

Does Experience Rating Improve Obstetric Practices? Evidence From Geographical Discontinuities in Italy*

Sofia Amaral-Garcia

ETH Zurich

Paola Bertoli

University of Economics, Prague

CERGE-EI Affiliate Fellow

Veronica Grembi

Copenhagen Business School

Baffi Center - Bocconi University

CEIS - Tor Vergata

Abstract

Using data from 2002 to 2009 inpatient discharge records on deliveries in the Italian region of Piedmont, we assess the impact of an increase in malpractice pressure on obstetric practices, as identified by the introduction of experience-rated malpractice liability insurance. Our identification strategy exploits the exogenous location of public hospitals in court districts with and without schedules for noneconomic damages. We perform difference-in-differences analysis on the entire sample and on a subsample which only considers the nearest hospitals in the neighborhood of court district boundaries. We find that the increase in medical malpractice pressure is associated with a decrease in the probability of performing a C-section from 2.3 to 3.7 percentage points (7% to 11.6% at the mean value of C-section) with no consequences for a broadly defined measure of complications or neonatal outcomes. We show that these results are robust to the different methodologies and can be explained by a reduction in the discretion of obstetric decision making rather than by patient cream skimming.

JEL Classification: K13; K32; I13

Keywords: Experience rating, Medical liability insurance, C-sections, Scheduled damages

*We thank Owen O'Donnell and two anonymous referees for comments that significantly improved the paper. We also thank Bernard Black, Daniel Chen, Libor Dusek, Michael Frakes, Nuno Garoupa, Vardges Levonyan, Giovanni Battista Ramello, Federico Revelli, Bettina Sifinger, Miriam Wüst, the seminar participants at Copenhagen Business School, ETH Zurich - Center for Law & Economics, University of Verona, IEL Collegio Carlo Alberto, EMLE Midterm Meeting 2015, DIW Berlin, Paris II - CRED, University of Economics in Prague, Max Planck Institute for Research on Collective Goods, CEIS Tor Vergata, the 2014 Spanish Law and Economics Association Annual Meeting, the 23rd Health Econometrics European Workshop, and the Annual Conference of Empirical Legal Studies (Berkeley) for helpful comments. We are grateful to Giuseppe Carchedi for his study on the Piedmont insurance scheme. *Dors Piemonte* provided data on Natalty Certificates and the Italian Ministry of Health on National Hospital Discharge Records.

1 Introduction

Cesarean rates are high in several developed countries and have been increasing considerably over the past decades (OECD, 2014). However, no significant changes in the medical risk profiles of mothers explain this shift (Declercq *et al.*, 2006), which leads to the question whether these additional procedures are warranted. Since the failure to perform timely C-sections is one of the most common allegations in malpractice claims against obstetricians (Sachs, 1989), high C-section rates are often attributed to fear of litigation. According to this view, C-sections are used as a form of defensive medicine in response to intense medical malpractice pressure, which is the joint result of the probability of being sued, of all litigation costs, and of high liability insurance premiums. For decades the conventional wisdom has been that higher malpractice pressure leads obstetricians to perform unnecessary C-sections. In fact, Localio *et al.* (1993) found a positive association between performing a C-section and both physician malpractice premiums and claims history. Dranove and Watanabe (2009) showed that obstetricians perform more C-sections after they or their colleagues have been sued, and Shurtz (2013) confirmed this result.

Currie and MacLeod (2008) challenged this common idea that higher malpractice pressure induces obstetricians to perform excessive C-sections. They provided contrasting evidence and showed that higher malpractice pressure does not necessarily lead to a more frequent use of C-sections, as supported by other studies (Sloan *et al.*, 1997; Dubay *et al.*, 1999; Frakes, 2012). The finding of Currie and MacLeod (2008) is consistent with a model of doctor behavior that depends on both patient conditions and the probability of committing an error with potential legal consequences. In a nutshell, they prove that the ultimate effect of an increase in malpractice pressure on the choice to perform a C-section cannot be uniquely determined *ex ante* on theoretical grounds: it depends on the level of utilization of the procedure (see also Frakes, 2015). According to their model the probability of undergoing a malpractice claim is a function of how much the chosen delivery type differs from the appropriate type conditional on the patient's health condition. On average, the difference between the performed and appropriate delivery types is large whenever C-section rates are above or below medical appropriateness given patients' health status. In the case of C-sections being over-performed, the probability of an error leading to liability is greater if the C-section is performed than if the C-section is not performed. Hence, an increase in malpractice pressure is expected to induce practitioners to lower the rate of C-sections converging to its optimal level.

Relying on the theoretical framework presented by Currie and MacLeod (2008), we assess whether an increase in malpractice pressure has affected i) decisions to perform C-sections, ii) the incidence of maternal complications, and iii) neonatal outcomes in the Italian region of Piedmont. Specifically, we exploit the increase in malpractice pressure triggered by adopting experience-rated medical liability insurance (*i.e.*, the adjustment of premiums to claims history). In 2005, Piedmont introduced experience-rated liability insurance for all its public hospitals without modifying the liability system, which means that the probability of plaintiffs bringing suits was unaffected. By linking hospitals' insurance premiums to providers' risk exposure, the reform made hospitals more accountable for their claims experience, thus

more concerned with inducing their clinicians to reduce unnecessary risk exposure. At the time of this reform, Piedmont reported a high rate of C-sections with more than 30% of deliveries being a C-section. Therefore, the expectation is that an increase in medical malpractice pressure will decrease the incidence of the reference procedure, moving the C-section rate toward the medically optimal level.

As the introduction of experience rating involved all hospitals in Piedmont, we exploit the geographical variation in court districts within the region to identify the effect of the variation of malpractice pressure. In particular, our identification strategy relies on hospitals' geographical location, which determines the competent court in the case of litigation. Courts differ with respect to their adoption of noneconomic damage schedules; schedules mean not only that the maximum recoverable compensation is limited, but also that awards are more predictable and less varied.¹ Limits to noneconomic damages are regarded in the related literature as good policy tools to decrease malpractice pressure as they reduce the unpredictability and variability of damages (Studdert and Mello, 2005; Bovbjerg *et al.* 1989). As a result, the *interaction* between the introduction of experience rating in 2005 and the fact that only some courts apply schedules produces natural "treatment" and "control" groups, which are differently affected by the policy shift. The increase in malpractice pressure due to the new link between premiums and damage awards is larger for hospitals that face no limitations on compensations, as opposed to hospitals for which compensations are limited by courts' schedules. This is why we use a difference-in-differences (diff-in-diff) strategy, which contrasts hospitals in courts with or without schedules, before and after the policy shift. To increase internal validity we also implement a robustness check using only the hospitals in the nearest neighborhood across the court districts so that the common trends assumption is more plausible.

Experience rating implemented at the hospital level, as in this setting, directly affects hospitals through paid premiums, whereas it affects individual physicians only indirectly. As we study the impact of this reform on obstetrician practices, we are assuming that two conditions hold in a public healthcare system in which physicians are civil servants working for a single hospital. First, hospitals are sensitive to discounts on their insurance premiums. Second, hospitals are able to convey their priorities to their employees. Empirical evidence from the UK, which has a healthcare system similar to Italy's, supports these assumptions. Fenn *et al.* (2013) use hospital data on *methicillin-resistant staphylococcus aureus* infections in England and Wales to show that hospitals react to potential discounts on their liability risk-pooling contributions. Faced with the possibility of lower expenditures, hospitals manage to lower infection rates.²

¹*Damage schedules* are simply tables with entries for the degree of injury severity and the victim's age. For example, if a newborn in one hospital located in the Turin court district suffered a 25% disability as a consequence of negligence during the delivery the baby could be awarded a maximum of 124,325 euros for noneconomic damages, while her 32-year old mother would have received up to 105,054 euros for the same type of injury inflicted on her. However, if the same case had occurred in a hospital located in the nearby court district of Pinerolo, which did not adopt schedules, the noneconomic award would be decided at the discretion of a judge. See Appendix A.

²Fenn *et al.* (2007) provide further evidence on how public hospitals convey their priorities to their personnel. Specifically, the authors explored the relationship between malpractice liability and the use of

Combining a unique dataset of inpatient discharge records on 265,532 deliveries and neonatal records from 2002 to 2009, we show that an increase in malpractice pressure leads to a 2.3 to 3.7 percentage points (7% to 11.6% at the mean of C-section) reduction in the use of C-sections at the patient level and a 3 to 5.5 percentage points (8.5%-15%) reduction at the hospital level. Moreover, the observed decrease in C-sections is not associated with changes in maternal health as proxied by the incidence of complications. Nor it is associated with changes in neonatal health as proxied by the incidence of low Apgar scores (*i.e.*, below 7) or the use of resuscitation measures. The main results on C-sections are robust to several specifications and robustness checks (*i.e.*, the exclusion of newborns delivered on weekends and inclusion of hospital-year trends), and they are not due to anticipatory effects. Analyzing possible channels of the policy’s effect, we show that cream skimming of low-risk patients could not explain the decrease in C-sections. However, we detect a reduction of the role played by nonmedical factors in driving physicians’ final decisions on whether to perform C-sections or vaginal deliveries. This means that the policy principally affected decisions to perform C-sections on mothers with fewer indications for cesarean delivery.

Our findings have implications that support the design of hospital-level policies since they highlight the important role that hospitals can play in reducing inappropriate medical choices. An additional implication is that hospital experience rating should receive more consideration. By awarding lower premiums to healthcare providers, which are able to cope with risk exposure, such policies can reduce medical liability costs with no particular consequences on patients’ access to care.

The remainder of the paper is organized as follows. Section 2 provides background on the Italian healthcare system and the Piedmont experience rating policy. Section 3 explains the econometric strategy and data used. Section 4 discusses the results and their robustness, and investigates the channels of the detected effects. Section 5 concludes.

2 Institutional Background

The Italian National Health Service provides uniform and comprehensive medical care to all residents irrespective of their socio-economic status. Regional governments are charged with the local organization of healthcare services, the delivery of which depends mainly on public supply. Public hospitals are managed by independent public trusts or local health units. In 2010, over 95% of infants born in Piedmont were delivered in public hospitals (Ministero della Salute, 2013). Based on their municipality of residence, patients are enrolled in healthcare plans managed by local health units and are assigned to a public hospital (*i.e.*, their home hospital). Nevertheless, patients can opt to receive treatment at their preferred hospital.

Hospitals are responsible for providing liability insurance to their medical personnel. It is therefore impossible to have uninsured hospital-employed physicians. Malpractice insurance premiums are traditionally set as a percentage of the gross payroll paid to employees by hospitals and are not adjusted based on previous claims (Amaral-Garcia and Grembi, 2014). However, since 2005, Piedmont has had a regional malpractice liability fund for all

imaging and scanning diagnostic procedures by English hospitals.

public hospitals (*Fondo Speciale per i rischi di responsabilità civile delle ASL*). According to the regional insurance plan, the lion's share of each hospital's contribution to the fund (*i.e.*, premiums) depends on its claims history. In particular, 30% of each hospital's overall contribution is still based on the gross payroll paid by the institution, but the remaining 70% is computed on the basis of the hospital's average risk exposure in the previous three years, defined as claims received and compensations paid.³

The regional fund amounted to 45 million euros from 2005 to 2007 (15 million per year) and 60 million euros from 2008 to 2010 (20 million per year). From 2005 to 2009, the Piedmont plan paid 23,306,265 euros in compensation, with an average payment of 27,575 euros per claim.⁴ Between 2002 and 2004, local health unit expenditures for medical malpractice insurance cost between 5.5 and 7 euros per inhabitant per year. Once the regional fund was implemented, this cost decreased to 3.5 euros from 2005 to 2007 and 4.5 euros from 2008 to 2010.

3 Empirical Analysis

3.1 Data and Outcomes

We use a unique dataset of inpatient discharge records from the National Hospital Discharge Records (*Schede di Dimissione Ospedaliera - SDO*) on 265,532 deliveries from 2002 to 2009. For the analysis at the hospital level, we were able to collect data on neonatal outcomes for the same period through Natality Certificates (*Certificato di Assistenza al Parto*).

Our main outcome of interest is the decision to perform a C-section and we use a dummy *Csection*, which is equal to 1 if the method of delivery was a C-section and 0 otherwise. Aside from assessing the effect of higher malpractice pressure on C-sections, we need also to consider consequences for patients' health. Natural deliveries can pose relevant risks when improperly performed: mothers may suffer obstetric trauma, and newborns may suffer severe harm, namely brain damage (Jensen and Wüst, 2015). Consequently, increasing malpractice pressure could reduce the incidence of C-sections while increasing the incidence of negative maternal health outcomes. If this is the case, the policy would save money without improving obstetric practices. We thus include three outcomes: a proxy for adverse effects on mothers *Complications* and two to proxy those adverse effects on newborns *Apgar score* and *Resuscitation*.

³Despite the fact that experience rating is the norm in many lines of insurance (*e.g.*, health or car insurance), it is rarely adopted in medical malpractice insurance (Fournier and McInnes, 2001; Danzon, 2000; Weiler *et al.*, 1993). Several rationales have been used to oppose the introduction of experience-rated premiums at the individual physician-level (*e.g.*, a high variability of claims over short time periods, which makes it difficult to obtain a stable risk estimate). However, those arguments do not necessarily hold at the hospital level, and there is an ongoing debate over whether experience-rated premiums at the hospital level should be extensively adopted (Arlen, 2013; Mello, 2006; Sloan, 1990; Ellis *et. al.*, 1990).

⁴Data on claims received by hospitals before 2005 and those received by each hospital are not publicly available.

Complications is a dummy that captures the presence of delivery or post delivery problems from maternal fever to hysterectomy, as listed in Table 1. For newborns, we use dummy variables for the 5-min Apgar score below 7 *Apgar score* and an additional measure of resuscitation *Resuscitation*. Apgar scores are measured immediately after birth to determine the health of newborns and the need for resuscitation efforts. They are thus commonly used in the literature to check whether a change in the incidence of C-sections improves neonatal health outcomes (Frakes, 2012, Currie and MacLeod, 2008, Dubay *et al.*, 1999). *Apgar score* represents an Apgar score below 7, which indicates critical neonatal conditions (Casey *et al.*, 2001). Finally, the dummy *Resuscitation* controls for whether the newborn received any major (*e.g.*, ventilation) or minor (*e.g.*, aspiration) resuscitation. If the condition of both the mothers and newborns are not worse off after the introduction of an increase in malpractice pressure, this means that the policy helped to reduce inappropriate procedures at no cost to the health of the patients.

Table 1, about here

Descriptive statistics for Piedmont from 2002 to 2009 are shown in Table 2. On average, 33 out of 100 deliveries were C-sections in the entire sample, and 32 in the sample of the nearest neighborhood. These ratios can be considered high for at least two reasons. First, since 1985, the World Health Organization have argued that C-sections should not be higher than 10%-15%, and a recent goal set by Healthy People 2010 was to reduce the rate of C-sections in the US from 30% to 15% (World Health Organization). Second, significant risk factors, which would justify the use of C-sections, were reported for only 18 women out of 100 in the entire sample and for 16 for the nearest neighborhood sample.⁵ Given these numbers, we conclude that the difference between the appropriate and performed delivery types is substantial in the sample.

Overall, the majority of women were married, Italian, and delivered in their home hospitals. Approximately 17% reported complications associated with the delivery (15% for the sample of the nearest neighborhood). At the hospital level, 4 newborns out of 100 reported an Apgar score lower than 7 after 5 min (3 out of 100 for the nearest neighborhood sample). The values are similar for resuscitation efforts. The hospital-quarter mean of C-section rates is 35.4% for the entire sample and 37% for the nearest hospital sample.

Table 2, about here

3.2 Identifying the Effects of Increasing Malpractice Pressure

Hospitals operating in court districts that apply schedules face less malpractice pressure than do their neighbors in court districts that do not apply any constraints on compensation.⁶

⁵These statistics are consistent with the position of Italy in the international rankings on the use of C-sections: Italy counts the highest number of cesarean procedures in Europe, and it is among those OECD countries with the highest C-section rates (OECD, 2013; Meloni *et al.*, 2012; Ministero della Salute, 2011).

⁶We provide an indirect test of this assumption, which is based on the empirical literature on noneconomic damage caps and medical malpractice pressure, in footnote 13.

Consequently, when Piedmont implemented the 2005 experience rating policy, which in itself increased the accountability of healthcare providers, some hospitals ended up facing greater pressure than did others. The policy makes it more expensive to perform unnecessary C-sections. Our empirical analysis thus identifies the effect of an increase in the intensity of malpractice pressure triggered by the 2005 policy, which is given by the difference in the outcomes between the treated (hospitals operating in court districts without schedules) and control (hospitals operating in court districts with schedules) groups. Our identification relies on two assumptions: 1) a common trend in the outcomes of interest between the treated and control groups in the absence of the policy, and 2) the exogenous distribution of hospitals across court districts. While we test the common trend assumption in Section 4.2, to defend the exogenous distribution assumption we need to exclude at least three concerns. First, it is highly unlikely that hospitals with more medical malpractice cases have chosen to be located in a court district with schedules, operating a sort of forum shopping. In Italy, court districts were designated right after the creation of the Italian state (1861) and confirmed in 1941 by royal decree. The location of a hospital is constrained by several factors not related to schedules. For instance, the 1968 law n.132 set the minimum population size required to build a new hospital to 25,000 (Bertoli and Grembi, 2015). Before that year, hospitals' locations reflected the location of care centers that had been in place since the beginning of the previous century, quite before there was any debate on the adoption of schedules. Second, a court's decision to adopt schedules is not driven by hospitals operating in the same court district. Historically, the introduction of schedules was related to the need for compensating victims of car accidents and not injured patients.⁷ Third, there were no changes with respect to the adoption of schedules by any court district in the examined period. In Piedmont, between 2002 and 2009, 10 out of 16 courts of first instance were applying schedules to compute compensations to victims, as shown in Panel (a) of Figure 1. Courts shown in color are those that apply schedules. This figure also shows the distribution of hospitals across court districts.

Figure 1 about here

We define *Treated* as a dummy equal to 1 if the delivery took place in a hospital located in a court district without schedules and 0 otherwise. *Post05* is a dummy equal to 1 if the delivery was performed during or after 2005, when experience rating was introduced, and 0 otherwise. Finally, *Outcome_{iht}* represents either *Csection* or *Complications* for every mother *i* delivering in hospital *h* at time *t*. We estimate the following model:

$$Outcome_{iht} = \delta Treated_h * Post05_t + \gamma_t + \alpha_h + X1'_{iht}\sigma + X2'_{iht}\beta + X3'_{iht}\tau + \epsilon_{iht} \quad (1)$$

⁷The introduction of limitations on compensations is a discretionary decision of the courts, whose judges may vote in favor or against their implementation. The quantification of these ceilings is determined with both the help of medical experts and a consideration of previous cases (Sella, 2005). See Online Appendix A.

where γ_t is the year fixed effects to control for common shocks; for example, a change in the rate of reimbursement for different type of deliveries, which equally affected the treated and control groups. α_h is hospitals' fixed effects to control for unobservable characteristics at the hospital level from the number of practitioners to practice styles, and δ is the diff-in-diff estimator. The vectors of covariates represent controls for risk factors at the mother level $X1'_{iht}$, other characteristics of the mother such as her marital status $X2'_{iht}$, and socio-economic characteristics of the municipality where the mother resides $X3'_{iht}$, as described in Table 1.

Patient Flow is included in $X2'_{iht}$ and accounts for cases in which a woman delivers outside her home local health unit: through this, we control for those deliveries that represent an additional financial resource for hospitals so that we can capture the role of financial incentives among others. Hospitals belonging to the patient's local health unit are funded according to a prospective block budget based on the diagnostic-related group (DRG) system. Differently, services acquired by the local health unit from hospitals located in another local health unit are reimbursed on a pay-per-case basis (Fabbri and Robone, 2010). This means that those hospitals with higher levels of *Patient Flow* are more sensitive to the financial incentive provided by the higher reimbursement rate of C-sections versus vaginal deliveries (Francese *et al.*, 2014). Consequently, the inclusion of both year fixed effects and *Patient Flow* allow us to control for financial incentives.

We test the robustness of our specification on the sub-sample of deliveries in hospitals located in the nearest neighborhood across the court district border, as shown in Panel (b) of Figure 1. The rationale is to reduce unobservable heterogeneities between the treated and control groups and enhance the strength of the untestable common trends assumption. We define the nearest neighborhood as to include for each treated hospital the closest hospital in the control group. As a result, the nearest neighborhood comprises all hospitals within 23 km (12 miles) from the court district border.⁸ We also estimate Equation 2 at the hospital level, which is the equivalent of Equation 1 at the individual level. Using the sample of all hospitals, we estimate the following model for each of four outcome variables (*i.e.*, *Csection*, *Complications*, *Apgar*, and *Resuscitation*):

$$Outcome_{hq} = \delta Treated_h * Post05_q + \gamma_q + \alpha_h + X1'_{hq}\sigma + X2'_{hq}\beta + X3'_{ht}\tau + \epsilon_{hq} \quad (2)$$

where $Outcome_{hq}$ is the proportion of the respective outcome variable in hospital h in quarter q , which we generate using the information on the day of delivery. γ_q and α_h are the vectors of quarters' and hospitals' fixed effects, respectively. $X1$ and $X2$ are now at the quarter-hospital level. The controls in $X3$ remain at the hospital-year level because they are at the municipality level and therefore time invariant within the same year.

⁸This Euclidean distance, calculated using ArchGIS, can be translated into time by using a proxy for a speed of 50-70 km/h; 23 km amounts to a 20- to 27-min ride. Considering multiple neighborhoods around the court district border would not be meaningful, since we have six hospitals in the treated group.

4 Results

4.1 Effects of Malpractice Pressure on Obstetric Practices

Table 3 shows the results for Model 1 on the main outcomes at the patient level in Columns (1)-(4), on the whole sample and on the sample of the nearest neighborhood, estimated using a linear probability model (LPM) to facilitate the interpretation of the coefficients.⁹ Columns (5)-(10) of Table 3 report the results for Model 2 at the hospital level. We report the results when the sets of controls are added in a step-by-step manner in Online Appendix B. To cope with autocorrelation problems in the outcome variables of the diff-in-diff, the error term is clustered at the hospital level for Model 1, and at the court level for Model 2 (Bertrand *et al.*, 2004). The results include an additional robustness check: we run the same models dropping the deliveries that occurred over the weekend. One of the benefits of C-sections is that they can be scheduled during regular working days, from Monday to Friday, and during regular working hours, from 9:00 a.m. to 6:00 p.m. (Fabbri and Monfardini, 2008). Consequently, we expect that C-sections performed during weekends are most likely related to unplanned events. It is plausible that a natural delivery was expected to be performed during the weekend, but due to unforeseen complications, a C-section ended up being performed. Figure 2 shows the incidence of C-sections per day of the week, providing clear evidence of a significant drop in C-sections over the weekend.

Table 3 and Figure 2, about here

Consistent with our expectations, the increase in malpractice pressure determines a reduction of 2.3 percentage points in the utilization rate of cesarean deliveries (Column 1). This corresponds to a reduction of 7% at the sample mean of C-section (*i.e.*, 0.33) in the average C-section rate throughout the given period. The results are robust in the sample of the nearest neighborhood presented in Panel B, where the increased pressure turns out to have a stronger negative impact. The estimated decrease in the utilization rate of C-sections is 3.7 percentage points, which goes up to 4 when weekend deliveries are dropped. This estimate represents an 11.6% reduction in the average rate of cesarean deliveries in the reference sample. The results are confirmed by our robustness check on the sample that excludes weekend deliveries (Column 2), according to which the probability of receiving a C-section after the increase in malpractice pressure decreases between 2.7 (entire sample) and 4 percentage points (nearest neighborhood). This implies an average decrease of 7.5% at the sample mean of C-section without weekends (*i.e.*, 0.36) to 11.4%.¹⁰

The decrease in C-sections is not associated with a change in the incidence of complications for mothers, as apparent from both Columns (3) and (4). If C-sections were appropriate, which means due to real clinical conditions, we should observe an increase in complications

⁹We also run the same models using probit regressions; the marginal effects of the coefficients are equivalent to those we obtain with the LPM. Results are available on request.

¹⁰Results are also robust to the introduction of a hospital-year trend, as reported in Table 3B of Online Appendix B.

as C-sections are replaced by vaginal deliveries and this is not the case. No increase in *Complications* is registered in the sample of the nearest neighborhood (Panel B).¹¹

The results at the hospital level are equivalent to those at the patient level. Column (5) indicates a drop in the use of C-sections in a range between 8.5% at the sample mean (*i.e.*, 0.354) in Panel A and 15% at the sample mean (*i.e.*, 0.37) in Panel B. When weekend deliveries are excluded (Column (6)), the magnitude of the effect is stronger and shows a decrease from 10% to 18.7% at the sample mean. Complications are once again not statistically significant, as shown by Columns (7) and (8). Finally, Columns (9) and (10) report the estimates of the effect on neonatal outcomes. These findings suggest that there was no worsening in the health outcomes of newborns. There is no information on the delivery day for these outcomes, which makes it impossible to perform a further check that excludes weekend deliveries.

4.2 Validity Tests

Our results rely on the identifying assumption that the treated and control groups would follow a common trend in the outcomes of interest in the absence of the treatment. To test the common trend assumption, we use pre-treatment data and estimate the model described in Equation 3. This model introduces the leads and lags of the treatment and includes placebo tests of the real effect of the policy since it tests whether a common trend in the incidence of C-sections existed before the introduction of experience rating. If this holds, then the coefficients of the q leads should not be statistically different from zero. It also tests whether the policy was more or less effective in the immediate years of its implementation or later on, with the coefficients of the m lags expected to be statistically different from zero for different points in time. Through Equation 3, we can address the validity of our results against anticipatory effects on C-sections and provide a sharper assessment of the post-treatment effects.

$$Csection_{iht} = \sum_{t=-q}^{-1} Treated_h * D_t + \sum_{t=0}^m Treated_h * D_t + \gamma_t + \alpha_h + X1'_{iht}\sigma + X2'_{iht}\beta + X3'_{iht}\tau + \epsilon_{iht} \quad (3)$$

Figure 3 plots the results for Equation 3: each dot represents the estimated coefficient of a lead or lag and its relative 95% confidence interval. Our evidence supports a common trend or no anticipatory effects. Leads coefficients are not statistically different from zero, whereas lags coefficients are statistically significant starting from the year after the adoption of experience-rated insurance for medical liability, a reasonable time for hospitals to build up an incentive structure for their physicians.

Figure 3, about here

¹¹For a further check on complications for mothers which is more related to the precaution channel, see Online Appendix C.

4.3 Disentangling the Channels

The effect on the decision to perform a C-section triggered by the adoption of experience rating could be channeled by two mechanisms. On the one hand, healthcare providers could select their patients, thus imposing a form of cream skimming. Physicians may direct patients to the nearest hospital whenever they show substantial risk factors, implementing a so-called negative defensive medicine: the incidence of C-sections drops because the treated hospitals deal with more low-risk patients. This could be a sensible explanation, especially in the subsample of the nearest neighborhood. If the adoption of experience rating affects the decision to perform C-sections through this first channel, we would see a change in the probability of having a C-section as a function of risk factors, defined by the literature as the predicted probability of a C-section (Frakes, 2012 and 2013; Baiker *et al.*, 2006). The direction of the change in this channel should have the same sign of the effect observed on C-sections.¹²

Alternatively (or additionally), a rise in malpractice pressure may affect nonmedical rationales for performing C-sections (*e.g.*, time saving). The second channel represents the discretionary component in the choice of the delivery type. If the policy works mainly through this channel, it means that factors other than the patient’s medical condition become less important in shaping the delivery decisions of doctors, leading to a reduction in C-sections. We proxy the two channels through the models described in Equations 4 and 5.

$$Csection_{iht} = X1'_{iht}\sigma + X2'_{iht}\beta + \epsilon_{iht} \quad (4)$$

Equation 4 defines the probability of performing a C-section as a function of risk factors as grouped by $X1'_{iht}$ and $X2'_{iht}$. Using Equation 4, we derive 1) the predicted probability of receiving a C-section PPC_{iht} and 2) the residual $\hat{\epsilon}_{iht}$. These represent proxies for the first (PPC_{iht}) and second channel ($\hat{\epsilon}_{iht}$). Then, we test whether the reform had any impact on these measures using Equation 5. The parameters of interest are ρ and ϕ . From the sign and the significance of ρ and ϕ , we obtain a better view of the mechanisms in place once the reform is introduced.

$$Channels = \begin{cases} PPC_{iht} & = \rho Treated_h * Post05_t + \gamma_t + \alpha_h + \eta_{iht} \\ \hat{\epsilon}_{iht} & = \phi Treated_h * Post05_t + \theta_t + \lambda_h + \mu_{iht} \end{cases} \quad (5)$$

Table 4 shows the estimation results for Equation 5. There is no evidence of patient cream skimming. Interestingly, an increase in malpractice pressure does affect nonmedical rationales and thus reduces the discretionary component of the decision to perform a C-section. The

¹²Another check of patient cream skimming could be done on the average distance between the residence municipality of patients and the hospital municipality. On this measure, we run two checks. Using information on the municipality in which mothers reside, we calculate the average distance traveled by mothers delivering in a hospital in the treated group (14.41 km or 8 miles) as compared to the average distance they covered when delivering in a hospital in the control group (14.68 km or 9 miles). The difference is not statistically different from zero. Then, in the spirit of disentangling the channels, we use the average distance as an outcome of Equation 2. Results are shown in Table 4B; the treatment does not affect this measure.

magnitude of the effect is 7% of the average residuals in absolute terms on the entire sample and 11% in the nearest neighborhood sample.¹³

Table 4, about here

5 Concluding Remarks

We assess the impact of an increase in malpractice pressure on obstetric practices. Taking advantage of a pre-existing difference in the malpractice exposure of different healthcare providers, due to the possibility of courts' limiting noneconomic damages, we use a diff-in-diff strategy to evaluate the effects of the adoption of experience-rated insurance on the decision to perform C-sections. Our diff-in-diff results are also robust to a more local estimation run in a selected sample of hospitals. This sub-sample includes only the treated hospitals and their nearest control hospitals located directly across the geographical border of the court district. Our analysis shows that an increase in malpractice pressure decreases the average incidence of C-sections by 7 to 11.6%. Therefore, increasing malpractice pressure can reduce the use of inappropriate procedures.

Our results are robust to the inclusion of controls for financial incentives, once we include a measure for patient flow, to a drop in deliveries on weekends, when the occurrence of a C-section is far more likely related to an emergency procedure, and to controls for the hospitals-year trends. Finally, we investigate the possible channels of the detected effects. We prove that no patient selection mechanism (*i.e.*, negative defensive medicine) has been triggered by experience rating, though we do detect a decrease in discretion in the decision to perform a surgical procedure that is not related to medical factors.

Finally, our strongest and most robust finding is that this increase in malpractice pressure leads to a reduction in C-sections without worsening the health conditions of mothers and newborns. Physicians can be sensitive to policies implemented at the hospital level, and we do find an impact on treatment decisions made by physicians in such a setting. When properly structured, policies implemented at the hospital level may reduce unnecessary health care costs.

¹³In Table 5B of Online Appendix B, we present the results of a heterogeneity check on the differences between high- and low-level schedules. The rationale for this check is to provide indirect proof of the assumption that hospitals in court districts with schedules are facing less malpractice pressure compared to those in court districts without schedules. The starting assumption is that schedules reduce malpractice pressure because they decrease unpredictability in compensations. This has several advantages: for instance it is easier to find insurers once schedules are adopted (Bertoli and Grembi, 2013). However, if scheduled damages are very high, the decrease in pressure due to the high predictability is offset by the high awards. Compared to hospitals operating in low-schedule courts, those operating in high-schedule courts face more pressure. Consequently, we should see that the magnitude of the effect when we consider only high-schedule courts among the controls should be lower than when we consider low-schedule courts. Results in Table 5B confirm these expectations.

References

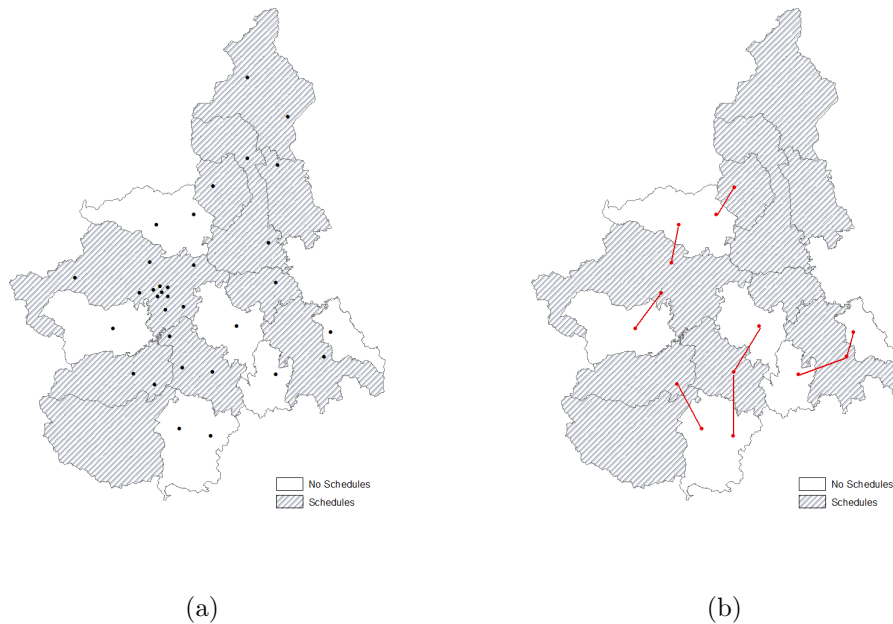
- Amaral-Garcia, S. and V. Grembi (2014). Curb Your Premium: The impact of monitoring malpractice claims, *Health Policy* 114: 139-146.
- Arlen, J. (2013). Economic Analysis of Medical Malpractice Liability and Its Reform. In: *Research Handbook on the Economic Analysis of Torts* (Jennifer Arlen, ed., Edward Elgar Press).
- Baicker, K., K. Buckles, and A. Chandra (2006). Geographic Variation in the Appropriate Use of Cesarean Delivery. *Health Affairs* 25:w355-w367.
- Bertrand, M., E. Dufflo, and S. Mullainathan, (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119 (1): 249-275.
- Bertoli, P. and V. Grembi (2013). Courts, Scheduled Damages, and Medical Malpractice Insurance, Baffi Center Research Paper No. 2013-145.
- Bertoli, P., and V. Grembi (2015). Life Saving Effects of Hospital Proximity, mimeo.
- Bovbjerg, R. R., Sloan F. A. and J. F. Blumstein (1989). Valuing Life and Limb in Tort: Scheduling "Pain and Suffering", *Northwestern University Law Review* 83(4), pp. 908-976.
- Casey, B. M., D. D. McIntire, and K. J. Leveno, (2001). The continuing value of the Apgar score for the assessment of newborn infants. *The New England Journal of Medicine*, 344(7): 467-471.
- 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 (2013). Diagnosis and Unnecessary Procedure Use: Evidence from C-sections. NBER Working Papers.
- Danzon, P. M. (2000). Liability for Medical Malpractice, *Handbook of Health Economics*, in A. J. Culyer and J. P. Newhouse (ed.), *Handbook of Health Economics*, Chapter 26, pp. 1339-1404, Elsevier.
- Declercq, E., F. Menacker and M. MacDorman (2006). Maternal risk profiles and the primary cesarean rate in the United States, 1991-2002. *American Journal of Public Health* 96(5), 867-872.
- Dranove, D., R. Subramaniam 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., R. Kaestner and T. Waidmann (1999). The Impact of Malpractice Fears on Cesarean Section Rates. *Journal of Health Economics* 18: 491-522.
- Dubay, L., R. Kaestner 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., C. L. Gallup and T. G. McGuire (1990). Should medical professional liability insurance be experience rated? *Journal of Risk and Insurance* 57(1): 66-78.
- Fabbri, D. and S. Robone (2010). The Geography of Hospital Admission in a National Health Service with Patient Choice, *Health Economics* 19(9): 1029-1047.
- Fabbri, D. and C. Monfardini (2008). Style of Practice and Assortative Mating: a Recursive Probit Analysis of Cesarean Section Scheduling in Italy. *Applied Economics*, 40: 1411-1423.
- Fenn, P., A. Gray and N. Rickman (2007). Liability, Insurance, and Medical Practice. *Journal of Health Economics* 26: 1057-1070.
- Fenn, P., A. Gray, N. Rickman, D. Vencappa 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. (2015). The Surprising Relevance of Medical Malpractice Law 82 *University Of Chicago Law Review*, 82(1): 317-391.
- 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. (2012). Defensive Medicine and Obstetric Practice. *Journal of Empirical Legal Studies* 9(3): 457-481.
- Francese M., M. Piacenza, M. Romanelli, and G. Turati (2014). Understanding inappropriateness in health spending: The role of regional policies and institutions in caesarean deliveries. *Regional Science and Urban Economics* 49: 262-277.
- 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.
- Jensen, V. M. and M. Wüst (2015). Can Caesarean section improve child and maternal health? The case of breech babies. *Journal of Health Economics*, 39: 289-302.
- Kachalia, A. and M. M. Mello (2011). New Directions in Medical Liability Reform. *New England Journal of Medicine* 364: 1564-1572.
- Kessler, D. P. and M. McClellan (1996). Do Doctors Practice Defensive Medicine? *Quarterly Journal of Economics* 111(2): 353-390.
- Localio, A. R., A. Lawthers, J. Begtson, L. Hebert, S. Weaver, T. Brennan and J. R. Landis (1993). Relationship Between Malpractice Claims and Cesarean Delivery. *Journal of the American Medical Association* 269: 366-373.
- Meloni *et al.* (2012). The role of caesarean section in modern Obstetrics. *Journal of Pediatric 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.

- Ministero della Salute (2013). Certificato di assistenza al parto (CeDAP), Analisi dell'evento nascita - Anno 2010.
- Ministero della Salute (2011). Taglio cesareo: una scelta appropriata e consapevole. Available at: http://www.snlg-iss.it/cms/files/LG_Cesareo_finaleL.pdf.
- OECD (2013). Health at a Glance 2013: OECD Indicators, OECD Publishing. Available at: <http://www.oecd.org/health/health-systems/health-at-a-glance.htm>.
- OECD (2014). Health Statistics 2014 - Frequently Requested Data. Available at: <http://www.oecd.org/els/health-systems/oecd-health-statistics-2014-frequently-requested-data.htm>.
- Sachs, B. (1989) Is the Rising Rate of Cesarean Sections a Result of More Defensive Medicine? In: Medical Professional Liability and the Delivery of Obstetrical Care. Volume II, 2740. An Interdisciplinary Review, Victoria P. Rostow and Roger J. Bulger (eds). Washington, D.C.: National Academy Press.
- 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., S. S. Entman, B. A. Reilly, C. A. Glass, G. B. Hickson and H. H. Zhang (1997). Tort Liability and Obstetricians' Care Level. *International Review of Law and Economics* 17: 245-260.
- Studdert, D.M. and M. M. Mello (2005). Options for Rational Scheduling and Valuation of Noneconomic Damages, Report to the Washington State Noneconomic Damages Task Force.
- Weiler, P. C., H. H. Hiatt, J. P. Newhouse, W. G. Johnson, T. Brennan and L. L. Leape (1993). A Measure of Malpractice: Medical Injury, Malpractice Litigation and Patient Compensation. Cambridge, Harvard University Press.
- World Health Organization (1985). Appropriate technology for birth. *Lancet* 2(8452): 436-437.

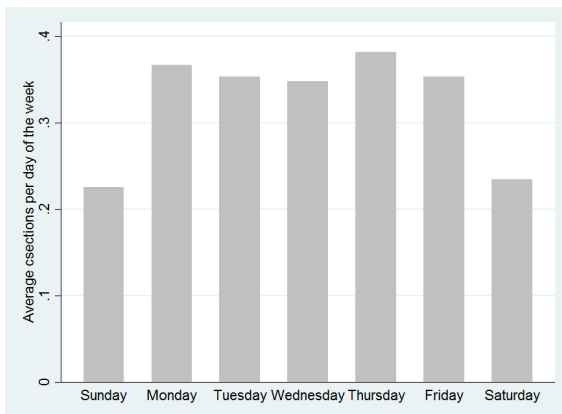
Tables and Figures

Figure 1: Schedules Application and Hospitals Distribution (2002-2009)



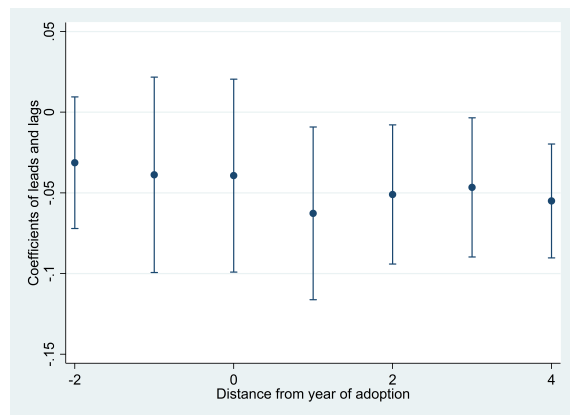
Notes: (a) Court districts' borders in black. White areas identify court districts that do not apply schedules of noneconomic damages. Grey striped areas identify court districts that apply schedules of noneconomic damages. Black dots represent the hospitals located in Piedmont. (b) Red dots represent treated hospitals and the closest hospitals in the control group.

Figure 2: C-sections per weekday



Note: Average proportion of C-sections per weekday.

Figure 3: Common Trend (Leads & Lags)



Note: Leads and lags coefficients controlling for X_1 , X_2 , X_3 , and excluding weekend deliveries.

Table 1: Outcomes and Controls

Outcomes		Controls			
	Mother	Babies ^b	X1 ^a	X2 ^a	X3 ^c
Delivery	Complications^a		Mother Level	Mother Level	Mother Municipality Level
C-section	Maternal fever	Apgar < 7	Previous C-section	Age	Average Income
	Maternal seizures	Resuscitation	Multiple Babies	Nationality	Sea level
	Rupture of the uterus		Breech	Marital status	Level of urbanization
	Coagulation problems		Diabetes	Patient flow	Education
	Hysterectomy		Eclampsia		
	Prolonged labor		Placenta		
	Dysfunctional labor		Disproportion		
	Anesthetic complications		Cancer		
	Fetal distress		Prolong pregnancy		
	Hemorrhage		Early-onset delivery		
	Perineal laceration		Cord prolapse		
	Cardiac complications		Use of alcohol/drug		
	Decompression sickness		Cephalopelvic disproportion		
	Renal failure		Hypertension		
	Embolism		Cardiac or lung conditions		
	Maternal distress		Obesity		
	Lung complications		Anemia		
	Other		RH Problems		
			Sexual transmitted diseases		
			Renal failure		

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). *Other* includes possible complications that differ from those explicitly mentioned, such as retained placenta and perineal hematomas among others. *Patient Flow* identifies deliveries occurring in a hospital not managed by the mother's enrolling local health unit. (b) These data are only available at the hospital-level through the Natality Certificates (CEDAP). (c) The controls at the mother's municipality level are consistent with those used at the county-level by Baicker *et al.* (2006). We include *Sea level* due to the territorial characteristics of Piedmont which has many mountain municipalities.

Table 2: Means and Standard Deviations of Key Variables

	Entire Sample			Nearest Neighborhood		
	All Deliveries	Treated	Control	All Deliveries	Treated	Control
Outcomes						
<i>C-section</i>	0.330 (-0.47)	0.336 (-0.472)	0.329 (-0.47)	0.317 (-0.465)	0.336 (-0.472)	0.301 (-0.459)
<i>Complications</i>	0.172 (-0.378)	0.129 (-0.336)	0.179 (-0.384)	0.148 (-0.355)	0.129 (-0.336)	0.164 (-0.37)
<i>Apgar < 7</i>	0.039 (-0.034)	0.026 (-0.015)	0.041 (-0.035)	0.031 (-0.016)	0.026 (-0.015)	0.035 (-0.015)
<i>Resuscitation</i>	0.036 (-0.028)	0.023 (-0.012)	0.039 (-0.029)	0.029 (-0.028)	0.023 (-0.012)	0.033 (-0.035)
Controls at the mother level						
<i>Risk Factors</i>	0.183 (-0.387)	0.172 (-0.377)	0.185 (-0.388)	0.162 (-0.368)	0.172 (-0.377)	0.153 (-0.36)
<i>Age (years)</i>	31.5 (-5.077)	31 (-5.197)	31.5 (-5.053)	31.1 (-5.103)	31 (-5.197)	31.2 (-5.023)
<i>Nationality</i>	0.837 (-0.369)	0.824 (-0.381)	0.839 (-0.367)	0.843 (-0.363)	0.824 (-0.381)	0.86 (-0.347)
<i>Marital Status</i>	0.669 (-0.47)	0.667 (-0.471)	0.67 (-0.47)	0.652 (-0.476)	0.667 (-0.471)	0.639 (-0.48)
<i>Patient flow</i>	0.365 (-0.481)	0.166 (-0.372)	0.398 (-0.489)	0.174 (-0.379)	0.166 (-0.372)	0.181 (-0.385)
Controls at the mother municipality level						
<i>Income (2012 euro)</i>	21,416 (2,846)	20,233 (2,335)	21,608 (2,875)	20,470 (2,480)	20,233 (2,335)	20,667 (2,579)
<i>Education</i>	0.063 (-0.029)	0.053 (-0.023)	0.065 (-0.03)	0.052 (-0.021)	0.053 (-0.023)	0.05 (-0.02)
<i>Low Level of Urbanization</i>	0.156 (-0.363)	0.256 (-0.436)	0.14 (-0.346)	0.214 (-0.41)	0.256 (-0.436)	0.18 (-0.384)
<i>Medium Level of Urbanization</i>	0.461 (-0.498)	0.725 (-0.447)	0.418 (-0.493)	0.621 (-0.485)	0.725 (-0.447)	0.534 (-0.499)
<i>High Level of Urbanization</i>	0.383 (-0.486)	0.02 (-0.138)	0.442 (-0.497)	0.165 (-0.371)	0.019 (-0.138)	0.286 (-0.452)
<i>Sea level (meters)</i>	687.525 (-609.158)	746.798 (-700.62)	677.871 (-592.367)	762.88 (-696.117)	746.798 (-700.62)	776.268 (-692.068)
Observations	265,532	37,190	228,342	81,864	37,192	44,672

Notes: Standard deviations in parentheses. *Apgar < 7* and *Resuscitation* are available through Natality Certificates (CEDAP) and are available only at the hospital-level (total number of observations is 1,054). *Risk Factors* captures the incidence of risk factors as described by *X1* in Table 1. *Nationality* is equal to 1 if the mother is Italian. *Patient flow* is equal to 1 if the hospital where the delivery occurred is not managed by the mother's local health unit. Variables at the mother-level are available through the Patient Discharge Records from the Ministry of Health. *Education* is the share of municipal residents with a college degree (2001 Census data). *Level of Urbanization* captures both population density per square kilometer and the municipality dimension. *High Level of Urbanization* refers to all municipalities with more than 500 inhabitants per square kilometer and at least 50,000 residents. *Medium Level of Urbanization* refers to all municipalities with more than 100 inhabitants per square kilometer that are located nearby a municipality with a high degree of urbanization or that have more than 50,000 residents. *Low Level of Urbanization* corresponds to what is not characterized by a *High Level of Urbanization* or a *Medium Level of Urbanization*. Variables at the mother's municipality level are provided by the Italian National Institute of Statistic. *Income* is available on an annual basis, while the other variables in *X3* are available as measured in the 2001 Census data.

Table 3: Estimated Effects on C-Sections, Complications, and Neonatal Outcomes

	Patient-Level				Hospital-level							
	Mothers Outcomes		Complications		Mothers Outcomes		Complications(r)		Babies Outcomes			
	C-sections	No Weekend (2)	All Weekend (3)	No Weekend (4)	C-sections(r)	All Weekend (5)	No Weekend (6)	All Weekend (7)	Complications(r)	All Weekend (8)	Apgar <7(r)	Resuscitation(r)
Panel A: Entire Sample												
δ	-0.023** (0.010)	-0.027** (0.011)	-0.041 (0.033)	-0.036 (0.032)	-0.030** (0.013)	-0.033** (0.013)	-0.029 (0.024)	-0.028 (0.025)	-0.008 (0.013)	-0.003 (0.005)		
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	265,532	202,273	265,532	202,273	1,054	1,054	1,054	1,054	1,054	1,054	1,054	1,023
Panel B: Nearest Neighborhood Sample												
δ	-0.037*** (0.015)	-0.040*** (0.016)	-0.043 (0.063)	-0.036 (0.060)	-0.055*** (0.017)	-0.060*** (0.017)	-0.041 (0.034)	-0.036 (0.036)	-0.009 (0.007)	-0.003 (0.013)		
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	81,864	62,043	81,864	62,043	447	447	447	447	447	447	447	432

Notes: δ is the coefficient of *Treated*Post05*. *Controls* are listed in Table 1. Columns (2), (4), (6) and (8) use the sample of deliveries performed from Monday to Friday. *Patient Level*: Results from linear probability model regressions. All regressions include years and hospitals fixed effects; *Controls* are listed in Table 1. Robust standard errors clustered at the hospital-level in parenthesis. *Hospital Level*: OLS regressions. r : the measure is a ratio out of the total number of deliveries per period. All regressions include both quarters and hospitals fixed effects; Robust standard errors clustered at the court-level in parenthesis. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 4: **The Channels: Patient Selection and Non-medical Cesareans**

Coefficient	Predicted Probability of C-Section	Residuals
Panel A: Entire Sample		
<i>Treated * Post05</i>	-0.008 (0.011)	-0.024** (0.010)
Year FE	Yes	Yes
Hospitals FE	Yes	Yes
Observations	265,532	265,532
Panel B: Nearest Neighborhood Sample		
<i>Treated * Post05</i>	-0.016 (0.016)	-0.035** (0.014)
Year FE	Yes	Yes
Hospitals FE	Yes	Yes
Observations	81,864	81,864

Notes: Results are from linear probability model regressions. Predicted probability of performing a C-section on the basis of risk factors; and the related residuals coincide with *Residuals* (equation 5). Robust standard errors clustered at the hospital-level in parenthesis. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Online Appendixes

Appendix A: Damages Schedules in Italy

Noneconomic damages for personal injuries resulting from non-criminal acts are a recent possibility in the Italian legal system. Legal scholars began to debate the relevance of this practice during the Sixties¹ and, until then, victims of personal injuries had the right to receive noneconomic damages in addition to economic damages *only if* their injuries had been caused by a criminal act.

At the beginning of the Eighties (1981-1984), the Court of Cassation intervened as the court of last appeal for both criminal and civil jurisdictions, ruling that it would be possible to receive noneconomic damages in non-criminal cases (Scarso, 2009). Once this right to noneconomic damages had been granted for all types of legal cases, the main issue became the assessment of damages, given the lack of economic guidelines that courts could follow. The need to provide guidance to judges with respect to the quantification of noneconomic losses was further stressed by a sharp rise in the number of injuries resulting from car accidents, which highlighted a high variability in victims' awards even within the same court.

Italian courts followed the experience of other European countries, such as France, and started to develop and adopt damages schedules. The aim was to reduce the variability of awards while guaranteeing both horizontal equity and vertical inequality (*i.e.*, victims suffering higher levels of injury should receive higher damages; and victims suffering similar injuries should receive an equivalent indemnity amount). In particular, Italian courts elaborated the criteria for evaluation by points, and a subsequent formula based on injury severity and victim age to quantify noneconomic damages. As a result, schedules are constructed as matrices with entries for the injury severity level and victim's age. A monetary value is assigned to every given level of permanent harm, and an age-related coefficient is assigned to any given age. Thus, different combinations of age and injury severity result in distinct compensation amounts (Table 1A).

Monetary values vary unevenly and increase more rapidly with injury severity, so that the greater the harm suffered by the victim (*i.e.*, the percentage points of disability), the higher the monetary value to be awarded. Differently, the rationale for the age-related coefficients is that a victim harmed at a younger age will bear the consequences of her physical impairment for a longer period of time than an older victim would.²

During this process, disability percentage points were defined by medical experts, who assigned different points to each possible level of harm (Comandè, 2005). Moreover, in order to guarantee consistency within courts' decisions, monetary values were defined according to previous cases (Sella, 2005). In practice, schedules impose a cap on the amount of damages that a victim can recover. Hence, the introduction of schedules was expected to increase certainty, but not to decrease deterrence with respect to the past.

¹Yet, a 1967 decision of the Court of First Instance in Florence refused to grant compensation to a 70-year-old retired victim on the basis that *there might be people without value according to the law* (*Nuovi orientamenti per la determinazione del danno* 1989).

²For more information on the development and history of damages schedules in Italy, see Bertoli (2014).

Table 1A: Example of Schedules

			Age	
			1	32
			Age Index	
Point value			1.000	0.845
Disability	25%	4,386	109,650	92,654

Notes: Values are expressed in 2012 euro and taken from the reference table adopted by the Court of Turin in 2009. In the case of a 25% disability suffered by a 1 years old victim (or younger), the reference noneconomic compensation amounts to 109,650 euros. This value is obtained by multiplying the monetary percentage point value (4,386 euros) by 25 and by the age index (1.000). Differently, if the same disability is suffered by a 32-year-old victim, the reference compensation amounts to 92,654 euros.

In 1986, the Constitutional Court ruled that schedules (*tabelle per il danno biologico*) were constitutional for setting noneconomic damages.³ Nevertheless, the adoption of schedules was voluntary and some courts did not adopt them, thus leaving the quantification of damages at the discretion of judges and thereby not imposing any cap. The decision to adopt scheduled damages depends on the judges practicing in a given court: judges should vote on the implementation of schedules. If adopted, schedules must be applied to compute damages for all types of injuries (*e.g.*, injuries resulting from car accidents, work accidents, or medical malpractice).

Finally, considering that schedules apply to the entire civil system and their conception aimed to deal with the difficulties of assessing damages for the victims of motor vehicle accidents, it is clear that the decision to implement them is exogenous with respect to the phenomenon of medical malpractice.

³In the sentence n.184/1986, for the assessment of noneconomic damages, the Constitutional Court recognizes the validity of schedules by identifying their fundamental elements. Specifically, these elements are (i) the specification of monetary values of general application, that is, values that can be applied to any case of personal injury, and (ii) the possibility to adapt these values according to the severity of the injury suffered by the victim.

Appendix B: Additional Tables

This Appendix provides additional information, which is also discussed in the paper. In particular, we present:

- the results of Equations 1 and 2 without controls and adding vectors of controls in a stepwise manner (Tables 1B and 2B);
- the robustness check adding a hospital-year trend to Equation 1 (Table 3B);
- the results of using the average distance between the patient residence and the delivery hospital as an outcome of Equation 1 (Table 4B);
- the heterogeneity using the low and high level of schedules (Table 5B).

Table 1B: Results at the Patient-level

	All Deliveries			Vaginal Deliveries					
	C-sections	Complications		Complications					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Entire Sample									
δ	-0.032*** (0.012)	-0.022** (0.010)	-0.023** (0.010)	-0.038 (0.031)	-0.041 (0.033)	-0.041 (0.033)	-0.058 (0.047)	-0.060 (0.049)	-0.060 (0.048)
X1	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
X2	No	No	Yes	No	No	Yes	No	No	Yes
X3	No	No	No	No	No	No	No	No	No
Observations	265,537	265,537	265,537	265,537	265,537	265,537	177,961	177,961	177,961
Panel B: Nearest Neighborhood Sample									
δ	-0.051*** (0.009)	-0.038** (0.015)	-0.037** (0.015)	-0.039 (0.059)	-0.042 (0.064)	-0.043 (0.063)	-0.065 (0.079)	-0.067 (0.082)	-0.068 (0.082)
X1	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
X2	No	No	Yes	No	No	Yes	No	No	Yes
X3	No	No	No	No	No	No	No	No	No
Observations	81,865	81,865	81,865	81,865	81,865	81,865	55,923	55,923	55,923

Notes: Results are from linear probability model regressions. δ is the coefficient of *Treated * Post05*. All regressions include years and hospitals fixed effects; and the complete set of covariates, X1, X2, and X3 are listed in Table 1. Robust standard errors clustered at the hospital-level in parenthesis. The specifications estimated in columns 7, 8 and 9 use the sample of vaginal deliveries. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 2B: Results at the Hospital-level

		Mother Outcomes			Vaginal Deliveries			Babies Outcomes		
		All Deliveries			All Deliveries			All Deliveries		
		Complications (r)			Complications (r)			Apgar <7 (r) Resuscitation (r)		
		All	No	All	All	No	All	All	All	All
		Deliveries	Weekend	Deliveries	Weekend	Deliveries	Weekend	Deliveries	Deliveries	Deliveries
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(7)	(8)
C-sections (r)										
All Deliveries										
δ	-0.033** (0.014)	-0.035** (0.013)	-0.041 (0.028)	-0.37 (0.027)	-0.067 (0.046)	-0.065 (0.049)	-0.008 (0.012)	0.000 (0.006)		
<i>Controls</i>	No	No	No	No	No	No	No	No	No	No
Observations	1,054	1,054	1,054	1,054	1,054	1,054	1,054	1,054	1,054	1,023
Panel B: Nearest Neighborhood Sample										
δ	-0.060*** (0.016)	-0.063*** (0.014)	-0.040 (0.048)	-0.035 (0.046)	-0.067 (0.070)	-0.059 (0.069)	-0.010 (0.009)	-0.001 (0.016)		
<i>Controls</i>	No	No	No	No	No	No	No	No	No	No
Observations	447	447	447	447	447	447	447	447	447	432

Notes: Results are from OLS regressions. δ is the coefficient of *Treated * Postf05*. All regressions include quarters and hospitals fixed effects. The complete set of covariates, X1, X2, and X3, listed in Table 1, has been computed as hospital-quarter incidence ratios (*e.g.*, percentage of Italian mothers out of the total of mothers per quarter/hospital). Columns 2-4-6 use the sample of deliveries performed from Monday to Friday. The specifications estimated in columns 5 and 6 use the sample of vaginal deliveries. Robust standard errors clustered at the court-level in parenthesis. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 3B: Robustness Checks - C-sections

	C-sections	
	All Deliveries	No Weekend
Panel A: Entire Sample		
δ	-0.030*** (0.000)	-0.081*** (0.000)
Years FE	Yes	Yes
Hospital FE	Yes	Yes
Year*Hospital	Yes	Yes
X1	Yes	Yes
X2	Yes	Yes
X3	Yes	Yes
Observations	265,537	202,278
Panel B: Nearest Neighborhood Sample		
δ	-0.095*** (0.000)	-0.265*** (0.000)
Years FE	Yes	Yes
Hospital FE	Yes	Yes
Year*Hospital	Yes	Yes
X1	Yes	Yes
X2	Yes	Yes
X3	Yes	Yes
Observations	81,865	62,044

Notes: Results are from linear probability model regressions. δ is the coefficient of *Treated * Post05*. All regressions include the complete set of covariates, *X1*, *X2*, and *X3*, listed in Table 1. Robust standard errors clustered at the hospital-level in parenthesis. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 4B: Distances

	All deliveries (1)	No weekend (2)
Panel A: Entire Sample		
δ	-0.237 (0.206)	-0.154 (0.221)
X1	Yes	Yes
X2	Yes	Yes
X3	Yes	Yes
Obs	265,532	202,273
Panel B: Nearest Neighborhood Sample		
δ	0.086 (0.252)	0.202 (0.267)
X1	Yes	Yes
X2	Yes	Yes
X3	Yes	Yes
Obs	81,864	62,043

Note: The outcome variable is patient distance in km. δ is the coefficient of $Treated * Post05$. All regressions include the complete set of covariates, X1, X2, and X3, listed in Table 1. Column 2 uses the sample of deliveries performed from Monday to Friday. Robust standard errors clustered at the hospital level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5B: Heterogeneity

	C-sections
Panel A: Entire Sample	
Low schedule	-0.133*** (0.009)
High schedules	-0.024* (0.011)
<i>Difference between the two subsamples</i>	0.109*** (0.008)
Observations	202,273
Panel B: Nearest Neighborhood Sample	
Low schedule	-0.150*** (0.012)
High schedule	-0.037* (0.016)
<i>Difference between the two subsamples</i>	0.113*** (0.009)
Observations	62,043

Notes: Results are from linear probability model regressions, excluding weekend deliveries. All regressions include years and hospitals fixed effects; and the complete set of covariates, $X1$, $X2$, and $X3$, listed in Table 1. *High schedule* is a dummy equal to 1 if the schedule adopted by the reference court foresees a monetary value for a 25% disability to be higher than 4,500 euros. Robust standard errors clustered at the hospital-level in parenthesis. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Appendix C: Alternative Definition of *Complications*

In this Appendix, we consider an alternative measure of *Complications* and compute Patient Safety Indicators (PSIs) as suggested by the Agency for Healthcare Research and Quality (AHRQ, 2003). These indicators, based on ICD-9-CM codes for diagnosis and procedures, have the benefit of identifying potentially *preventable* complications or medical errors that occur in specific types of deliveries.⁴ As far as preventable injuries to the mother in vaginal deliveries are concerned, the AHRQ recommends two PSIs: obstetric traumas to the mother in the sample of vaginal deliveries with instruments (PSI 18) and obstetric traumas to the mother in the sample of vaginal deliveries without instruments (PSI 19).⁵ Therefore, both flag potential preventable traumas during vaginal delivery, and the difference is the population at risk for the adverse event (*i.e.*, vaginal deliveries with and without instrument). The AHRQ considers only cases of adverse events that can develop during a hospital stay, and that can generally be identified through the secondary code of diagnosis or procedure. This ensures that the adverse events being flagged were developed in the hospital instead of being present already at admission.

Table 1C shows the list of traumas included in PSI 18 and PSI 19, and Table 2C reports the descriptive statistics of these outcomes. There are no relevant differences between the treated and control groups.

Table 1C: Patient Safety Indicators (PSIs)

Vaginal Deliveries with Instruments (18)	Vaginal Deliveries without Instruments (19)
Obstetric traumas	Obstetric traumas
Perineal lacerations	Perineal lacerations
Vaginal traumas	Vaginal traumas
Perineal traumas	Perineal traumas

Note: Patient safety indicators (PSIs) are defined according to the classification proposed by the Agency for Healthcare Research and Quality (AHRQ). *Instruments* refers to the use of forceps or vacuum during the delivery.

Results of the estimation of Equation 1 using PSIs 18 and 19 as outcomes at the patient level are shown in Table 3C.⁶ The results for the entire sample (Panel A) show a reduction

⁴See AHRQ (2003) for more details. Iizuka (2013) describes how these indicators have been developed and tests the impact of medical malpractice-related policies on PSIs using data from the US.

⁵Differently from Iizuka (2013), we do not include obstetric traumas in cesarean deliveries. This category can be highly affected by emergency C-sections, meaning deliveries that began as natural ones but ended up as cesareans.

⁶We estimated the same specifications using a negative binomial model, and the results are consistent

Table 2C: Means and Standard Deviations of PSIs

	Entire Sample			Nearest Neighborhood		
	Treated	Control	Vaginal Deliveries	Treated	Control	Vaginal Deliveries
Outcomes						
<i>PSIs 18 and 19</i>	0.005 (0.070)	0.004 (0.066)	0.006 (0.080)	0.005 (0.070)	0.004 (0.061)	0.006 (0.078)
<i>PSI 19</i>	0.005 (0.069)	0.004 (0.065)	0.006 (0.080)	0.005 (0.069)	0.004 (0.061)	0.006 (0.078)
Observations	37,192	228,345	177,961	37,192	44,673	55,923

Notes: Standard deviations are in parentheses. *PSIs 18 and 19* are calculated with respect to the sample of all vaginal deliveries, while *PSI 19* is calculated with respect to the sample of vaginal deliveries without instruments.

in preventable adverse events, but the results are not statistically significant. If we consider the results in the subsample of the nearest neighborhood (Panel B), we can see a statistically significant reduction in preventable adverse events. According to these numbers, increasing providers' accountability reduces preventable complications in the sample of vaginal deliveries. Overall, the results at the hospital level tend to be in line with those obtained at the patient level and show higher levels of statistical significance (Table 4C). We find evidence in favor of the deterrent effect at the hospital level: higher malpractice pressure tends to decrease preventable medical complications for the vaginal procedures that we discussed. Similarly to Currie and MacLeod (2008) and Iizuka (2013) these results provide support to the claim that complications can be prevented by physicians' effort. In our setting, physicians' effort is influenced by the incentives provided at the hospital level.

Yet, we have to remember that two channels may be in place for this effect: an increase in precaution when a vaginal delivery is performed; or a change in the sample of women delivering vaginally (selection bias). In fact, the composition of the resulting vaginal sample might be affected. Consequently, the incidence of preventable complications may decrease not only because precaution levels have increased, but also because low-risk mothers might be transferred from the cesarean-delivery to the vaginal-delivery sample. Sample selection concerns are, however, mitigated by the fact that we find a reduction of C-sections (*i.e.*, an increase in the vaginal-delivery sample).

with those shown in the tables. Results are available upon request.

Table 3C: Injuries to Mother (Patient Safety Indicators)

	All Vaginal Deliveries (PSI 18-19)					Vaginal Deliveries Without Instruments (PSI 19)				
	All (1)	All (2)	All (3)	All (4)	No Weekends (5)	All (6)	All (7)	All (8)	All (9)	No Weekends (10)
δ	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.003 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.003 (0.002)
X1	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
X2	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
X3	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	177,961	177,961	177,961	177,957	129,268	176,696	176,696	176,696	176,692	129,268
Panel B: Nearest Neighborhood Sample										
δ	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.005* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.004* (0.002)	-0.005* (0.002)
X1	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
X2	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
X3	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	55,923	55,923	55,923	55,922	40,481	55,472	55,472	55,472	55,471	40,168

Notes: Results are from linear probability model regressions. δ is the coefficient of *Treated * Post05*. All regressions include years and hospitals fixed effects. X1, X2, and X3 are listed in Table 1. *PSI 18 and 19* (columns 1 to 5) are computed using the sample of all vaginal deliveries, while *PSI 19* (columns 6 to 10) is computed using the sample of vaginal deliveries without instruments. Columns 5 and 10 use the sample of deliveries performed from Monday to Friday. Robust standard errors clustered at the hospital level in parenthesis. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 4C: Results Hospital-level

		All Vaginal Deliveries (PSI 18-19)				Vaginal Deliveries Without Instruments (PSI 19)			
	All (1)	No Weekend (2)	All (3)	No Weekends (4)	All (5)	No Weekend (6)	All (7)	No Weekends (8)	
Panel A: Entire Sample									
δ	-0.005** (0.002)	-0.004** (0.002)	-0.005** (0.002)	-0.004* (0.002)	-0.005** (0.002)	-0.004** (0.002)	-0.005** (0.002)	-0.004* (0.002)	
<i>Controls</i>	No	No	Yes	Yes	No	No	Yes	Yes	
Observations	1,054	1,054	1,054	1,054	1,054	1,054	1,054	1,054	
Panel B: Nearest Neighborhood Sample									
δ	-0.005*** (0.001)	-0.004* (0.002)	-0.005*** (0.001)	-0.004** (0.002)	-0.005*** (0.001)	-0.004* (0.002)	-0.005*** (0.001)	-0.004** (0.002)	
<i>Controls</i>	No	No	Yes	Yes	No	No	Yes	Yes	
Observations	447	447	447	447	447	447	447	447	

Notes: Results are from linear probability model regressions. δ is the coefficient of *Treated * Post05*. All regressions include quarters and hospitals fixed effects. The dependent variable represents the quarterly incidence of the outcome considered. *Controls* includes the quarterly incidence at the hospital level of X1, X2 and X3 (e.g., percentage of Italian mothers out of the total of mothers per quarter/hospital) as listed in Table 1. Columns 2-4-6-8 use the sample of deliveries performed from Monday to Friday. Robust standard errors clustered at the court level in parenthesis. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

References Online Appendices

- Agency for Healthcare Research and Quality (AHRQ) (2003). *AHQR Quality Indicators: Guide to Patient Safety Indicators*. Rockville, Md.
- Iizuka, T. (2013). Does Higher Malpractice Pressure Deter Medical Errors? *Journal of Law and Economics* 56: 161-188.
- Bertoli, P. (2014). *Medical Malpractice in Public HealthCare System: An Empirical Investigation of Scheduled Damages*. PhD Dissertation EDLE program.
- Comandè G. (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.
- Scarso, A. P. (2009). Punitive Damages in Italy, in *Punitive Damages: Common Law and Civil Law Perspective*, H. Kuziol and V. Wilcox, Springer, Wien/New York.
- Sella, M. (2005). I sistemi tabellari, available at: <http://www.personaedanno.it>.