HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 20/09

Fees equalization and Appropriate Health Care

Emilia Barili; Paola Bertoli and Veronica Grembi

May 2020

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

Fees equalization and Appropriate Health Care*

Emilia Barili
University of Milan
University of Pavia

Paola Bertoli†
University of Verona
University of Economics, Prague

Veronica Grembi
University of Milan

This Draft: April 2020

Abstract

Fees equalization in health care brings under a unique tariff several medical treatments, coded under different Diagnosis Related Groups (DRGs). The aim is to improve healthcare quality and efficiency by discouraging unnecessary, but better-paid, treatments. We evaluate its effectiveness on childbirth procedures to reduce c-section overuse by equalizing the DRGs for vaginal and cesarean deliveries. Using Italian data and a difference-in-differences approach, we show that setting an equal reimbursement decreased c-sections by 2.5%. This improved the appropriateness of medical decisions with more low-risk mothers delivering naturally and no significant changes in the incidence of complications for vaginal deliveries. Our analysis supports the effectiveness of fees equalization in averting c-sections, but highlights the marginal role of financial incentives in driving c-section overuse which went back to normal in the short run. We found a stronger reduction in low-quality, and more capacity constrained hospitals. Moreover, the effect is driven by districts where the availability of Ob-Gyn specialists is higher and where women are predominant in the gender composition of the specialty-group.

JEL Classification: K13; K32; I13
Keywords: Fees equalization, Cesarean Sections, Difference in Differences

*We thank Davide Dragone, Osea Giuntella, Mario Macis, and the participants to the 24th annual conference of the Italian Health Economics Association, to the 2019 Young Economists' Meeting in Brno, and to the seminars at the University of Milan and at the University of Economics, Prague for their useful comments on a very early stage of this project. Support from the GACR Grant 17-20175S is gratefully acknowledged.

†CERGE-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. Corresponding author, email: paola.bertoli@univr.it.

1 Introduction

The economic driver is considered one of the main causes of distortion in the provision of appropriate health care: over-prescription of inappropriate medical treatments would be due to financial incentives represented by differences in reimbursement levels between equally effective alternatives. Given the agency role played by physicians in a patient-physician relation, the underlying assumption is that, at the margin, when weighting costs and benefits of a medical treatment on a (hypothetically) purely medical basis, the better paid treatment (or drug) is chosen. It seems obvious to argue that once the economic driver is removed, we should witness a decrease in unnecessary treatments: the greater the gap between payments before the equalization, the higher the expected benefits in terms of appropriateness. This intuition is the rationale for proposing the adoption of fees equalization across treatments which target the same disease or condition (OECD, 2016). Yet, even if the theoretical expectations are clear, empirical evaluations of these policies are rarely available.

We analyze the introduction of a fees equalization policy between vaginal deliveries and cesarean sections. The high c-section rates observed in many countries are often mentioned as a clear example of treatment overuse: while the World Health Organization defines acceptable a 15% incidence (WHO, 2015), the actual rates have well surpassed both this figure and what is generally suggested by obstetric indications. Different c-section rates across health care systems have been often put in relationship with economic incentives given to health care practitioners.[1] Generally, c-sections are paid better than vaginal deliveries and the difference in prices has often been used to explain the magnitude of the incidence of c-sections. The seminal work by Gruber and Owings (1996) shows that worse financial outcomes for US obstetricians caused by a decline in fertility in the 1970s led to a 16% higher c-section rate during the 1970-1982 period. Gruber et al. (1999) study the effects of price differentials between vaginal and cesarean reimbursements on the Medicaid population, concluding that each $100 increase in the absolute differential produces a 4% increase in c-section rates. Similarly, Johnson and Rehavi (2016) exploit the price differential between c-sections performed in insurance-owned hospitals and non-insurance-owned hospitals to show that women delivering in insurance-owned hospitals are 5 percentage points less likely to receive a c-section. Foo et al. (2017) also provide evidence of the effect that price differentials have on treatment

[1]The economic incentive argument also works in association with other explanations. First, doctors' inexperience and inadequate skills produce a mismatch between patient needs and the chosen delivery method. Hence, women who need a c-section may not receive it, while women whose medical conditions call for a c-section may deliver naturally (Currie and MacLeod, 2017). Second, the fear of litigation would induce doctors to perform unnecessary c-sections. As a reaction to the pressure of being involved in a malpractice lawsuit (even if unsuccessful) in the case of an error, doctors may perform a c-section to protect themselves rather than patients (Currie and MacLeod, 2008; Shurtz, 2014; Bertoli and Grembi, 2018).

decisions, regardless of the identity of the receiver of the payment. They indeed find that a one standard deviation higher physician price differential ($420) causes a 12% increase in the odds ratio for c-sections, while the same increase ($5,805) in the hospital price differential leads to a 31% higher odds ratio.

We focus on the fees equalization implemented in 2005 in Lombardy, a northern Italian region with a population of approximately 10 million inhabitants. The form of payment is standard for the type of policy: hospital reimbursements are adjusted only for medical needs (*i.e.*, with or without complications) and do not depend on the delivery method (*i.e.*, vaginal or cesarean). The same approach is followed by the payment reform for uncomplicated births adopted in Minnesota in 2009 (*i.e.*, Minnesota's Medicaid Program, 2009), for maternity care in England (UK) in 2013 (*i.e.*, Maternity Care Pathway, 2013), by the California Blue Cross reform in the 1990s, and by the Taiwan's National Health Insurance (NHI) in 2005. However, our case study has three main strengths that reinforce the external validity of our results. First, we do not focus on a specific type of patient (*i.e.*, Medicaid patients or patients insured under a specific company) but address standard prices (DRGs) enforced for all patients interested in a treatment (*i.e.*, pregnant women). Second, our reference policy consists of an increase in the price for vaginal deliveries (*i.e.*, 70 euros equal to +4.2% for vaginal deliveries without complications and 168 euros equal to +9.6% for vaginal deliveries with complications), rather than in a decrease in the price for c-sections. This difference is remarkable since, theoretically, a change in the relative price of delivery methods achieved by reducing the c-section tariff can trigger two contrasting effects: an income effect (*i.e.*, to compensate the lower payment per c-section, physicians end up performing more c-sections) and a substitution effect (*i.e.*, since vaginal delivery becomes more convenient, physicians can perform fewer c-sections), where the net effect is the combination of the two. In our case, since the price of c-sections is unchanged, the income and substitution effects move in the same direction by making vaginal deliveries more profitable compared to c-sections. Third, the reference policy targets hospitals operating in a public health care system with salaried physicians, rather than directly affecting the prices paid to individual physicians as in previous studies on fees equalization. This allows us to address the role of hospitals in directing physicians' choices.

Using patient level data for the period 2002-2008 and the neighboring region of Veneto as a control group, we apply a difference-in-differences approach and show that in the absence of a price difference between vaginal deliveries and cesarean sections, the latter decrease by 2.5% at the mean of c-section. This effect is totally driven by a reduction in c-sections without complications (-2.6%) and is a short run effect as shown by an event study approach. We take a further step to investigate the distribution of the reduction. While an average decrease in

c-sections, in the context of c-section overuse, is welcome, a more insightful approach should consider whether the reduction is achieved on low-risk rather than high-risk mothers: when needed, a c-section can improve the health of both the mother and the newborn (Card et al., 2019; Jensen and Wüst, 2015). Defining classes of medical risk based on observable clinical conditions, we estimate a reduction of 2.5% on low-risk mothers which does not translate into an increase in complications during vaginal deliveries. These results appear to be consistent with the findings produced by Kozhimannil et al. (2018) in the 2009 Minnesota case (-3.24% c-sections), by OECD (2016) for the 2013 English case (less than a -2%), and by Keeler and Fok (1996) for the California data (-0.7%). However, in these three cases, the fees equalization was obtained by simultaneously lowering the c-section price and increasing the price for vaginal deliveries, which makes more difficult to disentangle the role of the financial incentives through the usual channels of income and substitution effects.

We show that our results could not be explained by possible unintended effects of the policy as the selection of patients according to the severity of their condition (*i.e.*, cream skimming), mismatch between actual and reported health status (*i.e.*, upcoding), or a shift to neighboring regions of Lombard mothers seeking a c-section (*i.e.,* patient sorting). Finally, we exploit the characteristics of hospitals and physicians to provide an analysis of the factors channeling the reaction to the reform, improving the understanding of which characteristics make hospitals more or less responsive to price differentials. We estimate that the reduction in c-sections is stronger in hospitals facing greater capacity constraints (*e.g.*, limits to longer hospitalizations), defined as the number of used beds weighted on different measures. Moreover, c-sections decrease more in low-quality hospitals defined according to several proxies; that is, in hospitals where medical decisions are more likely to be affected by factors other than patient medical conditions. The patients' socio-economic status do not actually matter, whereas the characteristics of the involved physicians (*i.e.*, Ob-Gyn specialists) are relevant. We observe stronger effects in districts with a greater availability of Ob-Gyn specialists, which suggests the presence of spillover effects in the adoption of different practice styles as stated by Molitor (2018). The reduction is also driven by districts where there is a higher incidence of women within the specialty-group of Ob-Gyn. The gender of the physician may indeed determine a different response to incentives or perception over patients' needs. As a placebo, we replicate the analysis with these two dimensions over several other specialty-groups (*e.g.*, Psychiatry), showing no significant estimates.

To appreciate the full implications of fees equalization for delivery methods, we need to keep in mind that, since the average pre-policy incidence of c-sections was 28% in Lombardy, even if the policy was successful and practitioners responded to financial incentives, this response can explain only a small fraction of c-section overuse. This means that the economic

driver does not play the lion share of inappropriate use of medical treatments, at least as far as c-sections are concerned and with respect to the Italian case. In addition, although the estimated effects are difficult to monetize, we perform some back-of-the-envelop calculations to understand what is the approximate cost of this intervention. Considering that each of the 278,472 natural deliveries without complications occurring in the post-policy period received an extra 70 euros reimbursement and that the 126,025 c-sections without complications would have been 2.6% higher in number in the absence of the policy, Lombardy paid an extra cost of 19,493,040 euros to avert 3,364 c-sections. This means that each avoided c-section costed an extra 5,795 euros per delivery. To have the complete picture, one must also quantify the costs and benefits associated with the health outcomes enjoyed by the patients. Since we cannot observe the medium-long run effects of avoiding a c-section (*e.g.*, improved immune system and lower respiratory morbidity for infants and lower risks of hysterectomy or abnormal placentation for women), we cannot perform a proper cost-benefit analysis of the policy. Nevertheless, these figures provide some initial evidence to reason about the opportunity of this type of intervention.

Our work also contributes to the general literature on the role of financial incentives on the supply of health care, which usually identifies the economic motivation as the main driver in treatment selection, especially in the substitution between more and less intensive treatments. For example, exploiting the price shocks caused by the 1997 geographical consolidation of Medicare payment regions, Clemens and Gottlieb (2014) show that a 2% increase in payment rates is associated with a 3% increase in the supply of care and that this reaction is much stronger among relatively elective, but intensive, treatments. In the context of obstetric practices, Ho and Pakes (2014) find that insurers allocate patients across hospitals in their network based on the prices paid to the hospitals themselves and this effect is stronger the more highly capitated is the insurer.

The remainder of the paper is organized as follows. Section 2 provides an overview of the institutional setting and the theoretical expectations associated with fees equalization. Section 3 describes the data and defines the outcomes of interest. Section 4 presents both our identification strategy and the results, while Section 5 describes the validity and robustness checks. The analysis of the elements affecting the degree of responsiveness of hospitals to the monetary driver is presented in Section 6 and Section 7 concludes.

2 Institutional Setting and Theoretical Expectations

Italian health care services are managed at the local level, with regions (*i.e.*, 21 independent authorities) in charge of the provision of medical assistance to their resident population. Although bounded to the national standards set by the central government, regions enjoy a wide discretion to regulate and organize health care delivery within their borders. This has resulted in the creation of public hospital networks based on different combinations of hospitals managed by local health authorities (LHAs), independent hospitals (*e.g.*, teaching hospitals) and private accredited institutions (Anessi-Pessina et al., 2004). In 2015, 89% of infants born in Italy were delivered in public hospitals (Ministero della Salute, 2018), and this trend has been stable throughout the country (*e.g.*, 88.3% in 2004 according to Ministero della Salute (2007)). Each region is organized in LHAs, which administer the health plan to the residents in their catchment area. Patients are assigned to a hospital based on their municipality of residence (*i.e.*, their home hospital) but they can freely choose the hospital where to be treated, leading to both intra- and inter-regional mobility. However, out-of-region flows are quite limited with respect to childbirth as Italian mothers, on average, cover approximately 10.6 miles (*i.e.*, 17 km) to deliver their babies and often prefer their home hospital (Amaral Garcia et al., 2015). Since 1995, the financing of all hospitals operating within the Italian NHS has worked through a DRG-based system, which is enforced for every patient (Cavalieri et al., 2013).[2] The national government releases and updates a list of tariffs that serves as a benchmark and regions are free to decide their own rates and to differentiate them by type of provider (Bertoli and Grembi, 2017a).

To address the overuse of c-sections, Lombardy implemented a payment reform in 2005 that equalized the DRG tariff for vaginal and cesarean deliveries with complications and the DRG tariff for vaginal and cesarean deliveries without complications for all its hospitals. In both cases, the price equalization has been obtained by increasing the DRG tariff for a vaginal delivery to match that of a c-section. This change in price implied an increase in reimbursement for a vaginal delivery without complications by 4.2% (*i.e.*, 70 euros) and in the reimbursement for a vaginal delivery with complications by 9.3% (*i.e.*, 168 euros).[3] The

[2]DRGs are a common mechanism to group procedures by similar medical conditions and resource utilization to express hospital activity in standardized units comparable across providers. Based on cost data usually related to a set of chosen hospitals, a fixed (average) rate is assigned to each DRG and this rate is meant to cover the average expenses incurred in treating patients within each DRG. Therefore, DRG tariffs are not connected to the actual costs sustained for a given case by a specific hospital.

[3]Before the policy, deliveries were reimbursed according to six different DRG fees, two referred to c-sections (*i.e.*, c-section w/ complications –370–, c-section w/o complications –371–), four to vaginal deliveries (*i.e.*, vaginal w/ complications –372–, vaginal w/o complications –373–, vaginal w/ sterilization –374–, vaginal w/ operation room procedure except sterilization –375–). After the reform, they were grouped into delivery w/ complications (*i.e.*, DRGs 370, 372) and delivery w/o complications (*i.e.*, DRGs 371, 373, 374, 375).

rationale was to remove the potential economic incentive for c-sections by making vaginal deliveries more economically attractive (Regione Lombardia, 2010).

2.1 The expected effects of tariffs equalization

Fees equalization looks like an effective policy answer to the potential distortions in the use of c-sections caused by price differentials (MACPAC, 2019). According to the standard model of supply induced demand by McGuire and Pauly (1991), physicians get utility from income and leisure, whereas performing unnecessary treatments causes them disutility due to ethical or reputational concerns. This reasoning implies that when the marginal benefits of a treatment exceed its marginal costs, physicians take advantage of their agency relation with patients to increase the use of the treatment. The extensions of this model to childbirth (Keeler and Fok, 1996; Gruber et al., 1999; Fabbri and Monfardini, 2001; Johnson and Rehavi, 2016) assume that the profit rate between cesarean and vaginal deliveries is always positive: c-sections are paid more than vaginal without requiring greater physician input since they can be scheduled, they take less time, and the cost of their greater complexity does not overcome the benefit of their shorter duration. As a result, a negative shock in the profit rate due to a decrease in c-section price triggers two offsetting effects: physicians perform fewer c-sections as they became less profitable (*i.e.*, substitution effect), while the decreased income of physicians makes inducing c-sections more desirable to make up the lost earnings (*i.e.*, income effect).

The existing evidence on the effect of removing the price differential between obstetric practices argues for dominance of the substitution effect over the income effect to explain the overall small decrease in c-sections. For instance, Keeler and Fok (1996) examine an insurance reform under the California Blue Cross that equalized the fees between vaginal and cesarean deliveries by decreasing the prices of the latter by 18% and increasing the price of the former by 3%. As long as the income of the average doctor in their sample depends strongly on the patients covered by the Blue Cross plan, there should be a sizable income effect and related induction. Yet, the authors observe a marginal decrease in c-sections of 0.7%; thus, the income effect appears to be dominated by the substitution effect. The reduction reported by Kozhimannil et al. (2018), who study the 2009 price equalization implemented by Minnesota's Medicaid program, is slightly higher. The new payment scheme resulted in a change in the prices of both delivery methods (*i.e.*, +12.2% for vaginal and -33% for cesarean) and produced only a 3.24% reduction in c-sections. To the best of our knowledge, there is only another paper, Lo (2008), studying the effect of raising the fee for a vaginal delivery up to the level of a c-section. They analyze the policy introduced in Taiwan in 2005, finding no effect on the c-section rate during the immediate months following the implementation.

However, these analyses and the related theoretical models move from the perspective of

the individual physician and consider the fees directly paid to her for her own work, while the perspective of hospitals is ignored. When, as in the case of Italy, physicians are employed by a single hospital and are salaried, the hospital perspective becomes crucial. This does not mean that physicians are no longer responsive to economic incentive, but rather the hospital perspective questions the traditional assumption that the profit rate always breaks in favor of c-sections. Hospital physicians still have incentives to perform more generously reimbursed procedures because hospital funds are dependent on DRGs and departments with inadequate funds may be downscaled. The same holds for hospital managers who may also be interested in having surpluses to expand hospital capacity (Januleviciute et al., 2016). Moreover, hospitals have proven to be well able to convey their priorities to their employees (Dafny, 2005; Fenn et al., 2007, 2013; Amaral Garcia et al., 2015). However, without knowing the actual costs of each procedure, the mere fact that the DRG price for c-sections exceeds the price for vaginal deliveries may no longer be sufficient to conclude that the profit rate between cesarean and vaginal deliveries is positive in the first place. The DRG price is not meant to cover only the work of the doctor, as in the mentioned models with physician fees. The price must also fund the fixed costs (*e.g.*, equipments for the operating room) and the costs for additional inputs required (*e.g.*, materials –medications–, personnel –nurses and anesthetist–). Although the c-section price remains higher than the vaginal delivery price after the change, the net earnings for a c-section may be lower than those for a natural delivery.

In our case, Lombardy's policy achieves fee equalization by increasing the reimbursement for natural deliveries without changing the reimbursement for c-sections. As a consequence, under the new payment scheme, the income effect ends up working along the substitution effect: natural deliveries become relatively more remunerative, thus, physicians have the incentive to increase them (*i.e.*, income effect) preferring them to c-sections (*i.e.*, substitution effect). Hence, the reference policy should provide a perfect setting to study the role of economic incentives in driving medical decisions as both possible theoretical effects move in the same direction, that is, an increase in natural deliveries. However, the magnitude of such an increase is an empirical question. In fact, although monetary incentives should no longer benefit cesarean deliveries, these procedures still entail advantages in terms of shorter duration and scheduling possibilities. Understanding the relevance of these potential channels has important policy implications to promote a more appropriate choice of delivery methods.

The second strong assumption of these models is that patients must be homogeneous, so that performing a c-section is orthogonal to the patient type. In other words, reducing c-sections is good on average, as increasing c-sections is bad on average. Still, from a policy perspective, the main goal is to achieve a reduction in c-sections for low-risk mothers (*i.e.,*

women whose clinical condition are good enough not to require a c-section). In fact, as suggested by Currie and MacLeod (2017), better decision making by physicians (*i.e.*, less subject to non-medical incentives) should result in a reallocation of procedures from low-risk to high-risk patients. To the best of our knowledge, Fabbri and Monfardini (2001) provides the only theoretical framework that builds upon the work of McGuire and Pauly (1991) to account for the maternal risk profile. The basic intuition is that, other things being equal, the disutility of a c-section increases the less risky a patient is: for instance, the ethical and reputational concerns are greater. As the monetary advantage of c-sections over vaginal deliveries decreases, the marginal benefits of a c-section are less likely to offset the marginal costs the less risky a patient is. Therefore, one would expect to observe a reduction in c-sections for low-risk mothers and no effect for high-risk mothers.

2.2 Veneto as control region

To evaluate the impact of Lombardy's policy, we use the hospitals of this region as the treated group and those of the neighboring region of Veneto as control group. The specific Italian context, where each region is largely independent in the management of its health care services and, consequently, may implement different policies, requires the selection of a specific control area. Veneto is the Italian region that best serves this purpose mainly for three reasons.

First, Veneto and Lombardy are quite similar in terms of the performance of their regional health care systems. The regions are commonly ranked among the top 4 performers when health care is at stake (CREA sanitá, 2018), and they have proved to be among the most virtuous regions in the country as far as c-section rates are concerned. Their share of c-sections is consistently below the national average; for example, in 2004, 36.4% of childbirths in Italy occurred through a c-section (Ministero della Salute, 2007), while the statistics for Veneto and Lombardy were 27.2% and 26.5%. Moreover, as apparent from Figure 1, the two regions show analogous trends for what concern deliveries (panel (a)) and medical risk profile of expectant mothers, proxied by the risk prevalence (panel (b)).[4]

Second, pregnant women show an analogue socio-economic profile. As depicted in Figure 2, the age distributions are homogeneous between the two regions, even distinguishing by

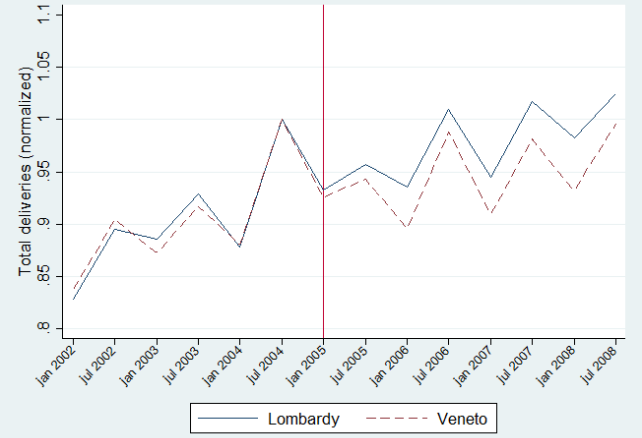
[4]The risk prevalence is defined as the evolution in the number of mothers suffering from at least one risk factor that may justify a cesarean intervention. The risk factors considered are anemia, hypertension, cardiovascular diseases, diabetes, sexually transmitted diseases, obesity, thyroid disfunction, drug addiction, eclampsia, cervix anomalies, previous c-section, immaturity or pre-term, prolonged pregnancy, renal failure, pelvic abnormality, precipitous labor, breech presentation, multiple delivery, antepartum hemorragia, abnormality in fetus heart rate, fetus rhesus isoimmunization, amniotic cavity problems, placenta previa, and fetus abnormality.

citizenship of the woman. In addition, we formally test the balance in the relevant maternal and fetus conditions, as well as in the characteristics of the delivery and labor, that might affect the performance of a c-section. Table 1 shows that there are basically no statistically significant differences between the two regions with the only exceptions of married mothers and mothers suffering from thyroid problems.

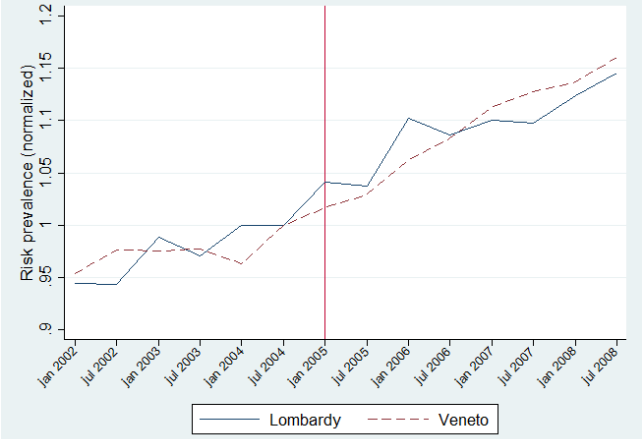
Finally, Veneto did not undergo any dramatic changes to its delivery payment policy during our observational period. In particular, the difference in reimbursement remained stable during the whole period with c-sections being paid 58% more than vaginal deliveries in all its public hospitals.

Figure 1: Trends by region

(a) Deliveries

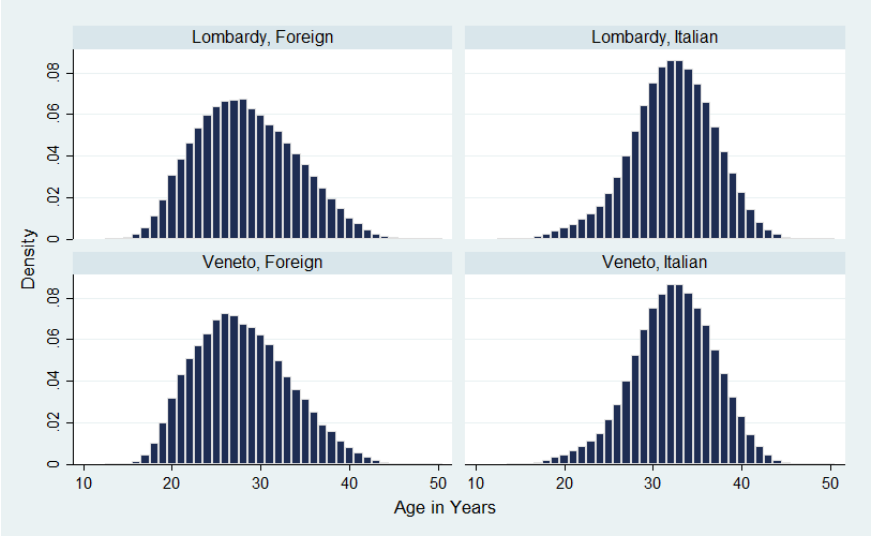


(b) Risk prevalence



Notes: Figures show the trends in deliveries and risk prevalence for the two regions, Veneto and Lombardy, normalized with respect to the last semester before policy implementation (2nd semester 2004). Figure (a) describes the evolution in the total number of deliveries registered in a hospital of the relevant region. Figure (b) shows the evolution of the risk prevalence; that is, the trend in the share of mothers suffering from at least one medical condition that would call for the performance of a c-section.

Figure 2: Age distribution by Region and Citizenship



Notes: Mothers' age distribution at the time of delivery by region of residence and citizenship. The average age is 28 for foreign women and 32 for Italian mothers, both in Lombardy and Veneto.

Table 1: Balance of the Observables

	Treated: Lombardy		Control: Veneto				
	Mean	Std. Dev.	Mean	Std. Dev.	Diff.	T-test	p-value
Women basic characteristics							
Age	31.531	4.984	31.536	5.008	-0.007	-0.302	0.763
Italian	0.838	0.369	0.838	0.368	-0.001	-0.525	0.599
Married	0.573	0.494	0.763	0.424	-0.190	-88.570	0.000
Delivery/labor characteristics							
Previous c-section	0.103	0.303	0.103	0.304	-0.000	-0.223	0.824
Multiple delivery	0.016	0.124	0.015	0.121	0.001	1.257	0.209
Ante-partum hemorrhage	0.001	0.028	0.001	0.030	-0.000	-0.922	0.357
Labor problems	0.008	0.090	0.009	0.092	-0.000	-0.934	0.350
Women medical conditions							
Obesity	0.000	0.015	0.000	0.016	-0.000	-0.299	0.765
Sexual transmitted diseases	0.000	0.013	0.000	0.015	-0.000	-1.114	0.265
Hypertension	0.000	0.013	0.000	0.011	0.000	0.728	0.467
Thyroid problem	0.001	0.033	0.002	0.050	-0.001	-6.526	0.000
Main complications	0.030	0.170	0.030	0.172	-0.000	-0.479	0.632
Reproductive system problems	0.236	0.425	0.237	0.425	-0.001	-0.257	0.797
Other problems	0.006	0.078	0.006	0.076	0.000	0.934	0.350
Fetus conditions							
Immaturity or Pre-term	0.000	0.016	0.000	0.017	-0.000	-0.608	0.544
Fetus problems	0.005	0.069	0.005	0.070	-0.000	-0.647	0.517

Notes: *Fetus problems* include fetus abnormality and rhesus isoimmunization related to the fetus. *Labor problems* refer to precipitous labor or prolonged labor/pregnancy. *Main complications* controls for placenta previa, cordon problems and eclampsia. *Reproductive system problems* refers to cervix anomalies and pelvic abnormalities. *Other problems* include diabetes, renal failure, addiction to drugs and cardiovascular problems.

3 Data and Outcomes

Our analysis relies on patient-level data taken from hospital discharge records (*SDO – Schede di dimmissione ospedaliera*) provided by the Italian Ministry of Health and referring to all childbirth events that occurred in public or private accredited hospitals in Veneto and Lombardy between 2002 and 2008.[5]

Since our focus is on the effect of the equalizing policy on the overuse of unnecessary c-sections, we restrict our sample as to exclude weekend deliveries. Since c-sections are generally scheduled during workdays, c-sections performed on weekends are likely to be emergency procedures (Amaral Garcia et al. (2015) and Bertoli and Grembi (2019)).[6] As a result, the final sample includes 716,445 observations: 490,599 for Lombardy and 225,846 for Veneto. Compared to the administrative data released by the Italian Institute of Statistics (ISTAT), our sample accounts for 30% of all childbirth events reported in the country (*i.e.*, all Italian births) between 2002 and 2008.

3.1 Outcomes of interest

The equalizing policy targets inappropriate c-sections, thus, the decision to perform a c-section is our main outcome of interest and is defined by a dummy *C-section*, which is equal to 1 if a woman delivered by c-section and 0 otherwise. As discussed in Section 2.1, we expect an overall decrease in the use of c-sections when the price differential between vaginal and cesarean deliveries is eliminated. However, from a policy perspective, the crucial aspect is to discourage the performance of c-sections on women whose medical conditions do not call for such a procedure. We test this aspect in two subsequent steps.

First, we look at the effect of the policy on both vaginal and cesarean deliveries, distinguishing between appropriate and inappropriate. Based on a logistic regression, we estimate the probability of receiving a c-section as a function of the maternal and infant observable risk factors (predicted probability of a c-section (PPC)). We use the resulting estimates to create *Inappropriate C-section* (*Appropriate C-section*) as a dummy taking value 1 if a c-section was performed on a woman with a predicted probability lower (equal to or higher) than 0.6 (Currie and MacLeod (2017) and Bertoli and Grembi (2019)). Similarly, *Inappropriate Vaginal* (*Appropriate Vaginal*) is equal to 1 if a woman with predicted probability equal to or higher (lower) than 0.6 delivers naturally.

[5]We do not extend the observation period beyond 2008 to avoid the possible noise caused by the great recession that since 2009 may have influenced the fertility choices of women, as well as the provision of hospital services (Bertoli et al., 2019).

[6]Still, we run all our analyses also on the whole universe of deliveries. Results are confirmed and available upon request.

Second, we estimate the effect of the policy on the use of c-sections by type of patient: low-, medium-, and high-risk patients. This further step allows us to provide more detailed insights into how the policy affects the matching between patient conditions and the chosen delivery method. Low-risk patients coincide with those women having a predicted probability of receiving a c-section equal to or less than 0.4, women with a probability between 0.4 and 0.7 represent medium-risk patients, while high-risk patients are those having a probability higher than 0.7.

To asses the impact of the policy on maternal and neonatal morbidity, we use two composite measures that are also estimated by type of delivery. *Preventable* indicates whether the patients (*i.e.*, mother and newborn) suffered any preventable delivery or post-delivery complications[7], while *Non-preventable* indicates whether patients reported a non-preventable delivery complication (including any fetus malposition, fetus or mother disproportion, complications related to the umbilical cordon, premature separation of placenta or rupture of the membranes, threaten premature labor, abnormal forces during labor).

Table A.1 reports the descriptive statistics of our sample. On average, 31% of women delivered by c-section and 14% of these procedures can be considered as inappropriate because they were performed on women with a low-risk profile. Overall, the vast majority of mothers were Italian; they were slightly older than 31 years and were married in 66% of the cases. C-sections registered a preventable problem in 3% of the cases and a non-preventable problem in 10% of the cases. Mothers delivering naturally were more likely to experience a problem that could have been anticipated (4%) than an unpredictable problem (3%).

4 Econometric Strategy

We apply a difference-in-differences (DD) estimation to identify the effect of the equalizing policy on $Outcome_{iht}$ for mother i giving birth in hospital h at time t, where t is the quarter-year. $Post_t$ is a dummy equal to 1 for t after the first quarter of 2005, and $Treated_h$ is a dummy identifying Lombard hospitals as the treated hospitals. Hence, δ is the DD estimator that captures the effect of the policy as shown in Equation 1:

$$Outcome_{iht} = \alpha + \delta(Treated_h * Post_t) + \beta_1 X1'_{iht} + \beta_2 X2'_{hit} + \beta_3 X3'_{hit} + \gamma_t + \omega_h + \epsilon \quad (1)$$

where γ_t are quarter-year fixed effects to control for common shocks and ω_h are hospital

[7]Consistent with Dubay et al. (1999), Dubay et al. (2001), Currie and MacLeod (2008), Dranove and Watanabe (2009), Dranove et al. (2011), Shurtz (2013), Shurtz (2014) and Bertoli and Grembi (2019), *Preventable* includes maternal fever, perineal laceration and infection, perineal hematomas, uterus laceration, infections, meconium, complications related to anesthesia, hemorragia, cardiac complications, embolism and retained placenta.

fixed effects to account for time-invariant unobservable characteristics across hospitals. We control for a wide range of factors that could affect the likelihood of receiving a c-section as listed in Table A.2. $X1'_{iht}$ groups the basic attributes of the mother, such as her age and marital status, $X2'_{iht}$ includes the health conditions that define her risk profile and $X3'_{iht}$ considers the characteristics of the pregnancy that may affect the choice of the delivery method and the incidence of complications.

4.1 Results

Table 2 shows the results for our primary outcomes. The adoption of an equalizing policy leads to a statistically significant reduction in the performance of c-sections. The magnitude of the average effect, as reported in Column (1), is a 0.8 percentage point decrease in c-sections, which corresponds to a 2.5% reduction at the mean of the variable.[8] According to columns 2 and 3, this overall decrease in c-sections is driven by a reduction in c-sections without complications (-2.6% at the mean of the variable), whereas we do not observe any change in the performance of c-sections with complications.

Although these findings confirm that hospitals respond to economic incentives, they also point out two additional aspects. First, economic incentives can explain only marginally the actual c-section overuse. Considering that the average c-section rate in Lombardy before the adoption of the policy was 28%, the economic motivation clearly explains only a very small portion of the gap with the traditional 10-15% target of policy makers. Eliminating the monetary advantage of c-sections over vaginal deliveries, is not sufficient to make the two procedures equally attractive conditional on patient conditions. Second, economic incentives seem to affect (even if marginally) the choice of the delivery method mainly in the less serious medical cases (*i.e.*, in the absence of complications).

If the equalizing policy were successful in removing the economic incentives to opt for an unnecessary c-section, one should observe an improvement in the matching between patient conditions and delivery methods. Hence, we check whether the estimated decrease in c-sections produces any change in the level of appropriateness of medical decisions. As apparent from Table 2, the policy leads to a reallocation of treatment decisions: the overall observed reduction in c-sections is driven by a drop in inappropriate c-sections that experience an overall 5% decrease (column 4). The decrease in inappropriate c-sections is not surprisingly associated with an increase in appropriate natural deliveries (+1.2%; column 7). Inappropriate vaginals and appropriate c-sections (columns 5 and 6) are instead not affected.

[8]The reduction in c-sections is robust to a logistic regression (Table A.3 in Appendix A) and to the replacement of the three sets of controls in Equation 1 with the predicted probability of receiving a c-section based on medical conditions (Table A.4 in Appendix A).

Table 2: DD - Delivery methods

	C-sections			**Inappropriate**		**Appropriate**	
	All	Without complications	With complications	C-section	Vaginal	C-section	Vaginal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
δ	-0.008***	-0.008***	0.013	-0.007***	-0.000	-0.001	0.008***
	(0.002)	(0.010)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Obs.	716,445	30,706	685,739	716,445	716,445	716,445	716,445
Mean	0.314	0.513	0.305	0.137	0.014	0.177	0.672
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Controls include $X1$, $X2$ and $X3$ as listed in Table A.2. *Inappropriate* and *Appropriate* procedures are identified based on the probability of receiving a c-section (PPC) estimated through a logistic regression as discussed in Section 3.1. Robust standard errors in parentheses. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

These findings are consistent with the effects registered by class of medical risk (Table 3). The improvement in the procedures' appropriateness is driven by a 2.5% reduction in c-sections on low-risk women, that is, patients the least in need of a cesarean delivery given their medical conditions. This is the risk class where one can reasonable expect to observe high response to the removal of the economic incentives. In fact, these are cases clearly not medically suitable for a c-section and for which the shift to a vaginal delivery should entail a lower risk of adverse outcomes. As a consequence of directly affecting the use and the appropriateness of delivery methods, the reference policy may impact the morbidity of the patients involved (*i.e.*, women and newborns). According to Table 4, no statistically significant effect on both preventable and non-preventable complications is observed, regardless of the delivery method.

We provide a more rigorous inspection of these results within an event study framework in Figures 3 and 4. We modify Equation 1 by adding half-yearly leads and lags and take the second half of 2008 as the reference semester. The plotted estimates for c-sections show the lack of anticipatory effects and a statistically significant decrease in the second semester after the policy implementation. The same holds for inappropriate c-sections and appropriate vaginal deliveries, while no anticipatory effects and no significant post-policy changes are observed for any of the remaining outcomes of interest. It is apparent that the main effect fades away after one year. This might be due to the fact that at the new equilibrium, once the new incentives scheme is incorporated in the decision making process of physicians, the other rationales driving the overuse of c-sections play a larger role.

Table 3: DD - C-sections by risk class

	Low Risk	**Medium Risk**			**High risk**
	ppc<=0.4	0.4<ppc<=0.5	0.5<ppc<=0.6	0.6<ppc<=0.7	ppc>0.7
	(1)	(2)	(3)	(4)	(5)
δ	-0.004**	-0.007	-0.009	0.025	0.001
	(0.002)	(0.032)	(0.324)	(0.028)	(0.003)
Obs.	569,502	5,243	4,855	6,999	129,846
Mean	0.163	0.458	0.588	0.696	0.937
Controls	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes

Notes: Controls include $X1$, $X2$ and $X3$ as listed in Table A.2. *Low-risk*, *Medium-risk* and *High-risk* patients are identified based on the probability of receiving a c-section (PPC) estimated through a logistic regression as discussed in Section 3.1. Robust standard errors in parentheses. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

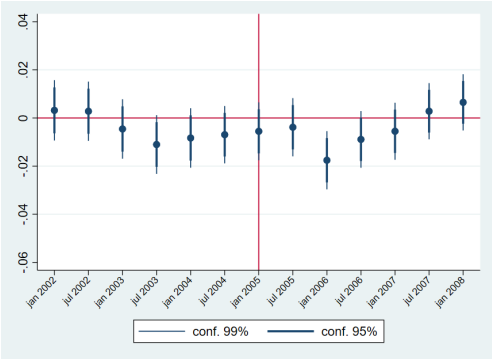
Table 4: DD - Complications

	Preventable		**Non-Preventable**	
	C-section	Vaginal	C-section	Vaginal
	(1)	(2)	(3)	(4)
δ	0.000	0.001	-0.001	0.001
	(0.001)	(0.001)	(0.002)	(0.001)
Obs.	716,445	716,445	716,445	716,445
Mean	0.039	0.043	0.102	0.028
Controls	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes

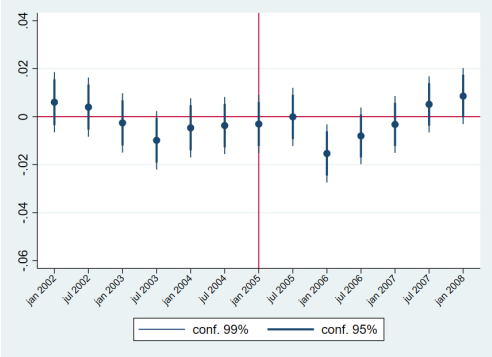
Notes: Controls include $X1$, $X2$ and $X3$ as listed in Table A.2. Robust standard errors in parentheses. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Figure 3: Leads and lags - Delivery Methods

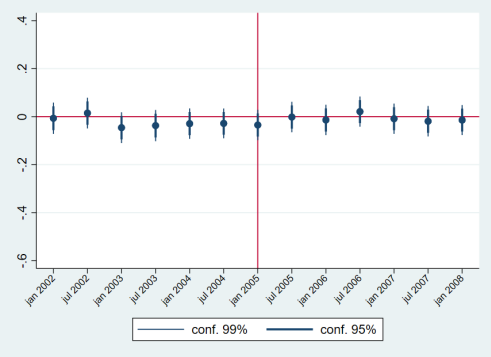
(a) C-sections



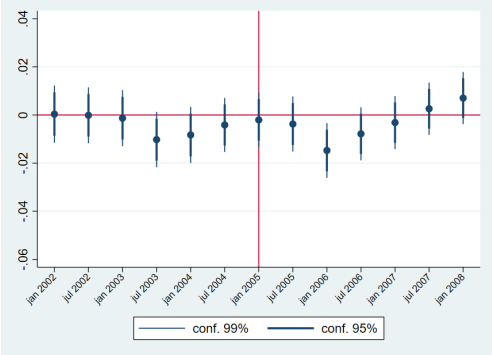
(b) C-sections w/o complications



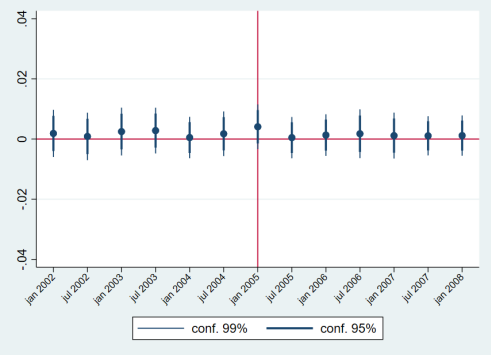
(c) C-sections w/ complications



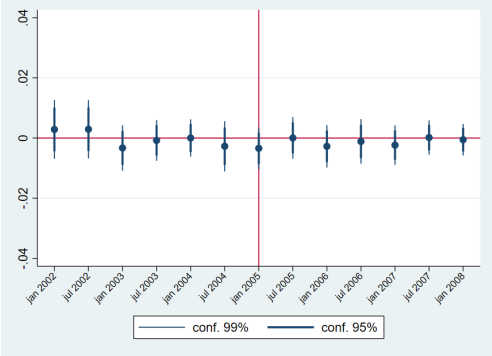
(d) Inappropriate c-sections



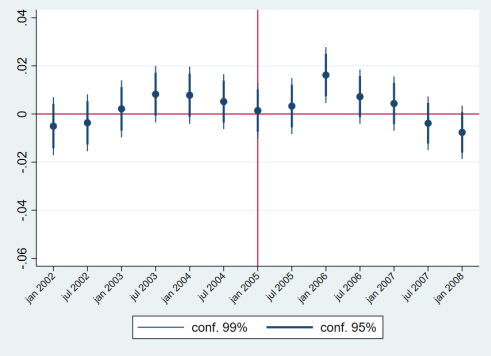
(e) Inappropriate vaginals



(f) Appropriate c-sections



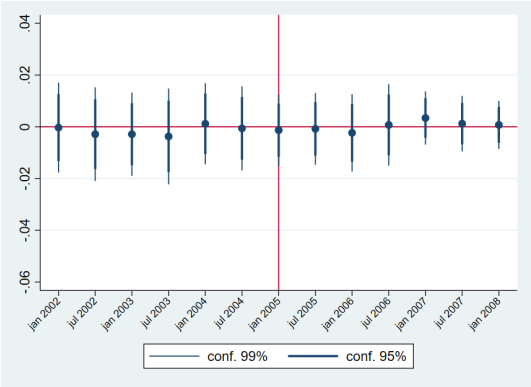
(g) Appropriate vaginals



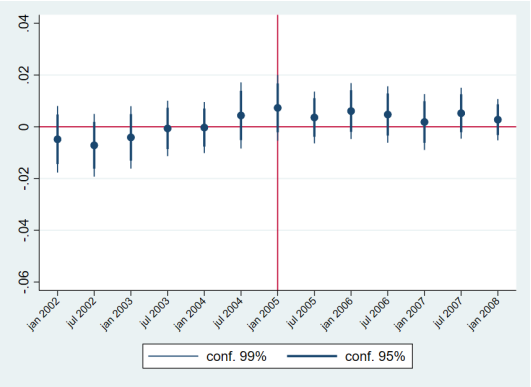
Note: The figures plot the leads and lags by semester taking Jul– Dec 2008 as the reference semester.

Figure 4: Leads and lags - Complications

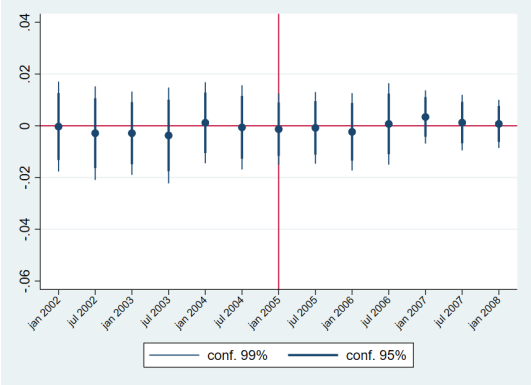
(a) Preventable complications: c-sec



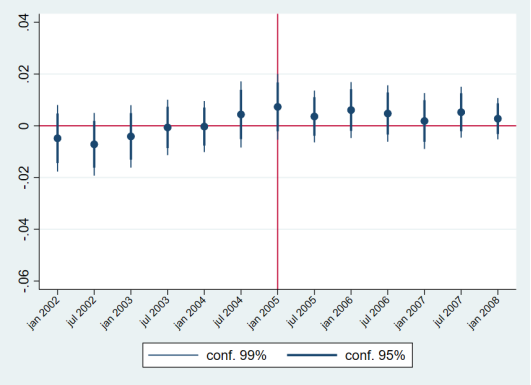
(b) Preventable complications: vaginal



(c) Non-preventable complications: c-sec



(d) Non-preventable complications: vaginal



Note: The figures plot the leads and lags by semester taking Jul– Dec 2008 as the reference semester.

4.2 Back-of-the-envelope calculations

To better appreciate the implications of the reference policy, we perform some back-of-the-envelope calculations of its monetary costs. In the years following the policy implementation (2005-2008), 278,472 natural deliveries without complications occurred in Lombardy, each of which cost 70 euros more, for an overall extra expenditure of 19,493,040 euros. Since cesarean deliveries amounted to 126,025, the 2.6% lower c-section rate due to the policy means that 3,364 c-sections were averted and Lombardy paid 5,795 euros to avoid each one of them (19,493,040/3,364). To provide a complete picture, the calculations should include changes in patients' health outcomes. We show that the observed reduction in c-sections is not associated with a change in patients' health performance during or immediately following delivery. However, the shift towards vaginal deliveries entails important potential health improvements in the medium/long term that we cannot assess. For example, the exposure to the maternal and surrounding environmental bacteria during vaginal delivery strengthens the

baby's immune system and intestinal colonization (Neu and Rushing, 2011; Dominguez-Bello, De Jesus-Laboy, Shen, Cox, Amir, Gonzalez, Bokulich, Song, Hoashi, Rivera-Vina, Mendez, Knight, and Clemente, Dominguez-Bello et al.; Scudellari, 2017), while the process of labor during vaginal delivery reduces respiratory morbidity (Hyde et al., 2012). Similarly, mothers delivering naturally are less likely to require a c-section in subsequent pregnancies, to receive a hysterectomy and to suffer from abnormal placentation and uterine rupture (Sandall et al., 2014). Hence, overall, we cannot conclude that equalizing payments was a bad investment policy-wise; rather, the provided findings should encourage a more thoughtful discussion among policy makers.

5 Validity and Robustness Checks

To defend the robustness of our results, we run a placebo by estimating Equation 1 on the sample of weekend deliveries. As explained in Section 3, c-sections occurring during weekends are more likely to be due to emergency reasons. Then, nullifying the economic incentive for c-sections should not impact or impact to a lower extent the performance of these procedures on weekends. Table 5 (column 1 and 2) shows that when we assess the reference policy on weekend deliveries, no effect is observed.

Besides this placebo, we also perform several tests to exclude the possible unintended effects that the equalization of the DRG tariffs of delivery methods could trigger and that could be driving our findings. First, our results could be explained by a change in the composition of treated patients. Increasing the DRG tariff of vaginal deliveries up to the tariff of c-sections could induce hospitals to attract healthier, thus less risky and costly, patients (*i.e.*, cream skimming). According to this explanation, we could observe a drop in c-sections also as a consequence of a change in the risk profile of deliveries. We investigate patient selection in two ways. We checked whether the equalizing policy modified the predicted probability of receiving a c-section (PPC) as in Baicker et al. (2006) and Amaral Garcia et al. (2015). If the removal of the price differential affects the PPC as a function of maternal risk conditions at the hospital level, then it follows that patient composition has changed. Then, we estimate Equation 1 on *Low–risk mothers*. This is a dummy that equals one if the woman does not present any pre-delivery risk conditions.[9] As apparent from columns 3 and 4 in Table 5, the policy did not affect both the probability of treating a low-risk mother and

[9]Following the medical literature, we define a low-risk mother as any woman who does not suffer from 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 (Bertoli and Grembi, 2019).

the predicted probability of receiving a c-section as a function of maternal risk conditions.

Second, if the reference policy reduces the incentive for Lombard hospitals to opt for a c-section, women residing in Lombardy and seeking such a procedure could choose to deliver in neighboring regions in the first place (*i.e.*, patient sorting). We test this scenario by checking whether treated hospitals experienced any change in the monthly number of managed deliveries, as well as in the monthly number of managed deliveries by women residing in Lombardy. As shown in Table 5 (columns 5 and 6), both measures of managed deliveries are not significantly affected.

Third, if the reference policy removes the price differential between vaginal and cesarean deliveries, it still entails a difference between deliveries with and without complications. Hence, this could create an incentive to favor vaginal deliveries with complications above all, that is, to "upcode" the severity of vaginal deliveries as to profit even more. However, as reported in column 7 in Table 5, there is no statistically significant effect of the equalizing policy on vaginal deliveries with complications as defined on the basis of the related DRG code (*i.e., 372*). Therefore, no upcoding episodes seem to play a relevant role.

Table 5: Robustness Checks

	Weekend deliveries		**PPC**	**Low-risk mothers**	**Deliveries**	**Resident deliveries**	**Vaginal with complications**
	All	without complications					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
δ	-0.001	-0.002	0.000	-0.000	4.063	4.083	-0.001
	(0.003)	(0.003)	(0.002)	(0001)	(3.011)	(2.947)	(0.001)
Obs.	225,548	216,871	716,445	716,445	6,374	6,374	716,445
Mean	0.189	0.180	0.304	0.637	112.401	109.516	0.021
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *PPC* is the predicted probability of receiving a c-section estimated through a logistic regression as discussed in Section 3.1. *Low-risk mothers* coincides with mothers who did not suffer from 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 (Bertoli and Grembi, 2019). *Deliveries* is computed as the monthly number of managed deliveries per hospital, while *Resident deliveries* is the monthly number of managed deliveries by women residing in the same region where the related hospital is located. *Vaginal with complications* identifies any vaginal delivery under the 372 DRG code. Controls include $X1$, $X2$ and $X3$ as listed in Table A.2. Robust standard errors in parentheses. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

6 When Is the Effect Stronger?

Finally, we focus on hospital and physicians' characteristics that can affect the intensity of the reaction to a change in financial incentives. This last step offers important policy implications, since it helps to understand when hospitals are more responsive to economic incentives and under which circumstances fees equalization can be more or less effective in reducing the overuse of medical treatments, as for the case of c-sections. When studying hospital features, we explore the role of capacity constraints, quality and patients' socio-economic background. While, concerning physicians, we analyze the characteristics of the specialized workforce. We focus on the availability of specialists and the incidence of women within the relevant medical specialty-group, comparing the results between Ob-Gyn and other groups of specialists. We generate dummies, D, for each characteristic and interact the dummies with $Treated_h * Post_t$. For each characteristic, 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.

6.1 Capacity constraints

Capacity constraints limit the degree of adjustment and discretion when deciding to keep a patient for an additional day or, by the same token, who requires more attention and care. Intuitively, this should constraint the decision to choose more aggressive treatments that require longer hospitalizations and more care. Hospitals facing more stringent capacity constraints should substitute more vaginal deliveries for c-sections, that, as surgical procedures, imply a longer occupancy of beds and more medical assistance during the recovery.

We proxy capacity constraints using the number of used beds weighted on the overall number of doctors, the overall number of nurses, and the number of available beds since these factors determine the capability of a ward to deal with patient turnover. We take the median value of each of these measures and define hospitals with above-median values as more constrained.

The results in columns 1, 2, and 3 in Table 6 confirm our prior: the tighter the constraint, the higher the shift to natural deliveries. Constraints on beds appears to be the most relevant. While for what concerns the personnel, constraints on nurses are way more significant than on doctors. Indeed, the first are the one more involved during the recovery of patients, thus, those that may benefit more from a shorter and smoother recovery of patients. It reinforces the idea that policies operating on financial incentives are effective even in a context where physicians are civil servants and receive a fixed wage. However, additional information seem to significantly affect the decision process (*e.g* overall characteristics of the ward in terms of available beds, used beds, personnel).

6.2 Quality

We consider hospital quality following the idea that better-quality hospitals are, by definition, those in which decisions are more adherent to patients medical needs. Hence, financial incentives are expected to be more relevant in low-quality hospitals, and the removal of financial incentives should have a stronger impact on the use of c-sections in this context.

We proxy high/low quality in three ways. Following the literature that identifies high levels of competition with higher quality (Bloom et al. (2013)), we label a hospital as high quality hospital if it faces an Herfindal-Hirschman Index (HHI) below the median HHI in our sample (*i.e.*, <3,882). As a second measure, we use the number of wards within a hospital. The presence of more wards signals a lower degree of hospital specialization. Therefore, better quality hospitals (obstetrically speaking) are expected to have a below-the-median number of wards (*i.e.*, <41). Finally, we use the estimated hospital fixed effects from our baseline specification to proxy quality. Fixed effects capture the unobserved time-invariant hospital characteristics, such as management or practice styles, which in the short run do not experience significant variation (McClellan and Staiger, 1999). Greater fixed effects are associated with lower quality (Bertoli and Grembi, 2017b); thus, high quality is represented by below-the-median fixed effects (*i.e.*, <-0.181).

In Table 6, columns 4 to 6 show that the effect is higher for low-quality hospitals, as defined according to the different proxies.

6.3 Patients' characteristics

Previous works underline how patients characteristics may interact with incentives given to physicians (Johnson and Rehavi (2016)). In the context of deliveries, physicians seem to be less likely to operate unnecessary c-sections on better educated and higher income mothers who tend to be more informed on the risk of alternative procedures and might select more carefully the hospital where to be treated.

We introduce two measures to proxy the wealth and education of patients: the rate of graduated individuals and the average income at the level of the patient's municipality of residence.[10] We define more educated or higher income individuals those coming from municipalities where the aforementioned indexes are above-the-median values (*i.e.*, above 28.8% graduated individuals; annual income above 21,394 euros).

In Table 6, columns 7 and 8 show that the response to the policy is not affected by the socio-economic background of patients.

[10]Due to data limitation, individual socio-economic information are not available. We use as proxies aggregated information, available at the municipality level. The municipality is the smaller administrative unit in Italy, with an average of 13,000 inhabitants in the sample considered.

6.4 Specialized workforce

Finally, we explore the role of specialized workforce characteristics. Our prior is that a greater availability of specialists may translate into spillover effects among colleagues: the greater the availability, the strongest should be the effect of the policy. While a different gender composition within the group may determine alternative reactions: males and females may weigh differently the factors that enter the decision process, such as financial incentives, reputational concerns, and patient needs.

With a focus on the LHA district where the hospital is located, we consider two measures for the main medical-specialty groups (*i.e.* Obstetric-Gynecologists, Surgery, Psychiatry, Pediatrics, and Rehabilitation). These measures describe the workforce characteristics by reporting the number of available specialists per 10,000 inhabitants and the incidence of women within the specialty-group. We consider the median values as the relevant thresholds.[11]

We find a stronger effect where the specialists' availability is higher and the incidence of women in the specialty-group is stronger (Table 7 columns 1-6). As a placebo, we show that the same measures are not significant if we consider specialty-groups not involved in the reform (Table 7 columns 2 to 5 and 7 to 10).

[11] With reference to Obstetric-Gynecologists, Surgery, Psychiatry, Pediatrics, and Rehabilitation, the median values for specialists' availability are respectively 1.79, 8.44, 1.58, 1.19, 0.62. The median values for the incidence of women within the specialty-group are 0.56, 0.29, 0.55, 0.69, 0.43.

Table 6: Heterogeneous effects by hospital characteristics

	Capacity constraint			**Hospital quality**			**Patient characteristics**	
	Doctors	Nurses	Beds	HHI	Wards	FE	Education	Income
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Less constrained	**Less constrained**	**Less constrained**	**High quality**	**High quality**	**High quality**	**Low education**	**Low income**
δ	-0.008***	-0.007***	-0.004***	-0.001	-0.000	0.003	-0.008***	-0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	More constrained	**More constrained**	**More constrained**	**Low quality**	**Low quality**	**Low quality**	**High education**	**High income**
δ	-0.011***	-0.011***	-0.014***	-0.012***	-0.006***	-0.019***	-0.008***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Difference	-0.003	-0.004*	-0.010***	-0.011*	-0.006***	-0.022***	-0.000	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Obs.	716,445	716,445	716,445	716,445	716,445	716,445	716,445	716,445
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *Less constrained* indicates that the number of used beds out of the total number of doctors (column 1), nurses (column 2), and available beds (column 2) in a hospital is a below-the-median value. *High-quality* indicates a hospital facing a below-the-median HHI (column 4), a hospital with a below-the-median number of wards (column 5), or a hospital with a below-the-median value of fixed effects (column 6). *High-education* indicates patients coming from municipalities where the rate of people having at least a high school diploma is above-the-median (column 7); *High-income* patients whose average income at the municipality level is above-the-median (column 8). Robust standard errors in parentheses. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 7: Heterogeneous effects by physician characteristics

	Availability of specialists (per 10,000 inhab.)					**Incidence of women specialists**				
	Ob/Gyn (1)	Surgery (2)	Psychiatry (3)	Pediatrics (4)	Rehabilitation (5)	Ob/Gyn (6)	Surgery (7)	Psychiatry (8)	Pediatrics (9)	Rehabilitation (10)
	Low	**Low**	**Low**	**Low**	**Low**	**Low**	**Low**	**Low**	**Low**	**Low**
δ	-0.006***	-0.006***	-0.010***	-0.006***	-0.009***	-0.005**	-0.010***	-0.009***	-0.010***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	High	**High**	**High**	**High**	**High**	**High**	**High**	**High**	**High**	**High**
δ	-0.010***	-0.010***	-0.007***	-0.009***	-0.006***	-0.010***	-0.007***	-0.008***	-0.007***	-0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Difference	-0.004*	-0.003	0.003	-0.003	0.003	-0.005**	0.002	0.001	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Obs.	716,445	716,445	716,445	716,445	716,445	716,445	716,445	716,445	716,445	716,445
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *High* availability of specialists indicates a number of physicians weighted per 10,000 inhabitants above the median in the LHA district where the hospital is located. *High* incidence of women identifies LHAs where the prevalence of women within each specialty group is above the median. Robust standard errors in parentheses. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

7 Conclusions

Financial incentives faced by health care professionals are traditionally considered to be a main driver of unnecessary treatments. By setting a single payment for a broader package of treatments, fees equalization may significantly discourage the overuse of health care.

We evaluate the introduction of this payment scheme in Italy, where the Lombardy region equalized the DRGs for vaginal and cesarean deliveries in 2005 to reduce the use of unnecessary c-sections. Using Veneto as control region and applying a difference-in-differences strategy, we find that the reference policy led to a 2.5% decrease in c-sections. This reduction is associated with an improvement in treatment selection as low-risk mothers are less likely to receive a c-section once the price differential is removed, while we do not find any statistically significant effects on patient health outcomes.

Overall, our analysis supports the potential of equalizing policies by showing that a relatively small increase in the price for natural deliveries (+4.2%) still produces a decrease in the incidence of c-sections. However, even though hospital physicians respond to the removal of the price differential, financial incentives turn out to play a limited role in driving the so-called treatment overuse. Elimination of financial incentives does not appear to be sufficient to make vaginal and cesarean delivery equally attractive, conditional on patient conditions.

Finally, the analysis of the channels of the effect on c-sections shows that the observed decrease is driven by a strong reaction of lower quality hospitals and hospitals facing stricter capacity constraints defined according to several proxies. A greater availability of Ob-Gyn specialists in the LHA district where the hospital is located increases the response to the policy, pointing at possible spillover effects, together with a higher presence of women within the medical-specialty group.

References

Amaral Garcia, S., P. Bertoli, and V. Grembi (2015). Does experience rating improve obstetric practices? Evidence from Italy. *Health Economics 24*(9), 1050–1064.

Anessi-Pessina, E., E. Cant, and C. Joninii (2004). Phasing out market mechanisms in the italian national health service. *Public Money and Management 24*(5), 309–316.

Baicker, K., K. S. Buckles, and A. Chandra (2006). Geographic variation in the appropriate use of cesarean delivery. *Health Affairs* (25), 355–367.

Bertoli, P. and V. Grembi (2017a). The political economy of diagnosis-related groups. *Social Science & Medicine 190*, 38–47.

Bertoli, P. and V. Grembi (2017b). The life-saving effect of hospital proximity. *Health Economics 26*(S2), 78–91.

Bertoli, P. and V. Grembi (2018). *Medical Malpractice: How Legal Liability Affects Medical Decisions*, Volume 14.

Bertoli, P. and V. Grembi (2019). Malpractice risk and medical treatment selection. *Journal of Public Economics 174*, 22–35.

Bertoli, P., T. L. B. Nguyen, and V. Grembi (2019). Birth in hard times: when your ethnic network can keep you up or drag you down.

Bloom, N., C. Propper, S. Seiler, and J. Van Reenen (2013). The impact of competition on management quality: Evidence from public hospitals. *Review of Economic Studies 82*(2), 457–489.

Card, D., A. Fenizia, and D. Silver (2019). The Health Impacts of Hospital Delivery Practices. *NBER Working Paper* (25986), 1–32.

Cavalieri, M., L. Gitto, and C. Guccio (2013). Reimbursement systems and quality of hospital care: an empirical analysis for italy. *Health Policy 111*(3), 273–289.

Clemens, J. and J. D. Gottlieb (2014). Do physicians' financial incentives affect medical treatment and patient health? *American Economic Review 104*(4), 1320–1349.

CREA sanitá (2018). Una misura di Performance dei SSR. . `https://www.creasanita.it/images/performancessr/Libro-Performance-V-Edition.pdf` [Online; accessed 21-May-2019].

Currie, J. and W. B. MacLeod (2008). First do no harm? tort reform and birth outcomes. *Quarterly Journal of Economics 123*(2), 795–830.

Currie, J. and W. B. MacLeod (2017). Diagnosing expertise: Human capital, decision making, and performance among physicians. *Journal of Labor Economics 35*(1), 1–43.

Dafny, L. (2005). How do hospitals respond to price changes? *The American Economic Review 95*(5), 1525–1547.

Dominguez-Bello, M. G., K. M. De Jesus-Laboy, N. Shen, L. M. Cox, A. Amir, A. Gonzalez, N. A. Bokulich, S. S. Song, M. Hoashi, J. I. Rivera-Vina, K. Mendez, R. Knight, and J. C. Clemente. Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer.

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), 1–33.

Dranove, D. and Y. Watanabe (2009). Influence and Deterrence: How Obstetricians Respond to Litigation against Themselves and Their Colleagues. *American Law and Economic Review 20*, 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.

Fabbri, D. and C. Monfardini (2001). Demand induction with a discrete distribution of patients. Dept. of Economics, University of Bologna.

Fenn, P., A. Gray, and N. Rickman (2007). Liability, insurance, and medical malpractice. *Journal of Health Economics* (26), 1057–1070.

Fenn, P., 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. *The Oxford Bulletin of Economics and Statistics 75*(3), 340–361.

Foo, P. K., R. S. Lee, and K. Fong (2017). Physicians prices, hospital prices, and treatment choice in labor and delivery. *American Journal of Health Economics 3*(3), 422–453.

Gruber, J., J. Kim, and D. Mayzlin (1999). Physician fees and procedure intensity: the case of cesarean delivery. *Journal of Health Economics 18*(4), 473–490.

Gruber, J. and M. Owings (1996). Physicians financial incentives and cesarean section delivery. *RAND Journal of Economics 27*(1), 99–123.

Ho, K. and A. Pakes (2014). Hospital choices, hospital prices, and financial incentives to physicians. *American Economic Review 104*(12), 3841–3884.

Hyde, M. J., A. Mostyn, N. Modi, and P. R. Kemp (2012). The health implications of birth by caesarean section. *Biological Reviews 87*(1), 229–243.

Januleviciute, J., J. E. Askildsen, O. Kaarboe, L. Siciliani, and M. Sutton (2016). How do hospitals respond to price changes? evidence from norway. *Health Economics 25*, 620–636.

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.

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.

Keeler, E. and T. Fok (1996). Equalizing physician fees had little effect on cesarean rates. *Medical Care Research and Review 53*(4), 465–471.

Kozhimannil, K. B., A. J. Graves, A. M. Ecklund, N. Shah, R. Aggarwal, and J. M. Snowden (2018). Cesarean delivery rates and costs of childbirth in a state medicaid program after implementation of a blended payment policy. *Medical Care 56*(8), 658–664.

Lo, J. C. (2008). Financial incentives do not always work-An example of cesarean sections in Taiwan. *Health Policy 88*(1), 121–129.

MACPAC (2019). Medicaid Payment Initiatives to Improve Maternal and Birth Outcomes. IssueBrief. Available at: `https://www.macpac.gov/wp-content/uploads/2019/04/Medicaid-Payment-Initiatives-to-Improve-Maternal-and-Birth-Outcomes.pdf` [Online; accessed 29-May-2019].

McClellan, M. and D. Staiger (1999). The Quality of Health Care Providers. *NBER Working Paper Series 7327*, 115–141.

McGuire, T. G. and M. V. Pauly (1991). Physicians response to fee chanes with multiple payers. *Journal of Health Economics 10*, 385–410.

Ministero della Salute (2007). Certificato di asistenza al parto (CeDAP). Analisi dell'evento nascita - Anno 2004. Available at: http://www.salute.gov.it/imgs/C_17_pubblicazioni_642_allegato.pdf [Online; accessed 21-May-2019].

Ministero della Salute (2018). Certificato di asistenza al parto (CeDAP). Analisi dell'evento nascita - Anno 2015. Available at: http://www.salute.gov.it/imgs/C_17_pubblicazioni_2766_allegato.pdf [Online; accessed 21-May-2019].

Molitor, D. (2018). The evolution of physician practice styles: Evidence from cardiologist migration. *American Economic Journal: Economic Policy 10*(1), 326–356.

Neu, J. and J. Rushing (2011). Cesarean versus vaginal delivery: Long-term infant outcomes and the hygiene hypothesis. *Clinics in Perinatology 38*(2), 321–331.

OECD (2016). Better ways to pay for health care. Technical report. Available at: https://www.oecd.org/els/health-systems/Better-ways-to-pay-for-health-care-FOCUS.pdf [Online; accessed 31-May-2019].

Regione Lombardia (2010). Il taglio cesareo: quando e perché. Technical report. Available at: https://www.ondaosservatorio.it/ondauploads/2015/02/Microsoft-Word-ducumento-per-la-regione-con-suggerimenti.pdf [Online; accessed 21-May-2019].

Sandall, J., R. M. Tribe, G. Mola, G. H. Visser, C. S. Homer, D. Gibbons, N. M. Kelly, H. Powell Kennedy, H. Kidanto, P. Taylor, and M. Temmerman (2014). Short-term and long-term effects of caesarean section on the health of women and children. *The Lancet 392*, 1349–1357.

Scudellari, M. (2017). New feature: Cleaning up the hygiene hypothesis. *Proceedings of the National Academy of Sciences 114*(7), 1433–1436.

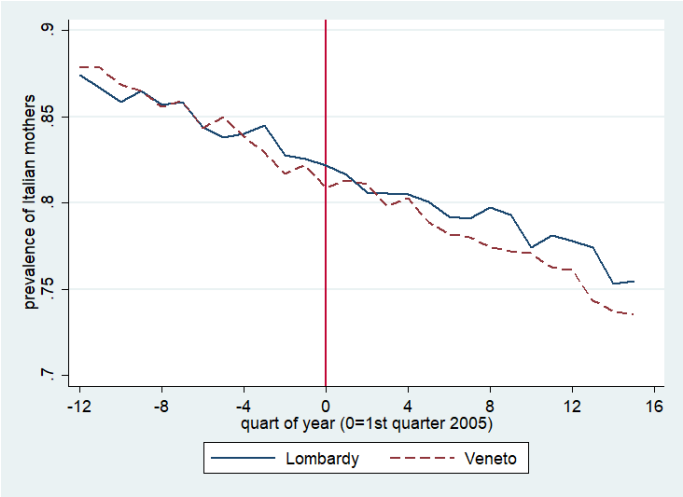
Shurtz, I. (2013). The impact of medical errors on physician behavior: Evidence from malpractice litigation. *Journal of Health Economics 32*, 331–340.

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

WHO (2015). WHO. Statement on Caesarean Section Rates. Available at: https://apps.who.int/iris/bitstream/handle/10665/161442/WHO_RHR_15.02_eng.pdf;jsessionid=719B32A5B0ECD74C2E40DEE3D8EFA1E7?sequence=1 [Online; accessed 21-May-2019].

Appendix A: Additional Tables and Figures

Figure A.1: Prevalence of Italian mothers by region



Note: The figures plots the trend in the prevalence of Italian mothers in both Lombardy and Veneto.

Table A.1: Summary Statistics

	Overall		Treated: Lombardy		Control: Veneto	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Primary Outcomes						
C-section	0.31	(0.46)	0.31	(0.46)	0.32	(0.47)
Inappropriate c-sec	0.14	(0.34)	0.13	(0.34)	0.15	(0.36)
Inappropriate vaginal	0.01	(0.12)	0.01	(0.12)	0.01	(0.11)
Appropriate c-sec	0.18	(0.38)	0.18	(0.38)	0.17	(0.38)
Appropriate vaginal	0.67	(0.47)	0.67	(0.47)	0.67	(0.47)
Secondary Outcomes - Maternal and Infant Morbidity						
Prevent c-sec	0.03	(0.17)	0.03	(0.16)	0.04	(0.19)
Prevent vaginal	0.04	(0.19)	0.03	(0.18)	0.05	(0.22)
Non-preventable c-sec	0.10	(0.30)	0.10	(0.30)	0.10	(0.30)
Non-preventable vaginal	0.03	(0.17)	0.03	(0.16)	0.03	(0.17)
Additional Outcomes						
Low-risk deliveries	0.64	(0.48)	0.63	(0.48)	0.66	(0.47)
PPC	0.30	(0.31)	0.30	(0.31)	0.31	(0.30)
Vaginal with complications	0.02	(0.14)	0.02	(0.15)	0.02	(0.13)
Deliveries	112.40	(76.39)	119.48	(82.52)	99.58	(61.78)
Resident deliveries	109.52	(74.87)	116.13	(80.72)	97.54	(61.12)
Demographics						
Age	31.56	(5.11)	31.57	(5.11)	31.54	(5.10)
Italian	0.81	(0.39)	0.81	(0.39)	0.81	(0.40)
Married	0.66	(0.47)	0.73	(0.44)	0.50	(0.50)

Table A.2: Controls

X1	**X2**[a]	**X3**[b]
Age	Anemia	Multiple delivery
Italian	Hypertension	Immaturity/preterm
Married	Cardiovascular problems	Prolonged labor
	Diabetes	Precipitous labor
	Sexually transmitted diseases	Placenta previa
	Previous c-section	Problems of the amniotic cavity
	Renal failure	Fetus rhesus isoimmunization
	Thyroid dysfunction	Fetus abnormality
	Obesity	Ante-partum hemorrhage
	Pelvic abnormality	Breech
	Drug addiction	Umbilical cordon problems
	Eclampsia	

Note: (a) Maternal risk factors are consistent with those used in Dubay et al. (1999), Dubay et al. (2001), Currie and MacLeod (2008), Dranove and Watanabe (2009), Dranove et al. (2011), Shurtz (2013) and Shurtz (2014).

Table A.3: Logistic Regression - Delivery methods

	C-sections		
	All	With complications	Without complications
	(1)	(2)	(3)
δ	-0.048***	0.041	-0.052***
	(0.015)	(0.067)	(0.016)
Obs.	716,445	30,706	685,739
Mean	0.314	0.513	0.305
Controls	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes

Notes: Controls include $X1$, $X2$ and $X3$ as listed in Table A.2. Robust standard errors in parentheses. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.4: PPC as Control - Delivery methods

	C-sections		
	All	With complications	Without complications
	(1)	(2)	(3)
δ	-0.007***	0.017	-0.008***
	(0.002)	(0.011)	(0.002)
Obs.	716,445	30,706	685,739
Mean	0.314	0.513	0.305
PPC as control	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes

Notes: Robust standard errors in parentheses. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$