subset of physicians adjust their practices in the direction of Y. In order to gauge the sensitivity of physician practices, and of the quality delivered, to the standards expected of physicians under the law, we explore the impacts of a liability reform of just this nature.

To understand the quasi-experimental framework that we construct, it is first important to note that liability systems in the United States have generally deferred to the customary

[6]The JCAHO and HQA have developed a rich set of process-based indicators that bear directly on the more nuanced practices that physicians undertake that have been scientifically linked to better health outcomes. However, these alternative indicators are designed for use with more clinically based medical records, as distinct from the kind of claims data represented by the NHDS (Hussey et al, 2007). As such, we emphasize that the quality indicators we use in our analysis do not represent the full universe of indicators employed by others, but they do attempt to paint as broad of a picture as possible within the limited set of information available with administrative hospital records.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

practices of the industry itself in order determine the standards it imposes. Essentially, parties in U.S. malpractice suits call upon expert witnesses to testify as to customary physician practices in order to determine the benchmarks to which the defendant physicians are held. Where the law has varied, however, is with respect to which physicians it looks in order to assess customary practices. Historically, state malpractice laws judged physicians against customary practices of physicians working in the same locality, thereby expecting physicians to follow the practices applied by those around them. Deviations in care from these customary standards that led to adverse medical events were judged as negligent. Over the latter half of the Twentieth Century, however, the majority of states amended their substantive malpractice laws to abandon locality rules in favor of rules requiring physicians to follow national standards of care, thereby geographically harmonizing clinical expectations under the law.[7]

These distinctions in liability rules are of particular relevance in light of the substantial geographical variations that have persisted in clinical practices over time, a phenomenon that has been the subject of a large literature in medicine and health economics (Chandra and Staiger, 2007; Skinner 2011). Different theories have been set forth as to why clinical practices have developed along distinct regional pathways. One reason could be that physicians operate off of a geographically-limited set of information (Wennberg and Gittelsohn 1973). A more nuanced story of productivity spillovers could also be present such that local conditions in one region lead to specialization in one practice style while local conditions in other regions call for specialization in a different type of practice (Chandra and Staiger 2007).[8] Regardless of the precise mechanism, locality rules may cement these regional pathways, either by discouraging physicians from deviating from local customs or by providing comfort to physicians wishing to maintain such customs. The adoption of national standard rules may break these forces and lead local physicians to follow the vastly distinct practices followed elsewhere in the nation. The fundamental hypothesis that we test in this paper is whether prevailing rates of health care quality shift in the direction of such altered expectations.

Importantly, this analysis affords us the opportunity to separately test how physicians respond to changes in malpractice standards which in some instances expect higher levels of quality and in other instances lower levels. For each of the quality metrics explored, a number of treatment states began the sample period with high quality levels while a number of others began with low quality levels, in which event the move towards a national standard represented a change in legal expectations in both directions depending on the pre-reform level of quality. In light of the possibility that physicians may respond differently to an

[7]Online Appendix B provides further details on the evolution of malpractice-standard rules. Roughly 16 stated abandoned the use of local standards in favor of national standards over the sample period. As discussed in Frakes (2013) earlier reforms (i.e., in the mid-20th Century) to the initial locality rules focused on making it easier for outside experts to testify as to local standards—e.g., allowing a New York physician to testify as to local South Carolina standards. The substantive reforms of interest in this study pick up after these earlier procedural changes and focus instead on changes in the expected practices—e.g., requiring that South Carolina physicians now follow the practices applied elsewhere.

[8]In another more nuanced evaluation of the development and persistence of practice variations, Chan (2016) demonstrates that changes in the relative influence of members of clinical teams—due to the nature of residency and internship programs—may disrupt otherwise continuous learning processes and pose a source of friction on convergence in behaviors.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

elevation of what is expected of them relative to a slackening of what is expected of them, we test for asymmetrical responses to the adoption of national-standard rules.

B. The effect of remedy-focused liability reforms

A key conceptual point of this paper is that a medical liability system is largely characterized—at least on a substantive level—by an underlying set of standards imposed on physicians. As such, exploring the impacts of a liability system on physician behavior, as the malpractice literature endeavors to do, requires understanding how physicians respond to the standards imposed by that system. In the previous subsection, we set forth one way to approach this analysis, asking what happens when those clinical standards themselves are modified. One may shed further light on this general inquiry by evaluating the marginal impact of a *given* set of liability standards, as distinct from a changing set of standards. In other words, for a particular set of standards, what happens to practice patterns when we diminish the influence of such standards on the margin?

A straightforward way to confront this alternative analysis is to explore the impacts of those penalty-reducing reforms most traditionally emphasized by the malpractice literature—e.g., dollar caps on non-economic damage awards.[9] Such reforms diminish the penalties associated with liability without altering the basis for liability.[10] A reduction in observed levels of quality upon the adoption of a damage cap might suggest that the present liability structure is otherwise encouraging higher levels of care—i.e., is deterring poor medical practices.

The validity of this latter approach of course rests on the assumption that damage caps do indeed meaningfully reduce the expected harm to be imposed upon physicians. At first blush, one may doubt this considering that physicians generally face limited immediate financial risk from associated damage awards insofar as they are insured against such losses with coverage that is typically not experience rated (Sloan 1990, Zeiler et al. 2007). However, despite this limited immediate financial risk, physicians may face a number of uninsurable costs as a result of malpractice liability–e.g., reputational and psychological damage (Jena et al. 2011).[11] Damage caps may therefore reduce the amount of such uninsurable harms to the extent that caps decrease the probability of suit—e.g., by leaving plaintiffs and/or plaintiffs' attorneys less inclined to file suit (Shepherd 2014). The estimated impact of caps on claims frequency / likelihoods has varied significantly across studies

[9]The seminal example of this approach is provided by Kessler and McClellan (1996). Non-economic damages generally represent over half of the typical malpractice award (Hyman et al. 2009). Furthermore, caps on such damages represent the tort-reform measure that has been most commonly associated with an observed change in certain malpractice outcomes: claims severity, physician supply and malpractice premiums. See Mello and Kachalia (2011) for a comprehensive review of relevant studies. Paik et al. (2013) provides a recent example. Nearly thirty states currently have non-economic damage cap provisions in place, most of which were adopted during the malpractice crisis of the 1980's. In most specifications, we also explore the association between observed health care quality and certain additional types of tort reforms, including reforms of the collateral source rule, caps on punitive-damages awards and other "indirect" reforms. Further descriptions on all reforms are provided in the Online Appendix.

[10]Not all studies that explore the marginal impacts of liability pressure do so by evaluating the impacts of damage caps and related reforms. Some evaluate the marginal influence of liability forces in the present liability system by drawing on variations in malpractice premiums or liability payments per physician. See, for example, Baicker et al. (2007).

[11]Subject to certain exceptions, payments made on behalf of physicians to settle claims or to satisfy judgments must, under federal law, be registered in the National Practitioner Data Bank (NPDB), an electronic repository which is made available to hospitals and certain other health care entities. The NPDB was established by the Health Care Quality Improvement Act of 1986, as amended (42 U.S.C. 11101 et seq.). This repository may reinforce any reputational consequences of malpractice liability.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Mello and Kachalia 2011), casting some doubt on the validity of using damage caps to assess the influence of the law on practices. Nonetheless, reinvigorating the promise of a damages-cap-based approach, a recent study by Paik et al. (2013) draws upon significant variation in state laws to find a roughly 29 percent drop in claims frequency upon a damage-cap adoption. This uncertainty in the first-stage evidence surrounding the damage-cap approach to a deterrence analysis lends support to the alternative approach embraced below that focuses instead on meaningful variations in standard-of-care formulations.

Moreover, these first-stage concerns aside, one may doubt that damage-cap adoptions will substantially influence prevailing quality levels in light of the fact that caps have generally been implemented in the United States in the face of a liability system that sets standards according to customary industry practices. In particular, if one merely diminishes the penalty facing physicians under a customary-standards system, physicians may on average simply continue to adhere to the customary industry practices that are the bedrock of those standards in the first place. After all, under a liability structure of this nature, the law exerts little independent force of its own to redirect practice patterns (assuming, of course, that the law is not altering the set of physicians it is looking to in order to assess custom, as discussed above). Systematically, under such rules, physicians may only alter their practices in response to liability fears due to uncertainty in their beliefs as to how courts will assess customary practices—i.e., they may aim to deliver higher quality than otherwise customarily desired over fear that courts will misjudge customary practices to entail such higher practices. Any such channel of influence may not be expected to induce substantially higher levels of quality insofar as it largely entails guesswork on the part of physicians, rather than the delivery of clear signals as to how to improve quality. Damage caps may therefore only induce changes in physician practices to the extent that they reduce the cost of uncertainty to physicians about whether their practice patterns deviate negatively from customary market practices. Caps otherwise do not alter the clinical expectations being placed upon physicians. In the empirical analysis below, we test the hypothesis that damage cap adoptions have modest impacts on levels of treatment quality chosen by physicians.

C. Persistence in Effects

Prior to the change in liability standards envisioned in subsection (A) above, one of the key forces that may have shaped the election of physicians to practice Technique X are their clinical beliefs regarding the appropriateness of this approach in this clinical circumstance. These beliefs may have been shaped during medical school and/or during their residency and early career training. However, physician beliefs may also update periodically due to changes in the environments in which they practice (Molitor 2015; Phelps and Mooney 1993). To the extent beliefs are heavily weighted by local observation of one's own practices and of the practices of one's immediate peers, one's beliefs over proper clinical practices may evolve over time in response to the evolution of the set of practices forming these observations. With considerations of updating in mind, it is possible that the effects of altered liability standards may be especially long-lasting. In other words, should physicians shift to a new equilibrium following a standard-of-care reform—i.e., from technique X to technique Y—physicians may come to assimilate this liability-induced equilibrium into their belief structures. Even if physicians are resistant to updating the practice styles they develop

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

early in their career, a shift in practices stemming from a change in liability standards may nonetheless exhibit signs of durability to the extent physicians from new cohorts establish their ongoing practice styles in the face of this post-liability-reform environment. The significance of this prediction is that physicians (new and/or existing) may retain their inclinations to practice Technique Y in the years ahead even if liability pressure is ultimately reduced. Estimating the impacts of damage-cap reforms provides us the ability to evaluate this prediction considering that such liability-reducing reforms almost universally occurred in states that had previously adopted national-standard rules.

II. Data and Methodology

A. Data and Quality Measures

We employ two data sources which allow us to study clinically validated measures of the quality of care provided by physicians in both inpatient and outpatient clinical settings. Primarily, we construct several quality measures using the 1979 to 2005 National Hospital Discharge Surveys (NHDS). The NHDS data, supplemented with geographic identifier codes, provides rich inpatient discharge records over a long enough period of time to allow us to draw on an extensive set of within-state variations in both liability-standard rules and damage-cap provisions. In certain secondary analyses considered below, we use data from the 1987 to 2008 Behavioral Risk Factor Surveillance System (BRFSS) to capture various rates of cancer screening.

Foremost among those organizations promulgating health care quality indicators is the Agency for Healthcare Research and Quality (AHRQ). AHRQ measures are particularly useful for the present study insofar as they are designed for use with administrative inpatient databases such as the NHDS. In constructing quality indicators, we build off of the three domains of quality indicators developed by the AHRQ, supplementing the AHRQ measures with those capturing cancer-screening practices. We provide a brief overview of each metric below, with additional details regarding the construction of the resulting quality metrics provided in the Online Appendix.

Inpatient mortality for selected conditions—Following the AHRQ's Inpatient Quality Indicators, we first construct a composite inpatient mortality rate for selected acute medical conditions using NHDS data. Unlike mortality rates computed over broad periods of time and over an entire area affected by a relevant legal regime—which may be influenced by a number of socioeconomic factors—IQI-inspired rates are designed to capture mortality events associated with clinical encounters themselves. To rule out selection concerns—i.e., concerns regarding the liability regime impacting the probability of patients appearing in the inpatient environment in the first place—this measure focuses on mortality among a subsample of discharges in which the primary diagnosis code indicates select medical conditions (e.g., acute myocardial infarction, stroke, etc.) that uniformly entail hospitalization upon their occurrence. Though perhaps formulated differently from this AHRQ design, mortality events represent the health outcome that has been most studied by medical malpractice scholars to date.[12] In an effort to illuminate a broader range of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

consequences of medical liability forces, we extend our focus beyond this conventional stopping point.[13]

One concern about estimating the relationship between medical liability forces and inpatient mortality rates is that changes in the liability environment may also impact the duration of inpatient stays, which in turn may impact the likelihood of the patient dying during the hospitalization. This is problematic insofar as it may compromise our ability to target a change in inpatient mortality rates stemming from the quality of medical care delivered. In the Online Appendix we present results of mortality-rate specifications in which we control for state-year mean changes in the number of bed days associated with the relevant hospitalizations. The results are virtually identical with and without such controls, perhaps alleviating these particular concerns.

Avoidable hospitalizations as a measure of outpatient quality—Second, we use the NHDS to capture the rate of avoidable hospitalizations (AH) within each state-year cell, a measure that is inspired by the AHRQ's Prevention Quality Indicators. Though constructed using inpatient data, AH rates are thought to reflect the quality of care prevailing in the associated outpatient community. Such measures identify conditions—e.g., asthma—with respect to which proper outpatient care would have prevented (or at least reduced the likelihood of) hospitalization. To alleviate any concern that such metrics are confounded by physician choice as to whether to hospitalize patients with the indicated conditions (which would threaten the ability to infer outpatient quality), we construct an AH rate that focuses only on a subset of AHs with little physician discretion over the decision to hospitalize—e.g., congestive heart failure.[14]

As explained in the Online Appendix, we form AH rates by normalizing AH counts in the NHDS records by an index of NHDS hospitalizations for certain medical conditions—e.g., acute myocardial infarction, stroke, etc.—with respect to which there is virtually no discretion over whether or not to admit the patient. This denominator captures the size of the relevant state-year cell without itself being sensitive to legal or financial incentives.[15]

---

[12]Lakdawalla and Seabury (2009), for instance, find that higher county-level malpractice pressure leads to a modest decline in county-level mortality rates. Other studies have focused on slightly more targeted populations. For instance, numerous studies have estimated the impact of malpractice reforms on infant mortality rates (or infant Apgar scores, which are seen as predictive of infant mortality), generally finding no relationship (Klick and Stratmann 2007; Frakes 2012; Currie and MacLeod 2008; and Dubay, Kaestner, and Waidmann 1999). Similarly, Kessler and McClellan (1996) estimate a trivial relationship between liability reforms and survival rates during the one year period following treatment for a serious cardiac event (e.g., acute myocardial infarction). Sloan and Shadle (2009) undertake a similar analysis.

[13]Indeed, few malpractice studies have investigated the link between malpractice law and metrics of the clinically-validated nature promulgated by agencies such as AHRQ. Two exceptions are perhaps provided by Iizuka (2013) and Currie and MacLeod (2008). The latter study finds that damage cap adoptions increase preventable complications of labor and delivery, suggesting that higher liability pressure improves patient safety. Iizuka (2013) finds that certain tort reforms—e.g., collateral source rule reforms and punitive damage caps—increase labor and delivery-related complications. Interestingly, Iizuka finds no such relationship with non-economic damage-caps, despite the fact that such caps arguably amount to the most significant reduction in liability pressure out of the four traditional reforms that he explores (Paik et al, 2013). While obstetrics has formed the canonical example of research in empirical malpractice, obstetricians themselves account for less than 3 percent of U.S. physicians. The health care quality processes that we study form the bread-and-butter' practices of generalist physicians which form the largest group of practicing physicians. Moreover, we analyze the quality of health care provided in outpatient settings, a setting which accounts for over 20% of the nation's total health care dollars (CMS 2011) and has received no special attention by the malpractice deterrence literature. One exception is perhaps Baicker and Chandra (2005), which documents notable sensitivity in mammography screening to changes in malpractice premiums.

[14]Nonetheless, we note that the low-discretionary AH-rate findings mirror those based on AH-rate constructions that are not limited by the degree of discretion in the admission decision.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Maternal Trauma and Complications—The AHRQ's Patient-Safety Indicators (PSIs) capture complications and adverse events that take place in inpatient settings following surgeries, procedures and deliveries. Using NHDS data, we focus our analysis of PSIs on those related to delivery / childbirth. Many PSIs reflect the quality of care provided during surgeries, rates of which may be a function of the liability environment, implicating issues of selection. Rates of childbirth, on the other hand, are unlikely to be impacted by malpractice pressures. More specifically, we construct a composite obstetric trauma indicator, grouping together cesarean trauma events with vaginal delivery trauma events (with and without instruments). To look at a broader, but related set of obstetric-related complications, we follow Currie and MacLeod (2008) and also consider the incidence of preventable delivery complications—e.g., fetal distress, excessive bleeding, precipitous labor, prolonged labor, dysfunctional labor, etc.

To the extent we evaluate these health care quality metrics separately on the cesarean delivery sample and vaginal delivery sample and to the extent a change in liability standards also impacts the rates of cesarean deliveries, it may be difficult to determine whether an observed change in quality is due to actual improvements in quality or to sample selection considerations. For these reasons, we focus in the main specifications on calculating the rates of maternal trauma and preventable delivery complications out of the full delivery sample. Nonetheless, in the Online Appendix, we also show results of the national-standard-adoption specifications separately for the vaginal and cesarean delivery samples. The distinction is actually most relevant for the preventable delivery complication specifications, to the extent maternal trauma rates largely arise only in the case of vaginal deliveries anyway (in which event, we do caution the reader that the maternal trauma specifications may suffer from some sample selection limitations).

Cancer Screening—To round out our assessment of outpatient care quality and to provide us with a process-based alternative to the AHRQ indicators (which are more directly outcomes based), we use patient self-reports from the Behavioral Risk Factor Surveillance System from 1987 to 2008 to compute incidences of mammography, physical breast exam, Prostate-Specific Antigen (PSA) testing, digital rectal exam, pap smear, and sigmoidoscopy / colonoscopy, used to screen for breast, prostate, cervical, and colon cancer, respectively. As explained in greater detail in the Online Appendix, we use national cancer screening guidelines to select the relevant age groups for the analysis and the window period of relevance for the exam—e.g., mammography within the previous two years for females. Cancer-screening has received little attention by the empirical malpractice literature to date, with the exception of Baicker and Chandra (2005), which observed sensitivity in mammography utilization to malpractice premium changes.[16]

[15]Frakes (2013) documents little sensitivity in the rate of occurrence of each of the components of this index to the adoption of national standard laws (specifically finding no statistically significant relationship between national standard laws and convergence between mean state/year rates and national rates for each of the components of this index, with average convergence point estimates that are also notably smaller in magnitude in comparison with the relationship between national standard law adoptions and regional convergence in various treatment intensity metrics).

[16]We acknowledge that the merits of cancer screening is a controversial topic and that 100 percent adherence is not so evidently ideal. We emphasize, however, that we focus on testing the determinants of screening only within those bounds determined by the medical and scientific communities as warranting screening in the first place and are thereby not viewing universal adherence as the benchmark. Nonetheless, we do caution that decisions over the precise parameters to employ in setting age and frequency cut-offs for

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Notes regarding quality metrics—Of course, particular quality measures may induce greater liability fears among physicians than others, given the frequency with which lawsuits arise in the associated medical contexts. For instance, cancer screening rates may be especially good measures to study the link between malpractice and health care quality given that missed cancer diagnoses constitute a frequent basis for malpractice lawsuits (Schiff et al. 2013). Our goal, however, was not to pre-select quality metrics based on some priors as to which quality domains are more implicated in liability settings. Rather, we have attempted to paint a more general picture of the link between liability forces and quality and have thus endeavored to collect as many quality metrics as possible, bearing in mind data limitations. We note, in addition, the possibility of spillovers in effects. That is, one might observe a link between liability forces and avoidable hospitalization rates even if few lawsuits are based on such events to the extent that AH rates are proxying for more general changes in outpatient practices and to the extent that liability-induced changes in outpatient behavior in other contexts—e.g., cancer screening—spill over into the management of chronic care conditions underlying avoidable hospitalizations.

Descriptive statistics—On average, each NHDS state-year cell contains roughly 424 discharges associated with the selected conditions used in the composite inpatient mortality rate measure, our first quality indicator. The average inpatient mortality rate among this sub-sample is 8 percent, as presented in Table 1. Likewise, each state-year cell contains an average of roughly 600 low-discretionary avoidable hospitalizations. After normalizing these low-discretionary AH counts by the non-discretionary medical events index discussed above, the average low-discretionary AH rate across state-year cells equals 1.0. Furthermore, each state-year cell in the NHDS contains on average roughly 600 deliveries.[17] Within this delivery subsample, maternal trauma (third or fourth degree lacerations) occurs nearly 4 percent of the time and preventable complications occur nearly 16 percent of the time. Finally, cancer screening rates among the relevant BRFSS participants ranges, on average, from 40 to 73 percent.

To describe the variation in quality of care across regions, Column 2 of Table 1 provides, for each quality indicator, a measure of the average gap over the sample period between the mean state level and the associated mean national level. More specifically, following Frakes (2013), we summarize this gap by calculating the mean absolute deviation between the state and national indicator levels (for each year) and normalizing this rate by the national level. For instance, on average over the sample period, the mean maternal trauma rate within a state differed from the national mean trauma rate by an amount equal to roughly 26 percent of the national level. Because this measure is computed over the entire sample period, it

screening guidelines are not without much debate and disagreement. The one screening metric that we employ that garners arguably the most controversy—regardless of the age bands and regardless of the frequency—is PSA testing, where frequent false positives arise (Atkins et al, 2005).

[17]Using the delivery sub-sample to demonstrate the distribution of observations across state-year cells, we note that the state-year delivery counts in the sample range from a low of 0 to a high of 3,090, with 113 deliveries at the 25th percentile and 1161 deliveries at the 75th percentile. While the number of observations represented by some states in the sample are small, we weight observations in each specification by the appropriate size of the state-year cell—e.g., the number of deliveries in the obstetrics analyses and the number of non-discretionary medical conditions in the inpatient mortality analyses. Nonetheless, the resulting findings are not driven by any one particular large (or small) state in the sample and are robust to dropping each treatment state one-by-one. Moreover, as demonstrated by Column 5 of Table 3, the results are also robust to the use of non-weighted specifications.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

somewhat understates the regional disparity measure that is most relevant to our analysis. In particular, in early years of the sample and among states which began the sample under a locality-rule regime, the average gap between the state and the national rate, for each of the listed indicators, is substantially larger than the figures provided in Table 1. For instance, in the pre-1982 period for those locality-rule jurisdictions with below-average levels of maternal trauma rates, the mean gap between the state and national rate of maternal trauma was nearly 68 percent of the national level. In the empirical analysis below, we explore whether these gaps are narrowed through the adoption of national-standard rules (approaching the inquiry separately from each side of the regional quality distribution).

B. Specifications

To explore whether the quality of health care provided by physicians is affected by the clinical malpractice standards expected of physicians under the law, we estimate the degree to which state mean rates for the relevant quality measures converge towards their respective national mean rates as states adopt national-standard rules. In this investigation, however, we allow for a differential convergent response from the top and the bottom of the regional quality distribution—that is, we allow for a different response when the law changes so as to expect a higher level of quality compared to when the law changes so as to condone a lower level of quality. Following Frakes (2013), we estimate:

$$Log(Q_{s,t})=\alpha+\boldsymbol{\gamma}_{\mathbf{s}}+\boldsymbol{\lambda}_{\mathbf{t}}+\boldsymbol{\varphi}_{\mathbf{s}}t+\boldsymbol{\beta}_{\mathbf{1}}\mathbf{X}_{\mathbf{s,t}}+\boldsymbol{\beta}_{\mathbf{2}}\mathbf{REMEDY_REFORMS}_{\mathbf{s,t}}+\beta_3 NS_{s,t}+\beta_4 HIGH_QUALITY_s*NS_{s,t}+\varepsilon_{s,t} \quad (1)$$

where s indexes state and t indexes year. State fixed effects, $\boldsymbol{\gamma}_{\mathbf{s}}$, and year fixed effects, $\boldsymbol{\lambda}_{\mathbf{t}}$, control for fixed differences across states and across years, respectively. $\mathbf{X}_{\mathbf{s,t}}$ represents certain demographic characteristics of the patient population, along with certain mean characteristics of the represented hospitals. Specifically, this matrix includes the percentage of patients in various age-sex categories, race categories (white, black and other), insurance categories (private, government, no insurance and other), along with the percentage of patients visiting hospitals of various bed sizes (0–100, 100–200, 200–300, 300–500 and 500+ beds) and of various ownership types (for-profit, non-profit and government).[18] $\mathbf{X}_{\mathbf{s,t}}$ also includes certain state-year characteristics from external sources—i.e., mean state-year physician concentration rates and, in the case of the obstetrics specification, fertility rates (and OB/GYN concentration rates).[19] Additional details on all covariates (including those estimated with the BRFSS sample) can be found in the Online Appendix.

[18]Following Frakes (2013), in the AH rate and mortality rate specifications, we form the relevant mean demographic incidences using the sample of discharges in which patients present themselves for acute myocardial infarction, stroke, gastro-intestinal bleeding or hip fracture. This subsample consists of patients that will almost universally seek hospitalization upon the occurrence of the event, in which case the sample itself is generally not sensitive to the prevailing legal environment. In the obstetrics specifications, we form all relevant incidences using the subsample of discharges associated with deliveries.

[19]Alternative specifications estimated in the Online Appendix also include HMO-penetration rates as a control, though such measures exclude Washington D.C. due to missing data prior to 1997 for D.C. Obstetric specifications also include controls for cesarean delivery and episiotomy utilization. The maternal trauma specifications also include a control capturing the risk-status associated with the delivery, specified following Frakes (2013) as the predicted probability of cesarean delivery (PPC). PPC values are calculated

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REMEDY_REFORMS$_{s,t}$ is a matrix representing a set of indicator variables for the incidence of additional tort reforms, primarily caps on non-economic damages awards (in addition to collateral source rule reforms, caps on punitive damages and reforms to the joint and several liability rule). In some specifications, we include state-specific linear time trends, $\boldsymbol{\varphi_s}$t, to control for slowly-moving correlations between the relevant quality measures in a state and the adoption of tort reforms by that state. In yet other specifications, we control instead for state-specific linear pre-treatment time trends (the construction of which is discussed in further detail in the Online Appendix). $Q_{s,t}$ represents the relevant healthcare quality measure –e.g., the composite inpatient mortality rate or the avoidable hospitalization rate.

$NS_{s,t}$ represents an indicator for a national-standard law. $HIGH_QUALITY_s$ is an indicator for a state that began the sample period with an initial rate below the national mean for the relevant quality indicator.[20] For all indicators other than cancer screening rates, high levels of the various indicators represent lower levels of quality (and vice versa). The coefficient of β_3 in this interaction specification can effectively be interpreted as the association between national-standard laws and quality indicator levels for states with initially low levels of quality (i.e., with initially above-average indicator levels). For states with initially higher than average levels of quality, this same association is captured by the sum of β_3 and β_4.

C. Additional Information on National-Standard Reforms

A causal interpretation of the associations estimated below may of course be challenged to the extent national-standard adoptions are enacted in response to trends in quality attainment in the associated states. To illuminate these concerns, we began by performing a qualitative review of the circumstances surrounding the various reforms. Perhaps given the critical nature of the local-versus-national distinction to the structure of the negligence regime, each treatment state reserved this decision for the state's high court (as opposed to the legislature and as opposed to allowing the reform to evolve via lower state court decisions). When deciding to abandon the locality rule, each state high court did so while ruling upon the equities of a specific case, which dealt with only a particular clinical context—e.g., the local versus national custom in dosing a particular anesthetic. As such, while the decision held precedential value for all future clinical contexts—be it pertaining to anesthetic dosing or obstetric treatment practices—the judges were likely only subjected to evidence pertaining to that single clinical context at issue. That likelihood is encouraging insofar as it suggests that national-standard decisions were tangential to the broad-based quality trends under investigation in our analysis.

using fitted values of a logit model (estimated annually) of the incidence of cesarean delivery on a set of individual risk factors and complications.

[20]In the Online Appendix, we show that the various covariates employed in the analysis below do not differ meaningfully between the initially high quality group of states and the initially low quality group of states. For the treatment states, we actually make the initially-high / initially-low classification by observing matters in the several years preceding the adoption of a national-standard rule. Nonetheless, we note that this classification does change if we assess initially high versus initially low at the very beginning of the sample period for the treatment states. The classifications are performed separately for each quality indicate, though they are similar across the various indicators. Of the 7 treatment states that have initially high avoidable hospitalization rates, 6 also have initially high inpatient mortality rates, and 5 have initially high avoidable delivery complication and maternal trauma rates.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Of course, even though making this precedential decision in the face of just one clinical context, it is possible that the judges were provided with evidence on prevailing health care trends at large through information provided via briefs submitted by either the litigants themselves or third parties. Regarding the latter possibility, our search of historical litigation records suggested that third parties filed virtually no amicus briefs in connection with these critical locality-rule abdication cases, suggesting few of the external interest group pressures that likely characterize many legislative enactments, such as non-economic damage caps. Reviews of the opinions themselves and of the available briefs filed by the parties suggests that the judges instead were driven by broader principles, as opposed to recent trends in the medical marketplace—that is, they felt more generally that maintaining the locality rule would be inequitable for the specific plaintiffs at hand in light of the fact that the more general, structural rationales behind the rule—e.g., decentralized medical education—had withered over a half a century ago. All told, while it cannot be said that adoptions were randomly assigned across jurisdictions, a review of the circumstances surrounding national standard adoptions raises few red flags that might leave us especially concerned over the parallel trends assumption underlying the difference in difference approach.

In Table 2, we demonstrate the initial quality distribution (in quintiles) for each AHRQ-inspired quality indicator at the beginning and end of the NHDS sample, separately for the control states and for those that ultimately adopt national-standard reforms over the sample. Table 2 further indicates the number of treatment states that fall into each quintile. Adoptions of national-standard reforms over the sample period fall all throughout this distribution. Of course, one might be concerned that reforms do not appear randomly distributed across states to the extent that adoptions are more concentrated on the tails of the distribution, with the greatest weight arguably at the low-quality end of the distribution. This latter fact alone does not necessarily undermine the validity of the difference-in-difference approach, but it does raise concerns that the mean difference-in-difference estimates possibly capture mean reversion in the relevant quality indicators over the sample period as opposed to a true policy response. To illuminate this concern, we estimate the following specification separately for initially-high and initially-low quality states:

$$Log(Q_{s,t})=\alpha+\boldsymbol{\gamma}_{\mathbf{s}}+\boldsymbol{\lambda}_{\mathbf{t}}+\boldsymbol{\varphi}_{\mathbf{s}}\mathrm{t}+\boldsymbol{\beta}_{\mathbf{1}}\mathbf{X}_{\mathbf{s,t}}+\boldsymbol{\beta}_{\mathbf{2}}\mathbf{REMEDY_REFORMS}_{\mathbf{s,t}}+\boldsymbol{\beta}_{\mathbf{3}}\sum_{m=-4}^{4}NS_{s,t-m}+\varepsilon_{s,t}$$

In other words, for each of these sub-samples—initially high and low quality—we estimate event-study counterparts to our difference-in-difference approach in which we include the national-standard indicator variable along with a series of lead and lag indicator variables for the adoption of a national-standard reform. As explained further below, the estimates for the lead indicator variables provide a sense of whether any improvement (decline) in quality associated with national-standard adoptions precipitated such adoptions, which might otherwise call into question a causal interpretation of the findings—e.g., instead signifying regression to the mean. The coefficients of the lagged indicators likewise provide a sense of the time path of the subsequent response in the outcome measures to the national-standard reforms. To further rule out mean reversion concerns below, we evaluate the magnitude of these estimated lagged coefficients in light of the overall level of mean reversion we estimate

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in the sample. To the extent the immediate or near-term responses in health care quality to the adoption of national-standard reforms trump the magnitude of the natural levels of convergence in outcomes in the sample, one might be left with even greater confidence in attributing the estimated coefficients to an actual effect of changed liability standards on physician behavior.

III. Results

A. Liability-Standards Analysis

AHRQ-Inspired Measures and Preventable Delivery Complications—In our first approach to exploring the link between malpractice forces and health care quality, we estimate the interaction specification indicated by equation (1) above and explore whether health care quality is influenced by reforms that directly alter the clinical standards of care expected of physicians, effectively separating the inquiry by whether the reform expects physicians to follow higher or lower standards of care. The results of this exercise are presented in Table 3.[21] Each of the 4 panels represents results for different AHRQ quality indicators.

The coefficients presented in the first row of each panel can be interpreted as the association between the given quality indicator and the adoption of a national-standard rule in those treatment states that began the sample period with initially low levels of quality (i.e., where the above-average quality variable equals zero), representing those states with initially high levels of the respective quality measure. In the case of inpatient mortality rates for selected medical conditions, the low-discretionary avoidable-hospitalization (AH) rate, the maternal trauma rate and the preventable delivery complication rate, we estimate that the adoption of a national-standard rule in such states is associated with a substantial and statistically significant (across nearly every specification) decrease in the respective indicator measure and thus a substantial *increase* in health care quality. More specifically, in the basic difference-in-difference specifications with only state and year fixed effects, we estimate an 8.1, 53.7, 39.6 and 40.3 log point decrease in the respective quality indicator in connection with national-standard adoptions. The magnitude and precision of these estimates are generally robust to the inclusion of various state-year covariates, state-specific linear time trends, and state-specific pre-treatment linear trends (Columns 2–4).[22] Moreover, the results hold whether we weight the dependent variable by the appropriate state-year cell size (e.g., the number of deliveries in the obstetrics specifications) or estimate unweighted specifications (compare Columns 1 and 5), while also hold (generally) whether we estimate specifications in logs or levels (compare Columns 1 and 6). The coefficients of the levels specifications—in light of the mean rates of the various quality indicators—suggest a 10, 52, 48 and 14 percent decrease in the respective quality indicator in connection with national-standard adoptions.

[21]Reported standard errors in Table 3 and in all subsequent tables are clustered at the state level to allow for arbitrary within-state correlations of the error structure.

[22]The inclusion of state-specific linear trends, however, increases standard errors enough in the case of the maternal trauma and inpatient mortality analyses that the results are left statistically insignificant.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Considering that a national-standard adoption in such initially-low-quality states entails a shift in clinical expectations in the direction of higher quality, the results from this exercise suggest that liability reforms that affirmatively elevate the standards expected of physicians may indeed succeed in inducing higher quality practices. These findings demonstrate a substantial closing of the gap between low-quality regions and other regions upon the retreat from a local-standard-of-care rule. For instance, considering the initial mean gap between maternal trauma rates in locality-rule states and national maternal trauma rates (equaling 68 percent after normalizing by the national rate, as stated above), the above findings imply that roughly half of this gap is closed upon the move from local to national-standard-of-care rules.

In Figures 1a, 2a, 3a and 4a, we shed further light on these findings by focusing on the initially low quality states and estimating the dynamic specification set forth in equation (2) above, which includes a set of lead and lag indicators for the adoption of a national-standard reform (the underlying tabular regression results are presented in Table B1 of the Online Appendix). The plotted line reflects the difference between the treatment and control states in the respective quality indicator in the years leading up to and following the adoption of national-standard rules (where this difference is normalized such that it equals zero at the time of reform). For the avoidable hospitalization and preventable delivery complications specifications (Figures 2a and 4a), the differential in quality between the treatment and control states remains roughly flat around zero in the years leading up to the reform, with the improvement in quality emerging following the reform itself. To the extent the improvement in quality suggested by Table 3 otherwise predated the reform, one might be left with less confidence in a causal interpretation of the documented associations and might instead infer that the primary estimates in Table 3 merely capture regression to the mean. The inpatient mortality results depicted in Figure 1a likewise do not evidence a pre-treatment improvement in quality; however, we note that the pre-treatment differential is not as flat as that depicted in Figures 2a and 4a. The point estimates for the maternal trauma specification perhaps suggest a declining level of quality in the years leading up to national-standard adoptions (though imprecisely). While this pre-reform trend is inconsistent with an assumption of otherwise parallel trends in quality between treatment and control states but for the reform, Figure 4a nonetheless suggests a reversal of this declining quality trend in the direction of quality improvements following the reform, consistent with the findings of the additional quality metrics and with expectations that a liability reform expecting physicians to delivery higher levels of quality may lead to a subsequent improvement in practices.[23]

[23]Again, the lack of pre-treatment trends indicating improvement in quality leading up to national-standard reforms in initially low quality states alleviates concerns that the findings in Table 3 are merely capturing regression to the mean. To further alleviate these concerns, we note that the magnitude of the immediate and near-term effects of national-standard adoptions—as demonstrated by Figures 1a—4a and by Table B1 in the Online Appendix—are considerable relative to the overall level of mean reversion in the relevant outcomes documented over the sample period (in the Online Appendix, we demonstrate how the various quality measures trend over the sample separately for the initially high and initially low quality areas). For instance, within 3 years following the national-standard adoption, the inpatient mortality rate for selected medical conditions fell by nearly 24 log points in the treatment states relative to the control states. In the nearly three decades of the sample itself, the inpatient mortality rate in all of the low quality states went down by 37 log points, representing around a 4 log point decline every three years. As such, national-standard reforms do appear to meaningfully accelerate a modest amount of background regression to the mean, consistent with an actual effect of the reforms themselves.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

While practices appear to improve upon a shift in clinical standards expecting higher quality, the results do not overwhelmingly suggest a corresponding decline in quality upon a shift in legal standards arguably condoning lower quality care. To assess this reverse question, we explore what happens to initially high quality states (states with initially low quality indicator levels) when they adopt national-standard rules, which, in the case of such states, arguably lower operable standards by expecting that physicians follow the lesser-quality practices applied elsewhere. These results can be obtained from Table 3 by adding the baseline effect in the initially low-quality states (row 1 in each panel) to the coefficient of the interaction term (row 2 in each panel), which captures the marginal alteration of this national-standard effect from instead beginning in an initially high quality state. Across the various indicators, this addition suggests that a national standard adoption in the initially high-quality states is associated with a 5.2, −0.9, 5.0, and an 11 log point change in the respective quality indicator. We do observe a mean decline in quality—that is, an increase in the respective indicator—for three of the four indicators upon this change in standards arguably condoning a lower level of quality. Even in those cases, however, these responses are more modest than the responses indicated above for the initially low-quality states. In the final row of each panel, we also indicate the statistical significance level for the sum of these coefficients.[24] Not only are the estimated associations between national-standard adoptions and the quality indicators smaller in magnitude when focusing on the initially high-quality states, they are also generally not significantly different from zero, except perhaps with respect to some of the specifications using inpatient mortality rates for selected conditions as the quality measure of interest. Even in the case of these morality-rate specifications, however, it appears from Figure 1b that the inpatient mortality rate response emerges largely in the period of time prior to the national-standard adoption, suggesting that it may not even be a true response to the law itself.

Cancers Screening Measures—For this liability standards analysis, our primary tables do not include results for the cancer screening measures. For some of these measures—e.g., PSA testing for prostate cancer—data are only available during the 2000s, affording no ability to draw upon relevant standard-of-care reforms. Likewise, with respect to sigmoidoscopy/colonoscopy screening for colon cancer, data are generally unavailable in the pre-reform years for the relevant treatment states to facilitate a difference-in-difference analysis. However, for the remaining cancer screening measures—e.g., those relating to breast and cervical cancer data—are available during a period of time—i.e., the 1990s—in which Indiana, Delaware and Rhode Island can be utilized as treatment states. Our intent, of course, is to separately test for the effect of national-standard adoptions for those treatment states with initially high and initially low cancer screening rates. For the breast-cancer-screening measures, this leaves only one state—Indiana—from which to explore the effect of a liability reform that entails a heightening of standards. In the case of pap smear testing, both Indiana and Rhode Island can be utilized as treatment states in exploring the effect of heightened standards. In either case, with only one or two treatment states, the point estimates from a difference-in-difference analysis are generally thought to be inconsistent

[24]In Table B5 of the Online Appendix, we estimate an identical specification to that set forth in equation (1), except replacing the *HIGH-QUALITY* indicator variable with an analogously specified *LOW_QUALITY* indicator, providing us an alternative way to depict the association between national standard law adoptions and health care quality in initially high quality areas.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Conley and Taber 2011), leaving us with arguably unreliable estimates (given a higher degree of chance that spurious developments explain the findings). As such, we do not include them alongside the primary results from this analysis, which draw upon much more extensive legal variation. Nonetheless, we present these results in the Online Appendix. Consistent with our other quality measures, the results from the cancer screening exercise likewise document an increase in quality attainment (in this case, an increase in cancer screening rates) upon a modification of standard-of-care rules that entail a heightening of expectations.

Specification checks—In addition to those performed above, we note that these liability-standards findings are robust to certain additional specification checks and alternative approaches. First, we note that the results are largely robust to a weighting of the observations using the inverse of the propensity of each treated/untreated observation to be treated/untreated based on the available covariates (Rosenbaum and Rubin, 1983). In an effort to better approach randomization in the assignment of treatment and control states and to achieve covariate balance, this approach effectively puts more weight on those observations whose treatment status is difficult to predict based on the observables. For instance, while Table 3 depicts a 35–55 log point reduction (i.e., improvement) in low-discretionary avoidable hospitalization rates in initially-low-quality states upon a national standard adoption, we estimate a statistically-significant 26–35 log point negative response when imposing inverse propensity weights (with similar patterns across the range of additional estimates).[25] Second, in the Online Appendix, we discuss the use of randomization inference in exploring the statistical significance of the effects of national-standard adoptions, a flexible approach that generates unbiased standard error estimates even in the face of a limited number of treatment groups. In essence, using the sample of observations from our control states, we simulate a set of placebo laws that match the distribution of the timing of actual reforms. We then estimate the association between the relevant quality indicator and the placebo laws, replicating this process 5,000 times. We observe that the actual coefficients of the national-standard law dummies from our primary specifications—which identify the association between the various quality metrics and national standard adoptions in initially low quality states—fall on the very bottom of the distribution of placebo coefficients generated through these simulations.[26]

Size of Response—The magnitudes of the estimated associations between national-standard adoptions and the various indicators in initially low quality states are considerable, giving some pause regarding the plausibility of any causal interpretation of such estimates. Even the smaller of the findings—an 8 log point decrease in inpatient mortality rates for selected medical conditions—are nonetheless noteworthy in comparison with other major public initiatives which have been shown to alter clinical behaviors and decrease mortality rates.[27] Of course, the magnitudes of the quality-indicator effects found above roughly

[25]The full set of inverse propensity weight results are available from the authors upon request. The results remain nearly unchanged when trimming the sample to those observations with propensity weights in the 0.1–0.9 range (Crump et al, 2009).

[26]This exercise suggests statistical significance at the 5 percent level (or better) for the low-discretionary avoidable hospitalization, maternal trauma rate and preventable delivery complication rate specifications, but only marginal significance (at the 10 percent level) for the inpatient mortality rate (for selected medical conditions) specification.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

comport with the treatment-utilization effects of national-standard adoptions found in Frakes (2013), which, for instance, estimated a roughly 17–18 percent increase (decrease) in the rate of use of intensive cardiac interventions for initially low (high) intensity areas following national standard adoptions, along with a nearly 7 percent increase (decrease) in the rate of cesarean deliveries for initially low (high) intensity areas (each representing roughly a 40 percent closing of the gap between national and state utilization rates following this geographic standardization of liability standards). While a perhaps implausibly large response in behavior is observed in association with these reforms, it is important to bear in mind that the retreat from the locality rule did represent a fundamental shift in the structure of the medical liability system, a shift that is featured prominently in many first-year Torts textbooks (see, for example, Farnsworth and Grady, 2008) and that is especially meaningful in light of the rampant regional disparities in medical practices that have intrigued scholars for decades. Should these estimates be taken as causal, the size of the response to national-standard adoptions perhaps suggests that the underlying mechanisms of action go beyond physicians altering behavior to comply with the new legal expectations over fears of otherwise violating the new negligence standard. For instance, it is possible that any such traditional response may be amplified to the extent third party payers or other non-physician parties alter their own reimbursement or oversight practices in light of the altered tort landscape. Unpacking these mechanisms is an important area for future research.[28]

B. Damage-Cap Analysis

Overview—To focus the above discussion on liability standards reforms and to facilitate the presentation of all of the various national-standards findings in a single table, the coefficients of the damage-cap variables and other traditional tort reform variables were omitted from Table 3 despite being included in those specifications in Table 3 where we indicate the inclusion of remedy-focused reforms.[29] Nonetheless, for the reasons stated in Part I, it may be of independent interest to explore the link between damage-cap adoptions and our various health care quality indicators. In Table 4, we present the corresponding coefficients for these various traditional tort reforms. In Columns 2 and 3 of Table 4, we estimate the same exact specifications indicated in Columns 2 and 3 of Table 3 but now present the previously omitted coefficients of the remedy-focused traditional reforms.[30] The specifications underlying Table 4 do depart from Table 3 in one instance, however. To show the baseline association between health care quality and non-economic damage cap reforms —representing the specific remedy-focused reform that garners the most attention by tort

[27]For example, Card, Dobkin, and Maestas (2008) estimate an effect of Medicare eligibility on 7-day mortality of 14–20 percent and on 28-day mortality of 7–9 percent. Similarly, Sommers, Baicker and Epstein (2012) estimate a 6.1 percent relative decline in the all-cause mortality rate among adults following the Medicaid expansions of several states during the 2000s.

[28]One dimension to this mechanism analysis relates to physician supply. Has quality improved in initially low-quality states because given physicians have updated their practices or because higher quality physicians have been attracted to this region? Questions of this nature are also subjects for future research. In an unpublished working paper, Michael Frakes, Matthew Frank and Seth Seabury have sought to ask a similar question in unpacking the findings of Frakes (2013), finding evidence suggesting both that given physicians have altered the intensity of their practices following national-standard reforms and that some portion of the increase in mean intensity following reforms in initially low intensity areas is due to a shift in the physician mix towards relatively more surgical specialists (available from authors upon request).

[29]Columns 1, 5, and 6 of Table 3 include only the national standard law indicator, not indicator variables for the remedy-focused reforms.

[30]To simplify the presentation of this secondary, damages-cap analysis, we present fewer robustness checks in these latter tables than that depicted in Table 3, though we note that our analysis below is fully robust to subjecting the damage-cap findings to the same array of robustness checks.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

reform analysts—Column 1 of Table 4 estimates a difference-indifference specification that includes *only* a dummy variable for the incidence of a non-economic damages cap, with no variables included for the incidence of a national-standard law or the incidence of the remaining remedy-focused reforms (e.g., collateral source rule reforms).

In Table 5, we follow the same approach as Table 4, but focus instead on the cancer-screening measures. While the cancer screening analysis is weakly suited for the liability-standards analysis given the date range of the BRFSS data and the timing of variations of the national-standard-rule reforms, we are able to rely upon a sufficient number of damage-cap reforms over the BRFSS sample period.

In general, we estimate an association between a non-economic damages cap and the relevant quality indicator that is statistically indistinguishable from zero, though relatively tightly bound around zero. As such, though we cannot rule out that greater malpractice pressure within the existing system of liability standards—as identified through the lack of a non-economic damages cap—induces higher quality health care on average for these quality domains, we can rule out that such forces induce substantially higher levels of quality. That is, while the above analysis suggests that an elevation in liability standards has the potential to deter medical errors and improve patient quality, the existing custom-focused system of liability standards does not appear to be substantially improving quality on the margin.

AHRQ-Inspired Measures and Preventable Delivery Complications—We begin by describing the results for the AHRQ-inspired health care quality indicators and the preventable delivery complications measure (Table 4), considering that these measures all reflect lower levels of quality as the relevant indicator level rises (and vice versa), whereas the cancer screening measures, which we discuss in subsection B(2) below (and Table 5), reflect higher levels of quality as the screening levels rise. We separate the discussions with this difference in mind to ease confusion in exploring the relevant associations.

Upon the adoption of a non-economic damage cap, we estimate mean changes in the inpatient mortality rate for selected conditions, the low-discretionary AH rate, the maternal trauma rate and the preventable delivery complication rate of 1.3, −0.8, −6.6, and −4.6 log points, respectively. This pattern of point estimates does not change meaningfully upon the inclusion of state-year covariates, other tort laws and state-specific linear time trends, as demonstrated by Columns 2 and 3 of Table 4. These estimates are generally not significant at the p=0.05 level of significance. One exception, however, is the avoidable hospitalization results in the specification with state-specific linear trends; however, even in that case, the estimated coefficient suggests that damage caps improve, rather than harm, avoidable hospitalization rates, contrary to a story in which the current system of liability standards is deterring harmful errors on the margin. With respect to the remaining measures, we cannot rule out that positive associations between damage caps and these various quality indicators exist—i.e., we cannot rule out that a deterrent effect does exist. However, as demonstrated by the 95 percent confidence intervals reported in Table 4, we can generally rule out that any such deterrent effects are substantial—e.g., at the top of those intervals, we find that the adoption of non-economic damage caps is associated with a 0–10 log point increase in the relevant quality indicator depending on the specification, with the upper end of this

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

confidence interval generally falling in the 4 log point range. That is, higher malpractice pressure within our given liability system—captured by the lack of a damage cap—can at most lead to a modest level of improvement in quality (i.e., a modest amount of deterrence).

In Table B3 of the Online Appendix, we present dynamic variants of the difference-indifference specifications estimated in Table 4, which include leads and lags of the damage-cap incidence variable, allowing us to explore how the differential in quality across treatment and control states evolves on a year-to-year basis. While the confidence bounds for each coefficient in this dynamic specification expand slightly with the inclusion of this additional set of policy variables, they continue to bound zero at a relatively tight rate confirming the conclusion of an at-most modest association between damage-cap adoptions and the various quality indicators. The Online Appendix likewise demonstrates the robustness of these findings to various additional specification checks, including, the consideration only of damage-cap adoptions that apply to tort contexts broadly, easing legislative endogeneity concerns—i.e., dropping states that adopted damage caps that apply only in the malpractice context.

In the Online Appendix, we also estimate specifications that take a novel, alternative approach to the codification of the damage-cap incidence variable. While the malpractice literature customarily codifies damage-cap adoptions in a simple binary fashion (0/1), non-economic damage cap provisions, in fact, take on a range of forms across jurisdictions. For instance, California imposes a flat, nominal $250,000 cap on non-economic damages awards, while Wisconsin imposes a $750,000 cap. Inspired by certain policy simulations performed in Hyman et al. (2009) and by the simulated instrument methodology employed in Currie and Gruber (1996), we codify caps by using closed-claims data from Texas during the period of time prior to the imposition of its non-economic damage cap (with information on the breakdown of economic versus non-economic damages associated with the claim) to simulate the degree to which the various damage-cap provisions across the various states reduce liability awards. We then use the results of this simulation exercise as the relevant damage-cap variable within the difference-indifference specification, as opposed to the simple binary approach. The estimated mean coefficients from those specifications using this alternative codification of damage-cap variables do not differ substantially from those derived from the traditional binary approach.

Finally, we note that the non-economic damage cap results generalize to the other traditional tort reforms included as covariates (joint and several liability reforms, collateral source rule reforms and punitive damages caps), suggesting a generally weak relationship between both inpatient and outpatient health care quality and a broader range of remedy-focused reforms. In the case of the inpatient mortality rate, avoidable hospitalization rate, and preventable delivery complications measures, the results of an F-test of joint significance of all remedy-focused tort measures fail to reject the hypothesis that the coefficients of the various tort reforms are all jointly equal to zero. In the case of the maternal trauma specifications, the estimated coefficient of the punitive damages cap dummy is negative and bounded away from zero, suggesting an improvement in quality in connection with such reforms and thus counter to any expectation that such reforms would relax malpractice pressures to the detriment of patient quality.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The above-estimated specifications include state-year controls for physician concentration rates (and OB/GYN concentration rates in the case of the obstetrics measures). Such controls may absorb any impact of the reforms that occur through changes in the physician population. However, these simple controls may not absorb all supply-related consequences of such reforms. One effect of non-economic damage cap adoptions sometimes hypothesized is that lower-quality physicians may be attracted to the jurisdiction subsequent to the reform (Seabury 2010), a development which could otherwise confound any attempt to isolate the impact of malpractice pressure on the quality provided by any given provider. Of course, to the extent that non-economic damage caps would attract low-quality physicians and lead to a decline in observed quality–e.g., to an increase in the quality indicators explored in Table 4–this omission could only help to explain any positive effects of such reforms on the indicators explored. That is, a correction for this bias would likely push the estimated impacts of the reforms on the observed indicators even lower, only lending further support to the claim that marginal increases in liability pressure within the current liability system do not appear to substantially improve the quality of care being delivered by physicians.

Cancer Screening Measures—As presented in Table 5, the pattern of results from the cancer-screening / damage-cap analysis mirrors that from the AHRQ-inspired quality measures (with even greater precision in the estimates). We estimate mean associations between damage-cap adoptions and the various cancer screening rates that are very nearly zero in magnitude. As above, we cannot rule out some level of reductions in quality—i.e., some reduction in screening rates—in connection with damage cap reforms that are designed to reduce liability pressure. However, the 95-percent confidence bounds for each rate suggest that we can rule out that substantial reductions in screening rates are associated with caps. Lower bounds for these intervals suggest a 2.1, 3.0, 4.0, 1.7, 0.2 and 3.2 percent reduction (and an even lower percentage-point reduction) in mammography, physical breast, sigmoidoscopy/colonoscopy, PSA testing, and digital rectal and pap smear examinations, respectively. To simplify the presentation of these results, we present only the results from the basic difference-in-difference specifications. In the Online Appendix, we demonstrate the robustness of these findings to the addition of a range of control variables, along with alternative constructions of the screening rates.

C. Interpretation of Damage-Cap Findings: Durability of Effect of Liability-Standards Changes

If our findings are taken to suggest that structural reforms to the way in which physicians are evaluated may substantially alter health care delivery practices, one may wonder whether subsequent reforms—e.g., caps—that blunt the impact of the now altered liability system may cause practices to revert back to where they were before the structural reforms. Informational models of physician behavior suggest why this may not be so. If physicians, especially newer physicians, form beliefs over proper practices to a large extent through their own past experiences or through the observation of the practices followed by others around them, then a shift in medical practices that arises in any manner—including that arising from fear over being out of compliance with changed legal expectations—may more gradually come to be assimilated into the belief structures of physicians over time. These considerations (combined with the fact that liability standards are generally based on

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

customary practices) may thus help us understand why damage cap adoptions—which primarily arose in states after previous retreats from the locality rule—were not associated with a substantial change in physician practices and thus did not appear to cause behaviors to revert back to their locality-rule-era levels. To assess this possibility more concisely, we also estimate the relationship between caps and the various quality metrics separately for initially high and initially low quality states. After all, it is with the initially low quality states where we see an improvement in quality following national standard reforms. As demonstrated by Table 6, we do not find evidence of a decline in quality in such states as damage caps are subsequently adopted, with coefficient estimates that are small in magnitude and statistically indistinguishable from zero. The estimates are also relatively tightly bound around zero. Even at the very upper end of the 95 percent confidence interval, the results suggest only a 2 log point increase in inpatient mortality rates in initially low quality states following the adoption of a non-economic damage cap (despite a mean estimate of an 8–10 log point decline in inpatient mortality rates in such states following previous shifts from locality rules to national standard rules). As such, malpractice-induced changes in practices may come to shape more durable physician norms and customs that may survive subsequent diminishment of liability forces.

D. Discussion and Conclusion

An extensive number of empirical malpractice studies have endeavored to test for the existence and scope of so-called "defensive medicine.[31]" While deterrence of medical errors can be viewed as a primary objective of the medical liability system, defensive-medicine is best characterized as a possibly unfortunate side-effect / cost of this system. Physicians may act defensively when they unnecessarily order costly tests, procedures and visits over fear of malpractice liability (OTA 1994). However, even if one's primary focus is to explore these side effects of liability, rather than to assess whether the law is achieving its stated goal of deterring medical errors, it is critical to bear in mind that labeling a response as "defensive" requires more than a mere understanding of whether liability encourages additional utilization of medical care. Since a defensive response is defined with reference to the necessity (or optimality) of the chosen level of treatment, this assessment requires a determination as to whether or not any malpractice-induced expansion in treatment is accompanied by corresponding improvements in quality or outcomes (Mello et al. 2010).

As such, whether the goal is to make an independent evaluation of the deterrent impact of medical liability—i.e., to simply determine if liability forces are encouraging medical providers to avoid the commission of harmful errors—or to properly diagnose a "defensive" response to liability, it is necessary to estimate the impact of the malpractice system on medical errors and health care quality. To date, however, most studies which assess the impact of malpractice pressure on health care quality have focused only the association between malpractice and broad measures of mortality. A contribution of our analysis is to use clinically validated measures of health care quality—bearing on mortality and non-mortality consequences—to estimate the effect of malpractice pressure on the care provided by physicians. In this process, it is also important to bear in mind the structure of the malpractice system itself, a factor generally overlooked in most empirical discussions of this nature. In estimating the impacts of remedy-focused / non-substantive reforms such as non-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

economic damage caps, one is effectively teasing out the marginal impacts of malpractice penalties under the current standards of care legally expected of physicians. The confidence bounds presented in our analysis suggest, at most, a modest degree of deterrence stemming from the present liability system. The mean point estimates suggest that under existing liability standards, malpractice penalties generate little to no benefits in health care quality. We caution that these findings should perhaps not be interpreted so as to suggest that medical liability forces are universally incapable of improving quality. Rather, they should be interpreted in light of the largely self-regulatory nature of our present malpractice system —i.e., in light of the fact that the law itself is presently not designed to impose independent expectations regarding quality.

The first half of our empirical analysis provides some hope, however, in the potential for medical liability to influence physician behavior. Drawing upon the one type of standard-of-care reform that states have experimented with to date—i.e., locality rule abdications—we investigate the impact of changing the clinical standards of care imposed upon physicians under the law, both in terms of elevated standards and slackened standards. All told, it appears that the relationship between health care quality and changes in clinical malpractice standards works in an expansionary direction only. That is, once physicians provide a high level of quality, they may maintain such practices even when the law may loosen its expectations at a later date. In contrast, physicians who provide a quality of care that is below what is expected by the law raise their practices to meet the higher expectations set by the law. Malpractice forces that alter the legal clinical standard to which physicians are held may therefore be effective in elevating the quality floor.

Ultimately, empirical malpractice investigations that fail to appreciate the structural considerations underlying tort law may misinterpret the findings derived from our experiences to date with traditional remedy-centric tort reforms. Such findings may suggest only a weak responsiveness to the law despite a potentially meaningful role for the law to play in shaping clinical practices and health care quality. Substantial work remains, of course, to understand the liability structure that will best serve society. Our analysis demonstrates that it would be premature to rule out medical liability from the health care quality discussion based on the limited findings that derive from damage-cap-focused studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Amy Finkelstein and Jon Gruber for their guidance and advice. We are likewise grateful for helpful comments received from Ronen Avraham, Bernie Black, Amitabh Chandra, David Hyman, Louis Kaplow, Justin McCrary, Michelle Mello, Mark Ramseyer, Max Schazenbach, Steve Shavell, Charlie Silver, Holger Spamann, Brad Wendel, Kathy Zeiler and seminar participants at the NBER Summer Institute Law and Economics Workshop, the NBER Summer Institute Health Care Workshop, the Harvard Law, Economics & Organization Seminar, the Northwestern Law and Economics Colloquium, the Northwestern Faculty Colloquium, the Midwestern Law and Economics Association Annual Meeting, the University of Texas Medical Malpractice Roundtable, the Cornell Law School Faculty Workshop and the Cornell Law School Faculty Retreat. We are also thankful to Negasi Beyene, Maria Owings and the staff at the Research Data Center at the National Center for

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Health Statistics for their help with the National Hospital Discharge Survey files and to Ronen Avraham for providing data on state tort laws. Frakes acknowledges support from the National Institute on Aging, Grant Number T32-AG00186 and Jena acknowledges support from the Office of the Director, National Institutes of Health (NIH Early Independence Award, 1DP5OD017897-01)

References

Atkins, David, Siegel, Joanna, Slutsky, Jean. Making Policy When the Evidence is in Dispute. Health Affairs. 2005; 24(1):102–113. [PubMed: 15647220]

Baicker, Katherine, Chandra, Amitabh. The Effect of Malpractice Liability on the Delivery of Health Care. In: Cutler, David M., Garber, Alan M., editors. Frontiers of Health Policy Research. Vol. 8. 2005.

Baicker, Katherine, Fisher, Elliott S., Chandra, Amitabh. Malpractice Liability Costs And The Practice Of Medicine In The Medicare Program. Health Affairs. 2007; 26(3):841–52. [PubMed: 17485765]

Bovbjerg, Randall, Berenson, Robert A. The Value of Clinical Practice Guidelines as Malpractice 'Safe Harbors.'. Urban Institute, Timely Analysis of Immediate Health Policy Issues; 2012.

Card, David, Dobkin, Carlos, Maestas, Nicole. Does Medicare Save Lives? Quarterly Journal of Economics. 2009; 124(2):597–636. [PubMed: 19920880]

Centers for Medicare and Medicaid Services. [accessed June, 17 2012] National Health Expenditure Accounts: Tables. 2011. Available at: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-andReports/NationalHealthExpendData/Downloads/tables.pdf

Chan, David. NBER Working Paper No. 21855. 2016. Informational Frictions and Practice Variation: Evidence from Physicians in Training.

Chandra, Amitabh, Staiger, Douglas O. Productivity Spillovers in Healthcare: Evidence From the Treatment of Heart Attacks. Journal of Political Economy. 2007; 115(1):103–40. [PubMed: 18418468]

Conley, Timothy, Taber, Christopher. Inference with 'Difference-in-Differences' with a Small Number of Policy Changes. The Review of Economics and Statistics. 2011; 1:113–25.

Crump, Richard, Joseph Hotz, V., Imbens, Guido, Mitnik, Oscar. Dealing with limited overlap in estimation of average treatment effects. Biometrika. 2009; 96(1):187–199.

Currie, Janet, Gruber, Jonathan. Health Insurance Eligibility: Utilization of Medical Care, and Child Health. Quarterly Journal of Economics. 1996; 111:431–66.

Currie, Janet, Bentley MacLeod, W. First Do No Harm? Tort Reform and Birth Outcomes. Quarterly Journal of Economics. 2008; 123:795–830.

Dewees, Don, Treblock, Michael J. The Efficacy of the Tort System and its Alternatives: A Review of Empirical Evidence. Osgoode Hall Law Journal. 1992; 30(1):57–138.

Dubay, Lisa, Kaestner, Robert, Waidmann, Timothy. The Impact of Malpractice Fears on Cesarean Section Rates. Journal of Health Economics. 1999; 18:491–522. [PubMed: 10539619]

Farnsworth, Ward, Grady, Mark. Torts: Cases and Questions. Wolters Kluwer Law & Business; 2009.

Fisher, Elliott, Wennberg, David, Stukel, Thrse, Gottlieb, Daniel, Lucas, FL., Pinder, Etoile. The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care. Annals of Internal Medicine. 2003; 138(4):273–87. [PubMed: 12585825]

Frakes, Michael. Defensive Medicine and Obstetric Practices. Journal of Empirical Legal Studies. 2012; 9(3):457–81.

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

Hussey, Peter S., Mattke, Soeren, Morse, Lindsey, Susan Ridgely, M. Evaluation of the Use of AHRQ and Other Quality Indicators. 2007. AHRQ Publication No. 08-M012-EF

Hyman, David, Black, Bernard, Silver, Charles, Sage, William. Estimating The Effect of Damage Caps in Medical Malpractice Cases: Evidence from Texas. Journal of Legal Analysis. 2009; 1:355–409.

Iizuka, Toshiaki. Does Higher Malpractice Pressure Deter Medical Errors? 56. Journal of Law and Economics. 2013; 1:161–88.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Jena, Anupam, Seabury, Seth, Lakdawalla, Darius, Chandra, Amitabh. Malpractice Risk according to Physician Specialty. New England Journal of Medicine. 2011; 365:629–636. [PubMed: 21848463]

Kachalia, Allen, Little, Alison, Isavoran, Melissa, Crider, Lynn-Marie, Smith, Jeanene. Greatest Impact of Safe Harbor Rule May Be to Improve Patient Safety, Not Reduce Liability Claims Paid by Physicians. Health Affairs. 2014; 33(1):59–66. [PubMed: 24395936]

Kessler, Daniel, McClellan, Mark. Do Doctors Practice Defensive Medicine? Quarterly Journal of Economics. 1996; 111:353–90.

Klick, Jonathan, Stratmann, Thomas. Medical Malpractice Reform and Physicians in High-Risk Specialties. Journal of Legal Studies. 2007; 36:S121–S142.

Lakdawalla, Darius, Seabury, Seth. NBER Working Paper No. 15383. 2009. The Welfare Effects of Medical Malpractice Liability.

Mello, Michelle M., Brennan, Troyen. Deterrence of Medical Errors: Theory and Evidence for Malpractice Reform. Texas Law Review. 2002; 80:1595–1647.

Mello, Michelle M., Chandra, Amitabh, Gawande, Atul A., Studdert, David M. National Costs of the Medical Liability System. Health Affairs. 2010; 29:1569–77. [PubMed: 20820010]

Mello, Michelle M., Kachalia, Allen. Evaluation of options for medical malpractice system reform: a report to the Medicare Payment Advisory Commission, (MedPAC, 2010). 2011. available at http://www.medpac.gov/documents/Apr10_MedicalMalpractice_CONTRACTOR.pdf

Molitor, David. The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration. Mimeo: University of Illinois; 2015.

National Center for Health Statistics. National Hospital Discharge Survey. Centers for Disease Control and Prevention; 1977–2005. [last accessed at NCHS Research Data Center on January 31, 2012]

Office of Technology Assessment. Defensive Medicine and Medical Malpractice. 1994. OTA-H--6O2

Paik, Myungho, Black, Bernard, Hyman, David. The Receding Tide of Medical Malpractice Litigation Part 2: Effect of Damage Caps. Journal of Empirical Legal Studies. 2013 forthcoming.

Phelps, Charles E., Mooney, Cathleen. Variations in Medical Practice Use: Causes and Consequences. In: Arnould, Richard J.Rich, Robert F., White, William D., editors. Competitive Approaches to Health Care Reform. Washington, DC: The Urban Institute Press; 1993. p. 140-78.

Schwartz, Gary T. Reality in the Economic Analysis of Tort Law: Does Tort Law Really Deter? UCLA Law Review. 1994; 42:377–444.

Seabury, Seth. RAND Working Paper No. WR-674-ICJ. 2010. Does Malpractice Liability Reform Attract High Risk Doctors?.

Polinsky, A Mitchell, Shavell, Steven. The Uneasy Case for Product Liability. Harvard Law Review. 2010; 123:1437–92.

Rosenbaum, Paul, Rubin, Donald. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.

Schiff, Gordon, Puopolo, Ann Louise, Huben-Kearney, Anne, Yu, Winnie, Keohane, Carol, McDonough, Peggy, Ellis, Donnie, Bates, David, Biondoillio, Madeleine. Primary Care Closed Claims Experience of Medical Malpractice Insurers. JAMA Internal Medicine. 2013; 173(22): 2063–2068. [PubMed: 24081145]

Shepherd, Joanna. Uncovering the Silent Victims of the American Medical Liability System. Vanderbilt Law Review. 2014; 67:151–195.

Sloan, Frank A. Experience Rating: Does It Make Sense for Medical Malpractice Insurance? American Economic Review (Papers and Proceedings). 1990; 80(2):128–133.

Sloan, Frank A., Shadle, John H. Is there empirical evidence for “Defensive Medicine ”? A reassessment. Journal of Health Economics. 28:481–491.

Sommers, Benjamin, Baicker, Katherine, Epstein, Arnold. Mortality and Access to Care among Adults after State Medicaid Expansion. New England Journal of Medicine. 2012; 367:1025–1034. [PubMed: 22830435]

Wennberg, John, Gittelsohn, Alan. Small Area Variations in Health Care Delivery: A population-based health information system can guide planning and regulatory decision-making. Science. 1973; 182(4117):1102–1108. [PubMed: 4750608]

Zeiler, Kathryn, Silver, Charles, Black, Bernard, Hyman, David, Sage, William. Physicians' Insurance Limits and Malpractice Payments: Evidence from Texas Closed Claims: 1990–2003. Journal of Legal Studies. 2007; 36

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- We explore the potential for medical liability to deter errors and improve quality.
- We find that health care quality is sensitive to medical liability standards.
- Quality improves when liability standards are reformed to expect better care.
- Quality does not decline when liability standards are reformed to condone worse care.
- Observed quality levels are not strongly associated with adoption of damage caps.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 1A.

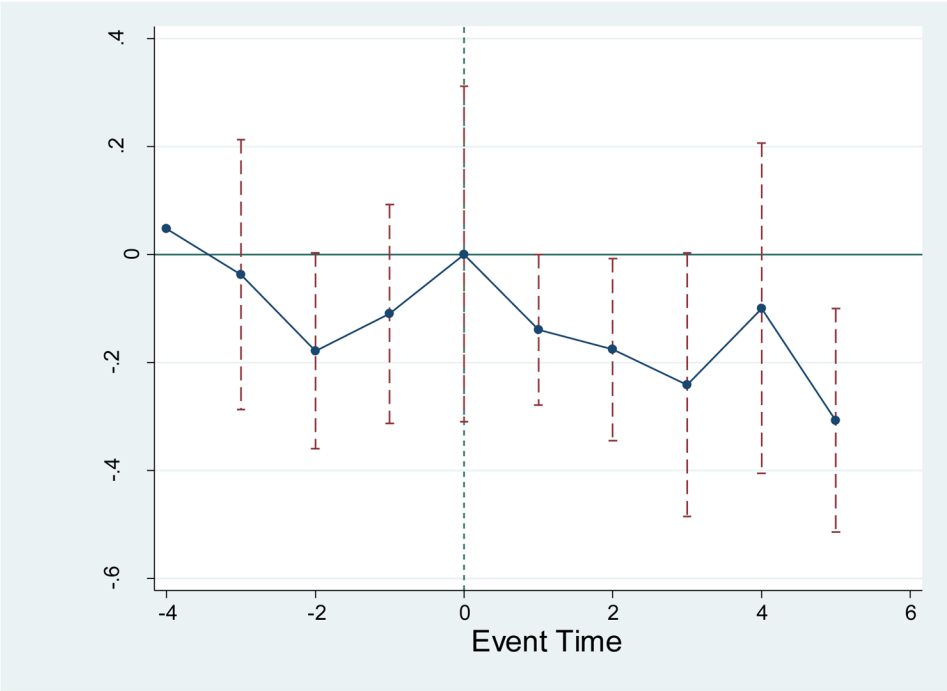
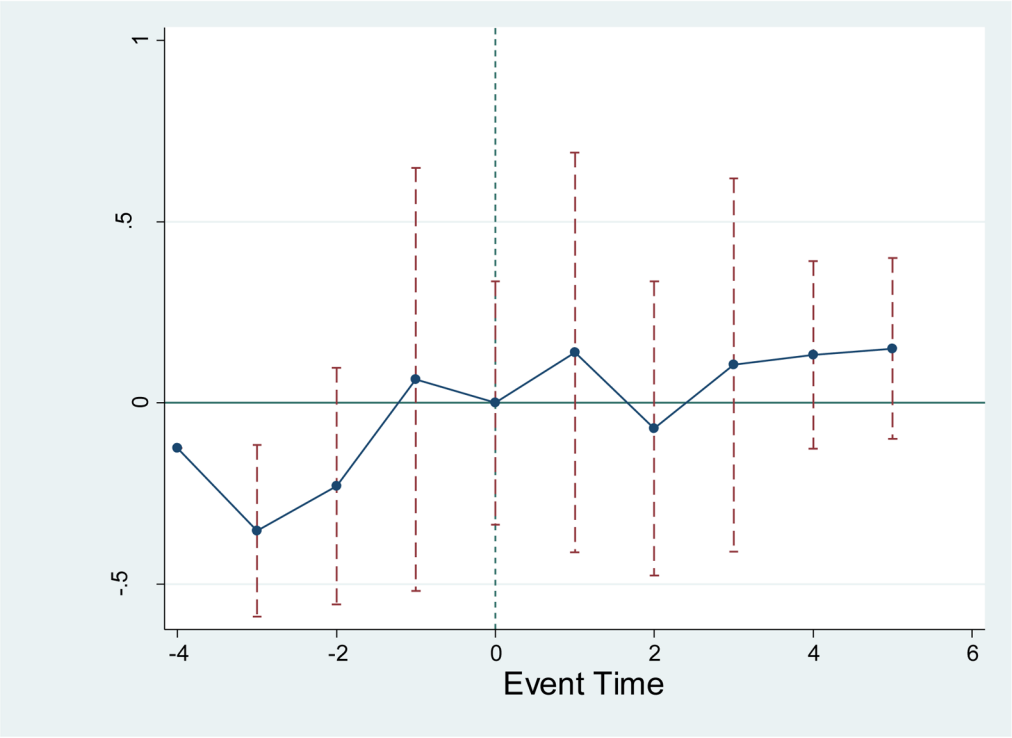


Figure 1B.



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 1.

Figure 1A. Relationship between National-Standard Adoptions and Inpatient Mortality Rates for Selected Medical Conditions in Initially Low-Quality States (Logged): Event-Study Analysis

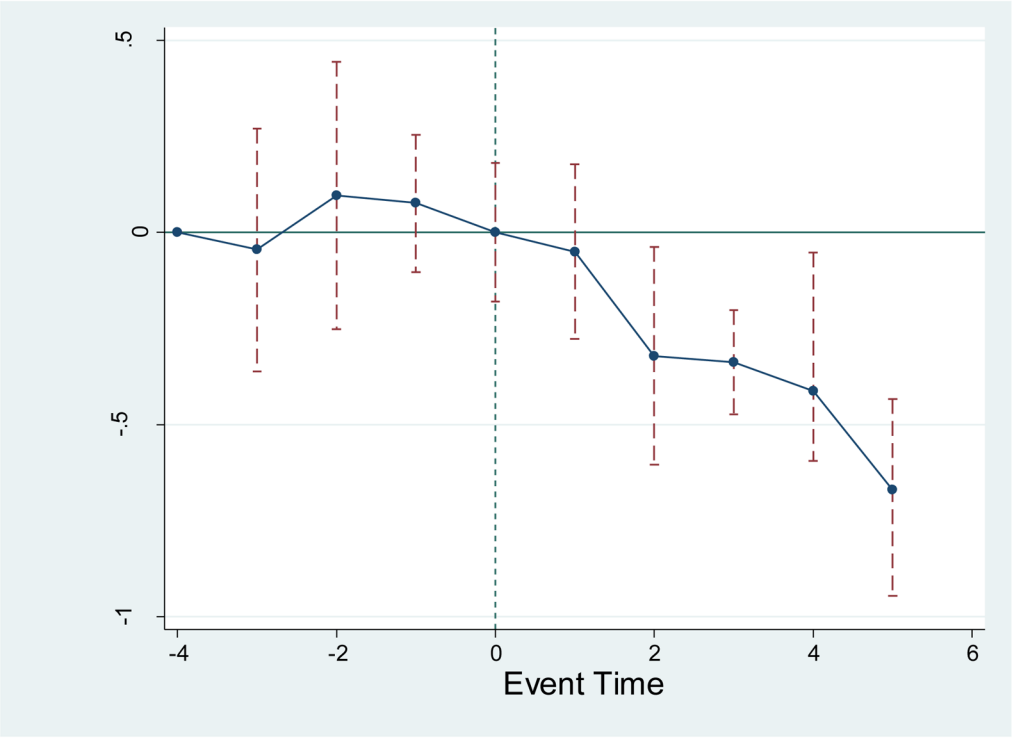
Notes: this figure plots the cumulative differential in the inpatient mortality rates for selected medical conditions between the treatment states and control states in the period of time preceding and following the adoption of national-standard rules. Figure 1A focuses only on initially low quality states. The estimates are derived from a dynamic difference-in-difference regression of the inpatient mortality rate—out of a sample of non-discretionary medical conditions—on a series of lead and lag indicators for the adoption of a national standard rule. The reference category in the regression is the time period prior to 4 years before the adoption of national standard laws. The coefficients and confidence bars were adjusted after this estimation in order to normalize the inpatient mortality-rate differential to 0 at the time of adoption. Regressions also include various state-year covariates and state-specific linear time trends and are weighted by the number of admissions (for the relevant state and year) in the sub-sample of discharges associated with the relevant selected conditions (e.g., acute myocardial infarction). Ninety-five percent confidence bounds (depicted by the dashed bars) are formed while correcting for within-state correlation in the error term.

Figure 1B. Relationship between National-Standard Adoptions and Inpatient Mortality Rates for Selected Medical Conditions in Initially High-Quality States (Logged): Event-Study Analysis

Notes: this figure plots the cumulative differential in the inpatient mortality rates for selected medical conditions between the treatment states and control states in the period of time preceding and following the adoption of national-standard rules. Figure 1B focuses only on initially high quality states. The estimates are derived from a dynamic difference-in-difference regression of the inpatient mortality rate (out of a sample of non-discretionary medical conditions) on a series of lead and lag indicators for the adoption of a national standard rule. The reference category in the regression is the time period prior to 4 years before the adoption of national standard laws. The coefficients and confidence bars were adjusted after this estimation in order to normalize the inpatient mortality-rate differential to 0 at the time of adoption. Regressions also include various state-year covariates and state-specific linear time trends and are weighted by the number of admissions (for the relevant state and year) in the sub-sample of discharges associated with the relevant selected conditions (e.g., acute myocardial infarction). Ninety-five percent confidence bounds (depicted by the dashed bars) are formed while correcting for within-state correlation in the error term.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 2A.



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 2B.

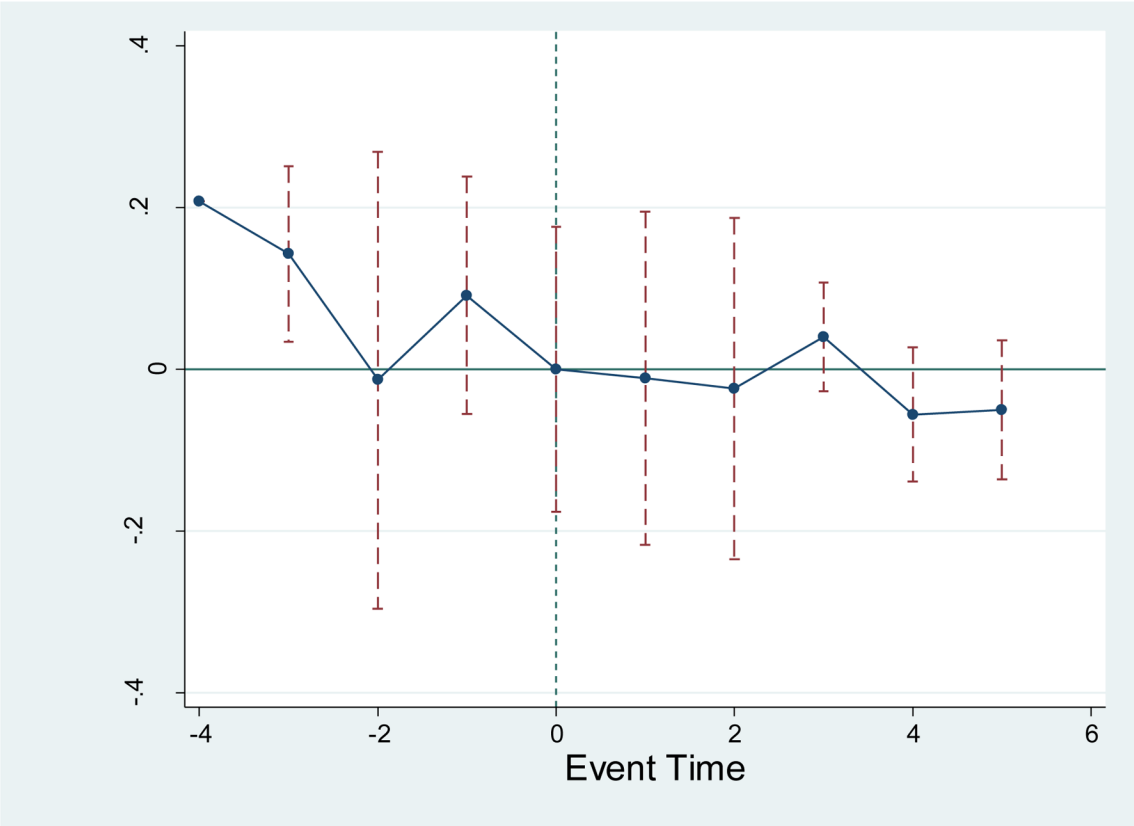


Figure 2.

Figure 2A. Relationship between National-Standard Adoptions and Low-Discretionary Avoidable Hospitalization Rates (Logged) in Initially Low-Quality States: Event-Study Analysis

Notes: this figure plots the cumulative differential in the low-discretionary avoidable hospitalization rates between the treatment states and control states in the period of time preceding and following the adoption of national-standard rules. Figure 2A focuses only on initially low quality states. The estimates are derived from a dynamic difference-in-difference regression of the low-discretionary avoidable hospitalization rate on a series of lead and lag indicators for the adoption of a national standard rule. The reference category in the regression is the time period prior to 4 years before the adoption of national standard laws. The coefficients and confidence bars were adjusted after this estimation in order to normalize the low-discretionary avoidable hospitalization-rate differential to 0 at the time of adoption. Regressions also include various state-year covariates and state-specific linear time trends and are weighted by the low-variation health index (i.e., the sum of discharges for acute myocardial infarction, stroke, hip fracture or gastrointestinal bleeding) associated with each state-year cell. Ninety-five percent confidence bounds (depicted by the dashed bars) are formed while correcting for within-state correlation in the error term.

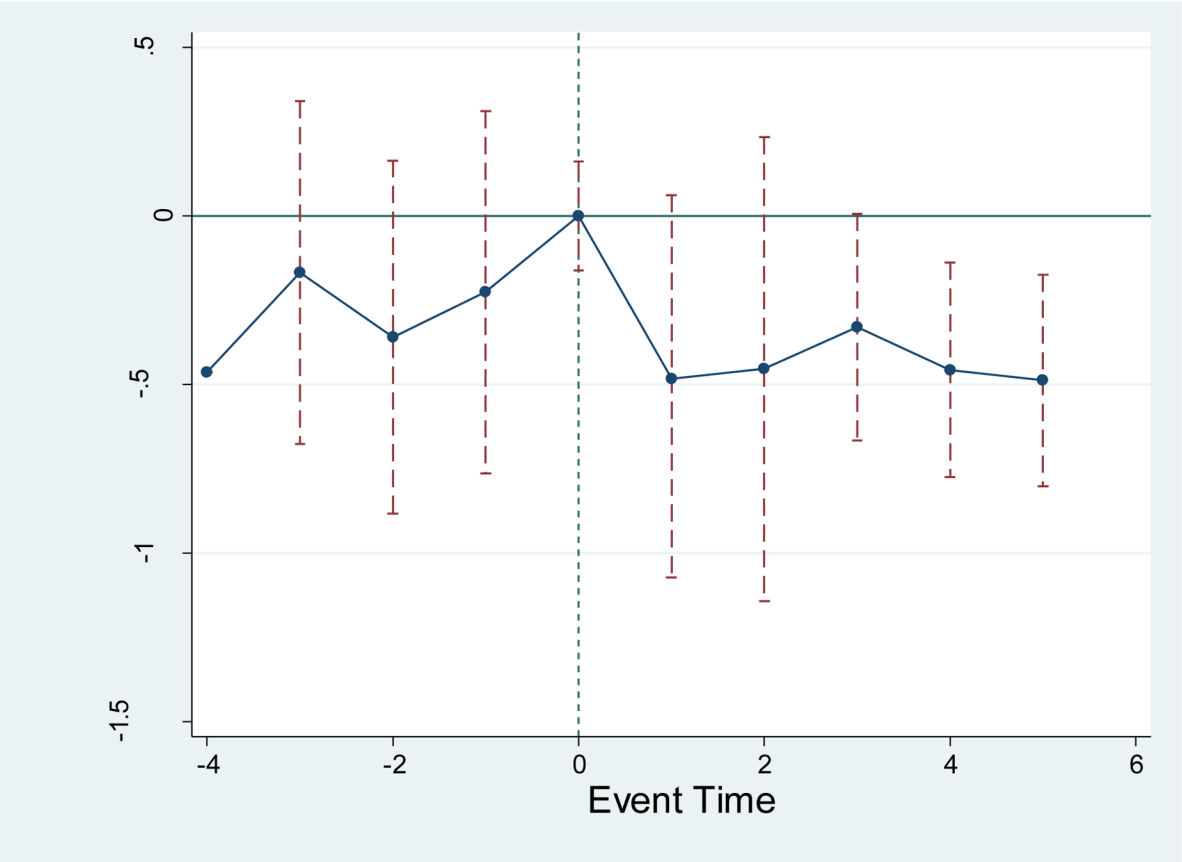
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 2B. Relationship between National-Standard Adoptions and Low-Discretionary Avoidable Hospitalization Rates (Logged) in Initially High-Quality States: Event-Study Analysis

Notes: this figure plots the cumulative differential in the low-discretionary avoidable hospitalization rates between the treatment states and control states in the period of time preceding and following the adoption of national-standard rules. Figure 2B focuses only on initially high quality states. The estimates are derived from a dynamic difference-in-difference regression of the low-discretionary avoidable hospitalization rate on a series of lead and lag indicators for the adoption of a national standard rule. The reference category in the regression is the time period prior to 4 years before the adoption of national standard laws. The coefficients and confidence bars were adjusted after this estimation in order to normalize the low-discretionary avoidable hospitalization-rate differential to 0 at the time of adoption. Regressions also include various state-year covariates and state-specific linear time trends and are weighted by the low-variation health index (i.e., the sum of discharges for acute myocardial infarction, stroke, hip fracture or gastrointestinal bleeding) associated with each state-year cell. Ninety-five percent confidence bounds (depicted by the dashed bars) are formed while correcting for within-state correlation in the error term.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 3A.



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 3B.

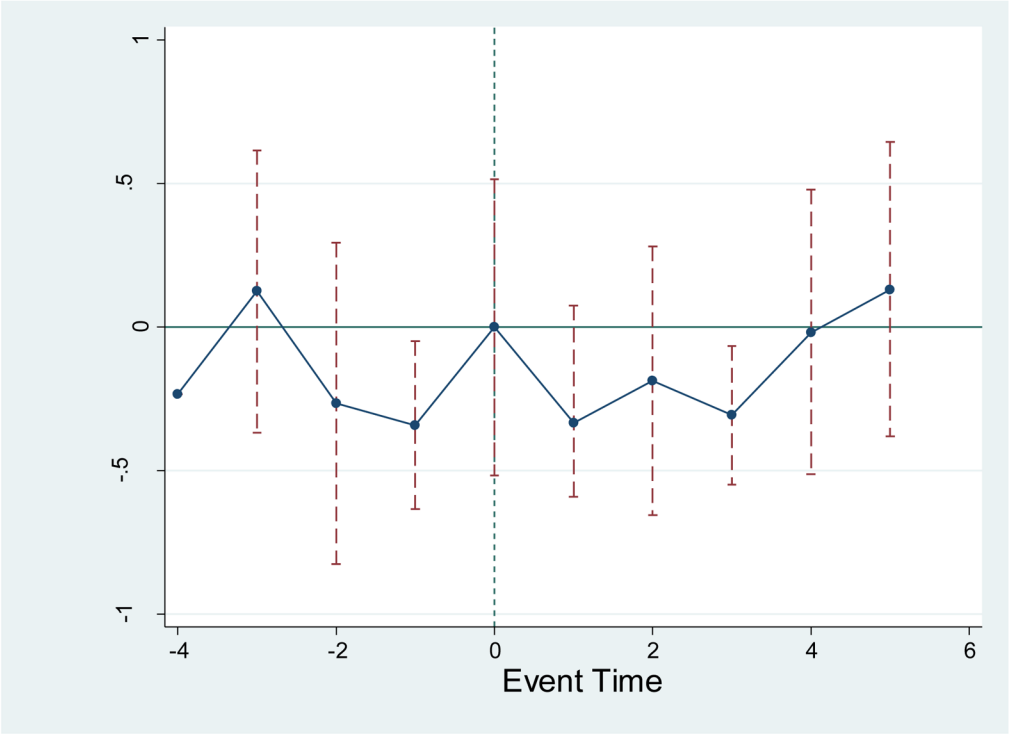


Figure 3.

Figure 3A. Relationship between National-Standard Adoptions and Maternal Trauma Rates (Logged) in Initially Low-Quality States: Event-Study Analysis

Notes: this figure plots the cumulative differential in maternal trauma rates between the treatment states and control states in the period of time preceding and following the adoption of national-standard rules. Figure 3A focuses only on initially low quality states. The estimates are derived from a dynamic difference-in-difference regression of the maternal trauma rate (out of the sample of deliveries) on a series of lead and lag indicators for the adoption of a national standard rule. The reference category in the regression is the time period prior to 4 years before the adoption of national standard laws. The coefficients and confidence bars were adjusted after this estimation in order to normalize the maternal-trauma-rate differential to 0 at the time of adoption. Regressions also include various state-year covariates and state-specific linear time trends and are weighted by the number of deliveries associated with each state-year cell. Ninety-five percent confidence bounds (depicted by the dashed bars) are formed while correcting for within-state correlation in the error term.

Figure 3B. Relationship between National-Standard Adoptions and Maternal Trauma Rates (Logged) in Initially High-Quality States: Event-Study Analysis

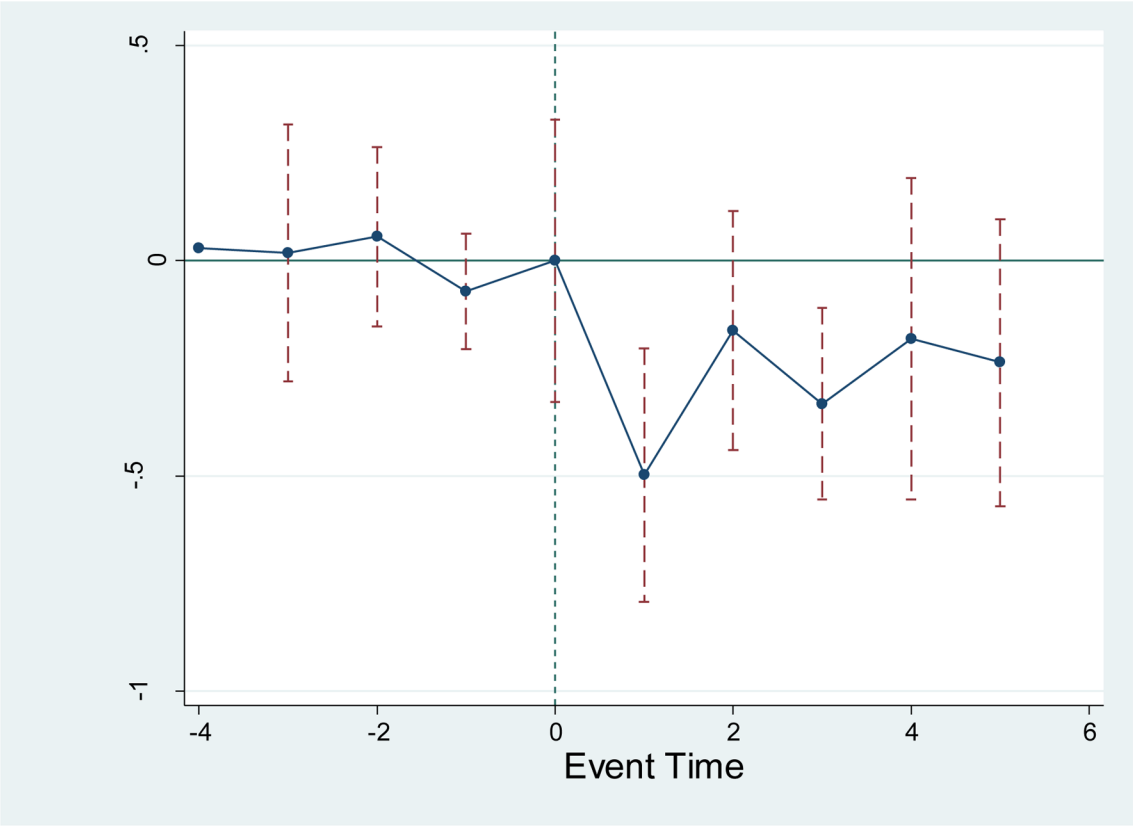
Notes: this figure plots the cumulative differential in maternal trauma rates between the treatment states and control states in the period of time preceding and following the adoption of national-standard rules. Figure 3B focuses only on initially high quality states. The estimates are derived from a dynamic difference-in-difference regression of the maternal

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

trauma rate (out of the sample of deliveries) on a series of lead and lag indicators for the adoption of a national standard rule. The reference category in the regression is the time period prior to 4 years before the adoption of national standard laws. The coefficients and confidence bars were adjusted after this estimation in order to normalize the maternal-trauma-rate differential to 0 at the time of adoption. Regressions also include various state-year covariates and state-specific linear time trends and are weighted by the number of deliveries associated with each state-year cell. Ninety-five percent confidence bounds (depicted by the dashed bars) are formed while correcting for within-state correlation in the error term.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 4A.



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Figure 4B.

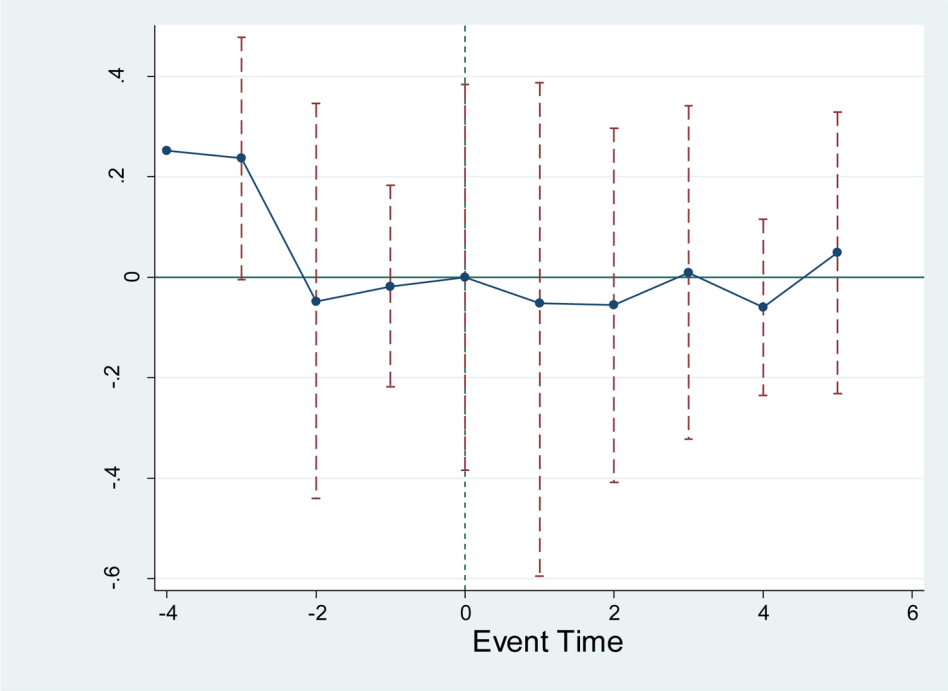


Figure 4.

Figure 4A. Relationship between National-Standard Adoptions and Preventable Delivery Complication Rate (Logged) in Initially Low-Quality States: Event-Study Analysis

Notes: this figure plots the cumulative differential in preventable delivery complication rates between the treatment states and control states in the period of time preceding and following the adoption of national-standard rules. Figure 4A focuses only on initially low quality states. The estimates are derived from a dynamic difference-in-difference regression of the preventable delivery complication rate (out of the sample of deliveries) on a series of lead and lag indicators for the adoption of a national standard rule. The reference category in the regression is the time period prior to 4 years before the adoption of national standard laws. The coefficients and confidence bars were adjusted after this estimation in order to normalize the preventable-delivery-complication-rate differential to 0 at the time of adoption. Regressions also include various state-year covariates and state-specific linear time trends and are weighted by the number of deliveries associated with each state-year cell. Ninety-five percent confidence bounds (depicted by the dashed bars) are formed while correcting for within-state correlation in the error term.

Figure 4B. Relationship between National-Standard Adoptions and Preventable Delivery Complication Rates (Logged) in Initially High-Quality States: Event-Study Analysis

Notes: this figure plots the cumulative differential in preventable delivery complication rates between the treatment states and control states in the period of time preceding and following the adoption of national-standard rules. Figure 4B focuses only on initially high quality states. The estimates are derived from a dynamic difference-in-difference regression of the preventable delivery complication rate (out of the sample of deliveries) on a series of lead and lag indicators for the adoption of a national standard rule. The reference category in the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

regression is the time period prior to 4 years before the adoption of national standard laws. The coefficients and confidence bars were adjusted after this estimation in order to normalize the preventable-delivery-complication-rate differential to 0 at the time of adoption. Regressions also include various state-year covariates and state-specific linear time trends and are weighted by the number of deliveries associated with each state-year cell. Ninety-five percent confidence bounds (depicted by the dashed bars) are formed while correcting for within-state correlation in the error term.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Descriptive Statistics

	Mean (Standard Deviation)	Percentage Absolute Deviation between State and National Mean
Panel A: Quality Measures (NHDS)		
Composite Inpatient Mortality Rate	0.08 (0.03)	0.16 (0.15)
Low-Discretionary Avoidable Hospitalization Rate (Scaled by Low- Variation Health Index)	1.00 (0.23)	0.15 (0.15)
Maternal Trauma Rate	0.04 (0.02)	0.26 (0.25)
Maternal Preventable Complications Rate	0.16 (0.06)	0.20 (0.20)
Panel B: Cancer-Screening Rates (BRFSS)		
Mammogram (within last year, female age 40–75)	0.73 (0.45)	-
Physical breast exam (within last year, female age 40–75)	0.64 (0.48)	-
Proctoscopic exam (sigmoidoscopy or colonoscopy within last 5 years, age 50–75)	0.40 (0.49)	-
PSA Testing (within last year, age 50–75)	0.53 (0.50)	-
Digital Rectal Exam for Prostate Cancer (within last year, age 50–75)	0.50 (0.50)	-
Pap smear (within last year, age 21+)	0.60 (0.49)	-

Notes: Standard deviations are in parentheses. Quality measures in Panel A are from a sample of 1190 state-year cells from the 1979–2005 NHDS files. Quality statistics in Panel A are weighted by the relevant denominator used in the state-year quality rate (e.g., the state-year delivery count or the state-year low-variation health index).

Source: Panel A: National Hospital Discharge Survey (1979–2005), Panel B: Behavioral Risk Factor Surveillance System (1987–2008).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

Distribution of AHRQ-Inspired Quality Metrics at Beginning and End of Sample, by Treatment Status (National Standard Rule Adoptions)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All States		Treatment States		Control States		Number of Treatment States (at Some Point) in 1979 Quintile?
	1979	2005	1979	2005	1979	2005	
Panel A: Inpatient Mortality Rate for Selected Conditions							
0–20th Percentile	0.000–0.111	0.000–0.041	0.056–0.110	0.027–0.048	0.000–0.110	0.000–0.042	3
20th–40th Percentile	0.111–0.134	0.041–0.052	0.110–0.133	0.048–0.057	0.110–0.133	0.042–0.051	3
40th–60th Percentile	0.134–0.150	0.052–0.058	0.133–0.164	0.057–0.062	0.133–0.147	0.051–0.057	1
60th–80th Percentile	0.150–0.169	0.058–0.073	0.164–0.186	0.062–0.089	0.147–0.160	0.057–0.064	2
80th–100th Percentile	0.169–0.342	0.073–0.164	0.186–0.308	0.089–0.125	0.160–0.342	0.064–0.164	6
Panel B: Low-Discretionary Avoidable Hospitalization							
0–20th Percentile	0.428–0.545	0.573–0.944	0.482–0.545	0.573–0.944	0.428–0.640	0.597–0.974	3
20th–40th Percentile	0.545–0.777	0.944–1.090	0.545–0.777	0.944–1.090	0.640–0.790	0.974–1.160	2
40th–60th Percentile	0.777–0.927	1.090–1.237	0.777–0.930	1.090–1.110	0.790–0.917	1.160–1.240	2
60th–80th Percentile	0.927–1.110	1.237–1.383	0.930–1.110	1.110–1.125	0.917–1.120	1.240–1.390	4
80th–100th Percentile	1.110–1.487	1.383–3.555	1.110–1.487	1.125–1.180	1.120–1.400	1.390–3.555	4
Panel C: Maternal Trauma							
0–20th Percentile	0.000–0.011	0.006–0.020	0.000–0.000	0.020–0.020	0.000–0.014	0.006–0.020	4
20th–40th Percentile	0.011–0.026	0.020–0.022	0.000–0.016	0.020–0.027	0.014–0.026	0.020–0.023	2
40th–60th Percentile	0.026–0.036	0.0220.029	0.016–0.036	0.027–0.029	0.026–0.031	0.023–0.029	2
60th–80th Percentile	0.036–0.049	0.029–0.035	0.036–0.050	0.029–0.030	0.031–0.048	0.029–0.037	2
80th–100th Percentile	0.049–0.090	0.035–0.050	0.050–0.090	0.030–0.045	0.048–0.085	0.037–0.050	5
Panel D: Preventable Delivery Complications							
0–20th Percentile	0.000–0.022	0.136–0.174	0.000–0.021	0.136–0.160	0.000–0.033	0.150–0.174	4
20th–40th Percentile	0.022–0.037	0.174–0.199	0.021–0.034	0.160–0.201	0.033–0.037	0.174–0.196	2
40th–60th Percentile	0.037–0.043	0.199–0.221	0.034–0.043	0.201–0.224	0.037–0.041	0.196–0.220	2
60th–80th Percentile	0.043–0.056	0.221–0.267	0.043–0.074	0.224–0.260	0.041–0.047	0.220–0.279	2
80th–100th Percentile	0.056–0.108	0.267–0.351	0.074–0.108	0.260–0.280	0.047–0.100	0.279–0.351	5

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Notes: Quality measures are from a sample of 1190 state-year cells from the 1979–2005 NHDS files. Treatment states represent those states that adopted a national-standard rule at some point over the sample period. Control states indicate all other states.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

The Relationship between National-Standard Laws and Various Health Care Quality Metrics

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Dependent Variable: Inpatient Mortality Rate for Selected Conditions						
National-Standard (NS) Law Dummy	−0.081 ** (0.035)	−0.081 ** (0.036)	−0.104 (0.088)	−0.096 ** (0.043)	−0.142 * (0.072)	−0.008 ** (0.003)
NS Law * Initially High Quality State	0.133 * (0.071)	0.174 ** (0.063)	0.234 * (0.131)	0.134 * (0.073)	0.212 * (0.119)	0.016 ** (0.007)
N	1131	1074	1074	1074	1131	1165
Significance Level of Sum of Above 2 Coefficients (not significant indicated by dash)	**	*	*	-	-	**
Panel B. Dependent Variable: Low-Discretionary Avoidable Hospitalization Rate						
National-Standard (NS) Law Dummy	−0.537 *** (0.075)	−0.469 *** (0.065)	−0.336 *** (0.059)	−0.555 *** (0.126)	−0.488 *** (0.091)	−0.520 *** (0.076)
NS Law * Initially High Quality State	0.528 *** (0.125)	0.456 *** (0.109)	0.297 *** (0.079)	0.475 *** (0.139)	0.527 *** (0.110)	0.500 *** (0.114)
N	1166	1108	1108	1108	1166	1166
Significance Level of Sum of Above 2 Coefficients (not significant indicated by dash)	-	-	-	-	-	-
Panel C. Dependent Variable: Maternal Trauma Rate						
National-Standard (NS) Law Dummy	−0.396 *** (0.103)	−0.321 ** (0.142)	−0.297 (0.184)	−0.300 *** (0.079)	−0.377 ** (0.151)	−0.019 *** (0.006)
NS Law * Initially High Quality	0.446 ** (0.178)	0.247 (0.211)	0.390 (0.316)	0.285 * (0.169)	0.623 ** (0.251)	0.024 *** (0.008)
N	1053	1005	1005	1005	1053	1107
Significance Level of Sum of Above 2 Coefficients (not significant indicated by dash)	-	-	-	-	*	-
Panel D. Dependent Variable: Preventable Delivery Complication Rate						
National-Standard (NS) Law Dummy	−0.403 *** (0.086)	−0.401 *** (0.093)	−0.433 *** (0.113)	−0.279 ** (0.131)	−0.363 ** (0.187)	−0.023 ** (0.009)
NS Law * Initially High Quality	0.513 *** (0.135)	0.422 *** (0.124)	0.486 ** (0.187)	0.262 (0.176)	0.725 ** (0.282)	0.033 ** (0.15)
N	1089	1035	1035	1035	1089	1107
Significance Level of Sum of Above 2 Coefficients (not significant indicated by dash)	-	-	-	-	**	-
Control Variables? (Including Damage Caps and Other Remedy-Focused Reforms)	NO	YES	YES	YES	NO	NO
State-Specific Linear Trends?	NO	NO	YES	NO	NO	NO
State-Specific Linear Pre- Treatment Trends?	NO	NO	NO	YES	NO	NO
State-Year Cells Weighted by Cell Size?	YES	YES	YES	YES	NO	YES

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variables Logged or in Levels?	Logged	Logged	Logged	Logged	Logged	Levels

Notes: robust standard errors corrected for within-state correlation in the error term are reported in parentheses. All regressions include state and year fixed effects. Where indicated, regressions in Panel A are weighted by the number of admissions (for the relevant state and year) in the sub-sample of discharges associated with the relevant selected conditions (e.g., acute myocardial infarction). Where indicated, regressions in Panel B are weighted by the low-variation health index (i.e., the sum of discharges for acute myocardial infarction, stroke, hip fracture or gastrointestinal bleeding) associated with each state-year cell. Where indicated, regressions in Panels C and D are weighted by the number of deliveries associated with the relevant state-year cell. Inpatient mortality rates are risk-adjusted for the incidence (among the sub-sample) of each of the conditions comprising the sub-sample of selected conditions.

Source: 1979–2005 National Hospital Discharge Surveys.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 4

Relationship between Damage Caps and Various Health Care Quality Metrics

	(1)	(2)	(3)
Panel A. Dependent Variable: Inpatient Mortality Rate for Selected Conditions (Logged)			
Non-Economic Damage Cap	0.013 (0.030)	−0.013 (0.031)	−0.049 (0.034)
Collateral Source Rule Reform	-	0.016 (0.028)	0.011 (0.040)
Punitive Damage Cap	-	0.004 (0.041)	0.013 (0.045)
Joint and Several Liability Reform	-	0.027 (0.039)	−0.003 (0.038)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable	[−0.047, 0.074]	[−0.075, 0.049]	[−0.117, 0.019]
F-Statistic (Malpractice Variables Jointly = 0)	-	0.13	0.54
Prob > F (p value)	-	0.97	0.71
N	1158	1074	1074
Control Variables?	NO	YES	YES
State-Specific Linear Trends?	NO	NO	YES
Panel B. Dependent Variable: Low-Discretionary Avoidable Hospitalization Rate (Logged)			
Non-Economic Damage Cap	−0.008 (0.023)	−0.021 (0.017)	−0.043** (0.020)
Collateral Source Rule Reform	-	0.005 (0.031)	0.009 (0.035)
Punitive Damage Cap	-	−0.036 (0.032)	−0.064 (0.035)
Joint and Several Liability Reform	-	−0.000 (0.039)	0.026 (0.029)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable	[−0.054, 0.038]	[−0.055, 0.013]	[−0.084, −0.002]
F-Statistic (Malpractice Variables Jointly = 0)	-	1.19	0.91
Prob > F (p value)s	-	0.328	0.464
N	1193	1108	1108
Control Variables?	NO	YES	YES
State-Specific Linear Trends?	NO	NO	YES
Panel C. Dependent Variable: Maternal Trauma Rate (logged)			
Non-Economic Damage Cap	−0.066 (0.060)	−0.098 (0.070)	−0.064 (0.065)
Collateral Source Rule Reform	-	0.048 (0.081)	0.022 (0.087)
Punitive Damage Cap	-	−0.159*** (0.058)	−0.207*** (0.059)
Joint and Several Liability Reform	-	0.129 (0.114)	0.113 (0.106)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable	[−0.186, 0.054]	[−0.239, 0.044]	[−0.195, 0.066]
F-Statistic (Malpractice Variables Jointly = 0)	-	1.94	2.93
Prob > F (p value)	-	0.11	0.03
N	1105	1005	1005
Control Variables?	NO	YES	YES
State-Specific Linear Trends?	NO	NO	YES
Panel D. Dependent Variable: Preventable Delivery Complication Rate (Logged)			
Non-Economic Damage Cap	−0.046 (0.048)	−0.051 (0.049)	−0.003 (0.049)
Collateral Source Rule Reform	-	0.029 (0.044)	0.022 (0.049)
Punitive Damage Cap	-	−0.040 (0.033)	−0.053 (0.058)
Joint and Several Liability Reform	-	0.043 (0.051)	0.005 (0.065)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	(1)	(2)	(3)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable	[−0.142, 0.051]	[−0.151, 0.048]	[−0.103, 0.098]
F-Statistic (Malpractice Variables Jointly = 0)	-	0.24	0.12
Prob > F (p value)	-	0.92	0.97
N	1116	1035	1035
Control Variables?	NO	YES	YES
State-Specific Linear Trends?	NO	NO	YES

Notes: robust standard errors corrected for within-state correlation in the error term are reported in parentheses. All regressions included state and year fixed effects. Regressions in Panel A are weighted by the number of admissions (for the relevant state and year) in the sub-sample of discharges associated with the relevant selected conditions (e.g., acute myocardial infarction). Regressions in Panel B are weighted by the low-variation health index (i.e., the sum of discharges for acute myocardial infarction, stroke, hip fracture or gastrointestinal bleeding) associated with each state-year cell. Mortality rates are risk-adjusted for the incidence (among the sub-sample) of each of the conditions comprising the sub-sample of selected conditions. Regressions in Panels C and D are weighted by the number of deliveries associated with the relevant state-year cell. Inpatient mortality rates are risk-adjusted for the incidence (among the sub-sample) of each of the conditions comprising the sub-sample of selected conditions. Columns 2 and 3 of Table 4 are from the same specification underlying Columns 2 and 3 of Table 4, simply focusing on the presentation of the coefficients of the remedy focused reforms as opposed to the coefficients of the national-standard variables.

Source: 1979–2005 National Hospital Discharge Surveys.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 5

Relationship between Damage Caps and Cancer Screening Rates

	(1)	(2)	(3)	(4)	(5)	(6)
	Mammogram	**Physical Breast Exam**	**Proctoscopic Exam**	**PSA Testing**	**Digital Rectal Exam**	**Pap Smear**
Non-Economic Damage Cap	−0.003 (0.006)	−0.005 (0.007)	−0.006 (0.005)	0.002 (0.006)	0.014 (0.008)	−0.007 (0.006)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable (Percentage Point Impacts)	[−0.015, 0.008]	[−0.019, 0.009]	[−0.016, 0.003]	[−0.009, 0.013]	[−0.001, 0.030]	[−0.019, 0.005]
95% Confidence Band, scaled by mean screening rate (Percentage Impacts)	[−0.021, 0.011]	[−0.030, 0.014]	[−0.040, 0.008]	[−0.017, 0.025]	[−0.002, 0.060]	[−0.032, 0.008]
N	1009965	1155814	843960	252232	340931	1662616

Notes: robust standard errors corrected for within-state correlation in the error term are reported in parentheses. All regressions included state and year fixed effects.

Source: 1987–2008 Behavioral Risk Factor Surveillance System Records.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 6

The Relationship between Damage Caps and Related Reforms and Various Health Care Quality Metrics, Separately for Initially High and Initially Low Quality States

	(1)	(2)
	Initially Low Quality States	**Initially High Quality States**
Panel A: Dependent Variable: Inpatient Mortality Rate for Selected Medical Conditions		
Non-Economic Damage Cap	−0.043 (0.032)	0.041 (0.056)
Collateral Source Rule Reform	0.026 (0.040)	−0.022 (0.037)
Punitive Damage Cap	−0.013 (0.059)	0.158*** (0.040)
Joint and Several Liability Reform	0.030 (0.051)	0.009 (0.042)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable	[−0.110, 0.022]	[−0.075, 0.157]
Panel B: Dependent Variable: Low-Discretionary Avoidable Hospitalization Rate		
Non-Economic Damage Cap	−0.004 (0.020)	−0.031 (0.022)
Collateral Source Rule Reform	0.007 (0.039)	−0.064** (0.025)
Punitive Damage Cap	−0.008 (0.036)	−0.031 (0.029)
Joint and Several Liability Reform	−0.023 (0.049)	0.059 (0.031)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable	[−0.049, 0.039]	[−0.077, 0.014]
Panel C: Dependent Variable: Maternal Trauma Rate		
Non-Economic Damage Cap	−0.021 (0.072)	−0.000 (0.087)
Collateral Source Rule Reform	−0.165 (0.137)	0.154 (0.091)
Punitive Damage Cap	−0.089 (0.073)	0.044 (0.092)
Joint and Several Liability Reform	0.215 (0.214)	−0.030 (0.098)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable	[−0.173, 0.130]	[−0.179, 0.179]
Panel D: Dependent Variable: Preventable Delivery Complication Rate		
Non-Economic Damage Cap	0.019 (0.039)	0.042 (0.092)
Collateral Source Rule Reform	−0.031 (0.052)	0.062 (0.043)
Punitive Damage Cap	0.092 (0.062)	−0.004 (0.046)
Joint and Several Liability Reform	−0.000 (0.052)	0.035 (0.070)
95% Confidence Band for Coefficient of Non- Economic Damage Cap Variable	[−0.063, 0.100]	[−0.135, 0.219]

Notes: robust standard errors corrected for within-state correlation in the error term are reported in parentheses. All regressions include state and year fixed effects. Regressions in Panel A are weighted by the number of admissions (for the relevant state and year) in the sub-sample of discharges associated with the relevant selected conditions (e.g., acute myocardial infarction). Regressions in Panel B are weighted by the low-variation health index (i.e., the sum of discharges for acute myocardial infarction, stroke, hip fracture or gastrointestinal bleeding) associated with each state-year cell. Regressions in Panels C and D are weighted by the number of deliveries associated with the relevant state-year cell. Inpatient mortality rates are risk-adjusted for the incidence (among the sub-sample) of each of the conditions comprising the sub-sample of selected conditions.

Source: 1979–2005 National Hospital Discharge Surveys.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.