# Malpractice Risk and Medical Treatment Selection*

Paola Bertoli†
University of Economics, Prague

Veronica Grembi
University of Milan
University of Economics, Prague

**Abstract**

We study how legal and financial incentives affect medical decisions. Using patient-level data from Italy, we identify the effect of a change in medical liability pressure by exploiting the geographical distribution of hospitals across court districts, where some districts increase the predictability of expected damages per injury while others do not. Using a difference-in-differences identification strategy, we show that as certainty of compensation increases, c-sections increase by 6.5 percentage points. There is no statistically significant effect on secondary health outcomes of either mothers or newborns, but the increase is higher for low-risk than high-risk mothers. The increase is driven by hospitals that have lower quality, are governed by inefficient court districts, face lower expected damages, and are paid more per c-section.

*JEL Classification: K13; K32; I13*
*Keywords*: Scheduled Damages, Cesarean Sections, Difference in Differences

*We are thankful for their comments on previous and preliminary drafts of this work to Bernie Black, Omri Ben-Shahar, Amitabh Chandra, Jill Fish, Michael Frakes, Francesco Moscone, Joseph Sabia, Ity Shurtz, Valentina Tonei, Zenon Zabinski, and the participants at the seminars at Harvard University, the Institute of Economic Studies (Charles University), University of Lugano, University of York, Sapienza University of Rome, ETH Zurich, Lund University, the 2016 Prague workshop on the Law and Economics of Health Care, the 2016 European Law and Economics Association Conference, the 2016 Conference of Empirical Legal Studies, the 2017 Conference of Spanish Law and Economic Association, and the 2017 Summer Institute in Law and Economics (University of Chicago). The paper greatly benefited from the comment of three anonymous refereesF. The usual disclaimer applies. Support from Grant VSE IGS F5/68/2015 is gratefully acknowledged.

†CEREGE-EI Teaching Fellow. CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, P.O. Box 882, 111 21 Prague 1, Czech Republic.

# 1 Introduction

Health care systems struggle to reduce the burden of inappropriate treatments, which increase expenditures without improving the medical condition of patients. Inappropriate care might be exacerbated by the medical liability regimes (*e.g.*, medical malpractice pressure and fear of litigation), among other factors. Assessing the consequences of medical liability on the selection of treatments is not always easy (*e.g.*, Danzon 2000): to avoid the risk of being sued, providers would resort to excessive care and/or refuse to accept the riskiest patients and procedures. Tort reforms that reduce medical liability are the traditional response of scholars and policy makers to this problem, which is often labeled as *defensive medicine*. However, the evidence on the direction and magnitude of the effects of a reduction of medical liability on defensive practices is mixed. For example, a decrease in medical liability has been shown to be associated with fewer unnecessary treatments (Yang *et al.*, 2009; Esposto, 2012; Shurtz, 2013), more unnecessary treatments (Dubay *et al.*, 1999; Currie and MacLeod, 2008), and no impact on the selection of medical treatments (Baldwin *et al.*, 1995; Sloan *et al.*, 1997; Frakes, 2012).

We study whether and how a reduction in medical liability affected the use of cesarean sections, which is considered one of the most inappropriately used procedures in developed countries. We measure the reduction in medical liability as due to the introduction of schedules of non-economic damages in Italy.[1] Our work improves on the existing literature in at least three ways. First, our institutional setting allows us to study the effect of a change in medical liability at the hospital level, while previous studies mainly focused on a state level.[2] Second, our identification exploits the not perfect overlap between *Local Health Authority* (LHA) districts and court districts, where scheduled damages were adopted. As a result, some hospitals have been affected by the liability reduction, while other nearby hospitals, managed by the same health authority, have not. Third, exploiting the analysis at the hospital level, we offer insights into the channels driving provider reaction that constitute a new starting point for a policy debate around malpractice risks. In particular, we focus on the quality of hospitals, their financial incentives, and the heterogeneous levels of medical liability pressure that hospitals might face.

Schedules of non-economic damages increase the predictability of payouts being tables

[1] *Damage schedules*, *scheduled damage tables* or *schedules* are other terms for the same tables. For further information on scheduling damages, see, among others, Avraham (2006) and Mello and Kachalia (2010).

[2] Notable among the studies adopting a narrower focus than the state level one are Lakdawalla and Seabury (2012), who identify the impact of malpractice liability through variations in county-level jury verdicts; Paik *et al.* (2012), who exploits differences in claim rates among counties in Texas to study the effect of the implementation of a strict damages cap; and Frakes and Gruber (2019), who exploit the no-liability system granted to the Military Health System.

with entries for the injury severity level and the victim's age (Table A.1): different combinations of age and injury severity lead to different compensation amounts. Schedules decrease the variance of expected damages and, in doing so, make the average expected compensation more meaningful for each agent involved in a potential malpractice case.

A reduced variance might have several effects on malpractice risk. It makes the assessment of the risk exposure of healthcare providers more accurate, facilitating access to insurance coverage. It also facilitates the calculation of the expected payoff of a malpractice claim, making an out-of-court agreement more likely. At the same time, attorneys, who are paid on a fee-for-service basis, have lower incentives to accept simple and less time-consuming cases, and potential risk-loving plaintiffs are discouraged by the elimination of the large-prize lottery system. All these effects release pressure on health care providers. In principle, there might also be an additional effect on claims: risk-averse plaintiffs might find legal proceedings a less fearful option under schedules. Yet, empirical evidence shows that overall claims may have decreased in Italy.

The change in the average of the distribution is an empirical issue as, ex ante, there might be three outcomes: the mean of the compensation distribution can decrease, increase, or remain unchanged. In the case of Italy, we provide descriptive and anecdotal evidence of a decrease in compensation. In particular, two aspects of the history of schedules in Italy should be considered. On the one hand, insurance companies, the main interest group lobbying to implement schedules for resolving traffic-accident related claims, were clearly concerned about avoiding an increase in the average compensation. On the other hand, judges, the ones in charge to approve schedules adoption, were not interested in reducing legal certainty by reversing the court's precedents in terms of average compensations for the same injuries.

Based on a unique dataset of inpatient discharge records from the Italian Ministry of Health for the period 2001-2004, our identification strategy exploits the fact that, during this period, 18 courts switched to schedules only in 2002, and there were no other malpractice or tort reforms or any change in hospital payments. Since hospital location determines the court in which malpractice claims must be filed, we consider hospitals located under the jurisdiction of courts switching to schedules as the treated group. Hospitals located in court districts that did not switch are the control group. Having geolocated the hospitals, we can minimize the unobserved heterogeneity by confining our analysis to those that are managed by a single LHA but are overseen by two courts, only one of which switched to schedules. Our final sample includes 22 treated and 19 control hospitals.

The adoption of schedules leads to a 6.5-percentage-point increase in c-sections (17% at the mean of c-sections) that begins at quarter 5 after the introduction of the policy. Although we cannot rule out different from zero effects on secondary health outcomes of either mothers or newborns (*e.g.*, preventable complications), estimating our preferred specification on the subsamples of high-, medium-, and low-risk deliveries shows that the use of c-sections increases the most among low- and medium-risk mothers (19%), rather than on high-risk mothers (5%). Overall, this suggests a greater mismatch between the selected delivery method and the patient's medical needs. Our findings are consistent with the idea that reduced medical liability decreases the cost of errors, so providers have no incentive to reduce the incidence of inappropriate care (Currie and MacLeod, 2008). Back-of-the-envelope calculations performed using the national Diagnosis Related Group (DRG) values for c-sections and vaginal deliveries provide an economic approximation of the first-order costs of the increase in c-sections: 1,400 extra euros per delivery.

Our results are robust to the use of different types of clusters, to the inclusion of court fixed effects and hospital fixed effects, and to hospital-month trends. We also exclude any anticipatory behaviors and the possibility of a change in either the number of managed deliveries or in the composition of managed patients as proxied by the predicted probability of c-section (*i.e.*, as a function of risk factors), the incidence of low-risk mothers or of low-weight newborns across treated and control hospitals. Finally, we show that the observed effects cannot be attributed to other policies enforced during the same period by testing the adoption of schedules in 2002 for hospitals that either consistently used or never employed schedules between 2001 and 2004.

Too often, the analysis of the impact of changes in medical liability is done without considering other relevant context-driven factors, which can affect the selection of medical procedures. In examining the drivers of the response to the implementation of schedules, we show that low-quality hospitals are more prone to overuse c-sections, where low quality is proxied by low volumes of deliveries (*i.e.*, less than 500 per year). When we focus on the interactions with other sources of liability pressure, we find that a higher increase in c-sections is associated with lower levels of schedules (*i.e.*, lower expected damages) and a higher level of civil backlog for the court district in which the hospital is located. Finally, stronger financial incentives, measured as the difference in the reimbursement rate between a cesarean and vaginal delivery, predict the increase of unnecessary c-sections. This confirms that hospitals exploit the decreased malpractice risk to further pursue the more profitable procedures, as suggested by Shurtz (2014) and Frakes (2015). This evidence suggests that closing small birth centers, providing more monitoring of the appropriateness of procedures, and investing in an efficient judiciary system can have a positive effect on the selection of appropriate medical procedures.

This paper is organized as follows: Section 2 presents the institutional setting and provides more information on schedules and their expected effects, while Section 3 defines data, main outcomes, and empirical strategy. Section 4 presents the main results and describes the validity and robustness checks together with the analysis of the channels. Finally, Section 5 concludes.

# 2 Institutional Background and Theoretical Expectations

## 2.1 The Italian Healthcare System and Its Liability Pressure

In Italy, the central and regional governments share the responsibility for healthcare. The national government establishes the minimum level of care to be provided to the population, while regions are responsible for the actual delivery of health services. In particular, regional governments manage the local provision of healthcare using a network of Local Healthcare Authorities (LHAs), which are geographically based organizations. Public hospitals are distributed within each LHA district and patient mobility is allowed.[3] Between 2001 and 2004, 87% of newborns were delivered in public hospitals (Ministero della Salute, 2002 and 2003) and this trend has not weakened over time; in 2013, more than 88% of mothers chose to give birth in a public facility (Ministero della Salute, 2015). Moreover, 84% of mothers gave birth in their own LHA, and 62% of them used the nearest hospital, traveling 10.33 km (6.4 miles) on average.

Physicians are held jointly liable with the hospital they work for (Traina, 2008), and the negligence standards they are subject to are based on what is established by their medical specialty (Frakes, 2013). There is no evidence that there were any changes in these standards over the considered period. If there were changes, they should be assumed to have affected every ward practicing Ob-Gyn. Hence, we can exploit the hospitals not affected by the switch to schedules to check this point, as is done in the robustness checks in Section 4.2.

Notwithstanding the fact that the healthcare system in Italy is public and hospitals physicians are civil servants, the fear of being dragged into a malpractice claim is well perceived by both physicians and hospitals. There are no national official statistics on either medical errors or requests for compensation, but some regions publish independent reports on the aggregated claim histories of their hospitals: between 2004 and 2010, hospitals in the Northern regions reported 9.22 claims per 100 beds, while the figures for the Central and Southern regions were 12.44 and 12.70, respectively (Ronzoni, 2012). According to a re-

[3]There are also private hospitals, which can act as completely private facilities or through special agreements with the public system (private accreditation).

port based on data from Lombardy (1999-2011) and Piedmont (2005-2011), the wards facing the most claims were emergency rooms, orthopedics, surgery, and obstetrics and gynecology (*e.g.*, Lombardy 2012; Piedmont 2012). On average, 30% of claims are dismissed, so it is not easy to classify claims as good proxies for medical errors. Data collected by consumers' associations (*e.g.*, Cittadinanzattiva) partially confirm this evidence. Ob-Gyn ranks steadily as the third medical specialty for the most denounced errors. Among the overall claimed cases, more than 40% would have occurred during delivery. Among these, 34% specifically tackle problems with the performance of a c-section (Cittadinanzattiva, 2003). Moreover, while there is no bullet proof evidence of a high incidence of medical errors, the public feeling toward this issue is quite strong. For instance, according to a survey conducted in 2005 among EU members (Eurobarometer, 2006), 97% of the Italian respondents rated medical errors as being of *high* importance in their country, compared to the 78% average for the EU 25 countries; 53% of Italians surveyed reported having *often* read about medical errors in their country, as opposed to the European average of 34%. However, Italian respondents or members of their families experienced a serious medical error in only 18% of cases, a figure that is in line with those of the rest of the EU countries. More recently, in 2009, only 15% of those interviewed reported having experienced any form of adverse event when interacting with the healthcare system, compared to the European average of 27% (Eurobarometer 2009). Italians also appear to be more prone to legal action. When in a doubt about a medical error, respondents would seek help from hospital management only 18% of the time, compared to a European average of 37%, while they would seek legal assistance in 53% of cases, as opposed to the 48% European average. In terms of redress, Italians expect action to be taken against both the hospital (51% vs. 36% for the EU 27) and the person responsible for the error (48% vs. 37% for the EU 27).

As physicians feel the pressure to be involved in a legal claim, hospitals are legally required to provide malpractice insurance coverage to their employees. However, physicians typically look for additional coverage in the market. Descriptive statistics on insurance contracts from a set of insurance companies show that in 2012, for instance, hospitals paid premiums up to 288 million euros, while individual physicians paid 255 million euros to their own insurance companies (ANIA 2014). When requesting an insurance contract, physicians' individual claim history does matter. Among other things, physicians are asked to report the number of claims received in the last 5 years, whether she was previously insured with any other company, the reasons why she wants to change insurer (*e.g.*, the previous policy was not renewed, or the previous insurer withdrew from the contract), the premium paid to the previous insurer, and how far back she would like the retroactive coverage to go. Physicians must also specify their medical specialties and gynecologists are separated into two categories: those who are involved in deliveries and those who are not. The former is considered riskier than the latter.[4]

## 2.2 Scheduled Damages and Why They Matter

Scheduled damages quantify the award to be granted based on the severity of the injury (*i.e.*, percentage of disability) and the victim's age, thereby minimizing judges' discretion in the assessment of compensation.[5] For instance, the schedule applied by the court of Milan in 2002 sets the non-economic compensation at 19,704 euros for an 11% disability suffered by a 3-year-old child, whereas this figure would increase to 34,935 euros if the same child suffered a 16% disability.

Given their structure, schedules are expected to remove the uncertainty of claim payouts by affecting both their variance and average. So far, the literature on medical malpractice has stressed the importance of reducing the average expected damage, almost disregarding the dispersion of the damages distribution (Frakes and Jena 2016). Still, descriptive data show that average compensation is often not a good approximation of the real financial burden entailed by a medical error. For instance, in Lombardy, the average compensation granted for a surgical error in the period 1999-2011 was 52,436 euros with a standard deviation of 160,726 euros.[6] The standard deviation is 3 times the mean revealing a highly dispersed distribution. The same holds true for diagnostic errors, which had a 3.3 ratio (mean 71,499; std. 236,669), and therapeutic errors, which had a 3.7 ratio (mean 70,987; std. 264,441), while a 2.9 ratio is registered when considering only Ob-Gyn. The skewness of the distribution of malpractice compensation is also confirmed by other datasets, as shown by Jena *et al.* (2011).

A main strength of schedules is that by reducing the variance, they make the average compensation truly representative of the distribution and remove the uncertainty of final payouts. Theoretically as variance decreases, the average compensation could either increase, decrease, or remain unchanged. In the Italian context, we can rule out that there was an increase in the average compensation, as a main concern when implementing the policy was to not affect the averages.

Italy has 165 courts of first instance, which, in the 1990s began to adopt scheduled damages to provide clear guidance in setting non-economic compensation (*i.e.*, 70% of the final paid damages) for any case of personal injury (Comandé, 2005). Insurance companies

[4]This information has been recovered through standard contracts provided by the Lloyds.

[5]As stated by Mello and Kachalia (2010, p.23) "Schedules promote both horizontal equity (similar injuries received similar compensation) and vertical equity (more severe injuries receive higher compensation). Schedules also serve the goal of absolute fairness in compensation-that is, setting damages to match (and not exceed) societal expectations about what constitutes appropriate compensation for particular injuries."

[6]Lombardy is the largest and most populated Italian region. It has approximately 10,000,000 inhabitants, the equivalent of the combined population of Sweden and Finland.

welcomed the policy and lobbied in its favor, in particular, to liquidate road traffic accidents (Busnelli, 1988). Still, insurers were also clearly interested in avoiding an increase in the average compensation. The same concern was shared by courts, which are the authorities responsible for approving schedule implementation, as judges vote for their introduction. Judges did not want to harm legal certainty by reverting the court's precedents in terms of average compensations for the same injuries.

Once in place, schedules were never repealed and are very unlikely to be waived (Bertoli 2014, and Bertoli and Grembi 2017c). For this to happen, judges must provide proper justification and adapt their decisions to the greatest extent possible to the average compensation granted in previous cases. In the Italian jurisprudence, scheduled damages, themselves, are traditionally defined consistently with past decisions. Their level is set to be coherent with the compensations most commonly granted. As a result, current rulings should not necessarily reduce deterrence relative to past levels.[7]

Given the double effect on the mean and variance, schedules matter for both physicians and hospitals, because they can affect the availability of insurance coverage as well as in- and out-of-court litigation. Removing the uncertainty of payouts also means removing it from the insurers' estimations of their risk exposure. Hence, operating in the malpractice line should become less complex with positive effects on the availability of malpractice coverages. At the same time, an easier calculation of the expected payoffs increases the likelihood of an out-of-court agreement, which has the advantages of being faster completion and less publicity than a legal proceeding, thus representing a smaller source of stress for the parties involved. On the victims' side, schedules allow them to better evaluate the benefits of filing a claim. Frivolous claims and/or claims for minor injuries might no longer be worth filing for risk-loving plaintiffs, but trials could become a less fearful option for risk-averse plaintiffs. By contrast, attorneys, who are paid on a fee-for-service basis, have lower incentives to accept simple and less time-consuming cases.

To provide evidence of the impact of the policy on litigation and compensations, we recovered data on insurance claims filed with commercial insurers that closed between 2000 and 2004.[8] Still, such evidence has two main limitations. First, we rely on insurance claims

[7]Physicians continue to face criminal charges if any proof of professional misconduct is provided. Under Italian criminal law, prosecutors do not have any discretion in terms of the case they can prosecute. Criminal charges are mandatory in any case in which the prosecutor's office receives a complaint, and filing a criminal complaint is free. There was no change in the criminal liability of physicians over our observation period. Additionally, as pointed out by Mello and Kachalia (2010), among the negative implications of the unpredictability of damages awards, we count a weakened deterrence of medical errors. Hence, ex ante, it is hard to say whether we should expect a decrease in deterrence due to the schedules introduction.

[8]Claims targeting individual physicians without the involvement of a hospital are not available. However, under the Italian legal system, such a scenario is more likely in private practice cases than for deliveries performed in a public hospital.

because publicly available information on malpractice claims at either the court or out-of-court level is not available. Second, the evidence we provide is based on the results at the new equilibrium triggered by the policy. This means that the frequency of claims in the post-policy period can be affected by both the effect of the policy on the incentives to file a claim but also by the different behaviors adopted by the medical practitioners in the new scenario.

We focus on four outcomes at the hospital level normalized by 100 beds: 1) claim frequency; 2) the percentage of dismissed claims; 3) the average amount reserved per claim; and 4) the variance of the average reserve per claim. Furthermore, from a dataset on the public procurement procedures for malpractice insurance contracts, we collected the number of insurers bidding in the tendering processes run by the hospitals included in our sample between 2001 and 2004.[9]

Table 1 summarizes the results.[10] After schedules introduction, although physicians are expected to take higher risk potentially leading to more claims, hospitals have a lower probability of being sued, as claim frequency is negatively affected. A change in the composition of claims contributes to this reduction, as is clear from the decrease in the frequency of dropped claims. In fact, the result in column (2) suggests that claims are less likely to be dismissed; thus, schedules discourage frivolous claims. This suggests less pressure on physicians and hospitals, who are less likely to receive a claim and to go through a proceeding with low legal merit. Both average reserves and their variance are lower under schedules, while the insurance market responds to schedules introduction with a higher number of bidding insurers in tenders for malpractice coverage.

Finally, we also recovered the annual number of new civil cases filed at each court included in our observational sample and normalized this figure by the resident population at the court district level. We cannot distinguish claims by type of case at the court level, so we address all types of liability cases. The results in column (6) in Table 1 show a negative relationship between schedules adoption and the number of new court cases.

Table 1, about here

Even though the ex ante effects of schedules are not necessarily clear-cut, the provided

[9]Public hospitals cannot go freely to the market to get an insurance contract; they need to go through public procurement procedures. These procedures are somewhat demanding, to the extent that hospitals often use brokers to manage them. The complexity and expensiveness of these procedures are the reasons why insurance contracts have an average duration of 32 months.

[10]The coefficients have been estimated based on the following equation: $Outcome_{jt} = \delta(Treated_j * Post_t) + \tau Treated_j + \gamma_t + \pi_j + \epsilon_{iht}$ where $j$ can be either the court or the hospital. $Post_t$ is a dummy that captures the post-treatment period, $Treated_h$ is a dummy that identifies the treated hospitals, and $\delta$ identifies the effect of schedules. $\gamma_t$ represents year fixed effects to control for common shocks, while $\pi_j$ are hospital or court fixed effects.

evidence suggests that this policy ultimately leads to a decrease of malpractice pressure. Physicians face a lower likelihood of being involved in a claim, of facing claims with low legal merit, and of being required to attend to an open court hearing. At the same time, reserves are more easily predictable and insurers are more willing to offer malpractice coverage.

## 2.3 Theoretical Expectations

*Ex ante*, it is difficult to theoretically predict the effect of a change in medical liability on the selection of the delivery method. The main assumption is that healthcare practitioners are concerned about facing a legal claim. Even when providers can obtain medical liability insurance that reduces their financial risk, they, nevertheless, perceive malpractice claims as a serious threat because they entail non-insurable costs, including serious reputational damages (Sage, 2004) and significant psychological and time costs.[11] However, the risk of facing a claim cannot be eliminated, as it is correlated with the probability of committing an error, which cannot be nullified even when taking precautions (Arlen and MacLeod 2005). As a result, healthcare providers are commonly expected to make medical decisions while considering both patients' conditions and the risk of being sued in the event of mistakenly performing a treatment. Within this framework, the conventional wisdom is to associate lower medical liability with a decrease in the need for providers to protect themselves against the risk of litigation. Since c-sections are traditionally classified as defensive medicine against the risk of being sued,[12] this approach predicts a reduction in c-section rates whenever there is a decrease in medical liability. However, the existing evidence is mixed and challenges this interpretation (Baldwin *et al.*, 1995; Dubay *et al.*, 1999; Currie and MacLeod, 2008; Yang *et al.*, 2009; Esposto, 2012; Shurtz, 2013).

To explain the conflicting evidence, more recent contributions attempt to adopt a broader perspective. Three models best identify this attempt and are found in Currie and MacLeod (2008), Shurtz (2014) and Frakes (2015). Currie and MacLeod (2008) model doctors' decisions as directly dependent on both the patient's condition and on the probability of committing an error with possible legal consequences. In essence, they prove that the final effect of a tort reform on the utilization rate of a procedure cannot be uniquely determined *ex ante* on theoretical grounds. They provide a framework to interpret the observed changes in medical decisions in response to the relative risk between performing and not performing a procedure. If we move from a point at which the use of a procedure is excessive, which

[11]Seabury *et al.* (2013) show that doctors, on average, spend over 4 years of a 40-year career with an open malpractice claim. However, there is no clear evidence on the magnitude of reputational costs.

[12]By performing c-sections, doctors reduce the risk to babies (*i.e.*, the most expensive potential victim) and can better control what actually happens in the delivery room.

means that it is not related to medical factors, then the probability of committing an error or experiencing adverse outcomes is higher with the use of that procedure than without it. This means that whenever an increase in malpractice pressure strikes (*i.e.*, physicians are held more accountable), we expect a decrease in the incidence of that procedure rather than an increase.[13] As a result, the effects of tort reforms depend on the specific reform under consideration: while the Joint and Several Liability (JSL) rule increases physicians' accountability and leads to a reduction in both the c-section rate and the number of labor complications, caps on noneconomic damages increase them.

A further extension of this dynamic is provided by Shurtz (2014) and Frakes (2015) who both stress the role of the financial incentives faced by practitioners.[14] According to Shurtz (2014), a doctor will perform a treatment at the margin up to the point at which the monetary gains are offset by the risk of being sued in the event of an error. When c-sections are profitable and overused, a reduction in medical liability will not affect the financial incentive to perform a c-section on the marginal patient; instead, the reduction simply decreases the malpractice risk associated with this decision. As a result, lower medical liability provides doctors with greater discretion to follow financial incentives, and thus, the use of medically unnecessary treatments might increase. Similarly, Frakes (2015) notes that a decrease in liability may have the traditionally expected result (*i.e.*, fewer procedures) if doctors have weak financial motivations.

# 3 Empirical Analysis

## 3.1 Data

We use two unique datasets from the National Hospital Discharge Records (*Schede di Dimissione Ospedaliera - SDO*). The first dataset contains records of all deliveries that took place in Italian public hospitals between 2001 and 2004, whereas the second dataset provides information on newborns.[15]

According to national legislation, a malpractice claim against a hospital or its employees

[13]By contrast, if the non-performance of the treatment implies a higher risk of being sued in the event of an error than the performance of the same treatment, such a variation in malpractice pressure leads to an increase in the use of the treatment. Hence, studying outcomes other than delivery methods could lead to different results.

[14]A seminal work that discusses the interaction between liability and financial incentives, but in a less formalized way, is Kessler and McClellan (2002), who theorize that tort reforms decreasing liability, such as caps, are less effective at reducing unnecessary treatments when there are strong monetary incentives from managed care.

[15]Patient-level data on deliveries are not available before 2001. The two datasets come separately for privacy reasons.

must be filed in the court district in which the hospital is located. This institutional feature allows us to exploit the geographical distribution of hospitals to identify the treated and control groups. The treated hospitals are those located in a court district that adopted schedules in 2002, while hospitals within the court district that did not switch to schedules represent the control group. We identify our sample of interest in three ways. First, we use the entire population of Italian public hospitals between January 2001 and December 2004 and provide evidence of differences among their observable characteristics. Then, to reduce both observable and unobservable heterogeneities between the treated and control hospitals, we focus on the deliveries that occurred in hospitals managed by the same LHA but that were overseen by two different courts, of which only one switched to schedules in 2002. This means that our final sample includes only treated and control hospitals located in the same LHA,[16] as depicted in Figure 1 for the representative region of Sicily. This is a unique design in which hospitals treating very similar patients and managed by the same administrative unit face different levels of liability pressure due solely to their location. Finally, the available data do not allow us to distinguish between emergency and planned c-sections. However, there is a clear trend in the use of cesarean deliveries during the week. In Table B.1, we plot the results of an OLS model in which the average number of performed c-sections is regressed on the days of a week. The results confirm that c-sections are performed mostly during workdays—as expected; thus, cesarean deliveries occurring on weekends are most likely to be performed for emergency reasons (Amaral Garcia *et al.*, 2015). To cope with this bias, we exclude all weekend deliveries and provide the results with the inclusion of weekends in Online Appendix B.

Figure 1, about here

The initial population analysis includes 587 hospitals (95 treated and 492 control), where 1,788,630 deliveries of 1,743,884 newborns took place, and these values decrease to 1,359,504 and 1,332,170, respectively, once weekend deliveries are not considered.[17] When we focus only on hospitals belonging to the same LHA but ruled by two different courts, of which only one switched to schedules in 2002, we count 95,750 deliveries and 95,792 newborns in a total of 41 hospitals. After dropping weekend deliveries, our main sample of interest includes a total of 72,830 deliveries and 73,717 newborns from 22 treated and 19 control hospitals.

We perform t-tests on the main observable characteristics of both the overall population of treated and control hospitals and on the smaller sample of treated and control hospitals belonging to the same LHA. The objective is to check whether there are any major differences

[16] The average distance between treated and control hospitals is 29.5 miles.

[17] Data on mothers and newborns might differ in number of observations, due to twins and stillbirths.

that could confound the estimated effect of our policy. We focus on hospitals' operational characteristics in the form of the number and composition of medical staff and the number of wards and beds, as well as on the hospital's overall treated population (*i.e.*, case-mix and entropy indexes). [18] As shown in Table 2 Panel (A), treated and control hospitals are structurally different when we use the entire population of hospitals. However, when we restrict the analysis to those hospitals within the same LHA, the observable characteristics are balanced as shown in Panel (B) where none of the t-test values for the differences between treated and control hospitals are statistically significant.

Table 2, about here

Tables A.2 and A.3 report the descriptive statistics. On average, nearly 38% of women gave birth by c-section between 2001 and 2004.[19] Since 1985, the World Health Organization (WHO) has established a range of 10%-15% as an acceptable incidence of c-sections, and, more recently, *Healthy People 2010* confirmed this view by establishing a new target of 15% for the performance of c-sections in the US. Hence, this evidence indicates a very high use of cesarean deliveries that cannot be explained by the risk profile of mothers.[20] Only 11 women out of 100 present health conditions that would require the use of a c-section. Overall, the majority of women were Italian and married; they were 30 years old, on average, and approximately 8% reported preventable complications. Of newborns, 2.7% suffered from a congenital anomaly, and 4.5% were negatively affected by some maternal conditions. On average 5.8% of newborns exhibited complications due to the performance of a c-section, while 22% were harmed during a vaginal delivery. Only 2% of newborns needed a breathing intervention.

Figure 2 plots the monthly trends of the normalized— at month -12— incidence of cections of treated and control hospitals. Treated and control hospitals report similar growth before the adoption of schedules, while the trends begin to diverge after their adoption (*i.e.*, month 0).

[18]This information is available only on an annual basis. The *Case-Mix Index* reflects the diversity and disease severity in the overall population of patients in a given hospital. *Entropy Index* expresses the case-mix dispersion in a given hospital, and thus the higher this index is, the more dispersed the care provided by a given hospital.

[19]Italy, as well as Spain, is the country with the highest incidence of first time mothers older than 35 in Europe (EUROSTAT 2018). Additionally, Italy has a public healthcare system, which means that a mother has no way to be sure that her Ob-Gyn is going to be on duty during her delivery. A way to ensure the presence of your own physician might be to agree to a c-section. There is no evidence that there was a dramatic change in the age of mothers in the period 2001-2004 or a change in the attitude towards dealing with a previously unknown Ob-Gyn in the same period. Hence, these drivers should not be seen as alternative explanations for our results.

[20]These figures are in line with Italy's performance in international rankings: it has the highest number of births by c-section in Europe and one of the highest among the OECD countries (OECD, 2013; Meloni *et al.*, 2012; Ministero della Salute, 2011).

Figure 2, about here

## 3.2 Main Outcomes

The decision to perform a c-section represents our main outcome of interest; thus, we use a dummy $C-section$, which is equal to 1 if a woman gave birth by c-section and 0 otherwise. In line with the model of Currie and MacLeod (2008), we expect an increase in the use of c-sections, as the decrease in liability pressure triggered by schedules occurs in a context of c-section overuse. However, it is possible that the patient population benefits from that increase. If a higher cesarean rate is associated with better outcomes for mothers and/or newborns, then arguing of an overuse of c-sections could be far-fetched, as the increase could be explained by a change in medical needs. This is only indirectly detectable by checking the incidence of preventable complications, which are our secondary outcomes. If preventable complications do not decrease, then it becomes more difficult to argue that there are health benefits for patients associated with the use of c-sections.

To capture the possible effects on maternal health, we use the variable *Preventable*, which indicates whether the mother suffered any preventable delivery or post-delivery problems as listed in Table 3.[21] By the same token, we also include two outcomes that proxy for adverse consequences on newborn health: *C-section complications*, and *Breathing interventions*.[22] Although the most commonly used measure, the Apgar score, is not publicly available, these outcomes should be considered reasonable alternatives. The first indicates whether the newborn suffered any harm specific to the choice of a cesarean delivery, while the second captures whether any attempts were made to improve newborn respiratory function (*e.g.*, intubation, ventilation, and respiratory maneuver).

Table 3, about here

## 3.3 Econometric Strategy

We identify the effect of a decrease in medical liability driven by the adoption of schedules on $Outcome_{iht}$, for mother $i$ delivering in hospital $h$ at time $t$, using a difference-in-difference (DD) approach. Because the only year in the period 2001-2004 when 18 courts implemented

[21]We also use an alternative definition of delivery problems, *Traumas*, which captures whether the mother reported any preventable traumas usually associated with the type of delivery performed. In essence, we derived *Traumas* by focusing on the preventable *Patient Safety Indicators* (PSI) developed by the Agency for Healthcare Research and Quality (AHRQ) as in Iizuka (2013). Using this alternative measure does not produce any significant effect. The results are available upon request.

[22]*C-section complications* for newborns captures whether there was a premature birth due to miscalculation of gestational age, whether there was any infant respiratory distress syndrome, and whether there were any complications due to anesthesia, among other issues.

schedules was 2002, $Post_t$ is a dummy that captures the post-treatment period, $Treated_h$ is a dummy that identifies the treated hospitals, and their interaction identifies the effect of schedules using the DD estimator ($\delta$) as defined by the model in Equation 1.

$$Outcome_{iht} = \delta(Treated_h * Post_t) + \tau Treated_h + \gamma_t + \pi_{lha} + \lambda(lha * m) + \alpha_m + X1^{'}_{iht}\sigma + X2^{'}_{iht}\beta + X3^{'}_{mht}\rho + \epsilon_{iht} \quad (1)$$

where $\gamma_t$ are year fixed effects to control for common shocks. $\pi_{lha}$ are LHA fixed effects, while $\lambda$ captures the effect of the interaction between LHA fixed effects and monthly fixed effects $\alpha_m$. Standard errors are clustered at the hospital level to address both possible serial correlation problems (Bertrand *et al.*, 2004) and because we expect heterogeneity in response to the treatment at the hospital level (Abadie *et al.*, 2017). To cope with the possible inference problem due to a small number of clusters, we apply a wild bootstrap procedure. We include three sets of covariates, that could predict the probability of having a c-section as listed in Table 3, and we also present the results without any covariates. $X1^{'}_{iht}$ groups control variables for the risk profile of the mother, $X2^{'}_{iht}$ considers the characteristics of the mother other than her health conditions, such as her age, that may affect the type of delivery or the incidence of complications, and $X3^{'}_{iht}$ controls for the socio-economic characteristics of the municipality of residence of the mother, which are potentially correlated with her health status, such as her income level.[23]

When we perform the analysis on newborns, the model in Equation 1 is modified as follows:

$$Outcome_{iht} = \delta(Treated_h * Post_t) + \tau Treated_h + \gamma_t + \pi_{lha} + \lambda(lha * m) + \alpha_m + X3^{'}_{mht}\rho + X4^{'}_{iht}\beta + \epsilon_{iht} \quad (2)$$

where $X4^{'}_{iht}$ controls for any congenital anomaly suffered by the baby and for any maternal conditions complicating the risk profile of the newborn. By worsening the health status of the baby, these factors may affect the outcome of the delivery itself, that is, the probability of having complications, as well as the need for a breathing intervention. We also estimate alternative specifications using hospital fixed effects to control for time-invariant characteristics at the hospital level (*e.g.*, practice styles) and hospital-month trends. We test the robustness of our preferred specification in several ways, including the use of different levels of clustering, as shown in Section 4.2.

[23]The results in which we add the covariate sets one at the time are provided in Table A.4. It is apparent that the risk profile of the mother – $X1^{'}_{iht}$ – groups the variables that highly predict the event of a c-section.

Our DD identification relies on the exogeneity of the year of schedule introduction with respect to the trend of medical malpractice claims. This is a plausible assumption for several reasons. First, schedules apply to every case of personal injury, from car accidents to workplace injury compensation, and the need to introduce them stemmed from the necessity to help judges in assessing damage awards in road traffic accidents rather than in medical malpractice cases (Busnelli, 1988). Hence, it is very unlikely that a court's decision to implement scheduled damages depends on hospitals operating in the same court district or, more generally, on malpractice claims. Supporting evidence in this respect is provided by Figure A.1. If we divide court districts into quartiles based on the c-section rates, it is apparent that, out of the total number of courts implementing schedules, a similar share falls into each quartile (Figure A.1 (a)). The same holds if courts are assigned to different quartiles based on the distribution of claim frequency (b), and judicial inefficiency as measured by civil backlog (c).[24] Second, there is no possibility of public hospitals engaging in forum shopping. Hospitals always respond to any malpractice cases before the court of the district in which they are located. Finally, since these are public hospitals, there is no room for strategic location. Hospital location is not determined according to court district performance or policies but according to the resident population's needs and their accessibility. In addition, since 1968, the creation of a new hospital has been subject to a population requirement of a minimum of 25,000 inhabitants (Bertoli and Grembi 2017a).

# 4 Results

## 4.1 Effects of Schedules

Tables 4 and 5 show our average results. The likelihood of performing a c-section increases in a statistically significant way. The magnitude of the effect, as shown by the coefficient in Column (3), which is our preferred specification, is a 6.5 percentage-point increase in the use of c-sections, which, at the sample mean, translates to an increase in c-sections of 17%. This incidence is higher than that estimated by Currie and MacLeod (2008) and could depend on several factors. Our effects are estimated in a context of very similar facilities: the nationwide effects are much lower. Tables A.6 and A.7 show the results based on the entire population of deliveries during our observational period. The introduction of schedules produces a 2.2 percentage-point increase in the use of c-sections in our preferred specification in Column (3). This means that the probability of giving birth by c-section after the adoption of schedules increases by 6% at the sample mean of c-sections (*i.e.*, 0.37), which is very close

[24]Following Bertoli and Grembi (2017c), we use $Backlog_{jt}=\frac{New\ Cases_{jt}+Pending\ Cases_{jt}}{Closed\ Cases_{jt}}$ - 1.

to the 5.5% increase observed by Currie and MacLeod (2008). Second, we are estimating only one policy change in our context, whereas the US context is often characterized by the simultaneous implementation of several reforms, which in some cases (*i.e.*, Joint and Several Liability and Caps on Damages) are expected to have opposite effects (see Bertoli and Grembi, 2018). Moreover, we are considering the effects in the very short run. It might be possible that over a longer period of time (when other forces are also in place), the effect is watered down. Finally, as it clearly appears from the heterogeneities analysis, this large effect is mainly driven by certain types of hospitals that face specific incentive schemes and that are realistically expected to deviate the most from an appropriate use of c-sections.

In Figure 3, we test our baseline estimation more rigorously in an event study framework. We modify our preferred specification (Table 4 Column 3) to include monthly leads and the ($Treated_h * Post_t$) dummy, taking January 2001 as the reference month. Plotting the estimated coefficients shows the lack of any anticipatory effects, as we observe a statistically significant increase of almost 5 percentage points only in the post-treatment period. However, the jump in c-sections is not simultaneous with the policy introduction. Once we modify our preferred specification to include both quarter leads and lags and take the first quarter of 2001 as the reference quarter, it becomes apparent that it took some time before the policy generated practical effects. As shown in Figure 4, there is no statistically significant increase during any quarter of the first year of schedules implementation (*i.e.*, 2002), but the policy begins to significantly affect the use of c-sections in quarter 5.

Figures 3 and 4, about here

The results for the secondary outcomes, referring to maternal and neonatal health, are not statistically significant. However, these are somewhat noisy, and thus we cannot rule out a decrease/increase in both maternal and neonatal complications. As cesarean complications increase by 31% at the average mean, breathing interventions decrease by 20%. In Table A.8, we report the confidence intervals in absolute and percentage terms for these outcomes. For example, we cannot rule out that once c-sections increase, c-section complications for newborns decrease to 0.5 percentage points (-9.5%). The event study versions of these outcomes do not reveal any particular trends, as shown in Figure 5.

Tables 4 and 5 and Figure 5, about here

In addition to directly affecting the health outcomes for mothers and newborns, the observed increase in c-sections could imply a change in the level of appropriateness of these procedures. For example, using data from New Jersey between 1997 and 2006, Currie and MacLeod (2017) find that c-section rates were too low for high-risk mothers and too high for low-risk mothers. In the spirit of their paper, we estimate the probability of receiving a c-section as a function of the maternal risk factors and use the resulting estimates to define 10 classes of medical risk (from 0.1 to 1). Figure 6 shows the c-section rates per class of medical risk before and after the policy implementation. This comparison highlights that, while there is an overall increase across all risk classes, those most affected seem to be those between 0.2 and 0.3 and those between 0.4 and 0.5, namely the cases for which the choice of a c-section seems to be less appropriate. We take these results further dividing our observations into three subsamples (*i.e.*, high-, medium-, and low-risk deliveries) with which we separately test our baseline model. As Table 6 makes clear, c-sections increase in all samples. However, while high-risk mothers (5.6% of our population) have a 5.5% higher likelihood of receiving a c-section, this probability increases by 19% for both medium- (6.9% of our population) and low-risk (87.5% of our population) mothers. Overall, the mismatch between patient conditions and the choice of a c-section increases, which further reinforces the argument that the policy has primarily encouraged unnecessary procedures.

Figure 6 and Table 6, about here

Finally, we provide some back-of-the-envelope calculations of the first-order economic implications of the increased use of c-sections.[25] For the sake of simplicity, we assume that the unnecessary c-sections observed would otherwise have been vaginal deliveries without complications. Hence, to quantify the additional expenditures, we consider the DRG prices for both cesarean and vaginal delivery without complications. According to the national DRG list adopted in Italy during our period of interest, the weighted price of a c-section without complications was 2,371 euros, whereas the corresponding price for a vaginal delivery without complications was 956 euros. If we consider the impact of the policy as estimated by the DD, a 17% increase in the use of c-sections corresponds to 4,705 additional cesarean deliveries (*i.e.*, 0.38*72,830=27,675 cesarean deliveries during the period 2001-2004). Overall, these additional procedures cost 11,154,962 euros, while they would have cost 4,497,980 euros as vaginal deliveries. This is equal to an overall waste of 6,656,982 euros, which means 1,415 wasted euros per delivery.

[25]We focus solely on the monetary consequences of c-section overuse. We cannot assess the health implications for women who undergo unnecessary cesarean sections, such as complications during future pregnancies and deliveries. Another monetary consequence of schedules introduction could be related to a change in litigation costs given the discussed effects that the policy has on trends in litigation (Section 2.2.). Still, such a change is very difficult to measure. On the one hand, litigation costs could decrease as out-of-court resolution becomes easier and claim duration could shorten. On the other hand, as observed by Mello and Kachlia (2010), administrative costs might increase as, for example, there are taxes/fees that need to be paid when depositing a litigation agreement.

## 4.2 Validity and Robustness Checks

To defend the robustness of our results on cesarean deliveries, we perform several tests and estimate alternative specifications. First, we check whether there was any change in the number of deliveries managed by treated hospitals after the adoption of schedules, as the increase in c-sections could have been driven by more managed deliveries. We also check whether our results could have been driven by a change in the composition of the treated patients: even if we prove that there is no statistically significant change in the number of deliveries, there could have been a change in their composition. Second, we test the adoption of schedules on a different sample of hospitals. This falsification test considers the hospitals that belong to the same regions in which the hospitals in our preferred samples are located but never experienced a change in schedules adoption status between 2001 and 2004. Finally, we control for court fixed effects and we provide the results clustering at a more aggregate level than hospitals, to determine how more conservative clustering affects our results.

In Table 7, we report the estimate of $\delta$ on a sample of monthly data at the hospital level, where the outcome is the number of monthly hospital deliveries. Regardless of the specification, there is no statistically significant effect of schedules adoption on the number of managed deliveries. To rule out any reactions on the patients' side, we need to assess more than just the number of deliveries. The increase in the certainty of compensation could have attracted riskier patients who decided to deliver in the treated hospitals rather than in a nearby control hospital, because in the event of an adverse outcome, they knew what could be expected in terms of non-economic compensation. According to this explanation, we could detect more cesarean sections as a consequence of a change in the risk profile of deliveries. To check for patient selection, we proceed in two ways. We estimate Equation 1 using two dummies as outcomes of interest: $Low-risk$ mothers and $Low-weight$ newborns. $Low-risk$ mothers equals one if the mother does not present any pre-delivery risk conditions,[26] while $Low-weight$ newborns equals one if the newborn was underweight for her gestational age. Then, we checked whether the policy affects the predicted probability of receiving a c-section ($PPC_{iht}$) as in Baiker et *al.* (2006). If schedules adoption modifies the $PPC$ as a function of maternal risk factors at the hospital level, then it means that there is a change in patient composition.

Table 8 shows that there is no change in the probability of treating a low-risk mother, in the predicted probability of receiving a c-section as a function of maternal risk factors, or

[26] According to the medical literature, a low-risk mother is any woman who does not present any of the following pre-delivery risk conditions: fetus malposition, previous c-section, diabetes, prolonged pregnancy, early labor, poor or excessive fetal growth, multiple gestation, fetal abnormality, antepartum hemorrhage, placenta previa, pre-eclampsia, eclampsia, toxemia, hypertension, polyhydramnios, oligohydramnios, and infection of the amniotic cavity.

in the probability of facing a low-weight newborn for treated hospitals after the adoption of schedules. In fact, even though the results would point in the direction of a 4.4% decrease in low-risk mothers, a 1.6% increase in $PPC$ and a 13.6% increase in low-weight babies, none of these findings is statistically significant with our without controls. Hence, it is difficult to argue that patient selection explains our effect.

Tables 7 and 8, about here

We then estimate Equation 1 considering only hospitals overseen by courts not switching to schedules in 2002. Our analysis relies on the assumption that schedules are the only relevant policy affecting medical liability in the observational period. If this is the case, we should not detect any changes in the trends of cesarean sections for hospitals governed by courts that did not switch to schedules in 2002. We consider all the hospitals located in the same regions of our main sample. The treated are now hospitals operating in a court district that *always* applied schedules, while the control hospitals operate in court districts that *never* used schedules. As is apparent from Table 9, when we assess the effect of a false adoption of schedules in 2002, no effect is observed. In Online Appendix B, we estimate all the main specifications and validity tests without dropping the weekend deliveries. The results confirm our baseline effects, although, as expected, their magnitude is lower.

Finally, we test the results on c-sections by adding the court fixed effects and clustering the standard errors at levels different from hospitals, such as the LHA level and the court level. The idea of using court fixed effects is to control for the time-invariant characteristics that might have affected the decision to adopt schedules in the first place, such as the political backgrounds of judges belonging to the same court, among others. Schedules are believed to enhance the sense of social justice. Table 10 provides the results. All levels of clustering considered may again raise concerns about the potential problems associated with the low number of clusters. For this reason, we also apply the correction for the small number of clusters by using wild score bootstrapping.

Tables 9 and 10

## 4.3 Who Performs More Cesarean Sections?

The assessment of the average effect of a decrease in medical liability pressure provides unclear policy implications *per se*. Tort reforms do not specifically target medical injuries but refer to all types of personal injuries, as in our case. The richness of our dataset allows us to identify the drivers of hospital response and, thus, provide more meaningful policy implications. In particular, we focus on three dimensions that are expected to play a role: hospital quality, other sources of liability pressure, and financial incentives.

Other things being equal, the increase in medically unnecessary c-sections is expected to be lower in high-quality hospitals than in low-quality ones. High quality denotes appropriateness, effectiveness, and a critical mass of highly skilled doctors. In high-quality facilities, doctors should have less room for strategic behavior, and medical decisions should be less affected by factors other than patients' medical conditions. We proxy for hospital quality with the number of yearly deliveries, following the strand of literature that matches high volumes of a procedure to better quality, due to a learning-by-doing process (Nueld Institute for Health, 1996; Kizer, 2003; Sound, 2010; Kristensen *et al.*, 2014; Avdic *et al.*, 2019).[27] The higher the number of mothers giving birth in a hospital, the better the hospital should be at coping with both deliveries and their unexpected consequences, meaning that there is a better match between the type of delivery and the type of patient. We define high/low volumes following a 2006 Italian law, which established 500 deliveries per year as the minimum number of procedures that a birth center should manage to be authorized to operate by the National Health System.

Second, if we demonstrate that there is a change in the liability pressure, it is important to understand how this interacts with other factors affecting medical liability pressure. We identify two main factors: (1) the level of civil backlog of the court district in which a hospital is located, which proxies for inefficiency in the judiciary system, and (2) the actual level of schedules, as proxied by the monetary value of a 25% disability.

Italian courts rank poorly in world rankings of judiciary performance, with a significant degree of within-country heterogeneity. One common parameter for these rankings is the average duration of a first-instance decision. As the average length of civil trials is highly correlated with the level of backlog, we use this latter measure to proxy judicial performance. The level of backlog measures the pending caseload of a court in a given year against the court's actual capacity to dispose of cases during the same year.[28] When the backlog index equals 0, all the entering and pending cases at time $t$ are solved, which is our benchmark for efficiency. A higher level of court inefficiency in settling civil cases might capture a lower level of deterrence for both physicians and hospitals. Hence, we expect more use of c-sections when courts are inefficient. We take the mean value of civil backlog and define courts with above-mean values of backlog as more inefficient.

We provide two measures of liability pressure related to the level of schedules. Courts

[27]Buzzacchi *et al.* (2016) also provide descriptive evidence on the positive impact of volume on malpractice claims.

[28]Information on the duration of medical malpractice cases is not available. There is no available information on the number of pending malpractice cases per court.

construct their schedules; thus, different monetary values can be assigned to given percentages of disability. The introduction of low schedules allows healthcare providers to enjoy the same increase in certainty on what they have to pay, but overall, they need to pay less than those governed by a high-schedule court.[29] We define high (low) schedules as those above (below) the mean value of scheduled damages for a disability of 25% and we use two ages for the weights: 40 and 60 years old. As a consequence, we expect to observe a higher increase in c-sections in hospitals subjected to low schedules.

Finally, financial incentives are also expected to interact with the adoption of schedules. Doctors and hospitals respond to monetary incentives (Gruber *et al.*, 1999; Grant, 2009; Cavalieri *et al.*, 2014; Johnson and Rehavi, 2016). Consistent with the literature, we expect that the greater the difference in the reimbursement between a cesarean and a vaginal delivery, the greater the incentive to perform a c-section would be. We define a high (low) difference in the reimbursement as a difference above (below) the mean value of DRG prices. DRG tariffs in Italy differ across and within regions whenever regions decide to apply different reimbursements to adjust for differences across hospitals (Bertoli and Grembi 2017b).

We generate dummies, $D$, for each channel and interact them with $Treated_h * Post_t$. For each driver, we report the results for $Treated_h * Post_t$ in each subsample defined by $D$ and the significance of the difference between the two samples.[30] The results of this analysis are reported in Table 11. As expected, schedules trigger more medically unnecessary cesarean sections in hospitals with the following characteristics: lower quality standards due to smaller size, (2) lower perceived deterrence from the judiciary system due to inefficiency, (3) lower levels of schedules, and (4) higher financial incentives. We also plot the pre-trends for $\lambda$ in Figure A.2 using quarter leads and the last quarter of 2001 as the reference quarter.

Table 11, about here

# 5 Concluding Remarks

This paper makes an important contribution by highlighting that optimal regulation is likely to be a function of a jurisdiction's legal system. A problem with tort reforms, especially in

[29]These differences have been considered unfair to potential victims. This is why a new regulation was implemented in 2016 to promote the adoption of a national schedule for personal injuries.

[30]The significance of the difference between the coefficients of $Treated_h * Post_t$ in the two subsamples is the parameter $\lambda$ in the following model:

$$Csection_{iht} = \delta(Treated_h * Post_t) + \rho(Treated_h * Post_t * D) + \alpha D + \gamma_t + \omega_h + X1'_{iht}\sigma + X2'_{iht}\beta + X3'_{mht}\tau + \pi_{lha} + \lambda(lha * m) + \alpha_m + \epsilon_{iht}$$

where $D$ is the dummy for each channel. For instance, for the quality channel, $D$= 1 if the hospital performs more than 500 deliveries per year.

the most studied context of the US, is that they have been repealed on several occasions soon after their introduction and that reforms that are expected to produce different effects, such as joint and several liability and caps on damages, are often jointly introduced (Bertoli and Grembi, 2018). This is why the empirical evaluation we provide is unique. By exploiting the implementation of schedules at the court district level in Italy, we study whether a reduction in medical liability affects the use of delivery methods in the context of childbirth.

During our period of observation, 2001-2004, 18 courts implemented schedules in 2002. Since hospital location determines the court in which a claim must be filed, medical liability decreases only for those hospitals overseen by courts that switched to schedules. Applying a DD estimation, we show that hospitals react to a decrease in medical liability by increasing the use of the more intensive and less appropriate treatments. In particular, schedules incentivize the performance of c-sections, which increase by 17% after their introduction. Neither maternal nor newborn health turn out to be statistically affected by these additional c-sections, which can therefore be classified as unnecessary procedures. Based on the weighted DRG prices of a c-section and vaginal delivery without complications during our observational period, back-of-the-envelope calculations suggest that such an increase in c-sections would have led to an overall increased expenditure of 6,656,982 euros (*i.e.*, 1,415 additional euros per delivery).

Our results support the findings in Currie and MacLeod (2008) by confirming that, in a context of c-section overuse, a decrease in liability does not discourage c-sections; instead, it may provide doctors with additional incentives to perform them. The role played by these additional incentives appears to be even stronger in Italy than in the US. We provide an in-depth perspective on this point through the analysis of the channels which reveals that the opportunistic behavior of doctors also depends on the characteristics of the hospitals and of the environment in which they operate. In particular, the hospitals that are more prone to respond to lower liability by increasing the number of c-sections are low-quality hospitals, hospitals subject to a more inefficient court, those facing lower levels of schedules, and those able to benefit more from the financial incentives of performing a c-section.

# References


Abadie, A., Athey, S., Imbens, G., and J. Wooldridge (2017). When You Should Adjust Standard Errors for Clustering?, NBER Working Paper Series WP 24003.

Avdic, D., Lundborg, D. and J. Vikström (2019). Estimating Returns to Hospital Volume: Evidence from Advanced Cancer Surgery. Journal of Health Economics, 63: 81-99.

Amaral-Garcia, S. and V. Grembi (2014). Curb Your Premium! Evaluating State Intervention in Medical Malpractice Insurance, Health Policy, 114: 139-146.

Amaral-Garcia, S., Bertoli, P., and V. Grembi (2015). Does Experience Rating Improve Obstetric Practices? Evidence From Geographical Discontinuities. Health Economics, 24: 1050-1064.

Associazione Nazionale fra le Imprese Assicuratrici (ANIA).(2014). Malpractice il grande caos. available at `http://www.ania.it/export/sites/default/it/pubblicazioni/Dossier-e-position-paper/Dossier-ANIA-Malpractice-il-grande-caos.pdf`.

Avraham, R. (2006). Putting a price on pain-and-suffering damages: A critique of the current approaches and a preliminary proposal for change. Northwestern University Law Review, 100: 05-23.

Baicker, K., Buckles, K. and A. Chandra (2006). Geographic Variation in the Appropriate Use of Cesarean Delivery. Health Affairs 25:355.

Baldwin, L., Hart, L. G., Lloyd, M., Fordyce, M. and R. Rosenblatt (1995). Defensive Medicine and Obstetrics. Journal of the American Medical , 274(20): 1606-1610.

Bertoli, P. (2014). Medical Malpractice in Public Health Care System: An Empirical Investigation of Scheduled Damages. PhD Dissertation EDLE program.

Bertoli, P. and V. Grembi (2017a). The Life Saving Effect of Hospital Proximity. Health Economics, 26 Suppl 2: 78-91.

Bertoli, P. and V. Grembi (2017b). The Political Economy of the Diagnosis-Related Groups. Social Science & Medicine, 190: 38-47.

Bertoli, P. and V. Grembi (2017c). Courts, Scheduled Damages, and Medical Malpractice Insurance. Empirical Economics, DOI 10.1007/s00181-017-1279-5.AGGIORNARE

Bertoli, P. and V. Grembi (2018). Medical Malpractice: How Legal Liability Affects Medical Decisions. Chapter in *Health Econometrics* to appear in B. Baltagi and F. Moscone "Contributions to Economic Analysis". Emerald Publishing

Busnelli, F. D. (1988). Danno biologico e danno alla salute, in "La valutazione del danno alla salute", edited by Bargagna e Busnelli, CEDAM, Padua.

Buzzacchi, L., Scellato, G. and E. Ughetto (2016). Frequency of medical malpractice claims: The effects of volumes and specialties. Social Science& Medicine, 170:152-160.

Chandra, A. and D. O. Staiger (2016). Identifying Sources of Inefficiency in Healthcare. American Economic Review: Papers & Proceedings, 106(5): 383-387.

Comandè Giovanni (2005). Towards a Global Model for Adjudicating Personal Injury Damages: Bridging Europe and the United States, 19 Temp. Intl & Comp. L.J. 241: 241-349.

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

Currie, J. and W. B. MacLeod (2017). Diagnosis Expertise: Human Capital, Decision Making, and Performance among Physicians. Journal of Labor Economics, 35: 1-43.

Danzon, P. (2000). Liability for Medical Malpractice. Chapter 26 of A. Culyer and J. Newhouse (eds.) Handbook of Health Economics, 1: 1339-1404.

Dranove, D., Subramaniam R., and A. Sfekas (2011). Does the market punish aggressive experts? evidence from cesarean sections. The BE Journal of Economic Analysis & Policy 11 (2).

Dranove, D., and Y. Watanabe (2009). Influence and Deterrence: How Obstetricians Respond to Litigation against Themselves and Their Colleagues. American Law and Economics Review 12: 69-94.

Dubay, L., Kaestner, R., and T. Waidmann (1999). The Impact of Malpractice Fears on Cesarean Section Rates. Journal of Health Economics 18: 491-522.

Dubay, L., Kaestner, R., and T. Waidmann (2001). Medical Malpractice Liability and Its Effect on Prenatal Care Utilization and Infant Health. Journal of Health Economics 20: 591-611.

Ellis, R. P., Gallup, C.L. and T. G. McGuire (1990). Should medical professional liability insurance be experience rated? Journal of Risk and Insurance, pp. 66-78.

Eurobarometer (2006). Eurobarometer on Medical Error. Special Eurobarometer 241, available at `http://ec.europa.eu/health/ph_information/documents/eb_64_en.pdf`.

Eurobarometer (2009). Patient Safety and Quality of Healthcare. Special Eurobarometer 327, available at `http://ec.europa.eu/public_opinion/archives/ebs/ebs_327_en.pdf`.

Fenn, P., Gray, A., and N. Rickman. 2007. Liability, Insurance, and Medical Practice. Journal of Health Economics 26: 1057-1070.

Fenn, P., Gray, A., Rickman, N., Vencappa, D., and O. Rivero (2013). The impact of risk management standards on patient safety: the determinants of MRSA infections in acute NHS hospitals, 2001-2008.Oxford Bulletin of Economics and Statistics, 75(3): 340-361.

Frakes, M. (2012). Defensive Medicine and Obstetric Practice. Journal of Empirical Legal Studies 9: 457-481.

Frakes, M. (2013). The Impact of Medical Liability Standards on Regional Variations in Physician Behavior: Evidence from the Adoption of National-Standard Rules. American Economic Review, 103(1): 257-276.

Frakes, M. (2015). The Surprising Relevance of Medical Malpractice Law. University of Chicago Law Review, 82(1): 317 - 391.

Frakes, M. and J. Gruber (2019). Defensive Medicine: Evidence From Military Immunity. American Economic Journal: Economic Policy (forthcoming).

Frakes, M. and A. Jena (2016). Does Medical Malpractice Law Improve Health Care Quality? Journal of Public Economics, 143: 142-158

Fournier, G. M. and M. M. McInnes (2001). The case for experience rating in medical malpractice insurance: an empirical evaluation. Journal of Risk and Insurance, 68(2): 255-276.

Iizuka, T. (2013). Does Higher Malpractice Pressure Deter Medical Errors? Journal of Law and Economics 56: 161-188.

Kachalia, A., and Mello, M. M. (2011). New Directions in Medical Liability Reform. New England Journal of Medicine 364: 1564-1572.

Kessler, D. and M. McClellan (2002). Malpractice law and health care reform: optimal liability policy in an era of managed care. Journal of Public Economics, 84: 175-197.

Kizer, K.W. (2003). The Volume-Outcome Conundrum, The New England Journal of Medicine; 349: 2159-2161.

Kristensen, P.K., Thillemann, T.M. and S.P. Johnsen (2014). Is bigger always better? A nationwide study of hip fracture unit volume, 30-day mortality, quality of in-hospital care, and length of hospital stay. Medical Care, 52(12): 1023-9.

Johnson, E. M. and M. M. Rehavi. 2016. Physicians Treating Physicians: Information and Incentives in Childbirth. American Economic Journal: Economic Policy, 8(1): 115-141.

Jena, A. B. , Seabury, S., Lakdawalla, D., and A. Chandra. (2011) Malpractice Risk According to Physician Specialty. New England Journal of Medicine, 365(7): 629-636.

Jones, A. M. (2012) (ed.) The Elgar Companion to Health Economics. Second Edition. Aldershot: Edward Elgar.

Lakdawalla, D. and S. Seabury (2012). The welfare effects of medical malpractice liability. International Review of Law and Economics, 32(4): 356-369.

Localio, A. R., Lawthers, A., Begtson, J., Hebert, L., Weaver, S., Brennan, T., and J. R. Landis (1993). Relationship Between Malpractice Claims and Cesarean Delivery. Journal of the American Medical Association 269: 366.

Mapelli, V. (2012). Il sistema sanitario italiano, Il Mulino, Bologna.

Meloni, A., Loddo A., Martsidis K., Deiana S.F., Porru D., Antonelli A., Marongiu D., Piras B., Paoletti A.M. and G.B. Melis (2012). The role of cesarean section in modern Obstetrics. Journal of Paediatric and Neonatal Individualized Medicine 1:53-58.

Mello, M. M. (2006). Understanding medical malpractice insurance: A primer, Princeton, NJ, The Robert Wood Johnson Foundation, Research Synthesis Report n. 8.

Mello, M. M. and A. Kachalia (2010). Evaluation of Options for Medical Malpractice System Reform. A Report to the Medicare Payment Advisory Commission (MedPAC) 2010 January 29.

Ministero della Salute. 2014. Certificato di assistenza al parto (CeDAP): Analisi dell'evento nascita, Anno 2013. (Available from: www.salute.gov.it/imgs/C_17_pubblicazioni_2431_allegato.pdf) [accessed 5 May 2016].

Ministero della Salute (2013). Certificato di assistenza al parto (CeDAP), Analisi dell'evento nascita - Anno 2010.

Nys, H. (2009). The Factual Situation of Medical Liability in the Member States of the Council of Europe, Reports from the rapporteurs, Conference 'The ever-growing challenge of medical liability: national and European responses', Strasbourg 2-3 June 2008, pp. 17-28.

OECD (2013), Health at a Glance 2013: OECD Indicators, OECD Publishing.

Paik, M., Black, B., Hyman, D. and C. Silver (2012). Will Tort Reform Bend the Cost Curve? Evidence from Texas. Journal of Empirical Legal Studies, 9: 173-216.

Ronzoni, E. (2012). Malasanitá in Italia, in Sanitá pubblica e assicurazioni, Cedam, Padova.

Sella, M. (2005). I sistemi tabellari, available at: http://www.personaedanno.it.

Shurtz, I. (2013). The impact of medical errors on physician behavior: Evidence from malpractice litigation. Journal of Health Economics 32: 331-340.

Shurtz, I. (2014). Malpractice Law, Physicians' Financial Incentives, and Medical Treatment: How Do They Interact? Journal of Law and Economics 57: 1-29.

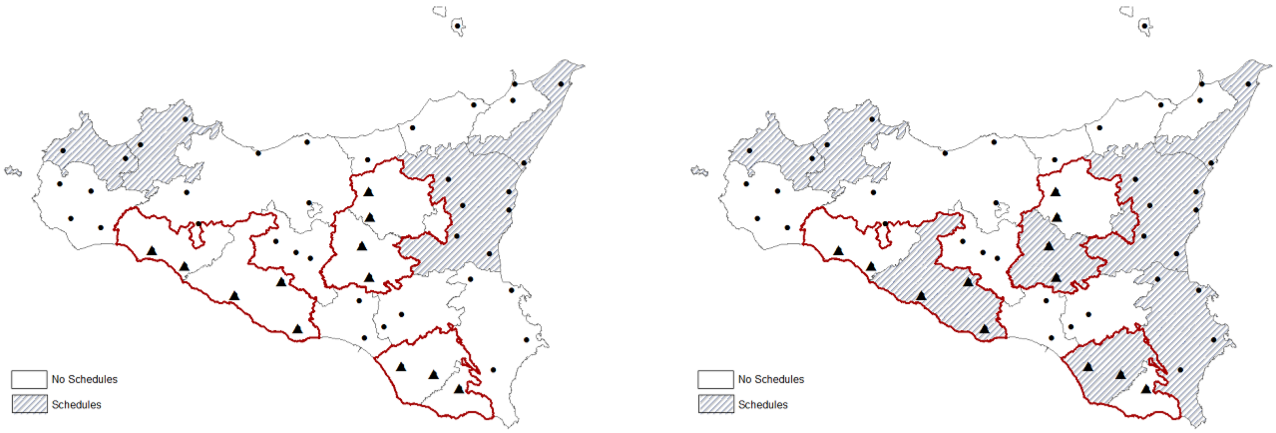
Sloan, F. A. (1990). Experience Rating: Does it Make Sense for Medical Malpractice Insurance? American Economic Review 80(2): 128-133.

Sloan, F. A., Entman, S. S., Reilly, B. A., Glass, C. A., Hickson, G. B., and H. H. Zhang (1997). Tort Liability and Obstetricians' Care Level. International Review of Law and Economics 17: 245-260.

Weiler, P. C., Hiatt, H. H., Newhouse, J. P., Johnson, W.G., Brennan, T.A., and Leape, L. L. (1993). A Measure of Malpractice: Medical Injury, Malpractice Litigation and Patient Compensation. Cambridge, Harvard University Press.

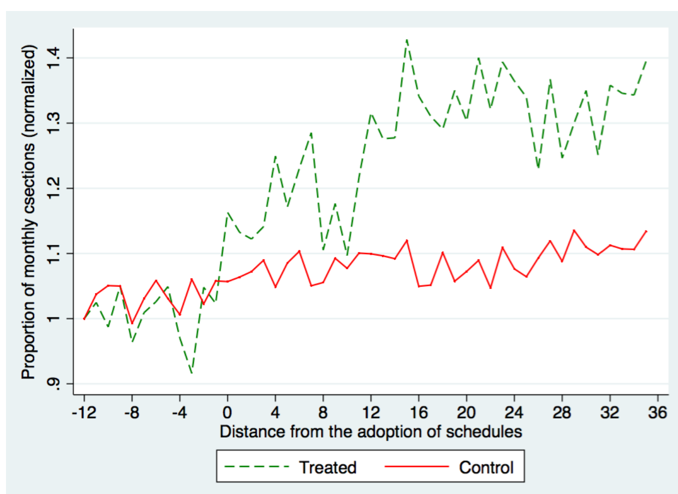
# Tables and figures

Figure 1: **Example of Schedules adoption and Hospital Distribution (2001-2004)**




Notes: The figure presents the region of Sicily as a representative example. Court districts' borders are in black. White areas identify court districts that do not apply schedules of non-economic damages. Grey striped areas identify court districts that apply schedules. Black dots and triangles represent public hospitals. Triangles are the unit used in the main dataset. Dots are the unit used in the falsification test. Thicker borders identify the borders of the LHAs governed by two different courts, one with and one without scheduled damages. On the left, we plot the figure with 2001 data, and on the right, we plot the figure with 2002 data.

Figure 2: **Monthly C-sections (Ratio)**




Notes: The figure shows the average cesarean section rate per month for treated and control hospitals. Month 0 represents the month in which schedules were adopted. The values of the series are normalized by the cesarean section rate at month -12.

Table 1: **Schedule and Malpractice Insurance Claim Outcomes**

| | **Frequency** | **Drop** | **Reserves** | **Var** | **Number of Insurers** | **New Claims** | **Compensation** |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Schedules | -0.011*** | -0.005*** | -132.077*** | -61,655.816*** | 0.165*** | -0.002* | -132.470*** |
| | (0.000) | (0.000) | (30.404) | (15,344.910) | (0.000) | (0.001) | (10.864) |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Provider FE | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Court FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Obs. | 99 | 99 | 83 | 78 | 30 | 92 | 49 |
| Mean | 0.188 | 0.083 | 248.983 | 259,681.700 | 0.500 | 0.007 | 123.885 |

Notes: For columns (1)-(4), we use the information collected on 4,422 malpractice insurance claims filed with commercial insurers in Italy and that closed between 2000 and 2004. *Schedules* is equal to 1 if the court district in which the hospital is located implemented schedules in 2002. All outcomes are expressed at the hospital level and normalized by the total number of beds in the hospital. *Frequency* represents the annual number of claims per hospital. *Drop* identifies the annual number of claims dropped before the end of the proceeding. *Reserves* represents the average reserve paid at the hospital level. *Reserves* and *Compensations* are in 2015 euros. For column (5), we use the information collected on the 30 bidding processes for malpractice insurance that were run between 2001 and 2004 by the hospitals included in our LHA sample. *Number of Insurers* represents the number of bidders in the tendering process. In model (5), we also control for LHA fixed effects, the type of awarding criterion and the presence of a broker. We also include the overall population, the population older than 65, the foreign population and the average income at the LHA level. In model (6), we use information on the courts belonging to our LHA sample between 2001 and 2004. *New Claims* is the ratio between the annual new claims received by a court and the resident population in the court district. Robust standard errors are in parentheses. These are clustered at the hospital level for models (1)-(5) and at the court level for model (6). Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 2: **Balance of the Observables**

| | Control | | | Treated | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Obs. | Mean | Std. Dev. | Obs. | Mean | Std. Dev. | Diff. | T-test | p-value |
| **Panel A: All Hospitals** | | | | | | | | | |
| *Nurse* | 1,836 | 398 | 428 | 349 | 259 | 201 | 139 | 5.930 | 0.000 |
| *Doctors* | 1,836 | 171 | 186 | 349 | 107 | 90 | 64 | 06.261 | .000 |
| *Personnel* | 1,836 | 913 | 1041 | 349 | 580 | 508 | 333 | 5.839 | 0.000 |
| *Beds* | 1,836 | 330 | 329 | 349 | 224 | 162 | 106 | 5.894 | 0.000 |
| *Used Beds* | 1,836 | 312 | 313 | 349 | 215 | 158 | 98 | 5.706 | 0.000 |
| *Wards* | 1,836 | 20 | 19 | 349 | 14 | 10 | 6 | 5.668 | 0.000 |
| *Used Wards* | 1,836 | 18 | 17 | 349 | 12 | 8 | 6 | 5.924 | 0.000 |
| *Discharges* | 1,836 | 12,303 | 12,051 | 349 | 9002 | 6,278 | 3,301 | 4.889 | 0.000 |
| *Case-Mix Index* | 1,836 | 0.928 | 0.139 | 349 | 0.056 | 0.108 | 0.072 | 9.229 | 0.000 |
| *Entropy Index* | 1,836 | 2.135 | 0.155 | 349 | 2.105 | 0.146 | 0.030 | 3.350 | 0.001 |
| **Panel B: Hospitals within the LHA Sample** | | | | | | | | | |
| *Nurse* | 76 | 281 | 269 | 88 | 271 | 219 | 9 | 0.243 | 0.809 |
| *Doctors* | 76 | 131 | 126 | 88 | 122 | 114 | 9 | 0.489 | 0.625 |
| *Personnel* | 76 | 667 | 689 | 88 | 621 | 594 | 46 | 0.459 | 0.647 |
| *Beds* | 76 | 240 | 222 | 88 | 246 | 189 | -6 | -0.200 | 0.842 |
| *Used Beds* | 76 | 231 | 220 | 88 | 236 | 180 | -5 | -0.173 | 0.863 |
| *Wards* | 76 | 17 | 16 | 88 | 15 | 13 | 1 | 0.587 | 0.558 |
| *Used Wards* | 76 | 15 | 16 | 88 | 14 | 12 | 0.8 | 0.340 | 0.734 |
| *Discharges* | 76 | 9,637 | 8,098 | 88 | 10,304 | 6,602 | -666 | -0.580 | 0.563 |
| *Case-Mix Index* | 76 | 0.877 | 0.164 | 77 | 0.864 | 0.138 | 0.013 | 0.532 | 0.595 |
| *Entropy Index* | 76 | 2.113 | 0.126 | 77 | 2.138 | 0.099 | -0.026 | -1.402 | 0.163 |

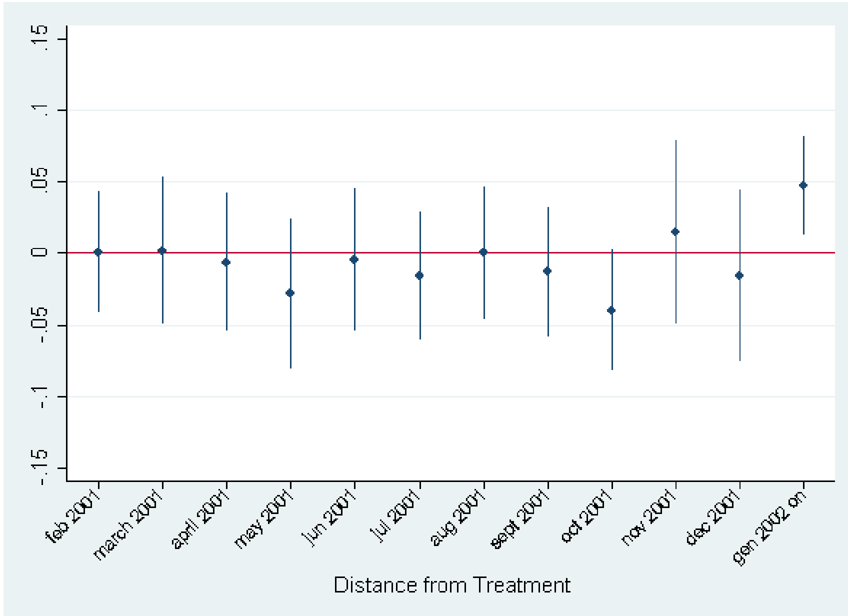
Notes: Here we perform t-tests because data at the hospital level are available only annually (2001-2004). *Case-Mix Index* reflects the diversity and disease severity in the overall population of patients in a given hospital. *Entropy Index* expresses the case-mix dispersion in a given hospital, thus the higher this index, the more dispersed the care provided by a given hospital.

Table 3: **Outcomes and Controls**

| | **Outcomes** | | **Controls** | | | |
|---|---|---|---|---|---|---|
| **Delivery** | **Mother Complications** | **Newborn Complications** | **X1**[a] | **X2**[a] | **X3**[b] | **X4** |
| | **Preventable**[a] | | Mother Level | Mother Level | Municipality Level | Newborn Level |
| C-section | Maternal fever<br>Perineal laceration<br>Perineal Infection<br>Uterus laceration<br>Infection<br>Prolonged labor<br>Anesthetic compl<br>Hemorrhage<br>Embolism<br>Cardiac compl<br>Hysterectomy<br>Maternal seizure<br>Obstructed labor<br>Other | Cesarean Complications<br>Vag Complications<br>Breathing interventions | Cervix<br>Multiple babies<br>Diabetes<br>Placenta<br>Eclampsia<br>Breech<br>Cord prolapse<br>Cancer<br>Use of alcohol<br>Early-onset delivery<br>Use of drugs<br>Hypertension<br>Lung conditions<br>Cardiac conditions<br>Obesity<br>Anemia<br>Renal failure<br>RH Problems<br>Sexually transmitted diseases<br>Previous Csec | Age<br>Nationality<br>Marital status | Average Income<br>Sea level<br>Level of urbanization<br>Education | Congenital Anomalies<br>Maternal Conditions |

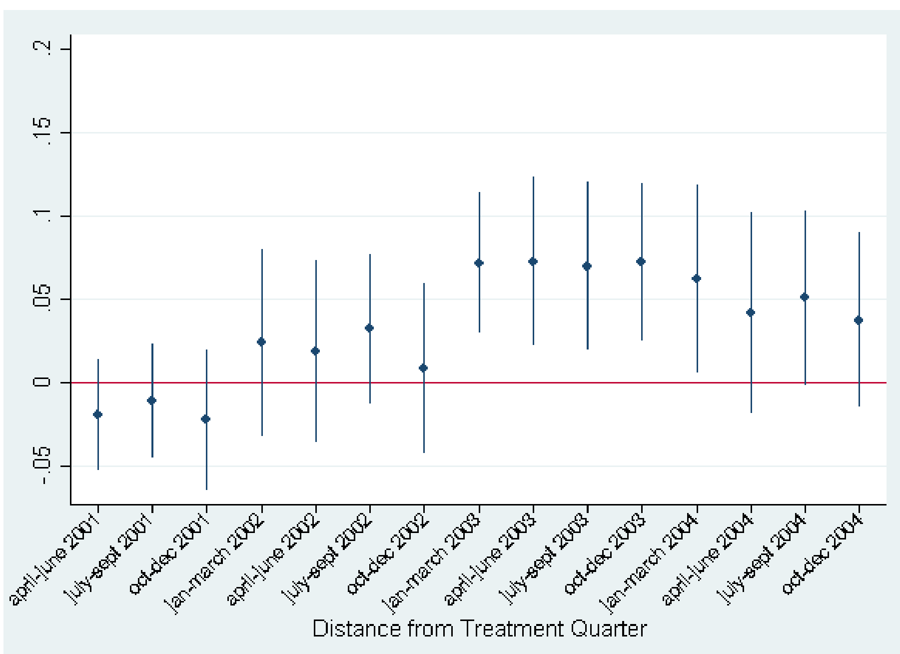
Note: (a) Both complications and controls for clinical risk factors are consistent with those used in Dubay *et al.* (1999), Dubay *et al.* (2001), Currie and MacLeod (2008 and 2013), Dranove and Watanabe (2009), Dranove *et al.* (2011), and Shurtz (2013 and 2014). Among the preventable complications, *Other* includes possible complications that differ from those explicitly mentioned, such as retained placenta and perineal hematomas among others. (b) The controls used at the mothers' municipality level are consistent with those used at the county level by Baicker *et al.* (2006). We also include *Sea level* due to the territorial characteristics of the country.

Figure 3: **Anticipatory Effects (Leads) and the Treatment**




Notes: The figure plots the monthly leads, with a confidence interval of 95%. January 2002 is the month in which schedules were adopted. The reference month is January 2001.

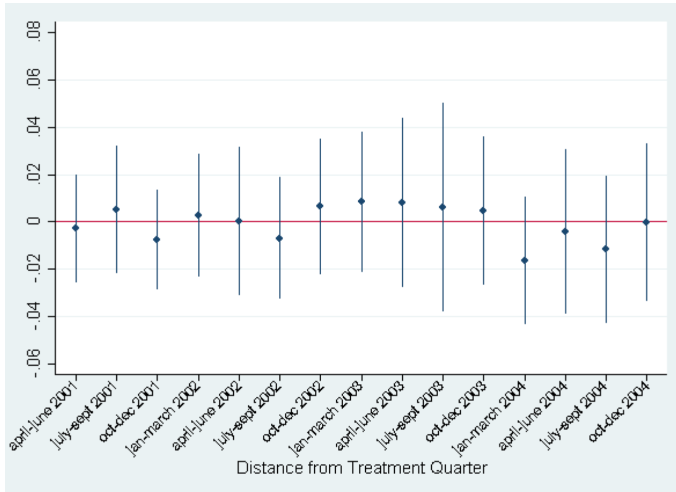
Figure 4: **C-sections: Leads and Lags**



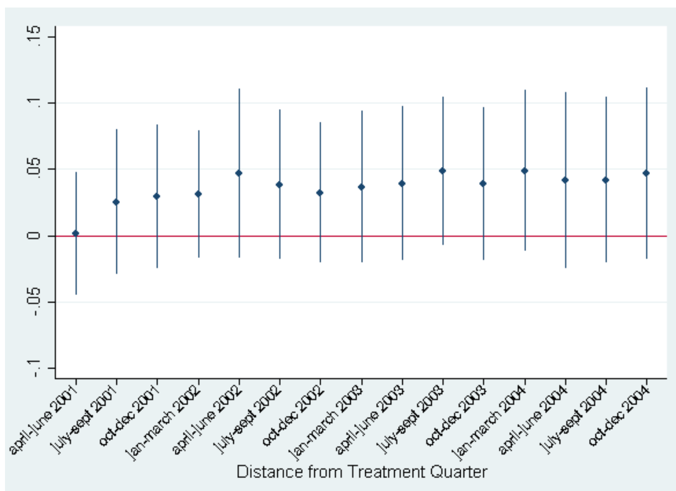

Notes: The figure plots the leads and lags by quarter, with a confidence interval of 95%. Jan-March 2002 is the quarter in which schedules were adopted. The reference quarter is Jan-March 2001.

Figure 5: **Secondary Outcomes: Leads and Lags**

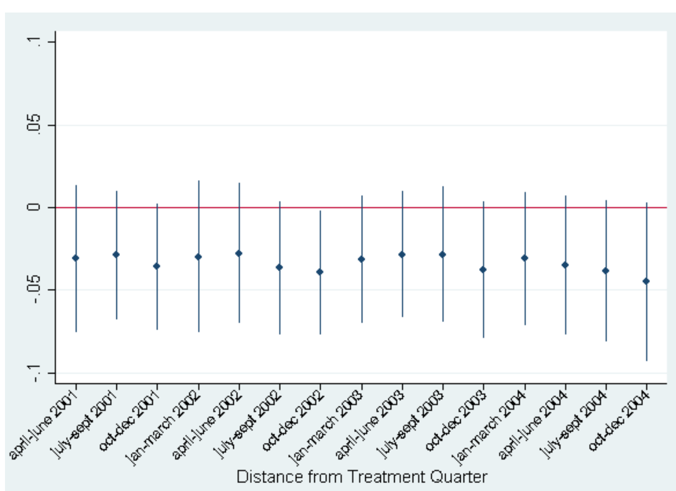
(a) Preventable complications




(b) C-section complications




(c) Breathing Interventions




Notes: The figure plots the leads and lags by quarter, with a confidence interval of 95%. Jan-March 2002 is the quarter in which schedules were adopted. The reference quarter is Jan-March 2001.

Table 4: **DD - Mothers Main Results**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Panel A: C-sections** | | | | | |
| $\delta$ | 0.079*** | 0.072*** | 0.065*** | 0.074*** | 0.065*** |
| | (0.026) | (0.000) | (0.000) | (0.024) | (0.021) |
| Obs. | 72,830 | 72,830 | 72,830 | 72,830 | 72,830 |
| Mean | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 |
| **Panel B: Preventable Complications** | | | | | |
| $\delta$ | 0.002 | 0.002 | 0.000 | 0.003 | 0.000 |
| | (0.016) | (0.015) | (0.056) | (0.011) | (0.004) |
| Obs. | 72,830 | 72,830 | 72,830 | 72,830 | 72,830 |
| Mean | 0.083 | 0.083 | 0.083 | 0.083 | 0.083 |
| Hospital FE | | ✓ | ✓ | ✓ | ✓ |
| Treated | ✓ | | | | |
| Year FE | ✓ | ✓ | ✓ | | |
| Month FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hosp-month trend | | | | ✓ | ✓ |
| Post | | | | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | | ✓ | | ✓ |

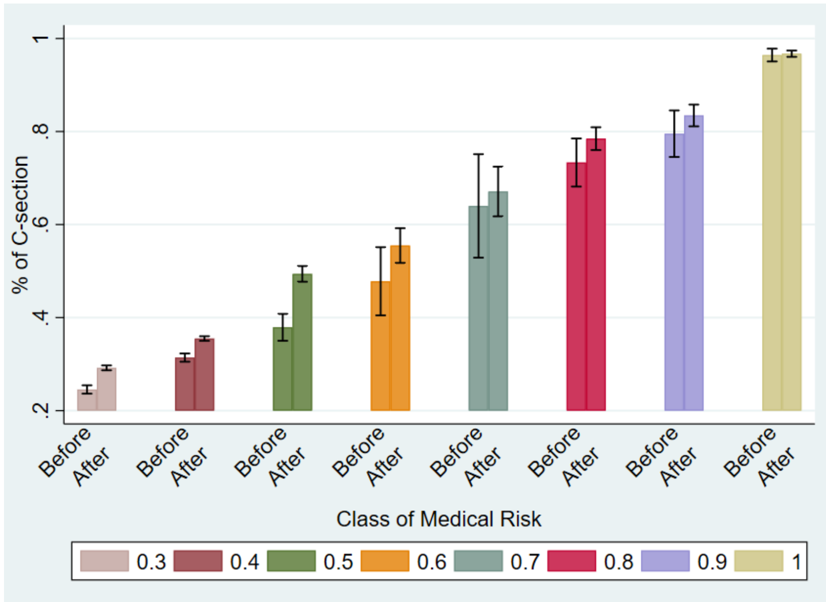
Notes: Controls include $X1$, $X2$ and $X3$ when outcomes refer to mothers, as listed in Table 3. Coefficients are estimated with a linear probability model. Standard errors are clustered at the hospital level in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 5: **DD - Newborns Main Results**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Panel A: C-sections Complications** | | | | | |
| $\delta$ | 0.020 | 0.021 | 0.018 | 0.019 | 0.017 |
| | (0.015) | (0.015) | (0.012) | (0.015) | (0.013) |
| Obs. | 73,717 | 73,717 | 73,717 | 73,717 | 73,717 |
| Mean | 0.058 | 0.058 | 0.058 | 0.058 | 0.058 |
| **Panel B: Breathing Intervention** | | | | | |
| | | | | | 6 |
| $\delta$ | -0.003 | -0.003 | -0.004 | -0.003 | -0.004 |
| | (0.006) | (0.006) | (0.006) | (0.007) | (0.006) |
| Obs. | 73,717 | 73,717 | 73,717 | 73,717 | 73,717 |
| Mean | 0.020 | 0.020 | 0.020 | 0.020 | 0.020 |
| Hospital FE | | ✓ | ✓ | ✓ | ✓ |
| Treated | ✓ | | | | |
| Year FE | ✓ | ✓ | ✓ | | |
| Month FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hosp-month trend | | | | ✓ | ✓ |
| Post | | | | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | | ✓ | | ✓ |

Notes: Controls include $X3$ and $X4$, as listed in Table 3. Coefficients are estimated with a linear probability model. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure 6: **Patient Risk Profile and the Use of C-section**




Notes: The figure plots the c-section rates per class of medical risk before and after the introduction of schedules in 2002.

Table 6: **DD - C-sections by Risk Classes**

| | **High risk** | **Medium risk** | **Low risk** |
|---|---|---|---|
| $\delta$ | 0.055*** | 0.094* | 0.063*** |
| | (0.021) | (0.052) | (0.000) |
| Obs. | 5,056 | 4,116 | 63,658 |
| Mean | 0.890 | 0.500 | 0.331 |
| Hospital FE | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ |
| Month FE | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ |
| Controls | ✓ | ✓ | ✓ |

Notes: Controls include $X1$, $X2$ and $X3$, as listed in Table 3. Coefficients are estimated with a linear probability model. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 7: **Number of Deliveries**

| | (1) | (2) | (3) |
|---|---|---|---|
| $\delta$ | 2.908 | 3.216 | 3.482 |
| | (2.268) | (2.200) | (2.238) |
| Obs. | 1,951 | 1,951 | 1,951 |
| Mean | 37.330 | 37.330 | 37.330 |
| Hospital FE | | ✓ | ✓ |
| Treated | ✓ | | |
| Year FE | ✓ | ✓ | ✓ |
| Month FE | ✓ | ✓ | ✓ |
| Hospital-month trend | | | ✓ |
| Post | | | ✓ |
| LHA FE | ✓ | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ |

Notes: For these results, we use a dataset at the hospital-month level, which explains the number of total observations. The monthly trend is a linear trend. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 8: **Patient Composition**

| | | |
|---|---|---|
| **Panel A: Low-Risk Mothers** | | |
| $\delta$ | -0.037 | -0.030 |
| | (0.027) | (0.024) |
| Obs. | 72,830 | 72,830 |
| Mean | 0.676 | 0.676 |
| **Panel B: Predicted Probability of a C-section** | | |
| $\delta$ | 0.007 | 0.006 |
| | (0.005) | (0.004) |
| Obs. | 72,830 | 72,830 |
| Mean | 0.365 | 0.365 |
| **Panel C: Low-Weight Newborns** | | |
| $\delta$ | 0.007 | 0.006 |
| | (0.009) | (0.008) |
| Obs. | 73,717 | 73,717 |
| Mean | 0.044 | 0.044 |
| Hospital FE | ✓ | ✓ |
| Year FE | ✓ | ✓ |
| Month FE | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ |
| LHA FE | ✓ | ✓ |
| Controls | | ✓ |

Notes: Controls include $X1$, $X2$ and $X3$ when the outcome is $Low - Risk$; $X3$, $Married$ and $Italy$ when the outcome is *Predicted Probability of a C-section*; or $X3$ and $X4$ when the outcome is $Low - Weight$, as listed in Table 3. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 9: **Hospitals Always Covered by Schedules**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **C-sections** | | | | | |
| $\delta$ | -0.001 | 0.000 | -0.002 | -0.002 | -0.003 |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| Obs. | 549,525 | 549,525 | 549,525 | 549,525 | 549,525 |
| Mean | 0.355 | 0.355 | 0.355 | 0.355 | 0.355 |
| Hospital FE | | ✓ | ✓ | ✓ | ✓ |
| Treated | ✓ | | | | |
| Year FE | ✓ | ✓ | ✓ | | |
| Month FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hosp-month trend | | | | ✓ | ✓ |
| Post | | | | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | | ✓ | | ✓ |

Notes: Here we consider all hospitals belonging to the same regions as in the sample used for the main analysis, but the treated are those always covered by schedules and the controls are those hospitals never applying schedules between 2001 and 2004. $\delta$ is the interaction between always being treated and the dummy for the period post-2002. Controls include $X1$, $X2$, and $X3$ as listed in Table 3. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 10: **DD - Robustness Checks on C-section**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Panel B: Cluster at the LHA Level** | | | | | |
| $\delta$ | 0.079** | 0.072** | 0.065** | 0.074** | 0.065** |
| | (0.039) | (0.032) | (0.033) | (0.031) | (0.032) |
| Obs. | 72,830 | 72,830 | 72,830 | 72,830 | 72,830 |
| Mean | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 |
| **Panel C: Cluster at the Court Level** | | | | | |
| $\delta$ | 0.079*** | 0.072** | 0.065** | 0.074** | 0.065*** |
| | (0.030) | (0.034) | (0.029) | (0.031) | (0.025) |
| Obs. | 72,830 | 72,830 | 72,830 | 72,830 | 72,830 |
| Mean | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 |
| **Panel D: Cluster at the Court Level & Court FE** | | | | | |
| $\delta$ | 0.079*** | 0.072*** | 0.065** | 0.074** | 0.065** |
| | (0.000) | (0.000) | (0.027) | (0.034) | (0.030) |
| Obs. | 72,830 | 72,830 | 72,830 | 72,830 | 72,830 |
| Mean | 0.380 | 0.380 | 0.380 | 0.380 | 0.380 |
| Hospital FE | | ✓ | ✓ | ✓ | ✓ |
| Treated | ✓ | | | | |
| Year FE | ✓ | ✓ | ✓ | | |
| Month FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hosp-month trend | | | | ✓ | ✓ |
| Post | | | | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | | ✓ | | ✓ |

Notes: Controls include $X1$, $X2$ and $X3$ when outcomes refer to mothers, as listed in Table 3. Coefficients are estimated with a linear probability model. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 11: **Drivers of C-section Overuse**

| | Quality | Liability Pressure | | | Financial Incentives |
|---|---|---|---|---|---|
| | **Delivery Volumes** | **Court Backlog** | **Schedules Level 40 yrs** | **60 yrs** | **Reimbursement** |
| | **Below 500** | **Low** | **Low** | **Low** | **Low** |
| $\delta$ | 0.104*** | 0.048*** | 0.117*** | 0.116** | 0.003 |
| | (0.000) | (0.018) | (0.044) | (0.044)) | (0.039) |
| | **Above 500** | **High** | **High** | **High** | **High** |
| $\delta$ | 0.046 | 0.123*** | 0.042 | 0.042 | 0.081* |
| | (0.033) | (0.036) | (0.061) | (0.058) | (0.047) |
| Diff. | -0.058** | 0.075** | -0.074* | -0.074** | 0.078*** |
| | (0.033) | (0.031) | (0.042) | (0.037) | (0.027) |
| Obs. | 72,830 | 72,830 | 72,830 | 72,830 | 72,830 |

Notes: *Below 500* indicates that the annual number of deliveries that occurred in a hospital is fewer than 500; *Above 500* indicates that this number is higher than 500. *40 yrs* and *60 yrs* indicate the level of the damage in case of a 25% disability suffered by a 40 and 60 year old person respectively. *Low Level* indicates a below-mean value, and *High Level* indicates an above-mean value. In each specification we control for the year fixed effects, LHA fixed effects, month fix effects, LHA-month interactions, the three sets of controls, and a dummy for the treated hospitals. Controls include $X1$, $X2$, and $X3$, as listed in Table 3. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

# Online Appendix A: Additional Figures and Tables

This Appendix provides additional tables and figures, which are also discussed in the paper. In particular, we present the following:

- Example of Schedules (Table A.1);
- Descriptive Statistics (Tables A.2 and A.3 );
- Covariate Sets Added One by One - Mother Results (Table A.4);
- Covariate Sets Added One by One - Newborn Results (Table A.5);
- Distribution of Courts Implementing Schedules (Figure A.1);
- Mothers Main Results at the National Level (Table A.6);
- Newborns Main Results at the National Level (Table A.7);
- Relationship Between Schedules and the Secondary Outcomes (Confidence Intervals) (Table A.8);
- Pre-trends of the Heterogeneous Effects (Figure A.2).

Table A.1: **Example of Schedules**

| | | | Age | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| | | | Age Index | | | | | | | |
| | | Point value | 1.000 | 0.995 | 0.990 | 0.985 | 0.980 | 0.975 | 0.970 | 0.965 |
| Disability | 10% | 1,729.83 | 17,298.34 | 17,211.84 | 17,125.32 | 17,038.86 | 16,952.37 | 16,865.88 | 16,779.39 | 16.692.89 |
| | 11% | 1,809.41 | 19,903.47 | 19,803.95 | 19,704.43 | 19,604.91 | 19,505.40 | 19,405.88 | 19,306.36 | 19,206.84 |
| | 12% | 1,888.40 | 22,660.82 | 22,547.52 | 22,434.21 | 22,320.91 | 22,207.60 | 22,904.30 | 21,981.00 | 21,867.69 |
| | 13% | 1,967.97 | 25,583.66 | 25,455.74 | 25,327.83 | 25,199.91 | 25,071.99 | 24,944.07 | 24,816.15 | 24,688.23 |
| | 14% | 2,046.97 | 28,657.58 | 28,514.29 | 28,371.00 | 28,227.71 | 28,084.43 | 27,941.14 | 27,797.85 | 27,654.56 |
| | 15% | 2,216.54 | 31,898.13 | 31,738.64 | 31,579.15 | 31,419.66 | 31,260.17 | 31,100.68 | 30,941.19 | 30,781.70 |
| | 16% | 2,205.54 | 35,288.61 | 35,112.16 | 34.935.72 | 34,759.28 | 34,582.83 | 34,406.39 | 34,229.95 | 34,053.51 |
| | 17% | 2,285.11 | 38,846.87 | 38,652.64 | 38,458.41 | 38,264.17 | 38,069.94 | 37,875.70 | 37,681.47 | 37,487.23 |
| | 18% | 2,364.11 | 42,553.91 | 42,341.14 | 42,128.37 | 41,915.60 | 41,702.83 | 41,490.06 | 41,277.29 | 41,064.52 |
| | 19% | 2,443.68 | 46,429.89 | 46,197.74 | 45,965.59 | 45,733.44 | 45,501.29 | 45,269.14 | 45,036.99 | 44,804.84 |
| | 20% | 2,522.67 | 50,452.48 | 50,201.21 | 49,948.95 | 49,696.68 | 49,444.41 | 49,192.14 | 48,939.88 | 48,687.61 |

Notes: Values are expressed in 2002 euros and taken from the reference table adopted by the Court of Milan in 2002. In the case of 10% disability suffered by a 3-year-old victim, the reference compensation amounts to 17,125.32 euros, which is obtained by multiplying the monetary percentage point value (1,729.83 euros) by ten by the age index (0.990). This mechanism foresees the simultaneous application of two criteria: (i) a progressive criterion for the determination of the monetary point values of the disability percentages, and (ii) a regressive criterion with respect to the age of the injured party. According to the first criterion, the compensation varies unevenly and more rapidly as the severity of the injury increases. In contrast, the regressive criterion reflects the fact that, considering the average possible lifetime of a person, a victim who has been harmed at a younger age would bear the consequences of the physical impairment for a longer period than an older victim (De Paola and Avigliano 2009).

Table A.2: **Descriptive Statistics**

| | **Total** | **Treated** | **Control** |
|---|---|---|---|
| **Outcomes at the mother level** | | | |
| C-section | 0.380 | 0.385 | 0.373 |
| | (0.485) | (0.487) | (0.484) |
| Preventable | 0.083 | 0.086 | 0.079 |
| | (0.276) | (0.281) | (0.269) |
| **Controls at the mother level** | | | |
| Age19 | 0.021 | 0.023 | 0.019 |
| | (0.143) | (0.149) | (0.135) |
| *Age*20_24 | 0.115 | 0.120 | 0.108 |
| | (0.319) | (0.325) | (0.310) |
| *Age*25_29 | 0.296 | 0.296 | 0.291 |
| | (0.455) | (0.456) | (0.454) |
| *Age*30_34 | 0.350 | 0.342 | 0.359 |
| | (0.477) | (0.474) | (0.480) |
| *Age*35_39 | 0.186 | 0.185 | 0.187 |
| | (0.389) | (0.388) | (0.390) |
| Age40 | 0.036 | 0.035 | 0.037 |
| | (0.185) | (0.183) | (0.188) |
| Italy | 0.931 | 0.933 | 0.927 |
| | (0.254) | (0.249) | (0.260) |
| Married | 0.763 | 0.790 | 0.731 |
| | (0.425) | (0.407) | (0.443) |
| Risk Factors | 0.113 | 0.111 | 0.115 |
| | (0.316) | (0.314) | (0.319) |
| **Controls at the mother municipality level** | | | |
| Average Income (2015 euro) | 17,104 | 16,904 | 17,350 |
| | (2,934) | (2,951) | (2,893) |
| Altitude (m) | 687.668 | 637.892 | 748.923 |
| | (491.707) | (445.861) | (536.543) |
| Low Urbanization | 0.269 | 0.298 | 0.232 |
| | (0.443) | (0.458) | (0.422) |
| Medium Urbanization | 0.636 | 0.652 | 0.617 |
| | (0.481) | (0.476) | (0.486) |
| High Urbanization | 0.095 | 0.050 | 0.151 |
| | (0.293) | (0.218) | (0.358) |
| Education | 0.060 | 0.059 | 0.061 |
| | (0.029) | (0.027) | (0.031) |
| Obs. | 72,830 | 40,180 | 32,650 |

Notes: *Outcomes* are described in Table 3. *Risk Factors* captures the incidence of risk factors, as described by dummies in *Cov*1 of Table 3. *Italy* is equal to 1 if the mother is Italian and 0 otherwise. *Income* is in 2015 euros. *Education* is the share of municipal residents with a college degree, as measured in the 2001 Census data. *Urbanization* captures both population density per square kilometer and the municipality dimension. It is provided by the National Institute of Statistics, as measured in the 2001 Census data. *Sea Level* is in meters. Variables at the mother level are available through the patient discharge records, while variables at the mothers' municipality level are available through the Italian National Institute of Statistics.

Table A.3: **Descriptive Statistics (Cont'd)**

| | **Total** | **Treated** | **Control** |
|---|---|---|---|
| **Outcomes at the newborn level** | | | |
| Cesarean Complications | 0.058 | 0.066 | 0.048 |
| | (0.234) | (0.248) | (0.213) |
| Breathing Interventions | 0.020 | 0.015 | 0.027 |
| | (0.140) | (0.121) | (0.162) |
| **Controls at the newborn level** | | | |
| Congenital anomalies | 0.027 | 0.026 | 0.029 |
| | (0.162) | (0.158) | (0.167) |
| Maternal conditions | 0.045 | 0.049 | 0.040 |
| | (0.208) | (0.216) | (0.196) |
| Obs. | 73,717 | 42,630 | 31,087 |

Notes: *Outcomes* is described in Table 3. *Income* is in 2015 euros. *Education* is the share of municipal residents with a college degree as measured in the 2001 Census data. *Urbanization* captures both population density per square kilometer and the municipality dimension. It is provided by the National Institute of Statistics as measured in the 2001 Census data. *Sea level* is in meters. Variables at the mother level are available through the patient discharge records, while variables at the mothers' municipality level are available through the Italian National Institute of Statistics.

Table A.4: **Covariate Sets Added One by One - Mother Results**

| | $X1'_{iht}$ | $X2'_{iht}$ | $X3'_{iht}$ |
|---|---|---|---|
| | Risk Factors | Mother Extra Characteristics | Municipality Characteristics |
| *C-section* | 0.065*** | 0.071*** | 0.071*** |
| | (0.000) | (0.000) | (0.000) |
| *Preventable Complications* | - 0.000 | 0.003 | 0.003 |
| | (0.000) | (0.013) | (0.049) |
| Obs. | 72,830 | 72,830 | 72,830 |
| Hospital FE | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ |
| Month FE | ✓ | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ |

Notes: Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

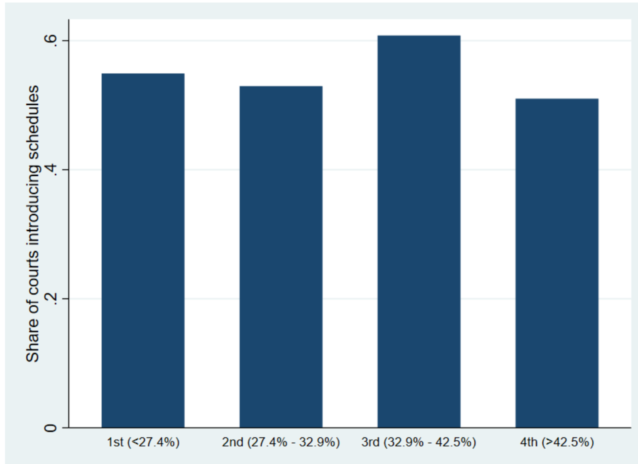
Table A.5: **Covariates Sets Added One by One - Newborn results**

| | $X3'_{iht}$ | $X4'_{iht}$ |
|---|---|---|
| | Municipality Characteristics | Newborn Risk Profile |
| *Cesarean complications* | 0.018 | 0.020 |
| | (0.013) | (0.014) |
| *Breathing complications* | -0.004 | -0.003 |
| | (0.006) | (0.007) |
| Obs. | 73,717 | 73,717 |
| Hospital FE | ✓ | ✓ |
| Year FE | ✓ | ✓ |
| Month FE | ✓ | ✓ |
| LHA-time interactions | | ✓ |
| LHA FE | ✓ | ✓ |

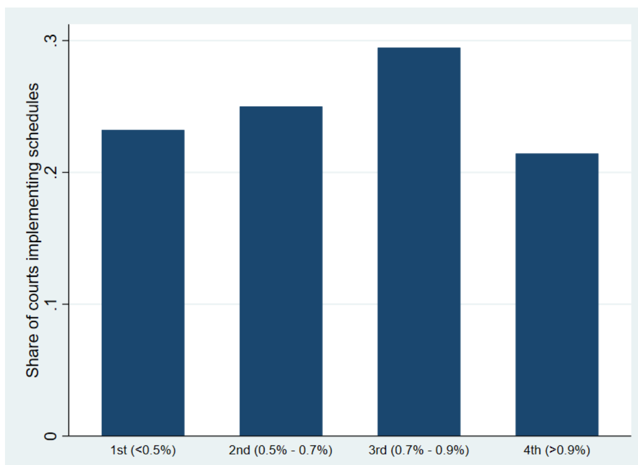
Notes: Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure A.1: **Distribution of Courts Implementing Schedules**

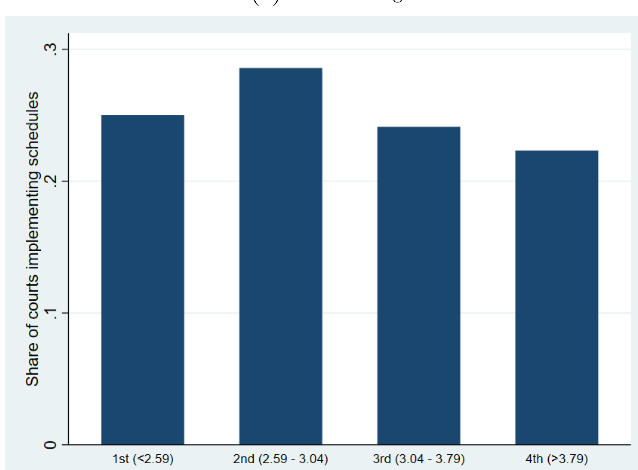
(a) C-section rates




(b) Claim frequency




(c) Civil backlog




Notes: The figure shows the share of schedules implementing schedules that falls into each quartile of the distribution of c-section rates (a), claim frequency (b), and court civil backlog (c).

Table A.6: **DD - Mothers Main Results at the National Level**

| | (1) | (2) | (3) |
|---|---|---|---|
| **Panel A: C-sections** | | | |
| $\delta$ | 0.023*** | 0.022** | 0.022*** |
| | (0.009) | (0.009) | (0.009) |
| Obs. | 1,359,504 | 1,359,504 | 1,359,504 |
| Mean | 0.371 | 0.371 | 0.371 |
| **Panel B: Preventable Complications** | | | |
| $\delta$ | -0.008 | -0.006 | -0.004 |
| | (0.006) | (0.006) | (0.006) |
| Obs. | 1,359,504 | 1,359,504 | 1,359,504 |
| Mean | 0.102 | 0.102 | 0.102 |
| Hospital FE | | ✓ | ✓ |
| Treated | ✓ | | |
| Year FE | ✓ | ✓ | ✓ |
| Month FE | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ |
| Controls | | | ✓ |

Notes: Controls include $X1$, $X2$ and $X3$ when outcomes refer to mothers, as listed in Table 3. Coefficients are estimated with a linear probability model. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A.7: **DD - Newborns Main Results at the National Level**

| | (1) | (2) | (3) |
|---|---|---|---|
| **Panel A: Csections Complications** | | | |
| $\delta$ | 0.003 | 0.003 | 0.003 |
| | (0.005) | (0.005) | (0.005) |
| Obs. | 1,332,170 | 1,332,170 | 1,332,170 |
| Mean | 0.046 | 0.046 | 0.046 |
| **Panel B: Breathing Intervention** | | | |
| $\delta$ | 0.003 | 0.003 | 0.003 |
| | (0.002) | (0.002) | (0.002) |
| Obs. | 1,332,170 | 1,332,170 | 1,332,170 |
| Mean | 0.021 | 0.021 | 0.021 |
| Hospital FE | | ✓ | ✓ |
| Treated | ✓ | | |
| Year FE | ✓ | ✓ | ✓ |
| Month FE | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ |
| Controls | | | ✓ |

Notes: Controls include $X3$ and $X4$, as listed in Table 3. Coefficients are estimated with a linear probability model. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

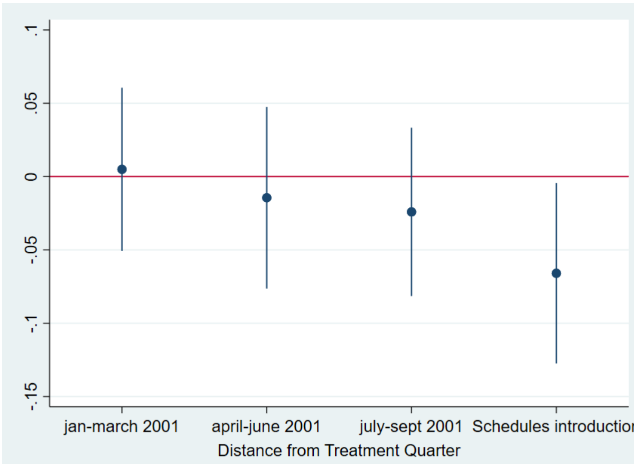
Table A.8: **Relationship Between Schedules and the Secondary Outcomes (Confidence Intervals)**

| | Mother | Newborn | |
|---|---|---|---|
| | **Preventable Complications** | **C-sections Complications** | **Breathing Interventions** |
| $\delta$ | 0.000<br>(0.056) | 0.018<br>(0.012) | -0.004<br>(0.006) |
| 95% Confidence Interval | [-0.110, 0.110] | [-0.005,0.041] | [-0.016, 0.008] |
| 95% Confidence Interval as Fraction of Mean | [-1.322, 1.322] | [-0.095, 0.716] | [-0.788, 0.388] |
| Obs. | 72,830 | 73,717 | 73,717 |
| Hospital FE | ✓ | ✓ | ✓ |
| Year FE | ✓ | ✓ | ✓ |
| Month FE | ✓ | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ |
| Controls | ✓ | ✓ | ✓ |

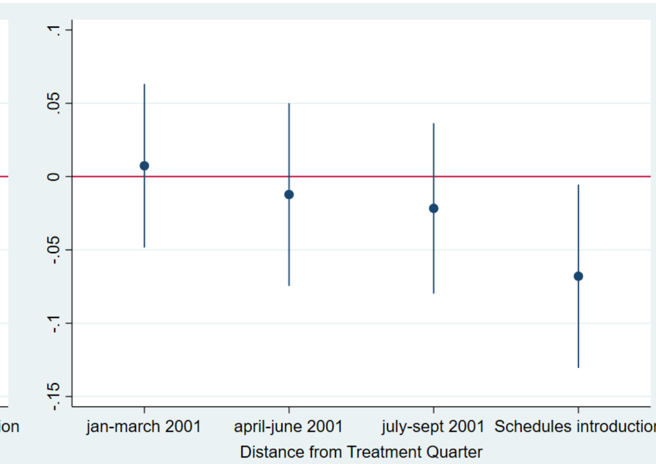
Notes: Controls include $X3$ and $X4$, as listed in Table 3. Coefficients are estimated with a linear probability model. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure A.2: **Pre-trends of the Heterogeneous Effects**

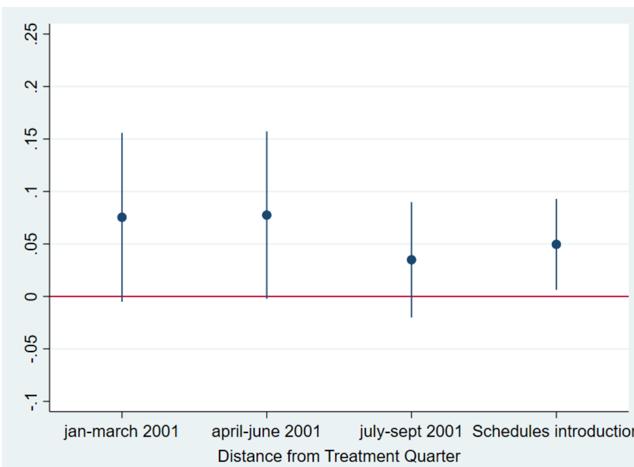
(a) Schedule 40 yrs



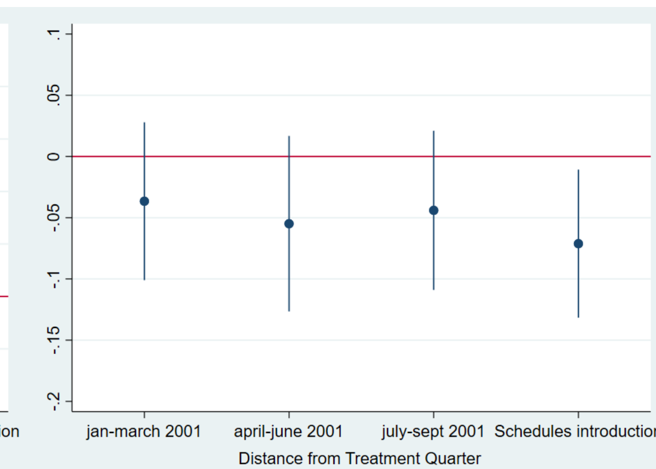

(b) Schedule 60 yrs



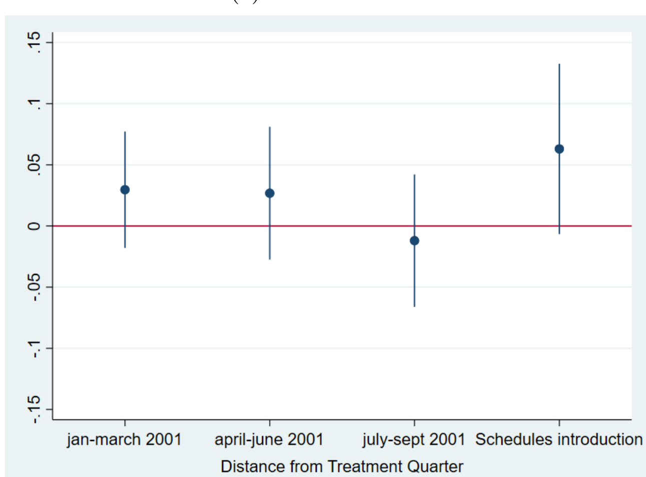

(c) Court backlog




(d) Delivery volumes




(e) Reimbursement




Notes: The figure plots the leads by quarter, with a confidence interval of 95%. Jan-March 2002 is the quarter in which schedules were adopted. The reference quarter is Oct-Dec 2001.

## Online Appendix B: Additional Figures and Tables

In this Appendix, we report the results of the specifications tested in the main text using also the deliveries taking place during the weekends. When including these observations, we also control for the weekday fixed effects. In particular, we present the following:

- Weekdays Deliveries (Table B.1);
- Mothers Main Results Including Weekend Deliveries (Table B.2);
- Newborns Main Results Including Weekend Deliveries (Table B.3);
- Robustness Checks on C-section Including Weekend Deliveries (Table B.4);
- Number of Deliveris Including Weekend Deliveries (Figure B.5);
- Patient Composition Including Weekend Deliveries (Table B.6);
- Leads and Lags Including Weekend Deliveries (Figure B.1);
- Hospitals Always Covered by Schedules Including Weekend Deliveries (Table B.7);
- Drivers of C-section Overuse Including Weekend Deliveries (Table B.8);
- Descriptive Statistics Including Weekend Deliveries (Tables B.9 and B.10).

Table B.1: **Weekday Deliveries**

| | **Deliveries** | **Cesarean Sections** |
|---|---|---|
| Tuesday | 35.779*** | 20.192*** |
| | (11.910) | (4.435) |
| Wednesday | 27.966** | 12.554*** |
| | (11.951) | (4.445) |
| Thursday | 32.812*** | 19.666*** |
| | (11.937) | (4.445) |
| Friday | 15.124 | 9.234** |
| | (11.882) | (4.430) |
| Saturday | -65.744*** | -55.388*** |
| | (11.882) | (4.435) |
| Sunday | -124.374*** | -100.852*** |
| | (11.980) | (4.456) |
| Obs. | 1,523 | 1,523 |

Notes: The reference sample was obtained by collapsing the dataset for the main analysis by weekdays of each year. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B.2: **DD - Mothers Main Results**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Panel A: C-sections** | | | | | |
| $\delta$ | 0.074*** | 0.067*** | 0.060*** | 0.068*** | 0.060*** |
| | (0.026) | (0.025) | (0.019) | (0.025) | (0.020) |
| Obs. | 95,750 | 95,750 | 95,750 | 95,750 | 95,750 |
| Mean | 0.363 | 0.363 | 0.363 | 0.363 | 0.363 |
| **Panel B: Preventable Complications** | | | | | |
| $\delta$ | 0.003 | 0.004 | 0.001 | 0.004 | 0.001 |
| | (0.013) | (0.011) | (0.023) | (0.011) | (0.007) |
| Obs. | 95,750 | 95,750 | 95,750 | 95,750 | 95,750 |
| Mean | 0.083 | 0.083 | 0.083 | 0.083 | 0.083 |
| Hospital FE | | ✓ | ✓ | ✓ | ✓ |
| Treated | ✓ | | | | |
| Year FE | ✓ | ✓ | ✓ | | |
| Month FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Weekdays FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hosp-month trend | | | | ✓ | ✓ |
| Post | | | | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | | ✓ | | ✓ |

Notes: Controls include $X1$, $X2$ and $X3$ when outcomes refer to mothers, as listed in Table 3. Coefficients are estimated with a linear probability model. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B.3: **DD - Newborns Main Results**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Panel A: Csections Complications** | | | | | |
| $\delta$ | 0.021 | 0.022 | 0.020 | 0.022 | 0.019 |
| | (0.014) | (0.015) | (0.013) | (0.016) | (0.013) |
| Obs. | 95,792 | 95,792 | 95,792 | 95,792 | 95,792 |
| Mean | 0.057 | 0.057 | 0.057 | 0.057 | 0.057 |
| **Panel B: Breathing Intervention** | | | | | |
| $\delta$ | -0.003 | -0.005 | -0.005 | -0.004 | -0.005 |
| | (0.006) | (0.007) | (0.007) | (0.006) | (0.007) |
| Obs. | 95,792 | 95,792 | 95,792 | 95,792 | 95,792 |
| Mean | 0.020 | 0.020 | 0.020 | 0.020 | 0.020 |
| Hospital FE | | ✓ | ✓ | ✓ | ✓ |
| Treated | ✓ | | | | |
| Year FE | ✓ | ✓ | ✓ | | |
| Month FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Weekdays FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hosp-month trend | | | | ✓ | ✓ |
| Post | | | | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | | ✓ | | ✓ |

Notes: Controls include $X3$ and $X4$, as listed in Table 3. Coefficients are estimated with a linear probability model. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B.4: **DD - Robustness Checks on C-section**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Panel B: Cluster at the LHA Level** | | | | | |
| $\delta$ | 0.074* | 0.067* | 0.060* | 0.068* | 0.060* |
| | (0.039) | (0.040) | (0.034) | (0.036) | (0.034) |
| Obs. | 95,750 | 95,750 | 95,750 | 95,750 | 95,750 |
| Mean | 0.363 | 0.363 | 0.363 | 0.363 | 0.363 |
| **Panel C: Cluster at the Court Level** | | | | | |
| $\delta$ | 0.079*** | 0.072** | 0.064*** | 0.068** | 0.060* |
| | (0.000) | (0.034) | (0.000) | (0.034) | (0.034) |
| Obs. | 95,750 | 95,750 | 95,750 | 95,750 | 95,750 |
| Mean | 0.363 | 0.363 | 0.363 | 0.363 | 0.363 |
| **Panel D: Cluster at the Court Level & Court FE** | | | | | |
| $\delta$ | 0.079*** | 0.072** | 0.064** | 0.068** | 0.060** |
| | (0.000) | (0.030) | (0.028) | (0.032) | (0.027) |
| Obs. | 95,750 | 95,750 | 95,750 | 95,750 | 95,750 |
| Mean | 0.363 | 0.363 | 0.363 | 0.363 | 0.363 |
| Hospital FE | | ✓ | ✓ | ✓ | ✓ |
| Treated | ✓ | | | | |
| Year FE | ✓ | ✓ | ✓ | | |
| Month FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hosp-month trend | | | | ✓ | ✓ |
| Post | | | | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | | ✓ | | ✓ |

Notes: Controls include $X1$, $X2$ and $X3$ when outcomes refer to mothers, as listed in Table 3. Coefficients are estimated with a linear probability model. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B.5: **Number of Deliveries**

| | (1) | (2) | (3) |
|---|---|---|---|
| $\delta$ | 3.591 | 3.899 | 3.974 |
| | (2.653) | (3.317) | (3.045) |
| Obs. | 1,953 | 1,953 | 1,953 |
| Hospital FE | | ✓ | ✓ |
| Treated | ✓ | | |
| Year FE | ✓ | ✓ | ✓ |
| Month FE | ✓ | ✓ | ✓ |
| Hospital-month trend | | | ✓ |
| Post | | | ✓ |
| LHA FE | ✓ | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ |

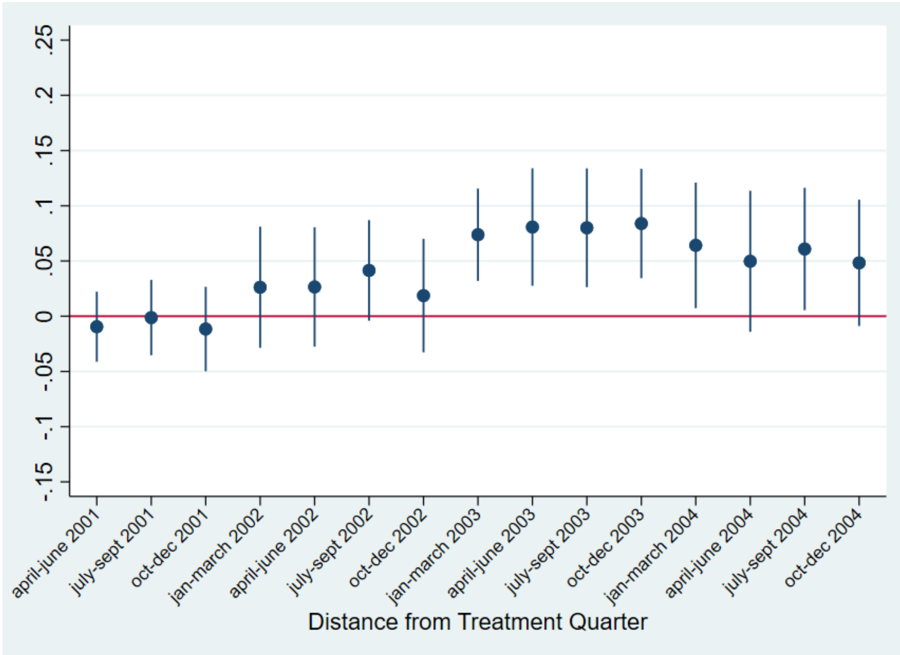
Notes: For these results, we use a dataset at the hospital-month level, which explains the number of total observations. The monthly trend is a linear trend. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B.6: **Patient Composition**

| | | |
|---|---|---|
| **Panel A: Low-Risk Mothers** | | |
| $\delta$ | -0.036 | -0.029 |
| | (0.025) | (0.023) |
| Obs. | 95,750 | 95,750 |
| Mean | 0.685 | 0.685 |
| **Panel B: Predicted Probability of a C-section** | | |
| $\delta$ | 0.006 | 0.006 |
| | (0.005) | (0.005) |
| Obs. | 95,750 | 95,750 |
| Mean | 0.363 | 0.363 |
| **Panel C: Low-Weight Newborns** | | |
| $\delta$ | 0.004 | 0.004 |
| | (0.008) | (0.007) |
| Obs. | 95,792 | 95,792 |
| Mean | 0.043 | 0.043 |
| Hospital FE | ✓ | ✓ |
| Year FE | ✓ | ✓ |
| Weekdays FE | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ |
| LHA FE | ✓ | ✓ |
| Controls | | ✓ |

Notes: Controls include $X1$, $X2$ and $X3$ when the outcome is $Low - Risk$ or $X3$, $Married$ and $Italy$ when the outcome is *Predicted Probability of a C-section*, while controls include $X3$ and $X4$ when the outcome is $Low - Weight$, as listed in Table 3. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure B.1: **Leads and Lags**




Notes: The figure plots the leads and lags, quarterly, with a confidence interval of 95%. Jan-March 2002 is the quarter of schedules adoption.

Table B.7: **Hospitals Always Covered by Schedules**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **C-sections** | | | | | |
| $\delta$ | -0.002 | -0.001 | -0.003 | -0.003 | -0.003 |
| | (0.007) | (0.008) | (0.007) | (0.008) | (0.007) |
| Obs. | 725,160 | 725,160 | 725,160 | 725,160 | 725,160 |
| Mean | 0.329 | 0.329 | 0.329 | 0.329 | 0.329 |
| Hospital FE | | ✓ | ✓ | ✓ | ✓ |
| Treated | ✓ | | | | |
| Year FE | ✓ | ✓ | ✓ | | |
| Month FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Hosp-month trend | | | | ✓ | ✓ |
| Post | | | | ✓ | ✓ |
| LHA-time interactions | ✓ | ✓ | ✓ | ✓ | ✓ |
| LHA FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Controls | | | ✓ | | ✓ |

Notes: Here we consider all hospitals belonging to the same regions as in the sample used for the main analysis. The treated hospitals are those that were always covered by schedules and the control hospitals are those that never applied schedules between 2001 and 2004. $\delta$ is the interaction between always being treated and the dummy for the period post-2002. Controls include $X1$, $X2$, and $X3$, as listed in Table 3. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B.8: **Drivers of C-section Overuse**

| | **Quality** | **Liability Pressure** | | | **Financial Incentives** |
|---|---|---|---|---|---|
| | **Delivery Volumes** | **Court Backlog** | **Schedules Level 40 yrs** | **Schedules Level 60 yrs** | **Reimbursement** |
| | **Below 500** | **Low** | **Low** | **Low** | **Low** |
| $\delta$ | 0.121*** | 0.025 | 0.111** | 0.112* | -0.002 |
| | (0.046) | (0.033) | (0.052) | (0.060) | (0.010) |
| | **Above 500** | **High** | **High** | **High** | **High** |
| $\delta$ | 0.046 | 0.100** | 0.042 | 0.041 | 0.076*** |
| | (0.063) | (0.043) | (0.064) | (0.055) | (0.000) |
| Difference | -0.075* | 0.075*** | -0.069* | -0.071* | 0.078*** |
| | (0.043) | (0.028) | (0.038) | (0.043) | (0.000) |
| Obs. | 95,750 | 95,750 | 95,750 | 95,750 | 95,750 |

Notes: *Below 500* indicates that the annual number of deliveries that occurred in a hospital is fewer than 500; *Above 500* indicates that this number is higher than 500. *Low Level* indicates a below-mean value, and *High Level* indicates an above-mean value. In each specification we control for the year fixed effects, LHA fixed effects, month fix effects, LHA-month interactions, the three sets of controls, and a dummy for the treated hospitals. Controls include $X1$, $X2$, and $X3$ as listed in Table 3. Robust standard errors clustered at the hospital level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table B.9: **Descriptive Statistics**

| | **Total** | **Treated** | **Control** |
|---|---|---|---|
| **Outcomes at the mother level** | | | |
| C-section | 0.363 | 0.371 | 0.353 |
| | (0.481) | (0.483) | (0.478) |
| Preventable | 0.083 | 0.085 | 0.080 |
| | (0.276) | (0.279) | (0.272) |
| **Controls at the mother level** | | | |
| Age19 | 0.021 | 0.022 | 0.019 |
| | (0.143) | (0.148) | (0.137) |
| $Age20\_24$ | 0.116 | 0.121 | 0.110 |
| | (0.321) | (0.326) | (0.313) |
| $Age25\_29$ | 0.294 | 0.296 | 0.292 |
| | (0.456) | (0.456) | (0.454) |
| $Age30\_34$ | 0.349 | 0.341 | 0.359 |
| | (0.477) | (0.474) | (0.480) |
| $Age35\_39$ | 0.185 | 0.185 | 0.185 |
| | (0.388) | (0.388 | (0.184) |
| Age40 | 0.035 | 0.035 | 0.035 |
| | (0.183) | (0.183) | (0.184) |
| Italy | 0.928 | 0.932 | 0.925 |
| | (0.256) | (0.253) | (0.264) |
| Married | 0.760 | 0.788 | 0.726 |
| | (0.427) | (0.409) | (0.446) |
| Risk Factors | 0.110 | 0.108 | 0.113 |
| | (0.313) | (0.310) | (0.317) |
| **Controls at the mother municipality level** | | | |
| Average Income (2015 euro) | 17,109 | 16,890 | 17,365 |
| | (2,933) | (2,951) | (2,890) |
| Altitude (m) | 686.902 | 649.903 | 743.444 |
| | (492.479) | (450.010) | (534.738) |
| Low Urbanization | 0.268 | 0.301 | 0.230 |
| | (0.443) | (0.458) | (0.421) |
| Medium Urbanization | 0.634 | 0.649 | 0.617 |
| | (0.482) | (0.477) | (0.486) |
| High Urbanization | 0.097 | 0.051 | 0.153 |
| | (0.296) | (0.220) | (0.360) |
| Education | 0.060 | 0.059 | 0.061 |
| | (0.029) | (0.027) | (0.031) |
| Obs. | 95,750 | 52,798 | 42,952 |

Notes: *Outcomes* are described in Table 3. *Risk Factors* captures the incidence of risk factors, as described by dummies in $Cov1$ of Table 3. *Italy* is equal to 1 if the mother is Italian and 0 otherwise. *Income* is in 2015 euros. *Education* is the share of municipal residents with a college degree, as measured in the 2001 Census data. *Urbanization* captures both population density per square kilometer and the municipality dimension. It is provided by the National Institute of Statistics, as measured in the 2001 Census data. *Sea Level* is in meters. Variables at the mother level are available through the patient discharge records, while variables at the mothers' municipality level are available through the Italian National Institute of Statistics.

Table B.10: **Descriptive Statistics (Cont'd)**

| | **Total** | **Treated** | **Control** |
|---|---|---|---|
| **Outcomes at the newborn level** | | | |
| Cesarean Complications | 0.057 | 0.064 | 0.047 |
| | (0.231) | (0.244) | (0.211) |
| Breathing Interventions | 0.020 | 0.015 | 0.026 |
| | (0.139) | (0.120) | (0.160) |
| **Controls at the newborn level** | | | |
| Congenital anomalies | 0.026 | 0.025 | 0.028 |
| | (0.160) | (0.157) | (0.164) |
| Maternal conditions | 0.046 | 0.051 | 0.040 |
| | (0.210) | (0.220) | (0.195) |
| Obs. | 95,792 | 59,930 | 40,862 |

Notes: *Outcomes* is described in Table 3. *Income* is in 2015 euros. *Education* is the share of municipal residents with a college degree as measured in the 2001 Census data. *Urbanization* captures both population density per square kilometer and the municipality dimension. It is provided by the National Institute of Statistics as measured in the 2001 Census data. *Sea level* is in meters. Variables at the mother level are available through the patient discharge records, while variables at the mothers' municipality level are available through the Italian National Institute of Statistics.