



HAL
open science

Impact of tariff refinement on the choice between scheduled C-section and normal delivery: Evidence from France

Alex Proshin, Alexandre Cazenave-Lacroutz, Lise Rochaix

► To cite this version:

Alex Proshin, Alexandre Cazenave-Lacroutz, Lise Rochaix. Impact of tariff refinement on the choice between scheduled C-section and normal delivery: Evidence from France. *Health Economics*, 2023, 32 (7), pp.1397-1433. 10.1002/hec.4672 . hal-04157204

HAL Id: hal-04157204

<https://hal.science/hal-04157204>

Submitted on 13 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Impact of tariff refinement on the choice between scheduled C-section and normal delivery: Evidence from France

Alex Proshin^{1,2}  | Alexandre Cazenave-Lacroutz³  | Lise Rochaix²

¹Canadian Center for Health Economics, Toronto, Canada

²Hospinnomics (PSE – École d'Économie de Paris, Assistance Publique Hôpitaux de Paris – AP-HP), 1 Parvis Notre Dame, Paris, France

³Centre de Recherche en Économie et Statistique, Institut National de la Statistique et des Etudes Economiques, Palaiseau, France

Correspondence

Alex Proshin.

Email: alks.proshin@gmail.com

Abstract

Studying quasi-experimental data from French hospitals from 2010 to 2013, we test the effects of a substantial diagnosis-related group (DRG) tariff refinement that occurred in 2012, designed to reduce financial risks of French maternity wards. To estimate the resulting DRG incentives with regard to the choice between scheduled C-sections and other modes of child delivery, we predict, based on pre-admission patient characteristics, the probability of each possible child delivery outcome and calculate expected differences in associated tariffs. Using patient-level administrative data, we find that introducing additional severity levels and clinical factors into the reimbursement algorithm had no significant effect on the probability of a scheduled C-section being performed. The results are robust to multiple formulations of DRG financial incentives. Our paper is the first study that focuses on the consequences of a DRG refinement in obstetrics and develops a probabilistic approach suitable for measuring the expected effects of DRG fee incentives in the presence of multiple tariff groups.

KEYWORDS

C-section, delivery, diagnosis-related groups, financial incentives, midwives, obstetricians, refinement, vaginal labor

JEL CLASSIFICATION

C52, I11, I12

1 | INTRODUCTION

The topic of how financial stimuli can affect the quality, quantity, and composition of healthcare services has stirred much debate. Substantial anecdotal and scientific evidence has been pointing to the fact that healthcare providers may take into consideration factors other than clinical aspects or patients' preferences when prescribing a medical treatment. Our paper estimates the responsiveness of healthcare providers to a change in hospital-level financial stimuli. We investigate the impact of a diagnosis-related group (DRG) refinement, leading to an increase in the number of associated tariffs. Such a change was introduced in France by adding additional criteria and diagnoses to pricing formulae in order to better account for differences in patient severity. The results of our study provide no evidence that there was a connection between hospital-level changes in financial incentives and the rate of scheduled C-sections performed by obstetricians in France between 2010 and 2013.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. Health Economics published by John Wiley & Sons Ltd.

The risk that financial incentives may lead healthcare providers to change their practices under prospective payment schemes (PPS) has been analyzed extensively in the theoretical literature. For example, Ellis (1998) considers healthcare under/overprovision and patient selection as major concerns that may arise under PPS. His model predicts that in a perfect information setting where competitive healthcare providers fix both a maximum severity for admitted patients and a bundle of services available at each severity level, high-severity patients will receive a socially suboptimal amount of treatment (“skimping”), while the opposite should occur to low-severity patients.

However, empirical findings show that the effects of DRG tariff refinements on health care provision are relatively small or insignificant. For example, attempting to decompose the effect of a DRG tariff refinement for HIV treatment into a moral hazard component (a change in the intensity of treatment induced by tariffs) and a selection component (or “cream-skimming”), Gilman (2000) found that financial incentives created by the DRG tariff refinement had only a minor effect on supply as reimbursement rates were already very generous. In a more recent study, Janulevicute et al. (2016) use fixed effects models to estimate the responsiveness of healthcare providers to changing stimuli across different DRG groups between 2003 and 2007. They found that a 10% increase in DRG reimbursement rates leads to a 0.8%–1.3% increase in the number of patients treated for medical DRGs in Norway. However, no such effect was found for surgical DRGs.

At the level of individual providers, part of the explanation for the substantial observed variation in C-section rates, both between countries and within countries' smaller “local markets,” can be found in the demand inducement hypothesis. For instance, Gruber et al. (1999) find a direct positive effect of the price differential between C-section and normal childbirth on the rate of C-sections in the US, which itself can be decomposed into an income effect and a substitution effect across services. However, Grant's (2009) replication of Gruber et al. (1999) concludes that much of the difference in C-section and normal delivery rates is in fact attributable to coincidental trends and a sample bias. More recently, Allin et al. (2015) study a sample of Canadian women who gave birth between 2006 and 2011 and conclude that increasing the fee differential between C-sections and normal deliveries does provide an incentive for physicians to favor C-sections over vaginal labor. However, the magnitude of the effect is relatively small: doubling the fee differential increases, on average, the probability of a C-section only by 5.6%. Johnson and Rehavi (2016) provide evidence that physicians perform fewer C-sections on parturient women who are themselves physicians, stressing the importance of the information asymmetry which exists between physicians and mothers when a decision is made regarding a child delivery option.

An important behavioral component deemed to mitigate the impact of an induced-demand channel on medical practice is provider altruism. Analyzing the general evidence available on provider altruism, Galizzi et al. (2015) concludes that although the existing research generally supports the hypothesis that physicians largely behave “altruistically” in their healthcare decisions, the underlying preferences for altruistic behavior are not uniform. In the obstetrics context, Lefèvre (2014) showed that the increase in C-section rates before holiday Mondays (e.g., Memorial Day or Labor Day) in the US was attributable to the scheduling effect of C-sections rather than to convenience or leisure-related grounds.

An alternative explanation for high C-section rates in some countries is fear of malpractice claims. As argued by Dubay et al. (1999), low-income families with lower education levels are more likely to sue their physician on malpractice grounds. Studies that investigated this hypothesis include Currie and MacLeod (2008), Shurtz (2013), and Yang et al. (2009). The evidence tends to support the hypothesis that heightened fears of malpractice claims (proxied by court claims), or reforms increasing physician liability, indeed make physicians resort to defensive medicine, resulting in higher C-section rates.

Tariff changes could have made “upcoding” a possible strategy for healthcare providers, generating more profits without affecting the volume or quality of services. Such an effect is observed by Dafny (2005), who finds that price changes resulting from suppressing age criteria in DRGs (DRG aggregation) in the US in 1988 and subsequent recalibration of tariffs led to an increase of the share of top-coded patients, while evidence that this price shock affected service volume or intensity is weak. A recent study by Di Giacomo et al. (2017) highlights the possibility of the “upcoding” channel in Italian obstetrics markets operating under DRGs. Conditional on the mode of child delivery, they find that hospitals experiencing the highest increases in DRG reimbursement rates (but not necessarily due to refinement) are more likely to subsequently code their patients as more complex in terms of severity. The discussion of the potential implications of this “upcoding” channel is provided in Sections 6.2 and 7.

In the French context, Milcent and Rochut (2009) analyze the variation in C-section probability across public non-profit, private for-profit and private non-profit hospitals, controlling for obstetrician/gynecologist and midwife availability, in the spirit of Gruber and Owings (1996). They find that private for-profit hospitals are significantly more likely to perform C-sections than both public and private non-profit hospitals. Moreover, the ratio of obstetrician/gynecologist per patient has a positive impact on C-section rates, which may be indicative of physician demand inducement. However, in contrast to our strategy, tariff incentives are not the focus of their analysis. In addition, this study considers the year of 2003 when the DRG system was not yet introduced and global budgets were still in use. Regarding “upcoding,” Milcent (2021) documents higher levels of DRG

severity being claimed by French hospitals following a DRG severity refinement in 2009 that concerned multiple specialties, the strongest “upcoding” response being found in private for-profit hospitals.

This paper contributes to the existing literature in three ways. First, methodologically, our contribution consists in proposing a more elaborate empirical strategy to model the impact of hospital-level financial incentives on obstetricians' choices between child delivery modes. In contrast to earlier studies that considered average or aggregated fees for delivery cases (as in Allin et al., 2015) to calculate the corresponding tariff differential, we construct these measures based on the expected DRG tariff for each delivery outcome, taking the uncertainty with regards to final patient severity into account. The fact that we solely rely on pre-admissions (i.e., ante-partum) characteristics for outcome and severity predictions makes it possible to circumvent a potential reverse causality issue—so far not explicitly addressed in the above-reviewed C-section literature—stemming from a likely bidirectional dependence between the chosen child delivery mode and the occurrence of co-morbidities. Second, despite the availability of empirical studies on financial incentives in obstetrics, to our knowledge, our study is the first to focus on the effects of a DRG tariff refinement in the context of child deliveries. Third, we shed more light on the potential mitigating effect of gynecologist concentration on scheduled C-section rates. We hypothesize that this effect results from a greater access to medical care and a second medical opinion, and we differentiate this channel from that of supplier-induced demand, which is potentially affecting all types of C-sections, as discussed, for example, in Gruber and Owings (1996) and Milcent and Rochut (2009).

Our paper is organized as follows: Section 2 addresses context and study motivation, Section 3 presents the data, Section 4 defines the empirical strategy. The main results are presented in Section 5, with robustness checks in Section 6. The results of the study and policy implications are discussed and summarized in Section 7.

2 | INSTITUTIONAL BACKGROUND

The French healthcare system is characterized by the presence of multiple types of healthcare institutions delivering care at highly regulated prices. Starting from 2004, a French equivalent of a DRG system (called Groupes Homogènes de Malades) was gradually introduced to fund hospitals. The system splits all diagnoses into major diagnosis categories in a mutually exclusive way. They are further classified by diagnosis roots, which give the main hospitalization cause. Finally, based on a predefined list of complications, a severity level is attributed to each diagnosis root. It reflects the extent to which the patient's general health status, along with delivery-related medical circumstances, can make medical procedures associated with the diagnosis more risky and therefore more costly for the provider.¹

Since its introduction, the French DRG classification underwent several revisions. The latest major revision in obstetrics DRGs occurred in 2012 with a significant increase in the number of diagnosis groups, in order to fit more narrowly defined groups of patients. In particular, normal deliveries became differentiated into both single/multiple and unipara/multipara.² For C-sections, however, the only differentiation introduced was that between single/multiple deliveries. The number of severity levels also grew substantially within each DRG group. In addition, a number of new diagnoses were introduced into pricing formulas, creating an additional variation in tariffs across DRG classifications. Lastly, DRG tariffs were adjusted by the mother's gestational age so that pre-term and post-term pregnancies are reimbursed to hospitals at a higher rate (see Tables B2–B5 in Appendix B for details).

In France, most medical procedures are fully covered by the national health insurance fund or only require minor out-of-pocket payments. Both public and private hospitals (whether for-profit or not) are reimbursed through nationally set tariffs. Hospitals are compensated for their incurred costs, based on the DRG diagnosis root, its related severity level, and the hospital status (public or private).

C-sections are considered to be among the most commonly researched medical procedures, possibly due to their growing popularity worldwide, with the average C-section rates in OECD countries rising from 20% in 2000 to 28% in 2017 (OECD, 2019). Optimal and medically justified C-section rates have been a subject for both scientific and political debates for decades. The World Health Organization statement on C-section rates issued in 2015 argued that “CS rates higher than 10% were not associated with reductions in maternal and newborn mortality rates,” and that a medically justified rate may be lying within the range of 10%–15% of all child deliveries. In this respect, the relatively stable rate of around 20% observed in France since the late 2000s follows the pattern of other industrialized and economically advanced countries, which tend to report figures well above the WHO's indication.

In France, general practitioners rarely perform child deliveries, unless no qualified midwife or obstetrician is available in the area. Encoding diagnoses and entering medical information in the French administrative medical data recording system (PMSI-Programme Médicalisé des Systèmes d'Information) is performed primarily by admissions' secretaries/nurses, and

occasionally by midwives. When deciding on the mode of child delivery, obstetricians and midwives generally rely on their previous experience and may follow recommended practices summarized, for example, in HAS clinical guidelines (HAS, 2012). The French government requires pregnant women to attend at least seven pre-natal medical appointments with their registered physician and/or a midwife, during which risks of each delivery mode are expected to be discussed. A relatively high emphasis is placed on mothers' preferences for the mode of child delivery, unless medical examinations reveal at least one of the relatively few strong counter-indications against either C-section or normal delivery.

In the non-profit sector, obstetricians and midwives have the status of hospital employees and receive a monthly salary, determined according to a state-mandated salary grid. This grid takes into consideration length of employment and qualifications, but it is not directly tied to employees' performance. The nature of the contract may differ for obstetricians and midwives. While obstetricians are normally hired on a permanent contract with a trial period of several months, it is not uncommon for midwives to be employed on a fixed-term contract, especially at early stages of their careers. This can potentially make midwives relatively more susceptible to hospital-level incentives relative to physicians and obstetricians. In the for-profit sector, obstetricians and, to a lesser extent, midwives, can sign more flexible contracts that include both a fixed salary amount and an income supplement based on a fee-for-service/per birth basis. Thus, the transmission of hospital-level DRG incentives to individual providers can be somewhat facilitated in private-for-profit facilities, inasmuch as healthcare professionals heed, at least partially, income considerations in their medical decisions.

As far as DRG payments are concerned, tariff schedules differ depending on the hospital status. In the private-for-profit sector, DRG tariffs are historically lower than in public hospitals, since DRG schedules in private-for-profit facilities do not incorporate the providers' salary component and only reflect the average material cost of a procedure. Overall, DRG payments in France are supposed to be disconnected from, or have limited impact, on providers' income.

In the French context, DRG reforms therefore offer an opportunity to test the extent to which hospital level financial incentives can be transmitted to individual providers, knowing that the financial situation of the latter is not directly impacted. These hospital level incentives may transfer to providers through explicit internal planning and volume targets/projections set by hospital management, as well as more informal communication channels. Although empirical studies have not systematically drawn the distinction between the levels at which financial incentives are generated, the existing evidence on pay-for-performance funding is suggestive of individual level incentives being more effective in affecting medical practice than hospital-level incentives. In addition, the size of these incentives relative to provider benefit and the nature of their implementation (e.g., redistribution of revenues vs. providing new sources of funding) may also impact the clinical practice (for systematic reviews see Christianson et al. [2008] and Van Herck et al. [2010]). In this respect, the fact that the French 2012 DRG reform was budget-neutral, with resulting incentives significantly affecting only a fraction of child delivery cases may have mitigated the impact of the reform on providers' behavior.

The 2012 DRG reform followed the general trend of increasing DRG tariff refinement in Europe (Busse et al., 2011), with an increase in both the number of obstetric DRGs and the degree of severity differentiation. As a result, the total number of tariff groups increased from 7 to 30 (see Table B2 in Appendix B).

Different impacts were expected on facilities providing obstetric care, depending on their case-mix. Since the most complicated cases started to be reimbursed at higher rates after the DRG tariff refinement, larger hospitals endowed with more advanced medical equipment and well-suited to provide care to mothers with high severity co-morbidities were expected to benefit most from the reform. Smaller maternity wards, on the other hand, may have experienced both lower revenue uncertainty (due to and an increased financial pressure as DRG reimbursement rates decreased for the simpler cases [Vanlerenberghe, 2015]).

From an administrative point of view, DRG systems are considered to be costly to operate since they require regular monitoring and have a high coding-related workload (Busse et al., 2011). After the French 2012 reform, the costs of running the DRG system likely increased along with the number of DRGs. From a societal point of view, this could have led to lower public expenditure efficiency and more transactional costs due to an increased regulatory burden, potentially outweighing the benefits of the reform.

3 | DATA

The primary source of data containing information on patients' hospitalization and diagnoses in France is PMSI dataset, which served as the main source of data in this study. It is an administrative dataset covering all French hospitals starting from 2006.

Actual PMSI data were split into three different datasets, which separately contained information about:

1. Performed medical procedures.
2. DRG main diagnosis, patient and hospital characteristics.
3. Complications associated with the main diagnosis.³

It covers all individuals living in both metropolitan and overseas French departments. Full PMSI data access was given for the years 2010–2013.

The resulting data contains a rich set of individual level characteristics (including age, region and community of residence, length of hospital stay, month and year of hospital leave, etc.) and hospital characteristics, which include a hospital unique identifier, ownership information (i.e., public or private), legal status (i.e., profit or non-profit), geographic location, as well as a list of diagnoses and complications which occurred during each hospital stay. DRG diagnoses and severity levels are encoded in the current DRG version and, in some cases, both current and previous DRG versions. Each hospital stay also has a unique patient identifier which allows us to track hospitalization cases across different years for a given patient. The data provides information on DRG reimbursements for every single child delivery case that occurred in French healthcare institutions. Thus, we were able to observe diagnoses, co-morbidities and in-hospital services registered during childbirth, both before and after labor begins. These variables are summarized in Table 1 and Table B1 in Appendix B.

Different tariff schedules are set for for-profit and non-profit hospitals. To obtain information about tariffs on the basis of which hospitals are compensated, we use publicly available data provided by ATIH.⁴ On average, for-profit and non-profit institutions were affected differently by the reform. As shown in Table 1, the mean value of $\Delta tariff$ was positive (negative) for for-profit (non-profit) hospitals, indicating that C-sections became relatively more (less) profitable to perform in these types of facilities.

Income data were collected by Institut National de la Statistique et des Etudes Economiques (INSEE) and are available on the level of PMSI pseudo postal codes. To control for medical obstetric care availability, we use a gynecologist availability index calculated for the year 2011 at the municipal level. It takes into account gynecologists' volume of activity, service use rates differentiated by population age structure, supply and demand factors in neighboring municipalities and other parameters.⁵

Despite the fact that PMSI is the most complete administrative dataset on hospital admissions available to researchers in France, it does not directly record several characteristics important in our context. For instance, the unipara and multipara status was only observable for women who delivered normally. This creates an obstacle for predicting normal delivery tariffs in the post-reform period for women who delivered via a C-section, since those tariffs depended on the multiparty status.

To recover the unipara/multipara status at individual level, we used variables containing DRG roots from previous obstetrics PMSI datasets available from 2005. Similar to Lo (2008), we trace back the history of all hospital stays for a given patient 5 years prior to actual child delivery, using each patient's unique identifier.

Let M_n denote the fact that a mother was multipara and had a child in the period n , and \tilde{M}_n is multiparity status observed in PMSI data. According to Enquête Périnatale, $P(M_{2010}) = 0.566$. To estimate the number of recovered multiparity cases, we assume that $P(M_t) = const = P(M_{2010})$, $t = \{2010, 2011, 2012, 2013\}$ which is plausible because multiparity is directly linked to demographic characteristics such as fertility and age at delivery. They are stable in countries that have already undergone a demographic transition, such as France. This assumption is also justified by available statistical evidence since, according to the same survey, $P(M_{2003}) = 0.568$.

Overall, the share of all multipara cases that were retrieved at individual level is $\frac{P(\tilde{M}_{2010-2013})}{P(M_{2010-2013})} = 0.947$. The missing 5.3% of cases only cover C-section cases, since the multiparity status is completely observed for women who had a normal delivery.

In order to minimize the potential effect of this missing data on the internal validity of our results, in main models we restrict our analysis to multipara mothers, whose status is derived directly from PMSI or retrieved following the above described procedure.

In addition, after encoding all hospital stays into both pre-and post-reform DRG classification versions, we find that around 1% of cases are misclassified. We exclude these cases from the sample.

4 | EMPIRICAL STRATEGY

The individual-level dataset constructed for this study offers a unique opportunity to investigate healthcare providers' reaction to changes in hospital-level financial stimuli, following the March 2012 reform. In this study, we leverage the fact that after the DRG refinement, clinically identical cases were reimbursed differently, relative to pre-reform periods. As evidenced by the example presented in Table 2, uncomplicated C-sections and normal deliveries with the gestational age between 37 and 44 weeks in the pre-reform period were reimbursed based on a single tariff in each child delivery method, regardless of the single/multiple birth and mother's multiparity status. However, after the 2012 DRG refinement, the single tariff for normal deliveries was split into four new tariff groups based on both of these characteristics, while the single C-section tariff was differentiated only based on the multiparity status, resulting in two DRG tariffs.

TABLE 1 Summary statistics for obstetrical hospital stays in 2010–2013, by hospital type.

Variables	All hospitals		Public		Private (non-profit)		Private (for-profit)	
	Mean	St.D.	Mean	St.D.	Mean	St.D.	Mean	St.D.
Age	29.7	5.4	29.4	5.5	30.6	5.0	30.1	5.1
Multiple pregnancy	0.016	0.13	0.019	0.14	0.011	0.11	0.011	0.11
Multiparity (recovered from diagnoses lists and previous years)	0.54	0.5	0.547	0.5	0.507	0.5	0.533	0.5
C-section	0.207	0.41	0.199	0.4	0.203	0.4	0.223	0.42
– Scheduled	0.073	0.26	0.064	0.25	0.074	0.26	0.095	0.29
– Urgent	0.036	0.19	0.04	0.2	0.028	0.177	0.028	0.17
– Unscheduled	0.096	0.29	0.094	0.29	0.1	0.3	0.1	0.3
Normal delivery	0.794	0.4	0.801	0.4	0.797	0.402	0.776	0.42
Number of co-morbidities	4.46	2.66	4.96	2.86	4.0	2.21	3.37	1.83
Number of previous C-sections	0.09	0.32	0.09	0.32	0.088	0.31	0.096	0.32
Gynecologist availability index	6.95	3.85	6.53	3.69	8.84	4.34	7.35	3.84
Average income per household per person (by pseudo-postal PMSI code), in thousand euro	19.51	4.41	19.13	4.06	21.45	5.89	19.92	4.5
Tariff incentive measures								
$\Delta tariff$	−42.5	274.3	−70.48	307.5	−63.81	280.4	34.64	131.7
$(\Delta^{post} - \Delta^{pre}) / \Delta^{pre}$	−70.34	441.8	−99.38	514.6	−96.53	317.73	11.01	190.8
$r^{post} - r^{pre}$	−0.022	0.25	−0.043	0.27	−0.047	0.26	0.037	0.19
$(r^{post} - r^{pre}) / r^{pre}$	−0.009	0.15	−0.019	0.19	−0.021	0.18	0.02	0.13
Gestational age								
22–35 weeks	0.034	0.18	0.042	0.2	0.023	0.15	0.016	0.13
36–39 weeks	0.513	0.5	0.502	0.5	0.498	0.5	0.544	0.5
≥40 weeks	0.452	0.5	0.454	0.5	0.477	0.5	0.44	0.5
Mother co-morbidities								
Infection during pregnancy	0.042	0.2	0.041	0.2	0.034	0.18	0.047	0.21
Diabetes	0.061	0.24	0.067	0.25	0.063	0.24	0.046	0.21
Complicated breech	0.027	0.16	0.027	0.16	0.025	0.16	0.027	0.16
Oligohydramnios	0.015	0.12	0.017	0.13	0.015	0.12	0.008	0.09
Hypertension/Preeclampsia	0.056	0.23	0.057	0.23	0.058	0.23	0.052	0.22
Fetal growth anomaly	0.043	0.2	0.053	0.22	0.037	0.19	0.019	0.14
Hemorrhage/uterine rupture	0.016	0.13	0.017	0.13	0.014	0.12	0.015	0.12
Premature rupture of membrane	0.047	0.21	0.046	0.21	0.052	0.22	0.048	0.21
Observations	3,041,425		2,004,269		243,155		794,001	

To express the effects of the DRG refinement in terms of changes in the financial profitability of child delivery modes, we introduce a probabilistic algorithm that takes into account the inherent uncertainty faced by mothers and their obstetricians with respect to the realized child delivery mode and its level of severity. The key point of this algorithm consists in estimating DRG financial incentives such that the financial reward associated with each outcome is weighed by its estimated probability. In doing so, one has to rely only on the information available to mothers and obstetricians at the time when the child delivery mode is being chosen. In the context of obstetrics, the period when all child delivery outcomes can still hypothetically realize corresponds to the time before labor starts (i.e., ante-partum).

We implement the following five-step empirical strategy to estimate the effect of financial incentives:

1. We estimate the probabilities of each delivery outcome (i.e., scheduled, urgent, unscheduled C-sections, and normal delivery) based on patient characteristics observed in ante-partum:
 - (a) The estimation is based on the pre-reform (training) dataset using nested logistic models with the tree structure provided in Figure 1;

TABLE 2 Comparison of normal delivery and C-section tariffs in publicly funded hospitals, before and after the 2012 reform.

Normal delivery DRG tariffs (no complication, gestational age 37–44 weeks)				
	2011		2012	
	Unipara	Multipara	Unipara	Multipara
Single birth	2187.13		2458.69	2070.55
Multiple birth			3301.83	2897.28
C-section DRG tariffs (no complication, gestational age 37–44 weeks)				
	2011		2012	
	Unipara	Multipara	Unipara	Multipara
Single birth	2792.71		2850.37	
Multiple birth			3852.63	

Abbreviation: DRG, diagnosis-related group.

- (b) Based on obtained estimates, predicted outcome probabilities \hat{p}_w are assigned to each child delivery mode $w: \{\text{scheduled CS, unscheduled CS, urgent CS, normal delivery}\}$ and to both pre- and post-reform observations. For example, in Figure 1, mother X has predicted outcome probabilities $\hat{p} = \{0.1, 0.05, 0.05, 0.8\}$, where $\hat{p}_w = 0.1$ if $w \in \{\text{scheduled CS}\}$. Model specifications are provided in Appendix D1.
- We classify each hospital stay into both pre- and post-reform DRG classifications (i.e., severity levels and resulting tariffs).
 - We estimate probabilities of each severity level for each child delivery outcome in both DRG classifications:
 - Based on the pre-reform (training) dataset, estimate ordered logit models with levels of severity as the outcome (see Appendix D1 for model specifications);
 - Assign predicted severity probabilities $\hat{d}_{w,s}^0$ and $\hat{d}_{w,s}^1$ to all outcomes in both DRG classifications, where $\hat{d}_{w,s}^0$ and $\hat{d}_{w,s}^1$ are estimated probabilities to observe a given severity level s for outcome w in pre- and post-reform periods (0 and 1, respectively). For example, in Figure 1, the estimated probability $\hat{d}_{CSsch,1}^0$ to observe severity level 1 if mother X delivers her child via a scheduled C-section in the pre-reform period is 0.65.
 - We calculate an incentive measure reflecting the attractiveness of scheduled C-sections:
 - Compute the expected tariff as a probability-weighted average of tariffs over the set of possible severity levels s in pre- and post-reform periods (0 and 1, respectively): $\mathbb{E}(\text{tariff}_w^0) = \sum_s \hat{d}_{w,s}^0 \cdot \text{tariff}_{w,s}^0$ and $\mathbb{E}(\text{tariff}_w^1) = \sum_s \hat{d}_{w,s}^1 \cdot \text{tariff}_{w,s}^1$. For example, if Figure 1 represents a pre-reform DRG tree and if pre-reform tariffs for scheduled C-sections in that period were 1000, 1500, 2000 and 4000 euros for severity levels 1–4, respectively, then the expected tariff of patient X for a scheduled C-section in the pre-reform DRG is $\text{tariff}_{CSsch}^0 = 0.65 \cdot 1000 + 0.25 \cdot 1500 + 0.08 \cdot 2000 + 0.02 \cdot 4000 = 1265$.
 - Compute the difference in expected tariffs between unscheduled C-section and a probability-weighted average tariff of other child delivery modes, separately for pre-reform and post-reform DRG classifications.

$$\Delta^{pre} = \mathbb{E}(\text{tariff}_{CSsch}^0) - \frac{\sum_{w \in D} \hat{p}_w \cdot \mathbb{E}(\text{tariff}_w^0)}{\sum_{w \in D} \hat{p}_w}$$

and

$$\Delta^{post} = \mathbb{E}(\text{tariff}_{CSsch}^1) - \frac{\sum_{w \in D} \hat{p}_w \cdot \mathbb{E}(\text{tariff}_w^1)}{\sum_{w \in D} \hat{p}_w}$$

where $D: \{\text{unscheduled CS, urgent CS, normal delivery}\}$ is a set of delivery outcomes other than scheduled C-sections. For example, in Figure 1, if expected tariffs in the pre-reform DRG $\mathbb{E}(\text{tariff}_{CSun}^0)$, $\mathbb{E}(\text{tariff}_{CSur}^0)$, and $\mathbb{E}(\text{tariff}_{ND}^0)$ were 1500, 2000 and 850 euros, respectively, the average expected tariff for delivery modes other than scheduled C-section is $\sum_{w \in D} \hat{p}_w \cdot \mathbb{E}(\text{tariff}_w^0) = (0.05 \cdot 1500 + 0.05 \cdot 2000 + 0.8 \cdot 850)/0.9 = 950$ euros. Taking the figure for expected scheduled C-section tariffs from bullet 4(a), $\Delta^{pre} = 1265 - 950 = 315$. Graphically, it represents the difference of expected tariffs between two edges at node one in Figure 1.

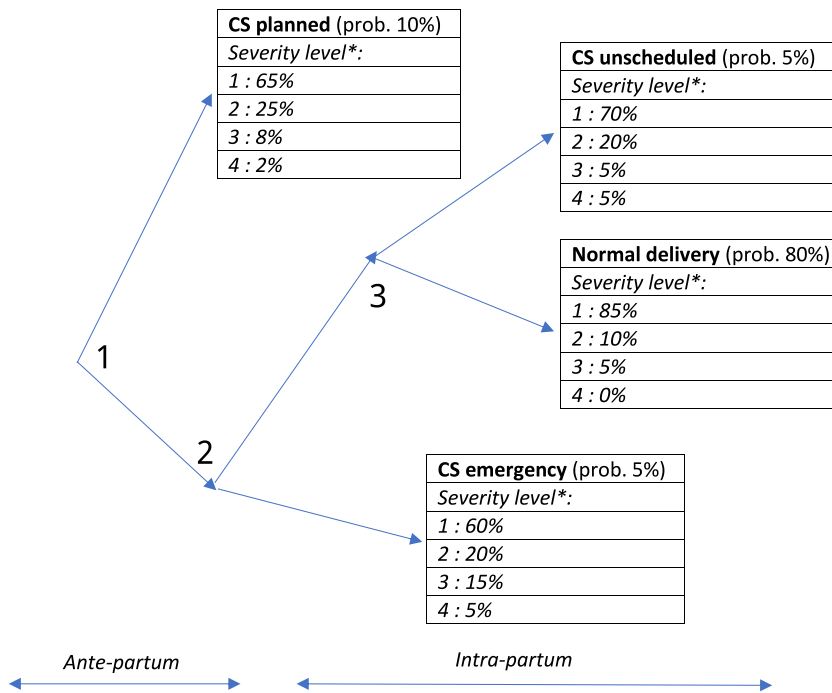


FIGURE 1 Decision tree for child delivery outcomes, with examples of diagnosis-related group levels of severity (probabilities given for illustrative purposes). [Colour figure can be viewed at wileyonlinelibrary.com]

- (c) Compute the difference between DRG classifications $\Delta tariff = \Delta^{post} - \Delta^{pre}$, which reflects the change in financial attractiveness of scheduled C-sections. For example, if, other things equal, we assume that due to a rise in C-section DRG tariffs $\Delta^{post} = 450$, the incentive value is $\Delta tariff = 450 - 315 = 135$ euros.
5. We run difference-in-difference (DiD) models testing the impact of incentives on the probability of scheduled C-section being chosen in ante-partum:
- Based on the value of $\Delta tariff$, separate observations into control (no change in incentives) and two treatment groups, for whom the resulting expected DRG tariff incentive was lower than -100 euros and above 100 euros.
 - Setting $\Delta tariff$ to zero for pre-reform observations (see Section 6.1 for event-study regressions, where pre-reform values of $\Delta tariff$ are kept as originally assigned).

In this study, we rely solely on pre-existing (i.e., ante-partum) patient characteristics to predict the probability of observing each delivery outcome, its severity level and, thus, the probability of belonging to a tariff group. This allows us to calculate the expected values of DRG tariffs that a hospital would receive at the point of time when a choice for the child delivery mode is being made in the ante-partum. The difference Δ^{pre} between the expected tariff for scheduled C-section delivery and probability-weighted average of other delivery modes in the pre-reform period subsequently serves as a reference point of the degree of financial attractiveness of scheduled C-section before the reform. Since tariffs changed in March 2012, so did the expected tariffs, assuming that factors other than the DRG refinement did not affect the underlying patterns of clinical practice. Thus, for every child delivery case, the difference between tariffs for scheduled C-sections and other child delivery modes in post- and pre-reform would reflect the strength of the financial incentive during the decision for or against a scheduled C-section in ante-partum.

At the individual level, the choice of child delivery mode in the post-reform period will not affect the value of the incentive measure, since all predictions are based on ante-partum diagnoses over which an obstetrician/gynecologist has no control. For example, in Figure 1 high blood pressure risks during pregnancy are much higher at node 3 than node 1. In contrast to other studies conducting economic evaluations of obstetric care choices, we also use predicted gestational age instead of the observed one, since the mode of delivery is likely to mechanically affect the length of pregnancy. It tends to be shorter in scheduled C-sections compared to normal deliveries and unscheduled C-sections, since the former are induced by external interventions. Thus, it mitigates, to the extent possible, the risk of coefficient bias due to reverse causality.

Time trends for cases associated with positive, negative and small changes of variable $\Delta tariff$ are presented in Figure 2. This variable is continuous, which makes it possible to estimate how different intensities of the change in financial attractiveness of scheduled C-section relative to other delivery options affected the probability of an obstetrician choosing the scheduled C-section.⁶

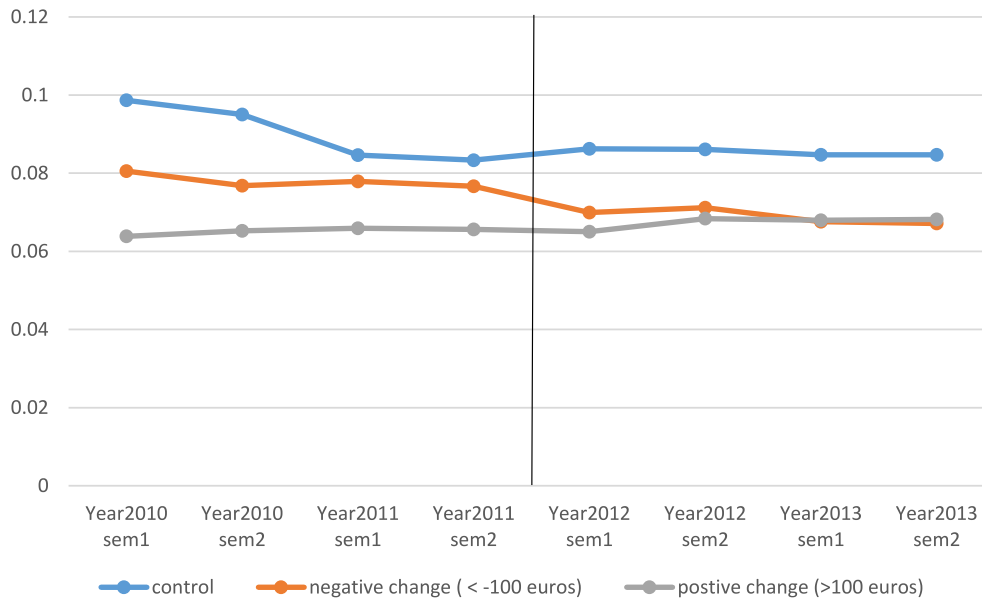


FIGURE 2 Time trend of scheduled C-section rate for patients with positive (>100 euros), negative (< −100 euros) and near-zero (control group) values of $\Delta tariff$, for all hospitals, 2010–2013. [Colour figure can be viewed at wileyonlinelibrary.com]

DiD specifications require that the pre-reform levels of an outcome variable follow parallel trends in order for the obtained estimates to reflect the causal impact of the studied intervention. In Figure 2, the orange line (the group associated with a decrease in $\Delta tariff$ variable by more than 100 euros) shows a minor decrease in scheduled C-section in the first semester of 2012, while the gray line (the group associated with a positive change) shows the opposite, albeit barely noticeable one semester later. In the meantime, C-section rates remained stable for observations in the control group (blue line) post-reform, although a moderate pre-reform decrease is present in the first semester of 2011, compared to the two other groups.

Shortly after the 2012 DRG tariff reform, the French National Authority for Health (HAS) issued obstetrics clinical guidelines and launched a pilot experiment to support their implementation. Maternity wards were invited to organize regular staff coordination meetings and peer reviews aimed at facilitating the adoption of HAS clinical guidelines and reducing care variation across teams and facilities. The participation in this experimentation was voluntary for hospitals and, in total, 165 French maternity wards participated out of 507, that is, 31% (HAS, 2014). Therefore, from 2013 and onward it is difficult to disentangle the pure DRG tariff reform effect from that related to HAS, leading to a possible upward bias in the estimates of interest if this experimentation led to higher scheduled C-section rates.

DiD specifications are commonly used to estimate the causal effect of policy changes. Since the outcome variable—scheduled C-section indicator—is binary, we refer to our model as a linear probability DiD.

To test the effect of the price incentive, we estimate:

$$CSsch_{iht} = \alpha + X'_{iht}\beta_1 + \beta'_2 \Delta tariff POST_t + d'_{iht}\xi + \gamma_t + \theta_h + \nu_t + \epsilon_{iht} \quad (1)$$

where $CSsch_{iht}$ is a binary outcome variable, that is, the probability of a scheduled C-section being performed for individual i , in healthcare institution h , in year t . Term α is constant, X_{iht} is a column vector of controls including age, secondary ante-partum diagnoses and other observed clinical characteristics, gynecologist availability index and pseudo-postal code average income, $\Delta tariff POST_t$ is the interaction term between the DRG refinement incentive and dummy variables for post-reform periods $POST_t$, d_{iht} is a column vector of dummy variables for treatment groups, and ϵ_{iht} is a random error term. Where indicated, the model also includes year, month and hospital fixed effects denoted γ_t , ν_t and θ_h , respectively. In all our models error terms are clustered at the hospital level.

The main coefficients of interest are contained in the β_2 vector. In these specifications, β_2 coefficients represent the percentage points by which the probability of a scheduled C-section increases if the DRG refinement incentive for scheduled

C-sections, scaled to thousand of euros, increases by one unit relative to the pre-reform DRG version. In other terms, if tariffs for other delivery modes are kept fixed, a 1000 euro increase in expected scheduled C-section tariff relative to the pre-reform DRG schedule would, on average, lead to a $100 \cdot \beta_2$ percentage point increase in the scheduled C-section probability.

Models are tested for public, private non-profit and for-profit care institutions, separately for single and multiple child deliveries.

5 | MAIN RESULTS

5.1 | DiD models with deterministic severity

We first present results from deterministic models to measure DRG tariff incentives for scheduled C-sections, in line with previous studies (e.g., Allin et al., 2015; Gruber et al., 1999). We therefore ignore the tree structure of obstetrician/midwife decisions about child delivery modes and consider that tariffs are known in advance with certainty, based on ante-partum characteristics.

Deterministic models rely on additional assumptions. First, it is assumed that the main co-morbidity (and thus, DRG severity levels) can perfectly predict the final DRG tariff for delivery modes based on the most severe ante-partum co-morbidity observed in ante-partum for a given mother. This assumption is likely to hold for scheduled C-sections and normal deliveries, where the most severe patient diagnosis is usually known during ante-partum. However, unscheduled C-sections are less predictable because they are normally performed when a vaginal delivery is no longer possible, due to high risks of severe health consequences. As a result, in this subsection, we restrict the analysis to the choice between a scheduled C-section and normal deliveries.

Second, we acknowledge that obstetricians and midwives would not prolong mothers' labor unnecessarily, leading to an increased severity (and subsequently a higher DRG tariff). Although violation of this assumption may lead to the reverse causality problem, making the incentive measure dependent on the outcome, we strongly believe that it holds for the overwhelming majority of obstetricians and midwives.

The deterministic measure $\Delta tariff$ follows the probabilistic algorithm described in Section 4, with the exception that in Step 3a the level of childbirth severity is assigned deterministically based on the most severe ante-partum co-morbidity and that the set of possible childbirth outcomes is restricted to scheduled C-sections and normal deliveries.

In Table 3, models (1)–(2) estimate the impact of financial incentives in public hospitals by semesters using all the controls specified in Section 3, with and without hospital fixed effects. The obtained estimates for the pre-reform semesters suggest the possibility of different pre-reform trends in control and treated groups due to statistically significant coefficients for the second semester of 2010 and the first semester of 2011. As for private non-profit and private for-profit maternity wards, the evidence of such trends is much weaker, which lends more credibility to the use of DiD specifications to inform the analysis. Post-reform coefficients β_2 in models (4)–(6) tend to be statistically insignificant and suggest that financial incentives associated with the 2012 DRG reform did not affect obstetric providers' practice. As for private for-profit maternity wards, the results are overall inconclusive, with coefficients for some semesters taking negative values and others remaining statistically insignificant at the 10% level. However, the statistically significant coefficients for the first semester in 2012 and the first semester in 2013 in models (7)–(8) only translate to a modest effect on the probability of scheduled C-sections: other things equal, a 1000 euro increase in the C-section tariff in the post-reform DRG schedule is expected to result in around 4 p.p. reduction in the scheduled C-section probability. The introduction of hospital fixed effects in specifications (2), (5), and (8) led to a very moderate increase in the share of explained variation and did not alter the main results. Gynecologist availability tends to have a moderating impact on the probability of scheduled C-sections, which is evidenced by negative coefficients across almost all studied subsamples and model specifications.

Table 4 reports estimated DRG tariff effects separately for groups with a positive or negative change in tariff incentives, based on the deterministically computed measure $\Delta tariff$. Overall, the results tend to support the hypothesis that the 2012 DRG reform did not significantly affect the probability of scheduled C-sections. Models (4)–(6) and (7)–(9) report statistically insignificant estimates, with the exception of the low $\Delta tariff$ group in the subsample of multiple pregnancy women giving childbirth in a private for-profit maternity ward. Although coefficients for the low $\Delta tariff$ group in the subsample of single pregnancy women delivering a child in public maternity wards are negative and statistically significant, the results should be interpreted with caution due to the probable existence of pre-reform trends. As earlier, higher gynecologist availability is generally associated with a lower probability of delivering a child via a scheduled C-section. Additional models that use the sample consisting of both unipara and multipara women are presented in Appendix C1.

TABLE 3 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta tariff > 100$ and $\Delta tariff < -100$) on scheduled C-section probability in multipara women in 2010–2013.

	Public			Private non-profit			Private for profit		
	Single	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple
	1	2	3	4	5	6	7	8	9
$\Delta tariff \times year$ 2010 semester 1	-0.0063 (0.0039)	0.006 (0.0039)	0.0095 (0.0173)	0.0003 (0.0133)	-0.0015 (0.0136)	-0.075 (0.113)	0.0052 (0.0138)	-0.0083 (0.0138)	-0.0488 (0.0498)
$\Delta tariff \times year$ 2010 semester 2	0.0069 (0.0032)**	0.0068 (0.0033)**	0.0156 (0.0142)	0.0052 (0.012)	0.0034 (0.0127)	-0.1759 (0.0846)**	0.0083 (0.142)	0.0046 (0.0144)	-0.0377 (0.0387)
$\Delta tariff \times year$ 2011 semester 1	-0.021 (0.0037)***	-0.0212 (0.0038)***	0.0079 (0.018)	-0.008 (0.0134)	-0.0093 (0.014)	-0.1144 (0.0864)	-0.0112 (0.0123)	-0.0149 (0.0129)	0.0165 (0.0426)
$\Delta tariff \times year$ 2012 semester 1	-0.0173 (0.0036)***	-0.0173 (0.0036)***	0.0091 (0.0164)	-0.0266 (0.0107)**	-0.0273 (0.011)**	-0.1748 (0.1204)	-0.0283 (0.0136)**	-0.0275 (0.0135)**	0.0396 (0.0522)
$\Delta tariff \times year$ 2012 semester 2	0.0214 (0.0036)***	0.00119 (0.0037)***	0.0093 (0.0168)	0.0037 (0.0144)	0.0021 (0.0148)	-0.1673 (0.1128)	-0.0082 (0.0153)	-0.0118 (0.0153)	0.0486 (0.0502)
$\Delta tariff \times year$ 2013 semester 1	-0.0159 (0.0035)***	-0.0162 (0.0035)***	-0.0035 (0.0158)	0.0105 (0.0123)	-0.0113 (0.0122)	-0.1307 (0.0895)	-0.0382 (0.0142)***	-0.0402 (0.0143)***	0.0108 (0.0603)
$\Delta tariff \times year$ 2013 semester 2	-0.0187 (0.004)***	-0.0188 (0.004)***	0.0145 (0.0192)	0.0062 (0.0148)	-0.0064 (0.0149)	-0.145 (0.1062)	-0.0155 (0.0157)	-0.0189 (0.0159)	0.0689 (0.0057)
Gynecologist access index	-0.001 (0.0001)***	-0.0005 (0.0001)***	-0.0002 (0.0009)	-0.0007 (0.0003)**	-0.0007 (0.000)***	0.0064 (0.0032)*	-0.0002 (0.0002)	-0.0006 (0.0001)***	-0.0026 (0.0018)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	955,342	955,342	15,356	117,178	117,178	1167	382,687	382,676	3692
<i>R</i> ²	0.28	0.28	0.17	0.33	0.33	0.28	0.39	0.39	0.21

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

Abbreviation: DRG, diagnosis-related group.

***1% sign., **5% sign., *10% sign.

5.2 | DiD models with probabilistic DRG tariff incentives

The results of various model specifications using the tariff incentive measure $\Delta tariff$ computed as described in Section 4 are presented in Tables 5 and 6. In Table 5, the coefficients for clinical characteristics reported in models (1)–(9) are in line with the general knowledge about obstetric interventions. Namely, having previously delivered via C-sections considerably increases the probability of a C-section in subsequent childbirths. C-sections in general are known to be associated with an increased risk of normal delivery failure, although they are not considered per se as a counter-indication to a normal delivery by the HAS.

The results also suggest that a higher concentration of gynecologists was consistently associated with a lower probability of a scheduled C-section. Although models presented in Milcent and Rochut (2009) and Gruber and Owings (1996) reported a positive coefficient for obstetrician and gynecologist concentration, their respective studies focused on C-sections regardless of whether or not they were scheduled.

For scheduled C-sections, a higher concentration of gynecologists may produce the reverse effect to that expected for unscheduled C-sections. While the latter are likely, according to Milcent and Rochut (2009) and Gruber and Owings (1996), to be driven by supplier-induced demand, scheduled C-section rates, on the other hand, are likely affected through the channel of reduced informational asymmetry between patients and care providers due to an easier access to a second medical opinion. This may serve as credible deterrence against unjustified C-section decisions.

Models (1)–(2) estimate the impact of DRG financial incentives on single birth multipara admitted to public facilities, using controls specified in Section 3, without and with hospital fixed effects, respectively. Although in both models the coefficient

TABLE 4 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta tariff > 100$ and $\Delta tariff < -100$) on scheduled C-section probability in multipara women in 2010–2013.

	Public			Private non-profit			Private for profit		
	Single	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple
	1	2	3	4	5	6	7	8	9
$T_{high} \cdot POST$	0.0169 (0.0108)	0.017 (0.0108)	-0.0312 (0.0391)	-0.0417 (0.0571)	-0.0436 (0.0576)	0.0208 (0.1850)	-0.0515 (0.0303)	-0.051 (0.0301)	0.0026 (0.021)
$T_{low} \cdot POST$	0.033 (0.0119)***	0.0334 (0.0118)***	-0.0292 (0.0414)	-0.0262 (0.0584)	-0.0287 (0.059)	0.1574 (0.1966)	-0.0253 (0.0315)	-0.0239 (0.0313)	-0.3779 (0.2052)*
Gynecologist access index	-0.0009 (0.0001)***	-0.0005 (0.0001)***	-0.0014 (0.0012)	-0.0006 (0.0003)*	-0.0007 (0.0001)***	0.0025 (0.0022)	-0.0001 (0.0001)	-0.0006 (0.0001)***	-0.0027 (0.0018)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
N	955,342	955,342	15,356	117,178	117,178	1167	382,687	382,676	3692
R^2	0.28	0.28	0.17	0.33	0.33	0.23	0.39	0.39	0.21

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

Abbreviation: DRG, diagnosis-related group.

***1% sign., **5% sign., *10% sign.

of interest β_2 is negative and statistically significant at 1%, its magnitude is small (0.7% decrease in C-section rate with every 1000 euro change in incentives). However, as we will explore in Section 6.1, public hospitals do fail the placebo test, warranting additional caution while interpreting this result. Assuming the magnitude of β_2 of 0.007 and the average size of incentives equaling 42 euros, the change in C-section rates attributable to the 2012 DRG reform is estimated to be lower than 0.1%. These findings are in line with conclusions from Allin et al. (2015) and Di Giacomo et al. (2017), that is, that the effect of the C-section to normal delivery fee ratio on C-section rates is likely to be modest and is statistically insignificant.⁷

Both private non-profit and private for-profit maternity wards show no significant reaction to DRG stimuli in terms of their propensity to perform scheduled C-sections.

In model (4), which applies to single birth mothers, the coefficient β_2 is significant, but this is no longer the case in model (5) after including hospital-fixed effects. As for the remaining models (6)–(9), β_2 is consistently insignificant. In general, the introduction of hospital-fixed effects in specifications (2), (5) and (8) led to only a negligible increase in the share of explained variation. Table 6 provides similar results for each post-reform semester separately.

In Tables 7 and 8 we present DiD estimation results for the effect of DRG incentives on groups that were affected in opposite directions. Thus, $\Delta tariff > 100$ indicates an incentive that favors scheduled C-section, while $\Delta tariff < -100$ indicates an incentive that favors other modes of childbirth. The set of included covariates is identical to specifications presented in Table 5.

Overall, across all hospital types, the results indicate no significant impact of DRG incentives supporting the hypothesis of revenue-maximizing care providers. In Table 7, coefficients β_2 for the group with a positive change in the tariff incentives tend to be either non-significant or, as in models (4)–(6) for private non-profit maternity wards, take a negative sign. Likewise, coefficients β_2 for the group with a negative tariff incentive change are non-significant and/or take a sign opposite to the one expected if profit-maximization incentives had played a significant role in clinical decisions. Similar results are obtained in Table 8 where effects are estimated by each post-reform semester. As discussed earlier, the size of the estimated impact in all cases remains very modest and unlikely to affect in a meaningful and noticeable way the pre-existing obstetric practices.

Compared to models with deterministically computed incentives, the probabilistic approach tends to generate fewer statistically significant estimates with smaller effect magnitudes. However, due to the limitations of the deterministic approach discussed in Section 4, its results should be interpreted with additional caution.

TABLE 5 DiD estimation of the impact of DRG refinement incentives ($\Delta tariff$, rescaled to thousands of euros) on scheduled C-section probability in multipara women in 2010–2013.

	Public			Private non-profit			Private for profit		
	Single	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple
	1	2	3	4	5	6	7	8	9
$\Delta tariff_{POST}$	-0.0071 (0.0026)***	-0.0074 (0.0024)***	-0.0121 (0.0274)	-0.0066 (0.0069)**	-0.0074 (0.007)	-0.29 (0.22)	-0.012 (0.01)	0.013 (0.01)	-0.034 (0.041)
Age	-0.012 (0.004)***	-0.02 (0.005)***	0.1295 (0.0843)	-0.036 (0.017)**	-0.03 (0.021)	-0.3 (0.26)	0.008 (0.008)	-0.006 (0.1)	0.091 (0.18)
ln(age)	-0.19 (0.06)***	-0.24 (0.06)***	-1.387 (1.1582)	0.52 (0.25)**	0.4 (0.26)	0.46 (0.38)	-0.11 (0.12)	0.03 (0.13)	-0.65 (2.85)
Age ²	<0.0001 (<0.0001)***	-0.015 (0.005)***	-0.008 (0.0007)*	-0.0013 (0.00014)**	0.0002 (0.0002)	0.002 (0.002)	-0.0006 (0.0007)	0.0001 (0.001)	-0.001 (0.001)
$E(gestational\ age)$	0.003 (0.005)	-0.04 (0.018)**	0.287 (0.2579)	-0.014 (0.017)	-0.004 (0.084)	0.1 (0.12)	-0.03 (0.008)***	-0.034 (0.03)	-0.011 (0.049)
Number of prev. CS	0.295 (0.005)***	0.296 (0.005)***	0.2762 (0.013)***	0.356 (0.011)***	0.353 (0.01)***	0.408 (0.041)***	0.426 (0.005)***	0.422 (0.005)***	0.372 (0.019)***
Median income (postal code)	-0.0001 (0.0002)	-0.0003 (0.0002)	0.0018 (0.0007)**	0.0004 (0.0003)	-0.0001 (0.0002)	0.0002 (0.0027)	-0.0006 (0.0003)**	-0.0004 (0.0001)***	0.003 (0.0015)*
Gynecologist access index	-0.0009 (0.001)***	-0.0005 (0.0001)***	-0.0002 (0.0009)	-0.0006 (0.0003)*	-0.0007 (0.0001)***	0.002 (0.0023)	-0.0002 (0.0002)	-0.0007 (0.0001)***	-0.027 (0.0018)
Treated_high	-0.0012 (0.0008)	-0.0022 (0.0008)***	0.038 (0.0336)	-0.01 (0.0067)	-0.012 (0.006)*	-0.089 (0.192)	-0.016 (0.005)***	-0.016 (0.005)***	0.0096 (0.024)
Treated_low	0.01 (0.002)	0.009 (0.0023)***	0.0055 (0.0297)	-0.024 (0.001)**	-0.026 (0.01)**	-0.406 (0.206)	-0.016 (0.005)***	-0.016 (0.006)***	-0.096 (0.069)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	955,342	955,342	15,356	117,178	117,178	1167	382,687	382,676	3692
<i>R</i> ²	0.28	0.28	0.34	0.33	0.33	0.37	0.39	0.39	0.26

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

Abbreviation: DRG, diagnosis-related group.

***1% sign., **5% sign., *10% sign.

6 | ROBUSTNESS CHECKS

6.1 | Event study analysis

To test whether the estimates obtained in Section 5 are not affected by pre-reform changes uncontrolled for in our models, we preform DiD placebo estimation. These models include as covariates the interaction terms between the probabilistic DRG incentive measure $\Delta tariff$ and dummies for all semesters in 2010–2013, except for the second semester of 2011 serving as the baseline pre-reform period. Estimation results are presented in Table 9 and illustrated in Figure 3 for each hospital status.

As mentioned in Section 5, results on single birth mothers who delivered birth in public maternity wards may not stand this placebo DiD test. In particular, in Table 9 model (1) shows 2 out of 3 statistically significant coefficients for pre-reform periods (for the first semesters of 2010 and 2011, with 5% and 1% level of significance, respectively). For the other hospital

TABLE 6 DiD estimation of DRG refinement on treatment groups ($\Delta\text{tariff} > 100$ and $\Delta\text{tariff} < -100$) on scheduled C-section probability in multipara women in 2010–2013.

	Public		Private non-profit		Private for profit	
	Single	Multiple	Single	Multiple	Single	Multiple
	1	2	3	4	5	6
$\Delta\text{tariff} \times \text{year 2012 semester 1}$	−0.0085 (0.0034)**	−0.0039 (0.044)	−0.0247 (0.0113)	−0.3711 (0.2547)	−0.0151 (0.015)	−0.0261 (0.0157)
$\Delta\text{tariff} \times \text{year 2012 semester 2}$	−0.011 (0.0041)***	0.0296 (0.043)	0.0096 (0.007)	−0.4534 (0.2404)*	0.0017 (0.0142)	0.0418 (0.0373)
$\Delta\text{tariff} \times \text{year 2013 semester 1}$	−0.0034 (0.0034)	−0.0422 (0.0254)*	−0.029 (0.015)	−0.0552 (0.2805)	−0.0272 (0.0158)*	−0.0365 (0.1305)
$\Delta\text{tariff} \times \text{year 2013 semester 2}$	−0.0065 (0.0043)	−0.0094 (0.0311)	−0.023 (0.015)	0.0816 (0.2941)	−0.0086 (0.0161)	−0.226 (0.1333)*
Gynecologist access index	−0.0005 (0.0001)***	−0.0002 (0.0009)	−0.0008 (0.0002)***	−0.0019 (0.0022)	−0.0006 (0.0001)***	−0.0028 (0.0018)
Other controls	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	NO	YES	NO
<i>N</i>	955,342	15,356	117,178	1125	382,687	3692
<i>R</i> ²	0.28	0.17	0.33	0.36	0.39	0.26

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

Abbreviation: DRG, diagnosis-related group.

***1% sign., **5% sign., *10% sign.

TABLE 7 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta\text{tariff} > 100$ and $\Delta\text{tariff} < -100$) on scheduled C-section probability in multipara women in 2010–2013.

	Public			Private non-profit			Private for profit		
	Single	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple
	1	2	3	4	5	6	7	8	9
$T_{\text{high}} \cdot \text{POST}$	0.015 (0.013)	0.0009 (0.0013)	0.0312 (0.0642)	−0.0155 (0.0076)**	−0.0159 (0.007)**	−1.086 (0.271)***	0.0075 (0.0088)	0.0065 (0.0088)	0.0637 (0.0333)*
$T_{\text{low}} \cdot \text{POST}$	0.0061 (0.0026)**	0.005 (0.0027)**	0.007 (0.0379)	−0.0074 (0.0125)	−0.0079 (0.0126)	−0.504 (0.2073)***	0.0147 (0.0114)	0.0141 (0.0112)	−0.0197 (0.0688)
Gynecologist access index	−0.0009 (0.0002)***	−0.0005 (0.0001)***	−0.0002 (0.0009)	−0.0006 (0.0003)*	−0.0007 (0.0003)***	0.0024 (0.0023)	−0.0002 (0.0002)	−0.0007 (0.0001)***	−0.0027 (0.0018)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	955,342	955,342	15,356	117,178	117,178	1167	382,687	382,676	3692
<i>R</i> ²	0.28	0.28	0.17	0.33	0.33	0.37	0.39	0.39	0.26

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

Abbreviation: DRG, diagnosis-related group.

***1% sign., **5% sign., *10% sign.

TABLE 8 DiD estimation of DRG refinement on treatment groups ($\Delta tariff > 100$ and $\Delta tariff < -100$) on scheduled C-section probability in multipara women in 2010–2013.

	Public		Private non-profit		Private for profit	
	Single	Multiple	Single	Multiple	Single	Multiple
	1	2	3	4	5	6
$T_{high} \times \text{year 2012 semester 1}$	0.002 (0.015)	-0.0875 (0.0889)	-0.01 (0.007)	-0.13 (0.07)*	0.001 (0.009)	-0.025 (0.048)
$T_{high} \times \text{year 2012 semester 2}$	0.002 (0.0015)*	0.1019 (0.1012)	-0.005 (0.007)	0.145 (0.107)	0.004 (0.009)	0.093 (0.051)*
$T_{high} \times \text{year 2013 semester 1}$	0.0003 (0.002)	-0.0306 (0.0667)	-0.029 (0.015)*	-0.134 (0.085)	0.008 (0.012)	0.064 (0.057)
$T_{high} \times \text{year 2013 semester 2}$	-0.002 (0.0019)	0.1233 (0.0799)	-0.023 (0.015)	0.011 (0.089)	0.01 (0.012)	0.033 (0.044)
$T_{low} \times \text{year 2012 semester 1}$	0.008 (0.005)*	-0.0016 (0.0486)	0.027 (0.018)	0.02 (0.044)	0.017 (0.014)	-0.082 (0.084)
$T_{low} \times \text{year 2012 semester 2}$	0.012 (0.005)**	-0.0154 (0.0477)	-0.027 (0.016)*	0.011 (0.043)	0.002 (0.015)	-0.049 (0.116)
$T_{low} \times \text{year 2013 semester 1}$	0.001 (0.004)	0.0307 (0.0458)	-0.0227 (0.019)	-0.061 (0.055)	0.021 (0.016)	-0.007 (0.143)
$T_{low} \times \text{year 2013 semester 2}$	-0.0005 (0.005)	0.0157 (0.0453)	-0.017 (0.022)	-0.078 (0.054)	0.143 (0.017)	-0.138 (0.175)
Gynecologist access index	-0.0005 (0.0001)***	0.0002 (0.001)	-0.0007 (0.0001)***	0.0022 (0.0022)	-0.0006 (0.0001)***	-0.0027 (0.0018)
Other controls	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	NO	YES	NO
N	955,342	15,356	117,178	1125	382,687	3692
R^2	0.28	0.17	0.33	0.36	0.39	0.26

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

Abbreviation: DRG, diagnosis-related group.

***1% sign., **5% sign., *10% sign.

types and mother groups, the results do not indicate that pre-reform periods were affected through other channels that would be uncontrolled for.

As for the estimated post-reform effects, the obtained results confirm the conclusion that DRG incentives did not significantly affect the probability of scheduled C-section in private non-profit and for-profit maternity wards. A few significant coefficients in models (3) and (6) do not exhibit a clear pattern. In public hospitals, two out of 4 coefficients corresponding to both 2012 semesters are negative and statistically significant. However, in light of the discussed pre-reform estimates, this should be interpreted with caution.

Finally, it is worth noting that similarly to the models presented in the previous section and even regardless of the results of placebo DiD tests on the significance of pre-reform coefficients, the magnitude of the post-reform estimated effects remains low and overall unlikely to impact the clinical practice. Appendix C presents additional models of this type, where hospitals are grouped by type/size (Tables C3, C4 and D1–D5) and different incentive thresholds are used in defining control and treated groups in Appendix D2. Conclusions corroborate the main results presented in this section.

TABLE 9 DiD placebo regressions on scheduled C-section probability in multipara women in 2010–2013 ($\Delta tariff$ rescaled to thousands of euros).

	Public		Private non-profit		Private for profit	
	Single	Multiple	Single	Multiple	Single	Multiple
	1	2	3	4	5	6
$\Delta tariff \times$ year 2010 semester 1	0.0051 (0.0018)**	-0.0315 (0.0338)	0.0055 (0.0096)	-0.415 (0.262)	0.016 (0.012)	-0.067 (0.212)
$\Delta tariff \times$ year 2010 semester 2	-0.002 (0.0026)	-0.0126 (0.0271)	0.013 (0.007)*	-0.28 (0.294)	0.013 (0.013)	-0.017 (0.175)
$\Delta tariff \times$ year 2011 semester 1	-0.012 (0.0037)***	-0.0236 (0.0305)	-0.003 (0.013)	-0.262 (0.236)	0.013 (0.014)	0.095 (0.171)
$\Delta tariff \times$ year 2012 semester 1	-0.0094 (0.0032)**	-0.0197 (0.0393)	-0.0212 (0.011)*	-0.6 (0.26)**	-0.0054 (0.017)	-0.037 (0.18)
$\Delta tariff \times$ year 2012 semester 2	-0.012 (0.0044)***	0.031 (0.0454)	0.0134 (0.0124)	-0.504 (0.25)*	0.0076 (0.015)	0.0045 (0.037)
$\Delta tariff \times$ year 2013 semester 1	-0.0043 (0.0038)	-0.0508 (0.024)**	0.0084 (0.0104)	-0.319 (0.31)	-0.019 (0.018)	-0.041 (0.143)
$\Delta tariff \times$ year 2013 semester 2	-0.0073 (0.0046)	0.002 (0.0341)	0.0019 (0.016)	-0.1599 (0.404)	-0.0036 (0.017)	0.228 (0.138)*
Gynecologist access index	-0.0005 (0.0001)***	-0.0014 (0.0012)	-0.0008 (0.0002)***	0.0017 (0.0022)	-0.0006 (0.0001)***	-0.0028 (0.0018)
Other controls	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	NO	YES	NO
<i>N</i>	955,342	15,356	117,178	1167	382,687	3692
<i>R</i> ²	0.28	0.17	0.33	0.23	0.39	0.21

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

6.2 | Upcoding by care providers

Since the 2012 DRG reform resulted in better reimbursement for providing care to more severe cases, the upcoding risk (i.e., coding patients as more severe without clinical justification), which has been documented in the empirical literature (Dafny, 2005; Di Giacomo et al., 2017), may have increased.

To test the plausibility of this channel, we run DiD models on the samples of patients delivering via different delivery modes. The tested outcome is the severity level for the corresponding child delivery mode in the post-reform DRG schedule. Post-reform DRG severity can serve as an immutable measurement of child delivery clinical complexity for mothers who delivered both before and after the DRG schedule was introduced. The rationale for choosing these outcomes lies in the fact that, if for a given child delivery it becomes significantly more profitable to perform a scheduled C-section compared to normal delivery (i.e., $\Delta tariff \gg 0$), the scheduled C-section will also be more likely to be upcoded since, probabilistically, there are more high-paying tariff groups corresponding to scheduled C-sections.

The estimation results for models testing the post-reform DRG severity level as the outcome variable are provided in Tables D1–D5 in Appendix D2. Overall, as expected, coefficients β_2 in Table D1 tend to be positive and significant in both scheduled C-sections and non-scheduled C-sections and take a negative sign in normal deliveries. However, the vast majority of these models do not stand placebo tests performed similarly to those presented in Subsection 6.1. For example, nearly all models for single birth mothers presented in Tables D2–D4 show signs of pre-existing nonparallel trends in the tested outcomes. This pre-existing trend and its causes are discussed in more detail in Section 7.

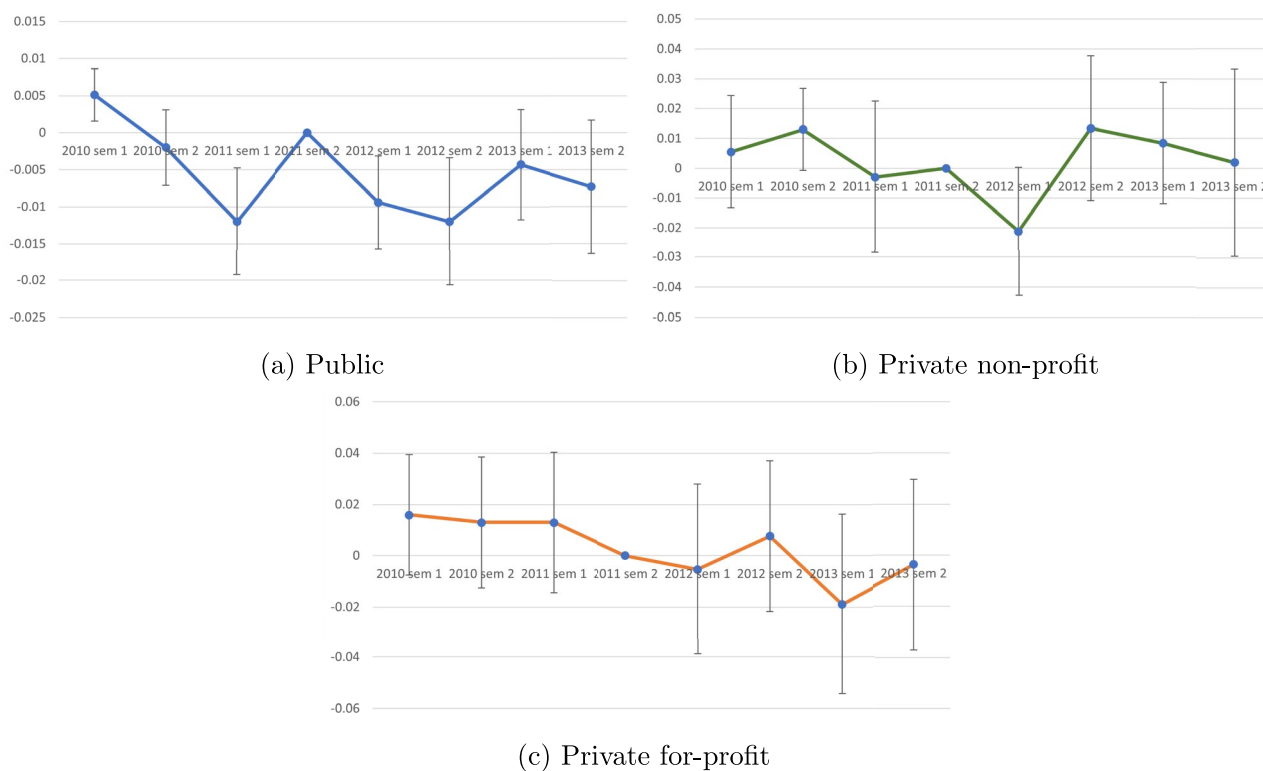


FIGURE 3 Diagnosis-related group reform effects for single birth multipara women (by semester, second semester 2011 as baseline), in public, private non-profit and private non-profit maternity wards, 95% confidence interval. [Colour figure can be viewed at wileyonlinelibrary.com]

7 | DISCUSSION AND CONCLUSION

In this study, we have shown the absence of a link between financial incentives and providers' choices between scheduled C-sections and normal deliveries. Obstetricians working either in non-profit or for-profit hospitals in France did not significantly respond to changes in tariffs caused by the 2012 DRG refinement. This main finding provides evidence that changes in financial stimuli arising at the centralized level do not seem to transfer from hospital management to clinicians' individual practice.

Three main channels are most likely to drive these results. First, healthcare providers' contracts make them insensitive to hospital-level financial objectives. As a result, it is difficult for hospital managers to ensure that financial reforms are implemented, since individual providers have substantial autonomy and legal guarantees. At providers' level, it may imply their adherence to “non-nocere” principles, or to altruistic considerations.

Second, due to the increasing complexity of hospital payment schemes, the time needed to evaluate the financial consequences of DRG reforms can be high for both healthcare management and individual practitioners. Since every single agent can be expected to use his own method to estimate changes in profitability between child delivery modes, the resulting profitability measures are bound to incorporate a degree of measurement error. As a result, our estimates are likely to be affected by the attenuation bias, making them biased toward zero.

Lastly, the reform eased the financial pressure on larger maternity wards since it resulted in an overall revenue increase. The 2012 DRG refinement mostly benefited those maternity wards that had a more than proportional share of high-severity cases, with complicated co-morbidities. Thus, after the reform, larger maternity wards did not have strong incentives to react to changes in terms of relative DRG tariffs because of their improved financial situation. On the other hand, overall revenues of smaller maternity wards generally remained stable or even decreased, making this channel less relevant for them.

Finally, some of the channels studied in the US or Canada to explain providers' choice of a delivery option, such as fear of malpractice claims or patient selection, are not likely to play a significant role in the French context. In addition, according to expert opinion collected from interviewed obstetric professionals, such exogenous factors as technical changes in child delivery practices were unlikely to play a significant role in our findings.

The evidence presented in this paper suggests that the main objective of the 2012 DRG refinement—decreasing financial risk for maternity wards—was successfully met without being associated with significant impacts on individual clinicians' practices. Yet the overall welfare impact of the DRG refinement reform is not clear. In line with Busse et al. (2011), transaction

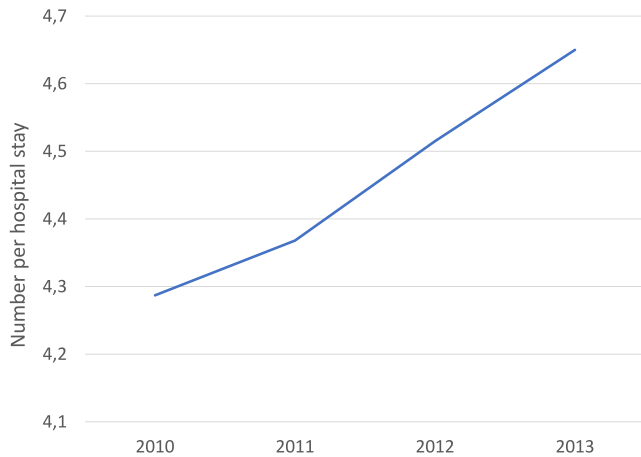


FIGURE 4 Average number of encoded co-morbidities per child delivery (2010–2013). [Colour figure can be viewed at wileyonlinelibrary.com]

and administrative costs associated with coding are likely to have increased as a consequence of the DRG refinement but the exact magnitude of this effect cannot be observed.

This study has several limitations. First, DiD strategy makes it possible to eliminate only the time-invariant component in group unobserved heterogeneity. A potential source of time-variant unincorporated variation could be the publication of practice guidelines by HAS in 2012 and subsequent 2013–2014 experimentation (see Section 4 for details), which were at least in part coincidental with the DRG reform. However, the HAS guidelines contained no strong indications or counter-indications for scheduled C-sections. Hence, the extent to which it could have affected clinical practices appears to be restricted to cases where obstetric patients presented with only relatively common co-morbidities, the final decision still remaining at the obstetrician's end, and in some cases, at the mother's discretion. In addition, the HAS experiment did not include the post-reform year 2012, for which no statistically significant effects of the DRG reform effects were found, and did not directly affect hospitals' monetary incentives.

Second, the upcoding channel discussed in Subsection 6.2, if present, can lead to non-random time-dependent measurement error in co-morbidity variables, which could cause estimates' bias and inconsistency. However, the percentage of top-coded patients did not change markedly over the study period. These patients with severe and major ante-partum co-morbidities represented only 0.21% of C-section cases in 2010 versus 0.26% in 2013, while corresponding figures for mothers who delivered normally were 0.31% and 0.37% in 2010 and 2013, respectively. There are two main reasons to believe that this “upcoding” channel did not significantly impact obtained estimates. First, as evidenced in Figure 4, the upward trend in the number of coded co-morbidities existed before the 2012 reform, making it unlikely for this channel to significantly bias the results of the estimates. In obstetrics, this trend was likely due to an increase in women's average age at child delivery (and, thus, higher frequency of age-related complications) and changes in co-morbidity classification in French DRG manuals. The latter occurred as a result of a consistently increasing number and level of thoroughness in co-morbidities coding with each subsequent versions of the DRG manual in France. Second, our identification strategy relies solely on ante-partum co-morbidities, which are less likely to be manipulated, due to patients' ability to better evaluate their health status in ante-partum and, in case of doubt, to seek a second medical opinion.

One interesting avenue for future research would be to study the effect of price incentives from the point of view of empirical industrial organization. In France there has been a recent trend toward merging or closing small maternity wards, leading to larger institutions belonging to different financial groups. The way in which these structural changes interacted with changes in price incentives remains an unanswered question. In addition, machine-learning prediction algorithms, such as neural networks, could be successfully applied in modeling the decision tree and tariff incentives, making it possible to test and compare the efficiency of various algorithms.

ACKNOWLEDGMENTS

We express our gratitude to Zeynep Or (IRDES) for her guidance throughout the project and for granting access to PMSI dataset for the duration of the study. We would also like to thank Françoise Nguen, Maripier Isabelle, Anne Penneau, H el ene Huber, Audrey Laporte, Andrew Clark, Gr egoire Mercier, Henri-Jean Philippe, Mark Stabile, Mark Dusheiko, Randall Ellis, Agn es Gramain, Marc-Arthur Diaye and Hospinnomics team for their invaluable comments and contributions.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Program de Médicalisation des Systèmes d'Information (PMSI) data can be accessed by all interested parties upon obtaining an approval from Commission Nationale de l'Informatique et des Libertés (CNIL) and authorized data holders. Other sources of data (INSEE, DREES and ATIH) are open-access and available online.

ORCID

Alex Proshin  <https://orcid.org/0000-0001-6517-6822>

Alexandre Cazenave-Lacrouz  <https://orcid.org/0000-0003-3488-6092>

ENDNOTES

- ¹ These classifications, which map a set of patient characteristics onto DRGs in a unique way, are established by a state-run agency, Agence Technique de l'Information sur l'Hospitalisation (ATIH), on an annual basis. For example, in DRG version 11d, the root 14Z14 stands for a vaginal single delivery by a multipara mother. In addition, it is followed by a letter which specifies the severity level of the diagnosis, for example, 14Z14 A, where A specifies that there was no significant complication.
- ² In medical terms, a multipara (or pluripara) mother designates “a woman who has given birth at least twice to an infant, live-born or not, weighing 500 g or more, or having an estimated length of gestation of at least 20 weeks” (Farlex Partner Medical Dictionary, 2012).
- ³ A full description of the datasets is given in Appendix A.
- ⁴ See the ATIH data on tariffs on the website <http://www.atih.sante.fr/tarifs-mco-et-had>.
- ⁵ For a detailed description of the index construction, interpretation and applicability see: <http://drees.solidarites-sante.gouv.fr/etudes-et-statistiques/open-data/professions-de-sante-et-du-social/l-indicateur-d-accessibilite-potentielle-localise-apl/article/l-indicateur-d-accessibilite-potentielle-localisee-apl#nb2-1>.
- ⁶ See Appendix C1 for specifications with alternative price incentive measures.
- ⁷ In particular, the authors find that increasing “remuneration for a C-section by 100 percent relative to the baseline fee of a vaginal delivery might increase the probability that a physician opts for C-section by 0.6%–1.1% points.” Assuming a French average inflation-adjusted vaginal delivery tariff of 2054 euros, French estimates would correspond to 0.9%–1.9% change in the scheduled C-section rate for an average delivery.
- ⁸ See *Stata, Nested logit regression* for additional reference.

REFERENCES

- Allin, S., Baker, M., Isabelle, M., & Stabile, M. (2015). *Physician incentives and the rise in C-sections: Evidence from Canada*. National Bureau of Economic Research.
- Busse, R., Geissler, A., & Quentin, W. (Eds.). (2011). *Diagnosis-related groups in Europe: Moving towards transparency, efficiency and quality in hospitals*. Wiley M. Open University Press.
- Christianson, J., Leatherman, S., & Sutherland, K. (2008). Lessons from evaluations of purchaser pay-for-performance programs a review of the evidence. *Medical Care Research and Review*, 65(6_suppl), 5S–35S. <https://doi.org/10.1177/1077558708324236>
- Currie, J. W., & MacLeod, B. (2008). First do no harm? Tort reform and birth outcomes. *Quarterly Journal of Economics*, 123(2), 795–830. <https://doi.org/10.1162/qjec.2008.123.2.795>
- Dafny, L. (2005). How do hospitals respond to price changes? *The American Economic Review*, 95(5), 1525–1547. <https://doi.org/10.1257/000282805775014236>
- Di Giacomo, M., Piacenza, M., Siciliani, L., & Turati, G. (2017). Do public hospitals respond to changes in DRG price regulation? The case of birth deliveries in the Italian NHS. *Health Economics*, 26(S2), 23–37. <https://doi.org/10.1002/hec.3541>
- Dubay, L., Kaestner, R., & Waidmann, T. (1999). The impact of malpractice fears on cesarean section rates. *Journal of Health Economics*, 18(4), 491–522.
- Ellis, R. (1998). Creaming, skimping and dumping: Provider competition on the intensive and extensive margins. *Journal of Health Economics*, 17(5), 537–555. [https://doi.org/10.1016/s0167-6296\(97\)00042-8](https://doi.org/10.1016/s0167-6296(97)00042-8)
- Farlex Partner Medical Dictionary. (2012). Retrieved February 8, 2023, from <https://medical-dictionary.thefreedictionary.com/multipara>
- Galizzi, M., Tammi, T., Godager, G., Linnosmaa, I., & Wiesen, D. (2015). Provider altruism in health economics. National Institute for Health and Welfare (THL). Discussion paper 4/2015. Helsinki, Finland.
- Gilman, B. (2000). Hospital response to DRG refinements: The impact of multiple reimbursement incentives on inpatient length of stay. *Health Economics, Policy and Law*, 1(1), 79–90. [https://doi.org/10.1002/1099-1050\(200006\)9:4<277::aid-hec513>3.0.co;2-1](https://doi.org/10.1002/1099-1050(200006)9:4<277::aid-hec513>3.0.co;2-1)
- Grant, D. (2009). Physician financial incentives and cesarean delivery: New conclusions from the healthcare cost and utilization project. *Journal of Health Economics*, 28(1), 244–250. <https://doi.org/10.1016/j.jhealeco.2008.09.005>
- Gruber, J., & Owings, M. (1996). Physician financial incentives and cesarean section delivery. *The RAND Journal of Economics*, 27(1), 99–123. <https://doi.org/10.2307/2555794>
- Gruber, J., Kim, J., & Mayzlin, D. (1999). Physician fees and procedure intensity: The case of cesarean delivery. *Journal of Health Economics*, 18(4), 473–490. [https://doi.org/10.1016/s0167-6296\(99\)00009-0](https://doi.org/10.1016/s0167-6296(99)00009-0)

- Haute Autorité de Santé. (2012). Indications de la césarienne programmée à terme. https://www.has-sante.fr/upload/docs/application/pdf/2012-03/indications_cesarienne_programmee_-_recommandation_2012-03-12_14-44-28_679.pdf
- Haute Autorité de Santé. (2014). Césarienne programmée à terme: Déploiement de la démarche. https://www.has-sante.fr/jcms/c_1726029/fr/cesarienne-programmee-a-terme-deploiement-de-la-demarche
- Janulevicuite, J., Askildsen, J., Kaarboe, O., Siciliani, L., & Sutton, M. (2016). How do hospitals respond to price changes? Evidence from Norway. *Health Economics*, 25(5), 620–636. <https://doi.org/10.1002/hec.3179>
- Johnson, E., & Rehavi, M. (2016). Physicians treating physicians: Information and incentives in childbirth. *American Economic Journal: Economic Policy*, 8(1), 115–141. <https://doi.org/10.1257/pol.20140160>
- Lefèvre, M. (2014). Physician induced demand for C-sections: Does the convenience incentive matter? Health, Econometrics and Data Group (HEDG) Working Papers 14/08, HEDG, c/o Department of Economics. University of York.
- Lo, J. (2008). Financial incentives do not always work – An example of cesarean sections in Taiwan. *Health Policy*, 88(1), 121–129. <https://doi.org/10.1016/j.healthpol.2008.02.013>
- Milcent, C. (2021). From downcoding to upcoding: DRG based payment in hospitals. *International Journal of Health Economics and Management*, 21, 1–26. <https://doi.org/10.1007/s10754-020-09287-x>
- Milcent, C., & Rochut, J. (2009). Tarification hospitalière et pratique médicale La pratique de la césarienne en France. *Revue Économique*, 60(2), 489–506. <https://doi.org/10.3917/reco.602.0489>
- OECD. (2019). Health at a glance 2019: OECD indicators. Caesarean sections. <https://www.oecd-ilibrary.org/sites/fa1f7281-en/index.html?itemId=/content/component/fa1f7281-en#>
- Shurtz, I. (2013). The impact of medical errors on physician behavior: Evidence from malpractice litigation. *Journal of Health Economics*, 32(2), 331–340. <https://doi.org/10.1016/j.jhealeco.2012.11.011>
- Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M., & Sermeus, W. (2010). Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research*, 10(1), 247. <https://doi.org/10.1186/1472-6963-10-247>
- Vanlerenberghe, J.-M. (2015). Rapport d'information fait au nom de la commission des affaires sociales sur l'enquête de la Cour des comptes relative aux maternités. No. 243. Sénat.
- WHO statement on caesarean section rates. (2015). https://www.who.int/reproductivehealth/publications/maternal_perinatal_health/cs-statement/en/
- Yang, Y., Mello, M., Subramanian, S., & Studdert, D. (2009). Relationship between malpractice litigation pressure and rates of cesarean section and vaginal birth after cesarean section. *Medical Care*, 47(2), 234–242. <https://doi.org/10.1097/mlr.0b013e31818475de>

How to cite this article: Proshin, A., Cazenave-Lacroutz, A., & Rochaix, L. (2023). Impact of tariff refinement on the choice between scheduled C-section and normal delivery: Evidence from France. *Health Economics*, 32(7), 1397–1433. <https://doi.org/10.1002/hec.4672>

APPENDIX A

The “Acts” dataset provides information on all medical procedures that were performed on a patient during their hospital stay in a given year. In addition, it includes characteristics that provide details about specific circumstances in which a medical act was performed (e.g., late in the night, on Sunday or on a holiday). These specific characteristics lead to a modification of the price associated with a procedure. For example, if a procedure was performed by an obstetrician between 0 and 6 a.m., this would lead to an increase in prices by 40 euros. Reflecting the two main delivery options: vaginal labor and C-section, medical procedures used to denote the fact of giving birth consist of two groups and are coded with prefixes “JQGD” and “JQGA” respectively. However, it is impossible to identify a given patient in this dataset since it only characterizes a hospital stay.

In CMD each diagnosis is associated with a medical procedure. The former are provided in the “RSA” dataset. In “RSA” each hospital stay has a unique patient identifier which allows us to track the cases of hospitalization across different years for a given patient. In addition, it allows us to find a corresponding medical activities contained in the previous dataset. It is worthwhile to note that since in France each version of DRG comes into effect on the 1st of March of each year, “RSA” dataset classifies diagnoses based on two versions of DRG, that were operational before and after the 1st of March. Apart from diagnoses, “RSA” dataset contains a rich set of individual level characteristics (including age, region and community of residence, length of hospital stay, month of exit, etc.) and hospital characteristics, which include hospital unique identifiers, ownership information (i.e., public or private), legal status (i.e., profit or non-profit), geographic location, etc.

The final “DIAG” dataset lists complications (or secondary diagnoses) that were observed during each hospital stay. These complications describe medical conditions that arise both before and after hospital admission. The data are presented in a way similar to “Actes” dataset. In DRG, these complications have different severity levels, the number of which may vary from 1 to 3 depending on year and diagnosis roots. It is worth noting that with the introduction of DRG version 11d in 2012, not only diagnosis roots underwent a major revision, but also severity levels.

APPENDIX B

TABLE B1 Correlation matrix of main variables.

	Sch. C-section	$r^{post} - r^{pre}$	$\Delta tariff$	Age	Num. of comorb.	Gest. age	Multiple pregn.	Num. of previous C-sect.	Sunday or holiday	Multipara	Average pseudo postal code income	Gyn. availability
Scheduled C-section	1											
$r^{post} - r^{pre}$	0.001	1										
$\Delta tariff$	-0.001	0.958	1									
Age	0.112	0.071	0.041	1								
Num. of comorb.	0.006	-0.164	-0.187	0.013	1							
Gest. age	-0.09	0.006	0.072	-0.016	-0.126	1						
Multiple pregn.	0.06	-0.067	-0.096	0.036	0.102	-0.27	1					
Num. of previous C-sect.	0.389	0.068	0.038	0.104	0.055	-0.057	-0.003	1				
Week-end	-0.07	0.009	0.01	-0.013	-0.01	0.014	-0.001	-0.025	1			
Multipara	-0.02	0.306	0.207	0.268	-0.098	0.024	-0.023	0.24	0.006	1		
Average pseudo postal code income	0.012	-0.01	-0.006	0.14	> -0.001	0.013	0.008	0.001	0.005	-0.033	1	
Gyn. availability	-0.0016	-0.018	-0.016	0.062	0.038	-0.009	0.003	-0.009	0.001	-0.05	0.171	1
Average pseudo postal code income												
Average pseudo postal code income	1											
Gyn. availability	0.171											1
Gyn. availability												

TABLE B2 Obstetrical DRGs before and after 2012 reform.

DRG before reform (versions 11b, 11c)			DRG after reform (versions 11d, 11e)		
DRG code	Num. of severity levels	Name of procedure	DRG code	Num. of severity levels	Name of procedure
14C02	3	C-section	14C06	4	C-section, with child dead
			14C07	4	C-section, multiple pregnancy
			14C08	4	C-section, single pregnancy
14C03	1	Normal delivery, with other interventions	14C03	4	Normal delivery, with other interventions
14Z02	3	Normal delivery	14Z10	2	Normal delivery, with child dead
			14Z11	2	Normal delivery, unipara mother with multiple pregnancy
			14Z12	2	Normal delivery, multipara mother with multiple pregnancy
			14Z13	4	Normal delivery, unipara mother with single pregnancy
			14Z14	4	Normal delivery, multipara mother with single pregnancy

TABLE B3 Effect of gestational age on severity level of a normal delivery with single pregnancy, after 2012 DRG reform.

Severity level	Gestational age				
	22–31	32	33–35	36	37–44
D	D	D	D	D	D
C	C	C	D	C	C
B	B	C	D	C	B
A	A	B	C	B	A

TABLE B4 Effect of gestational age on severity level of a C-section, after 2012 DRG reform.

Severity level	Gestational age				
	22–31	32	33–35	36	37–44
D	D	D	D	D	D
C	D	D	D	D	C
B	C	C	C	C	B
A	B	B	C	B	A

TABLE B5 Effect of gestational age on severity level of a normal delivery with multiple pregnancy, after 2012 DRG reform.

Severity level	Gestational age				
	22–31	32	33–35	36	37–44
B	B	B	B	B	B
A	A	A	B	A	A

APPENDIX C: ADDITIONAL DID MODELS

C1 | Additional DiD models with deterministic severity

A larger computational tractability of deterministic models allows us to test additional formulations of DRG incentives and makes it possible to apply them to a larger sample of patients, which includes both unipara and multipara mothers.

In this Appendix, the deterministic measure of tariff incentives of $\Delta tariff$ is computed similar to the method described in Section 5.1 of the main text. It is complemented by its three close analogs.

First, $\Delta tariff$ is divided by the baseline level of tariff differential Δ^{pre} , giving a measure of change in tariff incentives relative to the pre-reform period $(\Delta^{post} - \Delta^{pre})/\Delta^{pre}$.

Second, we construct a measure that relies on the ratio of DRG tariffs between scheduled C-sections and normal deliveries, instead of their difference. Using the notation from point 4(b) in Section 4, the ratio between scheduled C-section and normal delivery tariffs in pre- and post-reform periods is denoted as $r^{pre} = \mathbb{E}(tariff_{CSsch}^0)/\mathbb{E}(tariff_{ND}^0)$ and $r^{post} = \mathbb{E}(tariff_{CSsch}^1)/\mathbb{E}(tariff_{ND}^1)$, respectively. The alternative tariff incentive measure is defined as a difference of tariff ratios between post- and pre-reform periods, giving $r^{post} - r^{pre}$.

Third, similar to measures using tariff differences, we calculate the ratio-based incentive measure relative to the pre-reform level. Thus, dividing the above measure by r^{pre} gives the forth tariff incentive measure $(r^{post} - r^{pre})/r^{pre}$.

By definition, in 2010 all these price incentive measures have the value 0 because of the absence of tariff change.

The results for different measures of DRG refinement incentives are presented in Table C1. The tested models include a full set of covariates, including multiparity and multiple pregnancy indicators. The specifications also control for year, hospital, and treatment group fixed effects. As earlier, model errors are clustered at the hospital level. Estimation results are presented in Table C1.

Overall, the obtained estimated are in line with those presented in Section 5. The main coefficient of interest β_2 remains statistically significant only at 10% and 5% significance level in regressions run for public non-profit and private non-profit facilities, respectively. However, when measuring DRG refinement incentives using different approaches, even the minimum conventional 10% significance level of coefficients β_2 is no longer achieved, suggesting that DRG incentives did not have a statistically detectable effect on the scheduled C-section probability.

Similar to Section 5 with main DiD models, the magnitude of $\Delta tariff$ coefficients is not sufficient to have a meaningful impact on the outcome of interest. Of note, $\Delta tariff$ estimates seem to have even lower values than those presented in Subsection 5.1, which value may be caused by attenuation bias due to measurement error.

TABLE C1 DiD coefficients for deterministically computed incentive measures, by hospital type.

Measures of financial incentive	Hospital type		
	Public	Private non-profit	Private for-profit
$\Delta tariff$	-0.0008* (0.0005) (0.12)	-0.001** (0.0005) (0.03)	0.0046 (0.0038) (0.23)
$(\Delta^{post} - \Delta^{pre})/\Delta^{pre}$	-0.0003 (0.0004) (0.48)	-0.0006 (0.0004) (0.22)	0.0016 (0.0013) (0.23)
$r^{post} - r^{pre}$	-0.0012 (0.001) (0.66)	-0.002 (0.0014) (0.16)	0.004 (0.004) (0.26)
$(r^{post} - r^{pre})/r^{pre}$	-0.0009 (0.002) (0.62)	-0.0024 (0.0025) (0.23)	0.0064 (0.0056) (0.21)

Note: Coefficient std. errors and p -values are given in parentheses under the coefficient. DiD, linear probability model with year, treatment group, and hospital fixed effects.

***1% sign., **5% sign., *10% sign.

C2 | DRG effects, by hospital size and hospital status

Inasmuch as the 2012 DRG reform introduced a more refined risk adjustment mechanism for child deliveries, it is plausible that hospitals that had a higher share of more severe patients in their case-mix were also more likely to benefit from the reform since the refinement resulted, in particular, in higher reimbursements for this type of patients.

TABLE C2 DiD estimation of the impact of DRG refinement incentives ($\Delta tariff$, rescaled to thousands of euro) on scheduled C-section probability in single birth multipara women in 2010–2013, by maternity ward type.

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
$\Delta tariff \times POST$	0.0024 (0.0071)	-0.0072 (0.0048)	0.0083 (0.0037)**	0.0538 (0.0989)	—	—	-0.0187 (0.0174)	0.0061 (0.0147)	—
Gynecologist access index	<0.0001 (0.0004)	-0.0003 (0.0001)**	-0.0007 (0.0001)***	0.0001 (0.0003)	—	—	-0.0007 (0.0002)***	-0.0006 (0.0001)***	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	152,704	382,591	263,008	7053	—	—	180,612	197,941	—
<i>R</i> ²	0.32	0.29	0.27	0.35	—	—	0.4	0.38	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at hospital level. Maternity wards are of three types: Type 1 (for low-risk pregnancies), Types 2a/2b (for moderately complicated cases and most pre-term pregnancies), and Type 3 (well-equipped to treat high-risk patients with severe complications and co-morbidities).

***1% sign., **5% sign., *10% sign.

The French health authorities classify French maternity wards into three categories. Type 1 maternity wards tend to be smaller and treat patients who present with few co-morbidities and who are not expected to require specialized medical equipment during childbirth. Types 2a and 2b (for simplicity, called Type 2 hereafter) are deemed by the French authorities to be sufficiently medically equipped to manage more complicated childbirth cases, such as premature and low-weight births, and may have at their disposal an intensive care unit. Type 3 maternity wards, which typically are large university hospitals, are meant to provide obstetric and neonatal care in the most medically complicated childbirth cases. Thus, the share of mothers with severe co-morbidities is the lowest in Type 1 and the highest in Type 3 facilities.

While the DiD analysis may shed light on the differences in the DRG reform impact across maternity ward types, there are a number of practical difficulties that make this analysis challenging. In particular, PMSI has two identification codes for care facilities—a legal entity code (“*finess juridique*”) and a geographic entity code (“*finess géographique*”). While one legal entity code typically corresponds to several geographic entity codes, the maternity ward type is determined by the geographic entity code. However, due to PMSI data limitations, until 2012 only legal entity codes were provided, which makes it impossible to assign the maternity ward type to legal entities with multiple geographic entities. As a result, the merging of datasets results in a considerable and non-random loss of observations. In particular, the majority of Paris region public hospitals were dropped after the dataset merger. Similarly, most private non-profit maternity wards could not be included in the sample due to the impossibility to establish a correspondence between legal entity codes and maternity ward types.

The estimation results for models by hospital size and hospital status are provided in Tables C2–C4. Due to data missingness discussed above, 19.7%, 93.9%, and 1% of observations in public, private non-profit, and private for-profit hospitals, were dropped, respectively.

TABLE C3 DiD estimation of the impact of DRG refinement incentives ($\Delta tariff$, rescaled to thousands of euro) on scheduled C-section probability in single birth multipara women in 2010–2013, by maternity ward type.

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
$\Delta tariff \times year$ 2010 semester 1	0.0188 (0.0123)	-0.0026 (0.0048)	-0.001 (0.0042)	0.1339 (0.0148)**	—	—	0.0472 (0.018)	0.0214 (0.0204)	—
$\Delta tariff \times year$ 2010 semester 2	0.0101 (0.0122)	-0.0039 (0.006)	-0.0041 (0.0041)	0.1249 (0.0102)***	—	—	0.0586 (0.0219)***	0.0143 (0.0194)	—
$\Delta tariff \times year$ 2011 semester 1	0.0016 (0.0144)	-0.001 (0.0066)	-0.0196 (0.0074)**	0.1354 (0.0636)	—	—	0.0346 (0.0256)	0.0112 (0.0205)	—
$\Delta tariff \times year$ 2012 semester 1	0.0123 (0.0145)	-0.0035 (0.0072)	-0.0178 (0.0065)***	0.1466 (0.1242)	—	—	0.0066 (0.0291)	0.0025 (0.0232)	—
$\Delta tariff \times year$ 2012 semester 2	0.013 (0.0155)	-0.0126 (0.0073)*	-0.0147 (0.0063)**	0.167 (0.1448)	—	—	0.0402 (0.0284)	0.0133 (0.0209)	—
$\Delta tariff \times year$ 2013 semester 1	0.0054 (0.0139)	-0.0067 (0.0081)	-0.0077 (0.0062)	0.1553 (0.0656)	—	—	-0.0066 (0.0305)	-0.0077 (0.021)	—
$\Delta tariff \times year$ 2013 semester 2	0.0191 (0.0151)	-0.016 (0.0089)*	-0.0117 (0.0093)	0.2381 (0.0922)	—	—	0.023 (0.0296)	0.0039 (0.0259)	—
Gynecologist access index	<0.0001 (0.0004)	-0.0003 (0.0002)**	-0.0008 (0.0001)***	-0.0001 (0.0006)	—	—	-0.0007 (0.0001)***	-0.0002 (0.0002)	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	152,704	382,591	263,008	7053	—	—	180,612	197,941	—
<i>R</i> ²	0.32	0.28	0.27	0.35	—	—	0.4	0.38	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level. Maternity wards are of three types: Type 1 (for low-risk pregnancies), Types 2a/2b (for moderately complicated cases and most pre-term pregnancies), and Type 3 (well-equipped to treat high-risk patients with severe complications and co-morbidities).

***1% sign., **5% sign., *10% sign.

The estimates of interest are generally statistically insignificant in both Table C2 testing a single effect for the post-intervention period and in Table C4 presenting the results of the DRG reform impact on two treatment groups with a change in tariff incentives superior/inferior to 100/−100.

Table C3 estimates the impact of DRG incentives by semester and includes pre-reform periods. Estimation results for Type 1 and 2 public maternity wards and Type 2 private for-profit maternity wards suggest that the 2012 DRG tariff refinement had a minimal impact on the probability of recourse to the scheduled C-section. Although coefficients are sporadically significant in the subsamples of women delivering a child in Type 3 public facilities and Type 1 private for-profit maternity wards, the results overall point to the absence of meaningful long-lasting effects of the DRG refinement on the obstetric practice. In particular, the negative coefficients in Type 3 public maternity wards in the first and second semesters in 2012 subsequently lose their statistical significance in 2013.

TABLE C4 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta tariff > 100$ and $\Delta tariff < -100$) on scheduled C-section probability in single birth multipara women in 2010–2013, by maternity ward type.

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
$T_{high} \cdot POST$	0.0076 (0.0042)*	<0.0001 (0.0021)	-0.0023 (0.0022)	0.025 (0.0371)	—	—	-0.0031 (0.0145)	0.013 (0.0111)	—
$T_{low} \cdot POST$	-0.0008 (0.0101)	0.0019 (0.005)	0.0041 (0.0049)	-0.0287 (0.1294)	—	—	0.0058 (0.017)	0.0186 (0.0149)	—
Gynecologist access index	<0.0001 (0.0003)	-0.0003 (0.0001)**	-0.0008 (0.0001)***	0.0001 (0.0003)	—	—	-0.0007 (0.0001)***	-0.0006 (0.0001)***	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	152,704	382,591	263,008	7053	—	—	180,612	197,941	—
R^2	0.31	0.29	0.27	0.35	—	—	0.4	0.39	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level. Maternity wards are of three types: Type 1 (for low-risk pregnancies), Types 2a/2b (for moderately complicated cases and most pre-term pregnancies), and Type 3 (well-equipped to treat high-risk patients with severe complications and co-morbidities).

***1% sign., **5% sign., *10% sign.

APPENDIX D: REGRESSION MODELS AND SPECIFICATIONS

D1 | Nested logit and ordinal logit model specifications

Nested logit model specifications:

To estimate the probability of each child delivery mode, we use a nested logit model run on the sample of mothers who delivered their child via a scheduled C-section, urgent C-section, an unscheduled C-section, or a normal delivery in the pre-reform period (i.e., 2010–2011 fiscal years.) Due to the high computational intensity of this method, we rely on a random 3% sample (e.g., 95,000 observations) to produce estimates.

The nested logit tree structure illustrated in Figure 1 has three levels, each of which corresponds to decision nodes 1–3. In general, for nested logit models it holds that:

$$Pr(C_3 = k | C_1 = t, C_2 = j) = \frac{\exp(\eta_{ijk}/\tau_j)}{\sum_{l \in R_j} \exp(\eta_{ijl}/\tau_j)}$$

$$Pr(C_2 = j | C_1 = t) = \frac{\left\{ \sum_{k \in R_j} \exp(\eta_{ijk}/\tau_j) \right\}^{\tau_j/v_t}}{\sum_{l \in S_t} \left\{ \sum_{k \in R_l} \exp(\eta_{ilk}/\tau_l) \right\}^{\tau_l/v_t}}$$

$$Pr(C_1 = t) = \frac{\left[\sum_{j \in S_t} \left\{ \sum_{k \in R_j} \exp(\eta_{ijk}/\tau_j) \right\}^{\tau_j/v_t} \right]^{v_t}}{\sum_{l \in T} \left[\sum_{j \in S_l} \left\{ \sum_{k \in R_j} \exp(\eta_{ijk}/\tau_j) \right\}^{\tau_j/v_l} \right]^{v_l}}$$

where C_1, C_2, C_3 are outcome variables at each level, which in our analysis correspond to nodes 1–3; $T, S_t,$ and R_j are choice sets at level 1 (scheduled C-section vs. other modes), level 2 (urgent C-section vs. normal delivery or unscheduled C-section) and level 3 choices (normal delivery vs. unscheduled C-section), respectively. $\tau_j = \sqrt{1 - \rho_j}$, where ρ_j is a correlation coefficient between alternatives in nest j , is interpreted as a measure of dissimilarity between alternatives within nest j ,⁸ (Stata, 2022).

TABLE D1 DiD estimation of the impact of DRG refinement incentives on child delivery severity (in the post-reform DRG classification) in multipara women in 2010–2013, treatment group thresholds $\Delta tariff > 100$ and $\Delta tariff < -100$.

	Public			Private non-profit			Private for profit		
	Single	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple
	1	2	3	4	5	6	7	8	9
Scheduled C-sections									
$\Delta tariff \times year$	0.0456	0.0468	0.2655	0.1221	0.1271	-1.363	0.1587	0.1584	0.6485
2012	(0.0148)***	(0.014)***	(0.217)	(0.0355)***	(0.0359)***	(1.6641)	(0.059)***	(0.058)***	(0.2734)**
$\Delta tariff \times year$	0.0441	0.0629	0.0629	0.0488	0.0544	1.361	0.1286	0.0141	-1.9173
2013	(0.015)***	(0.0151)***	(0.2084)	(0.0432)	(0.0429)	(0.9675)	(0.0774)*	(0.0773)*	(0.717)***
Gynecologist access index	-0.0015	-0.0026	-0.0018	-0.0018	-0.0007	-0.0236	-0.0004	-0.0009	-0.0017
	(0.0005)***	(0.0006)***	(0.0013)	(0.0008)**	(0.0009)	(0.0092)**	(0.0003)	(0.0004)**	(0.0068)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	NO	NO	YES	NO	NO	YES	NO
<i>N</i>	59,439	59,439	2500	9106	9106	259	36,834	36,834	858
<i>R</i> ²	0.27	0.28	0.24	0.24	0.25	0.44	0.18	0.19	0.25
All C-sections									
$\Delta tariff \times year$	0.0386	0.0398	0.2038	0.076	0.0746	1.872	0.0858	0.0864	0.0931
2012	(0.0145)***	(0.0144)***	(0.1382)	(0.021)***	(0.0299)**	(0.733)**	(0.0465)*	(0.0459)*	(0.2894)
$\Delta tariff \times year$	0.0428	0.0434	0.0267	0.0207	0.0241	0.3154	0.0923	0.0911	-0.6776
2013	(0.0144)***	(0.0144)***	(0.1279)	(0.0327)	(0.0326)	(0.7757)	(0.045)**	(0.0447)**	(0.5203)
Gynecologist access index	-0.0018	-0.0037	-0.009	-0.0021	-0.0013	-0.0124	-0.0001	-0.0006	-0.0006
	(0.0007)**	(0.0008)***	(0.0035)***	(0.0007)***	(0.001)	(0.0072)*	(0.0005)	(0.0005)	(0.0051)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	122,854	122,854	6232	16,276	16,276	506	60,363	60,363	1690
<i>R</i> ²	0.32	0.33	0.23	0.22	0.23	0.23	0.21	0.21	0.15
Normal deliveries									
$\Delta tariff \times year$	0.0287	0.0319	-0.0829	0.0002	<0.0001	-0.104	-0.0561	-0.0504	-0.2648
2012	(0.0373)	(0.0386)	(0.0718)	(0.0256)	(0.0255)	(0.541)	(0.0627)	(0.0622)	(0.2697)
$\Delta tariff \times year$	-0.0056	<0.0001	0.0195	<0.0001	<0.0001	-0.1414	-0.0001	-0.0981	0.0772
2013	(0.0157)	(<0.0001)	(0.0645)	(<0.0001)	(0.0375)	(0.3914)	(<0.0001)	(0.0522)*	(0.0406)*
Gynecologist access index	-0.0001	-0.0008	-0.0008	-0.0003	-0.0001	0.0006	-0.0001	-0.0001	-0.0023
	(0.0002)	(0.0002)***	(0.0013)	(0.0003)	(0.0002)	(0.0025)	(0.0001)	(0.0001)	(0.0024)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO

(Continues)

TABLE D1 (Continued)

	Public			Private non-profit			Private for profit		
	Single	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple
	1	2	3	4	5	6	7	8	9
<i>N</i>	832,488	832,488	9124	100,902	100,902	661	322,324	322,324	2002
<i>R</i> ²	0.33	0.33	0.09	0.26	0.27	0.15	0.26	0.27	0.06

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

In a general three-level nested logit model, the utility function is expressed as $U_{ijk} = \eta_{ijk} + \epsilon_{ijk}$. In this case, $\eta_{ijk} = z_i\alpha_i + u_{ij}\gamma_j + x_{ijk}\beta_k$, where z_i , u_{ij} , x_{ijk} are sets of covariates that enter the model in first, second, and third levels of choice.

In our models, we include an identical set of regressors z_i across all three levels. Vector z_i includes: age of mother and its squared and log-values, expected gestational age (obtained as OLS-derived predicted values in the linear regression model with the same set of regressors), multiple birth dummy, ante-partum diagnoses, a dummy for private non-profit facilities, a dummy for private for-profit facilities, multipara birth dummy, with fixed effects for the year of child delivery, the month of child delivery.

Ordinal logit model specifications:

In order to predict the probability of a given severity level in each child delivery mode (i.e., scheduled C-section, urgent C-section, an unscheduled C-section, or a normal delivery) in step 3 of the estimation algorithm presented in Section 4, we estimate individual-level ordinal logit models that rely on mothers' clinically relevant characteristics that are expected to affect the occurrence of birth complications. In general, ordinal logit models are represented through a latent function of the type:

$$y_i^* = \alpha + x_i'\beta + u_i$$

where y_i^* is the underlying latent variable; α is a constant term, vectors x_i and β are a set of regressors and associated coefficients, respectively; u_i is a random term.

The outcome value is determined by the interval in which falls y_i^* for a given observation. In particular, for the outcome value j , $y_i = j$ if $y \in (\alpha_{j-1}; \alpha_j]$.

Ordinal logit models are estimated based on the complete sample consisting of women who delivered their child via a scheduled C-section, urgent C-section, an unscheduled C-section, or a normal delivery in the pre-reform period. Due to differences in the number of severity levels and mother characteristics used to define a severity level in each delivery mode, models and resulting outcome probabilities are estimated separately for single and multiple birth multipara mothers. The set of regressors x_i used for deriving β includes: age of mother and its squared and log-values, expected gestational age (obtained as OLS-derived predicted values in the linear regression model with the same set of regressors), ante-partum diagnoses, a dummy for private non-profit facilities, a dummy for private for-profit facilities, with fixed effects for the year of child delivery, the month of child delivery.

D2 | Additional models for upcoding and different DRG incentive thresholds

Tables D1–D5 present additional model specifications testing the upcoding channel discussed in Subsection 6.2. In Table D1, the coefficients of interest reflect the effect of DRG tariff incentives on the severity of scheduled C-sections, any C-sections, or normal deliveries in the post-reform years of 2012 and 2013. For scheduled C-sections and all C-sections, the estimates of interest in most instances are positive and statistically significant. However, as discussed later, the internal validity of DiD models in these sub-samples is likely to be compromised by pre-existing non-parallel trends in the treatment and control groups. As for mothers delivering via a normal delivery, DiD estimates, on the contrary, tend to be statistically insignificant and of the opposite—negative—sign.

Tables D2–D4 show estimation results by semester. The results indicate that models for single birth child deliveries—that is, models (1)–(2), (4)–(5), and (7)–(8)—fail placebo tests for the pre-reform effects in all of the three subsamples. As for multiple birth child deliveries, due to their much smaller sample size relative to single birth deliveries, the estimates of β are more volatile and, on average, less significant. In particular, in models (3) and (9) in Table D2 and models (3) and (6) in Tables D3 and D4, both pre-reform and post-reform coefficients of interest are generally statistically insignificant. In the few statistically significant cases, coefficient signs and significance seem sporadic.

TABLE D2 DiD estimation of the impact of DRG refinement incentives on scheduled C-section severity (in the post-reform DRG classification) in multipara women in 2010–2013, by pre- and post-reform semesters, treatment group thresholds $\Delta tariff > 100$ and $\Delta tariff < -100$.

	Public			Private non-profit			Private for profit		
	Single	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple
	1	2	3	4	5	6	7	8	9
Scheduled C-sections									
$\Delta tariff \times$ year	0.0632	0.0625	0.1803	0.1452	0.1436	-1.5947	0.283	0.2824	0.2451
2010	(0.0221)***	(0.0221)***	(0.3113)	(0.0668)**	(0.0661)**	(2.1027)	(0.0522)***	(0.0521)***	(0.9781)
semester 1									
$\Delta tariff \times$ year	0.0893	0.0883	0.1567	0.0845	0.0819	0.507	0.3017	0.3033	0.2607
2010	(0.0169)***	(0.0169)***	(0.251)	(0.0728)	(0.0721)	(2.0163)	(0.0409)***	(0.0397)***	(0.8412)
semester 2									
$\Delta tariff \times$ year	0.0873	0.0887	0.2949	0.1659	0.1668	-2.3298	0.2058	0.2042	0.6212
2011	(0.0197)***	(0.0196)***	(0.1974)	(0.0903)*	(0.0901)*	(1.0597)**	(0.0583)***	(0.0594)***	(1.1548)
semester 1									
$\Delta tariff \times$ year	0.086	0.0876	0.3238	0.0956	0.0967	-3.3043	0.3296	0.3299	-1.263
2012	(0.0217)***	(0.0219)***	(0.3033)	(0.0573)	(0.0566)*	(2.1338)	(0.0781)***	(0.0775)***	(1.1955)
semester 1									
$\Delta tariff \times$ year	0.091	0.0916	0.2751	0.1161	0.1217	-0.5204	0.2465	0.245	-2.2277
2012	(0.0196)***	(0.0196)***	(0.2429)	(0.07)	(0.0701)*	(2.1467)	(0.0733)***	(0.0736)***	(0.9428)**
semester 2									
$\Delta tariff \times$ year	0.0974	0.0971	0.6728	0.1207	0.1178	-3.3762	0.4183	0.4177	0.5552
2013	(0.0263)***	(0.0263)***	(0.3114)**	(0.0702)*	(0.0701)	(2.7417)	(0.0551)***	(0.0546)***	(1.0158)
semester 1									
$\Delta tariff \times$ year	0.0939	0.0959	0.6213	0.1223	0.1226	-0.9493	0.3503	0.347	0.0071
2013	(0.0231)***	(0.0232)***	(0.2527)	(0.078)	(0.0777)	(2.4704)	(0.0687)***	(0.0688)***	(1.0027)
semester 2									
Gynecologist	-0.0015	-0.0026	-0.0018	-0.0018	-0.0004	-0.0236	-0.0004	-0.0009	-0.0017
access index	(0.0005)***	(0.0006)***	(0.0043)	(0.0009)**	(0.0009)	(0.0009)**	(0.0003)	(0.0004)**	(0.0068)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum	YES	YES	YES	YES	YES	YES	YES	YES	YES
diagnosis FE									
Hospital FE	NO	YES	NO	NO	YES	NO	NO	YES	NO
<i>N</i>	59,439	59,439	2500	9106	9106	259	36,834	36,834	858
<i>R</i> ²	0.28	0.29	0.25	0.24	0.25	0.46	0.19	0.2	0.26

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

TABLE D 3 DID estimation of the impact of DRG refinement incentives on C-section (all C-section types) severity (in the post-reform DRG classification) in multipara women in 2010–2013, by pre- and post-reform semesters, treatment group thresholds $\Delta tariff > 100$ and $\Delta tariff < -100$.

	Public			Private non-profit			Private for profit		
	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple	
	1	2	3	4	5	6	7	8	9
All C-sections									
$\Delta tariff \times year$	0.0324	0.0333	0.2026	0.1557	0.154	-0.7708	0.294	0.2904	0.9199
2010 semester 1	(0.0152)**	(0.0151)**	(0.2171)	(0.0547)***	(0.0548)***	(0.9104)	(0.0574)***	(0.057)***	(0.5132)*
$\Delta tariff \times year$	0.0633	0.0629	0.2224	0.121	0.1228	-0.4759	0.3283	0.328	1.9669
2010 semester 2	(0.0157)***	(0.0154)***	(0.1572)	(0.0559)**	(0.055)**	(0.6745)	(0.0451)***	(0.0451)***	(0.541)***
$\Delta tariff \times year$	0.0652	0.0675	-0.0076	0.162	0.1611	0.7088	0.2161	0.2177	0.4282
2011 semester 1	(0.0158)***	(0.0157)***	(0.1139)	(0.0614)**	(0.0605)**	(0.5624)	(0.0481)***	(0.0476)***	(0.7659)
$\Delta tariff \times year$	0.0497	0.0509	-0.0012	0.1141	0.1152	-0.7461	0.2996	0.2963	-0.048
2012 semester 1	(0.0206)**	(0.0206)**	(0.1559)	(0.0415)***	(0.0405)***	(0.9752)	(0.0475)***	(0.0471)***	(0.6812)
$\Delta tariff \times year$	0.0857	0.0859	0.2067	0.0639	0.0698	-0.6319	0.2076	0.2064	-0.0178
2012 semester 2	(0.0158)***	(0.0158)***	(0.1226)*	(0.0582)	(0.0574)	(1.16)	(0.0562)***	(0.0561)***	(0.5117)
$\Delta tariff \times year$	0.0785	0.0784	0.1953	0.0876	0.0872	-2.3433	0.4202	0.4177	0.9668
2013 semester 1	(0.0159)***	(0.0159)***	(0.1787)	(0.0581)	(0.0701)	(0.7668)***	(0.0507)***	(0.0498)***	(0.2957)***
$\Delta tariff \times year$	0.0739	0.0741	0.2096	0.066	0.0641	-1.934	0.3699	0.3658	0.553
2013 semester 2	(0.0172)***	(0.0172)***	(0.1121)*	(0.0467)	(0.0777)	(1.0732)*	(0.0453)***	(0.0453)***	(0.3688)
Gynecologist access index	-0.0017	-0.0037	-0.0092	-0.0021	-0.0004	-0.0236	-0.0004	-0.0009	-0.0007
	(0.0007)**	(0.0008)***	(0.0035)***	(0.0007)***	(0.0009)	(0.009)**	(0.0003)	(0.0004)**	(0.0051)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
N	122,854	122,854	6232	16,276	16,276	506	60,363	60,363	1690
R ²	0.23	0.24	0.23	0.23	0.23	0.25	0.22	0.22	0.17

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

TABLE D4 DiD estimation of the impact of DRG refinement incentives on normal delivery severity (in the post-reform DRG classification) in multipara women in 2010–2013, by pre- and post-reform semesters, treatment group thresholds $\Delta tariff > 100$ and $\Delta tariff < -100$.

	Public			Private non-profit			Private for profit		
	Single	Single	Multiple	Single	Single	Multiple	Single	Single	Multiple
	1	2	3	4	5	6	7	8	9
Normal deliveries									
$\Delta tariff \times year$	0.1032	0.1051	-0.0637	0.1404	0.1449	0.0277	0.5133	0.5149	0.6384
2010	(0.0185)***	(0.0194)***	(0.0784)	(0.04)***	(0.04)***	(0.4205)	(0.0396)***	(0.0391)***	(0.4408)
semester 1									
$\Delta tariff \times year$	0.1183	0.1203	-0.0435	0.1172	0.1223	0.3488	0.5114	0.5148	0.2442
2010	(0.0187)***	(0.02)***	(0.0408)	(0.0438)**	(0.0436)***	(0.5526)	(0.0365)***	(0.0366)***	(0.2196)
semester 2									
$\Delta tariff \times year$	0.0743	0.0784	-0.0158	0.027	0.1611	-0.2815	0.0966	0.1062	0.4293
2011	(0.0368)**	(0.039)**	(0.0652)	(0.0334)	(0.0342)	(0.3177)	(0.0481)***	(0.0394)***	(0.4122)
semester 1									
$\Delta tariff \times year$	0.0522	0.0559	0.0188	0.0168	0.023	-0.3563	0.1966	0.2022	0.6902
2012	(0.0228)**	(0.0238)**	(0.0798)	(0.0457)	(0.0455)	(0.3402)	(0.0522)***	(0.0521)***	(0.32)**
semester 1									
$\Delta tariff \times year$	0.0568	0.0597	0.035	0.0744	0.0781	-0.2744	0.1814	0.1892	0.1206
2012	(0.0219)***	(0.0231)***	(0.0671)	(0.0473)	(0.0473)	(0.4324)	(0.055)***	(0.055)***	(0.0501)**
semester 2									
$\Delta tariff \times year$	0.0769	0.0796	-0.0271	0.0191	0.0247	-0.6717	0.2805	0.2857	0.4936
2013	(0.0242)***	(0.0251)***	(0.0482)	(0.0436)	(0.0435)	(0.4332)	(0.0458)***	(0.0455)***	(0.4221)
semester 1									
$\Delta tariff \times year$	0.053	0.0557	0.0588	-0.0028	0.0033	-0.2725	0.2411	0.2445	0.4908
2013	(0.0238)**	(0.0247)**	(0.0579)	(0.0513)	(0.0514)	(0.5272)	(0.0456)***	(0.045)***	(0.1781)***
semester 2									
Gynecologist access index	-0.0001	-0.0008	-0.007	-0.0003	-0.0004	-0.0236	-0.0001	-0.0009	-0.0021
	(0.0002)	(0.0002)***	(0.0013)***	(0.0003)	(0.0009)	(0.009)**	(0.00014)	(0.0004)**	(0.0024)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	832,488	832,488	9124	100,902	100,902	661	322,324	322,324	2002
<i>R</i> ²	0.34	0.34	0.09	0.27	0.27	0.15	0.27	0.27	0.07

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

Table D5 shows DiD estimation results for the sample of single births via scheduled C-sections in Type 1–3 hospitals, whose definition was introduced in Appendix C2. In line with earlier presented models, most specifications show signs of violation of the common trend assumption, as suggested by the presence of statistically significant pre-reform coefficients. In addition, models (1) and (4), in which this assumption might be more plausible, produce statistically insignificant estimates with magnitudes that can be considered as inconsequential from the health policy standpoint.

As expected, the coefficient for the gynecologist availability is of consistently negative sign. It is likely to reflect the fact that receiving pre-natal care reduces the risk of experiencing childbirth complications, regardless of the child delivery mode.

Tables D6–D9 present DiD estimates for the effects of the 2012 DRG refinement on the probability of scheduled C-section, where instead of the 100/–100 incentive threshold used in the main text, we test models with 300/–300 and 500/–500 thresholds. Similar to Table 9 presented in Subsection 6.1, models for single birth child deliveries in publicly funded non-profit hospitals show signs of non-parallel pre-reform trends when either of the two wider thresholds is applied (Tables D6 and D8). As for public non-profit facilities, although models (4) and (5) for single birth deliveries using the 300/–300 threshold yield 5% and 10% statistically significant coefficients for the first semester of 2010 and the first semester of 2011, respectively, these pre-reform effects disappear when the 500/–500 threshold is applied. In the meantime, regardless of the chosen threshold,

TABLE D5 DiD estimation of the impact of DRG refinement incentives ($\Delta tariff$, rescaled to thousands of euro, treatment group thresholds $\Delta tariff > 100$ and $\Delta tariff < -100$) on the encoded severity of scheduled C-sections in single birth multipara women in 2010–2013, by maternity ward type.

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
Scheduled C-sections									
$\Delta tariff \times year$ 2010 semester 1	0.0137 (0.0547)	0.0769 (0.0377)**	0.0647 (0.0399)	-0.2106 (0.5155)	—	—	0.1354 (0.0963)	0.3703 (0.057)***	—
$\Delta tariff \times year$ 2010 semester 2	0.0319 (0.0521)	0.0791 (0.0349)**	0.0987 (0.0318)***	-0.0127 (0.2829)	—	—	0.1572 (0.0721)**	0.3804 (0.0456)***	—
$\Delta tariff \times year$ 2011 semester 1	0.104 (0.051)**	0.0771 (0.0366)**	0.0931 (0.0325)***	-0.2874 (0.4335)	—	—	0.1409 (0.083)*	0.2445 (0.0716)***	—
$\Delta tariff \times year$ 2012 semester 1	-0.0045 (0.0656)	0.0929 (0.035)***	0.0576 (0.0417)	-0.1641 (0.2443)	—	—	0.0829 (0.1145)	0.4576 (0.076)***	—
$\Delta tariff \times year$ 2012 semester 2	0.0715 (0.0538)	0.1035 (0.0303)***	0.0538 (0.0378)	0.0652 (0.274)	—	—	0.0628 (0.1272)	0.3447 (0.0841)***	—
$\Delta tariff \times year$ 2013 semester 1	-0.0016 (0.0704)	0.0993 (0.0282)***	0.0452 (0.0666)	-0.1032 (0.1703)	—	—	0.3026 (0.066)***	0.5311 (0.0886)***	—
$\Delta tariff \times year$ 2013 semester 2	0.0891 (0.0623)	0.1102 (0.0327)***	0.0326 (0.0433)	-0.1917 (0.4587)	—	—	0.1939 (0.079)**	0.4709 (0.1047)***	—
Gynecologist access index	0.0001 (0.0011)	-0.0003 (0.0002)**	-0.0037 (0.001)***	-0.0002 (0.0018)	—	—	-0.0009 (0.0001)***	-0.0002 (0.0005)	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	NO	YES	YES	YES	YES	YES
<i>N</i>	10,341	23,443	15,6518	976	—	—	17,339	19,070	—
<i>R</i> ²	0.22	0.26	0.27	0.31	—	—	0.2	0.21	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at hospital level. Maternity wards are of three types: Type 1 (for low-risk pregnancies), Types 2a/2b (for moderately complicated cases and most pre-term pregnancies), and Type 3 (well-equipped to treat high-risk patients with severe complications and co-morbidities).

***1% sign., **5% sign., *10% sign.

the estimated post-reform DRG effects in both single and multiple birth samples have no signs of being statistically different from zero. Similarly, estimates for private-for-profit maternity wards consistently show no effect of the 2012 DRG reform on scheduled C-section probability. Tables D7 and D9 present DiD estimates separately for treatment groups with positive/negative changes in DRG tariffs above/below the specified thresholds. These models further corroborate the argument that the DRG refinement do not impact the rates of scheduled C-sections. Although models (1) and (2) for single births child deliveries in public non-profit hospitals yield statistically significant estimates at the 5% level, this effect is likely due to the lack of pre-reform parallel trend in treatment and control groups. Overall, estimates obtained from models in which the 500/–500 threshold is applied tend to produce the coefficients of interest with the least strong statistical significance.

Finally, Tables D10–D13 check the robustness of estimates across both of the aforementioned dimensions at once: the models feature both a wider DRG incentive threshold and are tested separately for Type 1–3 maternity wards. Their results go in line with the conclusions presented in this study. Tables D10 and D12 show estimates for pre- and post-reform semesters. Overall, the few sporadically and marginally significant coefficients provide no indication of any meaningful effect of the DRG refinement on scheduled C-section rates. Tables D11 and D13 showing model estimates separately for each treatment group point to the same conclusion. Finally, as earlier, the coefficients for the gynecologist availability index are generally negative and highly statistically significant, suggesting that a higher availability of gynecologists had a mitigating effect on scheduled C-section rates in France.

TABLE D6 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta tariff > 300$ and $\Delta tariff < -300$) on scheduled C-section probability in multipara women in 2010–2013.

	Public			Private non-profit			Private for profit		
	Single 1	Single 2	Multiple 3	Single 4	Single 5	Multiple 6	Single 7	Single 8	Multiple 9
$\Delta tariff \times year$ 2010 semester 1	0.0067 (0.0035)*	0.0069 (0.0034)**	-0.0631 (0.0438)	0.023 (0.0112)**	0.0207 (0.0236)	-0.2946 (0.2369)	0.0106 (0.0133)	0.0086 (0.0132)	0.0129 (0.2116)
$\Delta tariff \times year$ 2010 semester 2	0.0038 (0.0033)	0.0002 (0.0033)	-0.0294 (0.0389)	0.0328 (0.0102)***	0.0298 (0.0101)***	-0.2372 (0.2073)	0.0143 (0.0135)	0.0101 (0.0133)	0.0345 (0.1865)
$\Delta tariff \times year$ 2011 semester 1	-0.0117 (0.0042)***	-0.012 (0.0042)***	-0.0562 (0.0324)*	0.0012 (0.0012)	0.0093 (0.0127)	-0.1691 (0.2254)	0.0115 (0.0155)	0.0071 (0.016)	0.1628 (0.1793)
$\Delta tariff \times year$ 2012 semester 1	-0.0092 (0.0042)**	-0.0095 (0.0044)**	-0.0006 (0.0459)	-0.0042 (0.0128)	-0.006 (0.0129)	-0.4283 (0.2599)	-0.0125 (0.0189)	-0.0126 (0.0188)	0.0593 (0.1803)
$\Delta tariff \times year$ 2012 semester 2	0.0119 (0.0046)**	0.00119 (0.0046)**	0.012 (0.0491)	0.0319 (0.0015)**	0.0286 (0.0154)*	-0.3995 (0.2685)	0.011 (0.0173)	0.0061 (0.0173)	0.0875 (0.1032)
$\Delta tariff \times year$ 2013 semester 1	-0.0053 (0.0039)	-0.0058 (0.0039)	-0.0921 (0.0389)**	0.0116 (0.0124)	0.0093 (0.0124)	-0.2651 (0.2948)	0.022 (0.018)	-0.0248 (0.018)	0.059 (0.1579)
$\Delta tariff \times year$ 2013 semester 2	-0.0064 (0.0048)	-0.0069 (0.0048)	-0.02 (0.04)	0.023 (0.016)	0.021 (0.0157)	-0.1801 (0.3633)	0.0003 (0.0187)	-0.004 (0.0189)	-0.2078 (0.149)
Gynecologist access index	-0.0009 (0.0001)***	-0.0005 (0.0001)***	-0.0002 (0.0009)	-0.0018 (0.0008)**	-0.0007 (0.0009)	0.0214 (0.0021)	-0.0002 (0.0002)	-0.0009 (0.0004)**	-0.0028 (0.0018)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	955,342	955,342	15,356	117,178	117,178	1167	382,687	382,676	3692
<i>R</i> ²	0.28	0.28	0.17	0.32	0.33	0.37	0.39	0.39	0.38

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

TABLE D7 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta tariff > 300$ and $\Delta tariff < -300$) on scheduled C-section probability in multipara women in 2010–2013.

	Public			Private non-profit			Private for profit		
	Single 1	Single 2	Multiple 3	Single 4	Single 5	Multiple 6	Single 7	Single 8	Multiple 9
$T_{high} \cdot POST$	-0.0037 (0.0025)	0.003 (0.0025)	0.038 (0.0797)	-0.0013 (0.0105)	-0.003 (0.0105)	0.2992 (0.3067)	0.0082 (0.0118)	-0.0074 (0.0115)	-0.0356 (0.0711)
$T_{low} \cdot POST$	0.0057 (0.0026)**	0.0056 (0.0026)**	0.0201 (0.0263)	0.0097 (0.0093)	0.0094 (0.0094)	0.0964 (0.1524)	0.0107 (0.0079)	0.0109 (0.0079)	0.1014 (0.3408)
Gynecologist access index	-0.001 (0.0001)***	-0.0005 (0.0001)***	-0.0019 (0.0007)**	-0.0007 (0.0003)**	-0.0007 (0.0001)***	0.0025 (0.0022)	-0.0001 (0.0001)	-0.0006 (0.0001)***	-0.0028 (0.0018)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	955,342	955,342	15,356	117,178	117,178	1167	382,687	382,676	3692
<i>R</i> ²	0.28	0.28	0.17	0.33	0.33	0.23	0.39	0.39	0.21

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

TABLE D8 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta\text{tariff} > 500$ and $\Delta\text{tariff} < -500$) on scheduled C-section probability in multipara women in 2010–2013.

	Public			Private non-profit			Private for profit		
	Single 1	Single 2	Multiple 3	Single 4	Single 5	Multiple 6	Single 7	Single 8	Multiple 9
$\Delta\text{tariff} \times \text{year 2010}$ semester 1	0.0083 (0.0036)**	0.0085 (0.0036)**	-0.0478 (0.0442)	0.0097 (0.0109)	0.0069 (0.0107)	-0.2189 (0.279)	0.0073 (0.083)	0.0071 (0.0084)	0.0187 (0.2118)
$\Delta\text{tariff} \times \text{year 2010}$ semester 2	0.0019 (0.0032)	0.0018 (0.0031)	-0.0181 (0.0363)	0.0169 (0.0109)*	0.0298 (0.0109)	-0.2264 (0.2668)	0.0115 (0.096)	0.0089 (0.096)	0.0443 (0.1926)
$\Delta\text{tariff} \times \text{year 2011}$ semester 1	-0.0105 (0.0042)**	-0.0104 (0.0043)**	-0.0499 (0.0366)	0.0014 (0.0114)	-0.0014 (0.0114)	-0.0916 (0.2314)	0.0095 (0.0157)	0.006 (0.0163)	0.1702 (0.1837)
$\Delta\text{tariff} \times \text{year 2012}$ semester 1	-0.0076 (0.0041)*	-0.0084 (0.0044)*	-0.0399 (0.0381)	-0.0157 (0.0129)	-0.0177 (0.0132)	-0.2871 (0.2994)	-0.0146 (0.0164)	-0.0133 (0.0163)	0.0542 (0.2005)
$\Delta\text{tariff} \times \text{year 2012}$ semester 2	0.0105 (0.0043)**	0.0106 (0.0042)**	0.0266 (0.0041)	0.0214 (0.0146)	0.0177 (0.0147)	-0.35 (0.2647)	0.094 (0.0164)	0.0054 (0.0164)	0.0962 (0.1217)
$\Delta\text{tariff} \times \text{year 2013}$ semester 1	-0.0039 (0.0039)	-0.0044 (0.0039)	-0.0764 (0.0367)**	0.079 (0.0013)	-0.0026 (0.0013)	-0.1139 (0.3138)	-0.0238 (0.0238)	-0.0253 (0.0168)	0.0163 (0.1579)
$\Delta\text{tariff} \times \text{year 2013}$ semester 2	-0.0051 (0.0048)	-0.0056 (0.0048)	-0.0081 (0.0373)	0.0012 (0.0169)	0.0092 (0.0168)	-0.1305 (0.3842)	-0.0019 (0.0179)	-0.0046 (0.0182)	-0.1907 (0.0153)
Gynecologist access index	-0.001 (0.0001)***	-0.0005 (0.0001)***	-0.0002 (0.001)	-0.0006 (0.0003)**	-0.0007 (0.0002)***	0.0214 (0.0021)	-0.0002 (0.0002)	-0.0006 (0.0001)***	-0.0028 (0.0018)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	955,342	955,342	15,356	117,178	117,178	1167	382,687	382,676	3692
<i>R</i> ²	0.28	0.28	0.17	0.33	0.33	0.23	0.39	0.39	0.21

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

TABLE D9 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta\text{tariff} > 500$ and $\Delta\text{tariff} < -500$) on scheduled C-section probability in multipara women in 2010–2013.

	Public			Private non-profit			Private for profit		
	Single 1	Single 2	Multiple 3	Single 4	Single 5	Multiple 6	Single 7	Single 8	Multiple 9
$T_{high} \cdot POST$	-0.0036 (0.0066)	0.0034 (0.0066)	0.038 (0.0797)	-0.0086 (0.0246)	-0.0112 (0.0251)	0.2506 (0.231)	-0.0033 (0.0232)	-0.0005 (0.0228)	—
$T_{low} \cdot POST$	0.0059 (0.0028)**	0.0059 (0.0028)**	0.0201 (0.0118)	0.0075 (0.0098)	0.0065 (0.0097)	0.0564 (0.0604)	-0.0061 (0.0187)	0.005 (0.0187)	—
Gynecologist access index	-0.0009 (0.0002)***	-0.0005 (0.0001)***	-0.0002 (0.0009)	-0.0007 (0.0003)*	-0.0007 (0.0001)***	0.0021 (0.0023)	-0.0001 (0.0001)	-0.0006 (0.0001)***	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	NO	YES	YES	NO	YES	NO	NO	YES	NO
<i>N</i>	955,342	955,342	15,356	117,178	117,178	1167	382,687	382,676	—
<i>R</i> ²	0.28	0.28	0.17	0.33	0.33	0.23	0.39	0.39	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level.

***1% sign., **5% sign., *10% sign.

TABLE D10 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta tariff > 300$ and $\Delta tariff < -300$) on scheduled C-section probability in single birth multipara women in 2010–2013, by maternity ward type.

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
$\Delta tariff \times$ year 2010 semester 1	0.0236 (0.0122)	0.0036 (0.0049)	0.0033 (0.0049)	0.1765 (0.0564)*	—	—	0.0203 (0.0206)	-0.0014 (0.0171)	—
$\Delta tariff \times$ year 2010 semester 2	0.015 (0.0126)	-0.0027 (0.006)	-0.0043 (0.0051)	0.1547 (0.0615)*	—	—	0.0343 (0.0189)*	-0.0076 (0.0185)	—
$\Delta tariff \times$ year 2011 semester 1	0.0054 (0.0147)	-0.0096 (0.0068)	-0.0166 (0.0086)*	0.1676 (0.1093)	—	—	0.0215 (0.0245)	-0.0062 (0.021)	—
$\Delta tariff \times$ year 2012 semester 1	0.0169 (0.0145)	-0.0032 (0.0073)	-0.0149 (0.0073)**	0.181 (0.1299)	—	—	-0.0156 (0.0294)	-0.0116 (0.0251)	—
$\Delta tariff \times$ year 2012 semester 2	0.017 (0.0157)	-0.0126 (0.0073)*	-0.0125 (0.0071)*	0.1999 (0.1236)	—	—	0.022 (0.0283)	-0.0053 (0.0219)	—
$\Delta tariff \times$ year 2013 semester 1	0.0095 (0.0146)	-0.0065 (0.0089)	-0.0051 (0.0071)	0.2007 (0.1111)	—	—	-0.0276 (0.0299)	-0.0257 (0.0226)	—
$\Delta tariff \times$ year 2013 semester 2	0.023 (0.015)	-0.0156 (0.0088)*	-0.0089 (0.0094)	0.2761 (0.047)**	—	—	0.0054 (0.0282)	-0.0119 (0.0263)	—
Gynecologist access index	<0.0001 (0.0003)	-0.0003 (0.0002)**	-0.0008 (0.0001)***	-0.0001 (0.0006)	—	—	-0.0007 (0.0002)***	-0.0006 (0.0002)***	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	152,704	382,591	263,008	7053	—	—	180,612	197,941	—
<i>R</i> ²	0.32	0.28	0.27	0.35	—	—	0.4	0.39	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level. Maternity wards are of three types: Type 1 (for low-risk pregnancies), Types 2a/2b (for moderately complicated cases and most pre-term pregnancies), and Type 3 (well-equipped to treat high-risk patients with severe complications and co-morbidities).

***1% sign., **5% sign., *10% sign.

TABLE D11 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta tariff > 300$ and $\Delta tariff < -300$) on scheduled C-section probability in single birth multipara women in 2010–2013, by maternity ward type.

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
$T_{high} \cdot POST$	-0.0053 (0.0077)	-0.0091 (0.0041)**	0.0032 (0.0042)	0.0291 (0.0119)	—	—	-0.0192 (0.0207)	0.0002 (0.0142)	—
$T_{low} \cdot POST$	-0.0044 (0.0088)	0.0044 (0.0056)	0.0057 (0.0046)	-0.0277 (0.111)	—	—	0.0088 (0.0124)	0.0121 (0.0103)	—
Gynecologist access index	<0.0001 (0.0003)	-0.0004 (0.0002)**	-0.0007 (0.0002)***	0.0001 (0.0003)	—	—	<0.0001 (0.0003)	-0.0006 (0.0002)***	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES

(Continues)

TABLE D11 (Continued)

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	152,704	382,591	263,008	7053	—	—	180,612	197,941	—
<i>R</i> ²	0.32	0.29	0.27	0.35	—	—	0.4	0.39	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level. Maternity wards are of three types: Type 1 (for low-risk pregnancies), Types 2a/2b (for moderately complicated cases and most pre-term pregnancies), and Type 3 (well-equipped to treat high-risk patients with severe complications and co-morbidities).

***1% sign., **5% sign., *10% sign.

TABLE D12 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta tariff > 500$ and $\Delta tariff < -500$) on scheduled C-section probability in single birth multipara women in 2010–2013, by maternity ward type.

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
$\Delta tariff \times year$ 2010 semester 1	0.0171 (0.0108)	0.0047 (0.0052)	0.0058 (0.005)	0.1373 (0.039)*	—	—	0.015 (0.0119)	0.0013 (0.0116)	—
$\Delta tariff \times year$ 2010 semester 2	0.0085 (0.0117)	-0.0016 (0.0061)	0.0024 (0.0044)	0.1318 (0.0372)*	—	—	0.0296 (0.0012)**	-0.0075 (0.0141)	—
$\Delta tariff \times year$ 2011 semester 1	0.0039 (0.0139)	-0.0087 (0.0066)	-0.0132 (0.0081)	0.1481 (0.0874)	—	—	0.0191 (0.024)	-0.0065 (0.0221)	—
$\Delta tariff \times year$ 2012 semester 1	0.0107 (0.0135)	-0.0021 (0.0072)	-0.012 (0.0076)	0.1417 (0.1304)	—	—	-0.0186 (0.0266)	-0.011 (0.0212)	—
$\Delta tariff \times year$ 2012 semester 2	0.011 (0.0147)	-0.0112 (0.0073)	-0.0092 (0.0069)	0.1789 (0.1246)	—	—	0.0192 (0.0271)	-0.0048 (0.0208)	—
$\Delta tariff \times year$ 2013 semester 1	0.0031 (0.0141)	-0.0053 (0.008)	-0.0023 (0.0071)	0.1534 (0.0914)	—	—	-0.0293 (0.0264)	-0.025 (0.0225)	—
$\Delta tariff \times year$ 2013 semester 2	0.0169 (0.0135)	-0.0145 (0.0089)	-0.0063 (0.0095)	0.25 (0.0652)*	—	—	0.0033 (0.0249)	-0.0114 (0.0267)	—
Gynecologist access index	<0.0001 (0.0003)	-0.0003 (0.0002)**	-0.0007 (0.0001)***	0.0002 (0.0003)	—	—	-0.0007 (0.0002)***	-0.0006 (0.0002)***	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	152,704	382,591	263,008	7053	—	—	180,612	197,941	—
<i>R</i> ²	0.32	0.29	0.27	0.35	—	—	0.4	0.39	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level. Maternity wards are of three types: Type 1 (for low-risk pregnancies), Types 2a/2b (for moderately complicated cases and most pre-term pregnancies), and Type 3 (well-equipped to treat high-risk patients with severe complications and co-morbidities).

***1% sign., **5% sign., *10% sign.

TABLE D13 DiD estimation of the impact of DRG refinement incentives on treatment groups ($\Delta\text{tariff} > 500$ and $\Delta\text{tariff} < -500$) on scheduled C-section probability in single birth multipara women in 2010–2013, by maternity ward type.

	Public			Private non-profit			Private for profit		
	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3	Type 1	Types 2a/2b	Type 3
	1	2	3	4	5	6	7	8	9
$T_{high} \cdot POST$	-0.0089 (0.0248)	-0.0029 (0.0101)	0.0149 (0.0095)	0.0553 (0.0359)	—	—	-0.0003 (0.0497)	-0.0064 (0.0254)	—
$T_{low} \cdot POST$	-0.005 (0.0099)	0.0046 (0.006)	0.0052 (0.0047)	-0.0664 (0.1278)	—	—	0.0109 (0.0266)	-0.0101 (0.027)	—
Gynecologist access index	<0.0001 (0.0003)	-0.0003 (0.0001)**	-0.0008 (0.0001)***	0.0001 (0.0003)	—	—	-0.0007 (0.0002)***	-0.0006 (0.0002)***	—
Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Ante-partum diagnosis FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Hospital FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	152,704	382,591	263,008	7053	—	—	180,612	197,941	—
R^2	0.32	0.29	0.27	0.35	—	—	0.4	0.39	—

Note: Coefficient std. errors are given in parentheses under the coefficient. Error terms are clustered at the hospital level. Maternity wards are of three types: Type 1 (for low-risk pregnancies), Types 2a/2b (for moderately complicated cases and most pre-term pregnancies), and Type 3 (well-equipped to treat high-risk patients with severe complications and co-morbidities).

***1% sign., **5% sign., *10% sign.