Impact of Diagnosis Related Group Refinement on the Choice Between Scheduled Caesarean Section and Normal Delivery: Recent Evidence from France

Aleksandr Proshin, Alexandre Cazenave-Lacroutz, Zeynep Or, Lise Rochaix

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2018 – 25

Impact of Diagnosis Related Group Refinement on the Choice Between Scheduled Caesarean Section and Normal Delivery: Recent Evidence from France

Aleksandr Proshin
Alexandre Cazenave-Lacroutz
Zeynep Or
Lise Rochaix

JEL Codes: C52, I11, I12
Keywords : C-section, DRG, midwifes, obstetricians, refinement, tariffs, vaginal labor

Impact of Diagnosis Related Group Refinement on the Choice Between Scheduled Caesarean Section and Normal Delivery: Recent Evidence from France

Proshin Aleksandr, Cazenave-Lacroutz Alexandre, Or Zeynep, Rochaix Lise

Abstract

Studying quasi-experimental data from French hospitals from 2010 to 2013, we test the effects of a considerable diagnosis related group (DRG) refinement that occurred in 2012. As a result, the reform had a direct impact on hospital-level financial incentives but did not immediately concern individual providers. Using a difference-in-differences approach, controlling for multiple patient, hospital and regional characteristics and allowing for hospital and year effects, we show that introducing new 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 financial incentives, to restricting the sample to bigger (>15%) DRG tariff incentive changes and to analyzing policy effects for individual years following the reform. Our results suggest that the DRG refinement did not lead to a transmission of hospital-level stimuli to midwifes and obstetricians. Our paper is the first study that focuses on the consequences of DRG refinement in obstetrics and develops an approach suitable for measuring monetary incentives in this setting.

Keywords: C-section, DRG, midwifes, obstetricians, refinement, tariffs, vaginal labor
JEL: C52, *I11*, *I12*

1. Introduction and literature review

The topic of how financial stimuli can affect quality, quantity and composition of healthcare services has stirred much debate. Substantial anecdotal and statistical evidence has been pointing to the fact that health care providers may take into consideration factors other than patient's objective clinical picture when prescribing a certain medical treatment. Our paper estimates the responsiveness of healthcare providers to a shock in hospital level financial stimuli caused by a drastic increase in the number of diagnosis related groups (DRG) in obstetrics. The latter occurred as a result of the introduction of additional criteria into pricing mechanism, thus leading to a DRG refinement. The results of this study provide no evidence that there was a connection between the change in financial incentives and the rate of scheduled C-sections in France between 2010 and 2013.

The idea that financial incentives may make care providers alter their medical practices under prospective payment systems (PPS) has been reflected on extensively in the theoretical literature. *Ellis* (1998) considers healthcare under/overprovision and patient selection as major concerns that may arise under PPS. His model demonstrates that in a perfect information setting where competitive healthcare providers fix both a maximum severity for admitted patients and a schedule of services available at each severity level, high-severity patients will get a socially suboptimal amount of treatment (" skimping"), while the opposite should occur to low-severity patients ("cream-skimming"). *Siciliani* (2006) features a more probable information asymmetry setting and shows that the purchaser's lack of knowledge concerning patient's average severity level results in permanently higher levels of a more intensive treatment (i.e. surgery) for low-severity patients. In order to reduce the information rent, the public authorities need to pay lump sum transfers to providers who own information about their average patient severity level. To account for more complex motivations of practitioners, the assumption of self-interested practitioners was alleviated in *Hafsteindottir and Siciliani* (2009) by introducing a measure of doctors' altruism. It was shown that if prices are differentiated according to severity levels (i.e. refined), hospitals will overprovide a costly treatment. However, in the case prices where are homogeneous across severity groups, doctors with a low altruism level will underprovide costly treatment.

In a certain contradiction with the theory, the few papers that addressed specifically the effect of DRG refinement on health care provision yield results exhibiting only a relatively small effect of financial incentives, or are inconclusive. Following the approach proposed in *Ellis and McGuire* (1996), *Gilman* (2000) attempts to decompose the effect for refinement in DRGs for HIV treatment into moral hazard (change in intensity of treatment induced by tariffs) and selection effects (or "cream-skimming"). The study suggests that financial incentives created by DRG refinement had only a minor effect on the services supply. However, it is likely that this conclusion was caused by reimbursement rates that were already very generous or/and by clinical features specific to HIV treatment. In a more recent study, *Janulevicuite et al.* (2015) use fixed effects models to estimate the responsiveness of healthcare providers to changing stimuli within different DRG groups between 2003 and 2007. They find that a 10% rise in reimbursement rates leads to 0.8–1.3 % increase in the number of patients treated for medical DRGs in Norway. In line with our findings, they found no such effect for surgical DRGs.

The specific case of child delivery draws economists' attention because it perfectly fits into the framework of the theory of incentives, where an intrinsic conflict may exist between healthcare practitioners' and patients' preferences. The healthcare providers' choice between delivery modes may be biased in favor of the option reimbursed more generously, to the extent financial profits play an important role in their utility function. In addition, being an elective surgery and having relatively few strong indications, C-section arguably allows increasing hospital revenues with little risk of facing claims of malpractice.[1] On the other hand, patients would rather prefer to avoid unnecessary complications and side effects induced by a C-section, especially when the latter could easily have been substituted with a normal delivery. This conflict is a consequence of information asymmetry existing between healthcare professionals (i.e. agents) and their patients (i.e. principals). From a societal point of view, it can lead to non-optimal levels of medical care and, as a result, redistribution of wealth in favor of a potentially opportunistic agent at the expense of the less informed party.

A part of the explanation for a large observed variation in C-section rates both between countries and smaller "local markets" could be found in the induced demand theory. For instance, *Gruber and Owings* (1996) show that after a fall in fertility in the USA in 1970-1982 physicians increased C-section rates in response to falling revenues. *Gruber et al.* (1999) find a direct positive effect of price differential between C-section and normal childbirth on the rate of C-sections in USA, which can be decomposed into the income effect and substitution effects across services. However, it is worthwhile to note that *Grant* (2009) replication of Gruber et al. (1999) concludes that much of the difference in C-section and normal delivery rates

[1]See Section 2 for French healthcare context

is in fact attributable to coincidental trends and sample bias. More recently, *Allin et al.* (2015) studied a sample of Canadian women who gave birth from 2006 to 2011 and come to the same conclusion that increasing the fee differential between C-section and normal delivery provides an incentive for physicians to favor C-sections over vaginal labor. However, the magnitude of the effect is relatively small: doubling the fee differential will on average increase the probability of a C-section only by 5.6%. *Johnson and Rehavi* (2016) provide evidence that physicians perform fewer C-sections on other physicians, underscoring the importance of information asymmetry existing between physicians and patients when the decision in favor a delivery option is being made.

However, *Grytten et al.* (2011) and *Lo* (2008) come to the opposite conclusion. Studying a sample of Norwegian women, Grytten et al. (2011) find that the more expert and informed parents are, the higher C-section rate is, although the difference diminishes over time since 1967. The study of child deliveries in Taiwan conducted by *Lo* (2008) shows that the fee equalization policy did not affect C-section rates.

Another alternative explanation of higher C-section rates 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. Works that investigated this hypothesis include, but not limited to, *Currie and MacLeod* (2008), *Shurtz* (2013) and *Yang et al.* (2009). The evidence tends to support the hypothesis that higher malpractice fears proxied by court claims or reforms increasing physician liability make physicians resort to defensive medicine, resulting in higher C-section rates. Other explanations also include physicians' preferences for leisure and work (*Brown*, 1996) patients' idiosyncratic preferences (*Lo*, 2003; *Grytten et al*, 2013) and different rates of adoption of diagnostic technology (*Grytten et al*, 2012).

This paper contributes to the existing literature in two ways. First, methodologically, our contribution consists in proposing a more complex and, plausibly, more realistic approach to modeling financial incentives in the obstetrician's choice between delivery modes. In contrast to earlier studies that considered *average* fees for delivery cases (as in *Allin et al.*, 2015) to calculate the corresponding fee incentives, we construct these measures based on observed and counterfactual (i.e. predicted for the alternative delivery mode) DRG tariff groups for every patient both before and after the DRG reform. Thus, this allows us to both address the problem of unobserved actual hospital costs and take into account the fact that for a given patient severity level and comorbidities (and, consequently reimbursement rates) for C-section and normal delivery are neither constant nor independent. In other words, our approach reflects the fact more severe patients who delivered with a C-section would also likely have had a higher severity level and more comorbidities, had a normal delivery occurred. To make the prediction of DRG tariff group for the counterfactual procedure possible, we assume that the main comorbidity that would determine the counterfactual DRG tariff group is known to the physician in advance. Thus, to avoid cases where this assumption will likely be violated, we concentrate specifically on scheduled C-sections. Second, our study fills a relative lack in studies that address the effect of price refinement on health care provision in the DRG setting - currently the most widely adopted payment scheme in developed economies. In addition, to the best of our knowledge, the consequences of price refinement have never been studied in the realm of obstetrics.

The rest of paper is organized as follows: Section 2 defines the empirical context and explains the motivation of the study, Section 3 describes the data, Section 4 presents the empirical strategy. The main results are presented Section 5 and are followed by robustness checks in Section 6. The results of the study and policy implications are further discussed in Section 7. Section 8 concludes.

2. Motivation and study context

The institutional environment of French healthcare system is characterized by the presence of multiple types of healthcare institutions delivering care at highly regulated prices. In 1986, following international practice in disease classification, France adopted a diagnosis-related group (DRG) system - Groupe Homogène de Malades (GHM) - which was largely based on the DRG used in the USA. Similar to other countries that adopted DRGs, the system splits in a mutually excludable way all diagnoses between major diagnosis categories (CMD). They are further classified by diagnosis roots, which give the main reason for hospitalization. Finally, based on complication lists, a severity level is attributed to each diagnosis root. It reflects the extent to which the patient's general state of health, along with delivery-related medical circumstances, can make medical procedures associated with the diagnosis more risky and costly. Formally, these classifications are established by a state-run agency, Agence Technique de l'Information sur l'Hospitalisation

(ATIH). On an annual basis, it publishes DRG decision trees that map a set of patient characteristics onto DRGs in a unique way.[2].

Since its introduction, GHM classification underwent several revisions.[3] The latest major revision in obstetrics DRGs occurred in 2012 with the introduction of GHM version 14d. This version saw a drastic increase in the number of diagnosis groups, making them more differentiated to fit more narrowly defined groups of patients. In particular, normal deliveries became differentiated into single/multiple and primipara/multipara[4] , with severities adjusted with respect to gestational age. However, for C-sections only the distinction between single/multiple pregnancies was introduced (see Appendix 2).

In France most medical procedures for the general population are covered by social security or need minor out-of-pocket payments. Both public hospitals and private clinics are reimbursed by the state through state-run social security agencies[5]. Hospital expenditures are compensated based on the GHM diagnosis root, the severity level related to it and the status of hospital. French healthcare system features two types of hospitals – ex-OQN and ex-DG – depending on the way they are funded. With some minor exceptions, ex-OQN represent private clinics, and ex-DG represent public hospitals and private not-for-profit care institutions. We will use these terms interchangeably.

Compared to the USA and Canada, whose obstetric markets have been studied most extensively in the literature, in France most child deliveries are performed by midwifes. In general, they conduct less complicated child deliveries, while obstetricians get involved when a patient presents with complex

[2]In the specific case of obstetrics and gynecology, it is given CMD 14 in GHM version 11d. For example, to refer to a vaginal single delivery by a multipara mother with no significant complication, the root is 14Z13. In addition, it is followed by a letter which specifies the severity level of the diagnosis, e.g. 14Z13A. Thus, the letter A denotes that there was no significant complication.

[3]Note: each version of GHM comes into effect on the 1st of March of each year

[4]In medical terms, a multipara (or pluripara) mother designates "a woman who has given birth at least twice to an infant, liveborn or not, weighing 500 g or more, or having an estimated length of gestation of at least 20 weeks" (Farlex Partner Medical Dictionary, 2012).

[5]Caisse Nationale de l'Assurance Maladie des Travailleurs Salariés (CNAMTS) is the most notable among them. Overall, around 75% of healthcare expenditures are reimbursed by state affiliated agencies.

morbidities. In addition, in France it is very uncommon for ordinary physicians to perform child deliveries. The only exception to this rule is absence of a qualified midwife/obstetrician in a reasonable proximity.

The way in which healthcare providers are remunerated in France is different depending on which sector - ex-DG or ex-OQN - they belong to. In the ex-DG sector, healthcare professionals have the status of 'salarié', i.e. receive a monthly salary that is nearly insensitive to the amount of professional activity in which they were involved. In the ex-OQN sector, obstetricians' and midwifes' contracts are essentially fee-for-service, since they are remunerated for each individual procedure that they perform. However, DRG payments in this sector do not cover provider's renumeration and only reflect the average material cost of a procedure and, as a result, they tend to be lower than in the ex-DG sector. Overall, DRG payments are supposed to be institutionally disconnected from care providers' income.

Therefore, contrary to the US and Canadian obstetric care markets, in France changes in DRG reimbursement rates only directly affect the financial standing of healthare institutions and not that of individual care providers. Thus, in the French context, DRG reforms give an opportunity to test the extent to which hospital level financial incentives can be transmitted down to individual providers, while the the financial wealth of the latter is not likely directly concerned.

In the context of the French healthcare, some other channels that are believed to affect providers' decision for a delivery option, are not likely to be of significant importance. In particular, compared to the USA, malpractice claims against obstetrician constitute a rare event, due to legislation differences and, arguably, different legal culture. In addition, anecdotal evidence from French practitioners suggests that there has been no significant change in either child delivery technology or medical guidelines during the period 2010-2013.

At the institutional level, one of the objectives of DRG refinements usually consists in reducing the uncertainty of hospital revenues due to high cost patients. Contrary to fee-for-service schemes, material costs of a procedure are not necessarily fully covered by DRGs due to significant cost differentiation of cases within each DRG. This pattern was very pronounced in the French DRG system before March 2012, there being only one DRG for C-section and two DRGs for normal delivery, with the number of severity levels ranging from one to three. However, as a result of the reform, both the number of obstetric DRGs and DRG severity differentiation considerably increased.

In total, the number of tariff groups increased from 7 to 30 (see Appendix 2), following the general trend of increased refinement of DRG reimbursement rates in Europe (*Busse et al.*, 2011). As a consequence, the reform essentially brought the reimbursement scheme closer to the "cost plus" approach, where each delivery is compensated for based on the average material cost calculated by the ATIH, and a predetermined margin of expected profit.

The reform had different effects on institutions providing obstetric care. Since the most complicated cases were reimbursed at higher rates after the DRG refinement, bigger hospitals and clinics having at their disposal advanced medical equipment (and thus allowed to admit patients with high severity) likely benefited from most reform. Although the revenue uncertainty decreased for small maternity wards, some of them likely faced an increased financial pressure due to the fact that DRG reimbursement rates decreased for the simpler cases *(Vanlerenberghe, 2015).* However, the exact profitability figures are not available for majority of healthcare institutions since they are confidential.

In addition, political economy considerations can also play an important role in healthcare reforms. The existing evidence suggests that state agencies competing for budgets may have incentives to overprovide services and increase in size (*Niskansen* 1971, 1975; *Goddard et al.*, 2006). 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 DRG reform, the costs of running the DRG system very likely increased along with the number of DRGs. From the societal point of view, this could have led to lower efficiency of public expenditures and more transactional costs due to an increased regulatory burden, potentially compromising the benefits of the reform.

3. Data

The primary source of data containing information on patients' hospitalization and diagnoses in France is *Programme de médicalisation des systèmes d'information* (PMSI) dataset, which served as the main source of data in this study. It is administrative data used to pay hospitals in the PPS system, which covered all French hospitals starting from 2006. Actual PMSI data were split into 3 different datasets, which separately contained information about:

1. medical procedures performed on a woman
2. GHM main diagnosis, as well as various patient and hospital characteristics
3. patient‘s complications associated with the main diagnosis [6]

It covers all individuals who live in both metropolitan and overseas French departments. Full PMSI data access was given for the years from 2005 to 2013.

The resulting dataset contains a rich set of individual level characteristics (including age, region and community of residence, length of hospital stay, month and year 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, and lists diagnoses and complications (or secondary diagnoses) that were observed during each hospital stay. DRG diagnoses and severity levels are encoded in the current DRG version and, in some cases, both current and preceding DRG versions. Each hospital stay also has a unique patient identifier which allows us to track the cases of hospitalization across different years for a given patient. The data provides us with information about a detailed list DRG reimbursements for every single child delivery case that occurred in French healthcare institutions. Thus, we were able to observe diagnoses, comorbidites and in-hospital services registered during childbirth, both before (ante-partum) and after (post-partum) labor begins. These variables are summarized in Table 1.

To get the information about tariffs on the basis of which medical institutions are compensated, we used publicly available data provided by ATIH[7].

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 that largely resemble postal codes. However, in the cases where there are too few people living in the postal code area, for the sake of securing patients' anonymity, PMSI pseudo codes may contain observations from several postal codes. This variable is calculated as a local yearly median income per household per person.

To control for medical care availability, we use a gynecologist availability

[6] A full description of datasets is given in Appendix 1.

[7] See the ATIH data on tariffs on the website `http://www.atih.sante.fr/tarifs-mco-et-had`

Table 1: Summary statistics for obstetrical hospital stays in 2010-2013, by hospital type

	All hospitals		Non-profit hospitals		Private clinics	
Variables	Mean	St.D.	Mean	St.D.	Mean	St.D.
Age	29.687	5.4	29.539	5.47	30.104	5.16
Multiple pregnancy	0.016	0.12	0.018	0.132	0.011	0.103
Multipatity (recovered from diagnoses lists and previous years)	0.536	0.498	0.539	0.498	0.53	0.5
C-section	0.207	0.405	0.2	0.401	0.225	0.417
Scheduled C-section	0.074	0.261	0.066	0.248	0.097	0.293
Number of comorbidities	4.459	2.665	4.845	2.806	3.372	1.827
Number of previous C-sections	0.09	0.318	0.089	0.4	0.096	0.323
Gynecologist availability index	6.947	3.852	6.739	3.846	7.35	3.839
Average income per household per person (by pseudo-postal PMSI code), in thousand euro	19.512	4.412	19.4	4.374	19.983	4.503
Price incentive measures:						
$r_t - \overline{r_{2010}}$	-0.01	0.18	-0.021	0.192	0.02	0.134
$\left(r_t - \overline{r_{2010}}\right)/\overline{r_{2010}}$	-0.007	0.146	-0.019	0.15	0.027	0.126
$\Delta_t - \overline{\Delta_{2010}}$	-70.341	441.786	-99.24	498.4	11.01	190.79
Gestational age:						
22-35 weeks	0.034	0.181	0.04	0.192	0.016	0.125
36-39 weeks	0.513	0.499	0.502	0.5	0.544	0.496
≥40 weeks	0.452	0.497	0.456	0.498	0.44	0.496
Patient comorbidities :						
Infection during pregnancy	0.042	0.201	0.041	0.198	0.047	0.21
Diabetes	0.061	0.24	0.068	0.25	0.046	0.209
Complicated breech presentation	0.027	0.16	0.028	0.163	0.027	0.162
Oligohydramnios	0.015	0.122	0.017	0.13	0.01	0.092
Hypertension/Preeclampsia	0.056	0.21	0.063	0.243	0.052	0.173
Fetal growth anomaly	0.043	0.147	0.051	0.152	0.019	0.093
Hemorrage/uterine rupture	0.016	0.137	0.017	0.152	0.015	0.114
Premature rupture of membrane	0.047	0.212	0.047	0.211	0.048	0.215
Observations	2,878,958		2,124,230		754,728	

Table 2: Example if tariff groups split for deliveries without significant complication and no gestational age anomalies for ex-DG (non-profit) hospitals, in euro

<table>
<tr><th colspan="5">Normal delivery DRG prices
(no complication, gestational age 37-44 weeks)</th></tr>
<tr><td rowspan="2"></td><td colspan="2">2011</td><td colspan="2">2012</td></tr>
<tr><td>unipara</td><td>multipara</td><td>unipara</td><td>multipara</td></tr>
<tr><td>single birth</td><td colspan="2" rowspan="2">2187.13</td><td>2458.69</td><td>2070.55</td></tr>
<tr><td>multiple birth</td><td>3301.83</td><td>2897.28</td></tr>
</table>

<table>
<tr><th colspan="5">C-section DRG prices
(no complication, gestational age 37-44 weeks)</th></tr>
<tr><td rowspan="2"></td><td colspan="2">2011</td><td colspan="2">2012</td></tr>
<tr><td>unipara</td><td>multipara</td><td>unipara</td><td>multipara</td></tr>
<tr><td>single birth</td><td colspan="2" rowspan="2">2792.71</td><td>2850.37</td><td>2850.37</td></tr>
<tr><td>multiple birth</td><td>3852.63</td><td>3852.63</td></tr>
</table>

index constructed by IRDES and calculated for the year of 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[8]. These indexes are available both for gynecology and for general practice for the year of 2010.

In addition, to get figures about observed C-section rates and mother multiparity, we used Enquête Nationale Périnatale 2003 and 2010 data collected by Direction de la Recherche, des Études, de l'Évaluation et des Statistiques (DREES). [9]

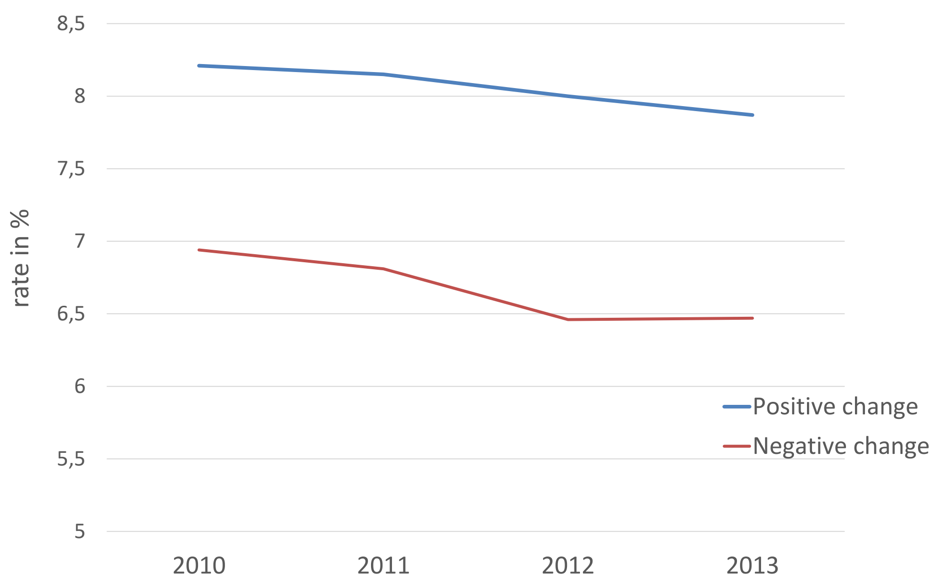
4. Empirical strategy

The individual-level data at our disposal provides us with a considerable opportunity to investigate the healthcare providers' reaction to changes in

[8] For a detailed description how the index was constructed, its 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

[9] These data are available upon request and can be found on the "Archives de Données Issues de la Statistique Publique" website https://www.cmh.ens.fr/greco/enquetes/XML/lil.php?lil=lil-0738

Figure 1: Scheduled C-section trend 2010-2013, for gropus with positive and negative changes in financial incentives $r_t - \overline{r_{2010}}$



hospital-level financial stimuli that ensued after a switch from GHM 11c to GHM 11d in March 2012. This empirical setting allows us to make use of the fact that after the refinement clinically identical cases were reimbursed differently, relative to periods preceding the reform.

To estimate the effect of financial incentives, we implement the following empirical strategy :

1. Given observed patient's characteristics and comorbidities, predict the DRG group and its severity level for pre- and post-reform DRG classifications both for C-section and normal delivery
2. Based on obtained counterfactual DRG reimbursement rates, evaluate the change in DRG tariff incentives proxied by the difference between DRG tariffs ratios for the current and the baseline year $r_t - \overline{r_{2010}}$, introduced and discussed further in this Section
3. run DID linear probability models

The first step essentially consists in predicting the counterfactual DRG and severity level, had the choice been made in favor of the other delivery option. If a mother delivered normally after the reform, the counterfactual outcome (scheduled C-section) corresponds to the lower right box in Table

Table 3: Example of observed and counterfactual outcomes (for a mother who delivered with a normal delivery after the DRG reform)

	DRG version before reform (11b, 11c)	DRG version after reform (11d, 11e)
Normal delivery	not observed (severity level and tariff predicted based on observed comorbidities)	observed DRG, comorbidities and tariff
Scheduled C-section	not observed (severity level and tariff predicted based on observed comorbidities)	not observed (DRG determinded based on mother characteristics, severity level and tariff predicted based on observed comorbidities)

3. To make this prediction, we use combidities observed during the actual delivery procedure (upper right box in example in Table 3), assuming that the main comorbidity for both modes of delivery is known to obstetrical care providers with certainty in advance. This assumption likely holds for scheduled C-sections and normal deliveries because few main comorbidies (i.e. the most severe for a given patient) should occur during labor. On the other hand, main comorbidities for unscheduled C-section are less predictable and will likely occur during labor. This leads us to focus specifically on the choice between a scheduled C-section versus other delivery outcomes. At the same time, we are aware of *Johnson and Rehavi* (2016) conclusion that information asymmetry may increase the probability of unscheduled C-sections. However, contrary to our study, in *Johnson and Rehavi* the presumed channel which affected the choice of delivery option is parental professional occupation (physician/non-physician) and not fee incentives.

The assumption that we make allows us to take into account the fact that a more complicated C-section is usually associated to more complicated normal delivery, and vice versa. This approach is different from one used, for example, in *Allin et al.* (2015) where the counterfactual fee for child delivery is assumed to be the average fee for a given type of delivery.

In order to implement the outlined strategy, several empirical issues need to be addressed. First, in the PMSI database before 2012 there is no information on whether women who gave birth via C-section were unipara or multipara. This makes it impossible to immediately attribute counterfactual normal delivery DRGs and severity levels to women who delivered via a C-section. In other words, this creates a situation where tariffs for both normal delivery and C-section for both pre- and post-reform reimbursement schedules can be either observed or predicted for women who delivered vaginally, while we cannot do the same for women who delivered with a C-section.

As a solution to this missing data problem, we used previous obstetrics PMSI data available from 2005. Since every patient has a unique identification number in this dataset, we are able trace back the history of all hospital stays for a given patient. Similar to *Lo (2008)*, to recover unipara/mutipara distinction on the individual level we look back five consecutive years [10] and see how many mothers already had a previous pregnancy that will subsequently allow classifying her as multipara.

Let us denote M_n the fact that a woman was multipara among patients she had a child in the year 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 make the assumption that $P(M_{2010-2013}) = const = P(M_{2010})$. This assumption is very plausible because multiparty is directly linked to demographic characteristics such as fertility and the age at giving birth. They are stable in countries that have already undergone a demographic transition, such as France. This assumption is also justified by statistical evidence since, according to the same survey, $P(M_{2003}) = 0.568$.

Overall, the share of all multipara cases that were retrieved on the individual level is $\frac{P(\tilde{M}_{2010-2013})}{P(M_{2010-2013})} = 0.947$. The missing 5.3% of cases are comprised of C-section cases only, since the multiparity status is completely observed for women who delivered normally. These cases likely originate from two different sources. First, they are comprised of women for whom the spacing between two consecutive childbirths exceeds 5 years. Second, this method ignores multipara women who had a child outside the French territory and, thus, were not registered in the PMSI dataset between 2005 and 2013. This category of patients can include immigrants and citizens with dual residence. According to the French Ministry of Interior, in 2010 immigrants represented at least 9 % of overall French population and tend to be both older (with some children being already born abroad) and have higher fertility than the population on average.[11] Thus, female immigrant population is highly likely

[10]Data limitations do not allow us to consider a larger spacing period.

[11]see report of French Ministry of Interior (Ministère de l'Intérieur) : Infos migrations, Vol.71, 2014. https://www.immigration.interieur.gouv.fr/Info-ressources/Donnees-

to have an increased proportion of multipara mothers.[12]

The second empirical issue arises from the fact that in our dataset, except for the year of 2012, a medical procedure can only be observed in either pre- or post reform DRG classification (see example given in Table 3 corresponding to DRG versions 11b and 11c). Without having the information on how the same activity would have been classified in different DRG versions in terms of diagnosis root and severity level it would impossible to predict what would have been DRG tariffs for these same cases across different DRG schedules. For the sake of this, we reconstructed the DRG decision tree for both pre-and post-reform periods available on ATIH website. The decision tree classifies cases into a specific DRG according to mother relevant medical characteristics (uniparity/multiparity, single/multiple delivery). The severity level is determined by the most severe comorbidity observed in a patient.

Hereafter, we will consider that observations belong to the same treatment group if they share a common diagnosis root and a common severity level for the observed and counterfactual child delivery mode in both pre-reform and post-reform DRG classifications. Since DRG tariffs are completely determined by diagnosis roots and severity levels, observations within one treatment group also have in common to the same amount of financial stimuli generated by the DRG reform. Overall, this gives rise to as many as 850 groups that received a treatment of different intensity. For the sake of simplicity, we excluded small groups with the number of cases inferior to 2500.

For each child delivery we predict the ratio $\frac{\overline{CS_{2010}}}{\overline{ND_{2010}}}$ of C-section and normal delivery reimbursement rates based on the DRG classification and reimbursement rates actual in the baseline year of 2010. Then this ratio is recalculated for the DRG schedule for the year when the child delivery actually occurred.

The difference of these two expressions is:

$$\frac{CS_t}{ND_t} - \frac{\overline{CS_{2010}}}{\overline{ND_{2010}}} \quad (1)$$

where CS_t and ND_t are tariffs for C-section and normal delivery respec-

statistiques/Etudes-et-publications/Publications/Numeros-parus-en-2014/Les-familles-des-immigres

[12]In an attempt to alleviate this bias, we tried restricting the sample only to younger women aged 20-25 who, thus, were not likely to have a child earlier. However, this did not significantly change the percentage of retrieved C-section multipara cases

tively, observed (or predicted for the counterfactual delivery outcome) using the DRG version of year t, and $\overline{CS_{2010}}, \overline{ND_{2010}}$ are baseline tariffs that would have been observed if the child delivery had taken place in 2010. For the sake of notational simplicity, we will further denote the ratio $\frac{CS_k}{ND_k}$ as r_k.

The change in variable (1) serves as the main measure of how relative profitability of two delivery options evolved in the course of time compared to 2010. Thus, for deliveries that occurred before the reform, $r_k - \overline{r_{2010}} = 0$, since both ratios are determined based on the pre-reform DRG versions and thus coincide [13]. When a child delivery occurred after the reform, the measure (1) is positive (negative) if and only if, compared to pre-reform tariffs:

1. C-section tariff increased (decreased)
2. normal delivery tariff decreased (increased)
3. a combination of both mentioned tariff changes occurred such that the C-section tariff increased (decreased) at a higher rate than one for normal delivery, compared to the baseline year[14]

This variable (1) is continuous, allowing us to track how different intensities of the change in relative financial profitability of scheduled C-section and normal delivery affected the probability of the obstetrical care provider's choice in favor of a delivery option in each child delivery case.[15]

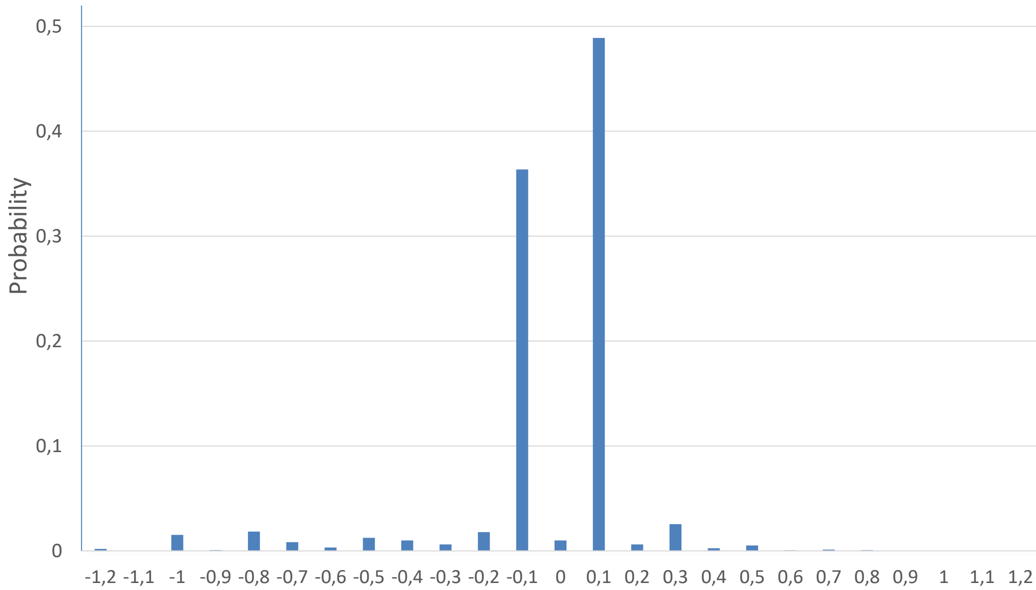
Every hospital incurs a financial cost for every C-section and normal delivery it performs, which is unobserved in the data. These costs are reimbursed to healthcare institutions to a certain degree (which may or may not surpass the real cost) by the state, private insurance organisms and out-of-pocket payments. From a financial point of view, C-sections are generally more expensive to perform than normal deliveries. However, in our context, the mere fact that a C-section is reimbursed more generously by DRG schedules than a normal delivery does not automatically mean that a C-section becomes financially more attractive than a normal delivery since this reimbursement

[13]Negligible ($r_{2011} - \overline{r_{2010}} < 0.015$) and unrelated to DRG classification yearly cost-adjustment occurred in 2011, which is for now not considered. We explore this change in robustness checks Section 6

[14]Inflation changes are eliminated in measure (1) since both tariffs enter the ratios

[15]See Section 6 for specifications with alternative price incentive measures.

Figure 2: Distribution of the change in financial incentive $r_{2012,2013} - \overline{r_{2010}}$, for all hospitals.

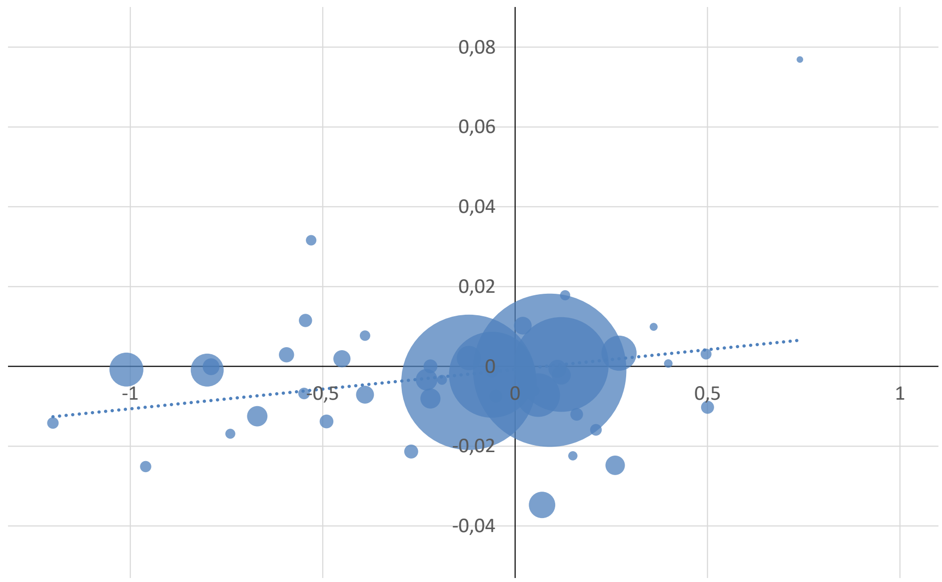


Legend: histogram width=0.1, bins=25.

may not cover the actual costs faced by each hospital. In other words, although every given secondary diagnosis is associated to a C-section and a normal delivery severity level (which may or may not coincide) the financial profitability of a specific case will ultimately depend on how well, for a given set of comorbidities and patient characteristics, the reimbursement covers the unobserved cost of the two delivery options. For each child delivery, the evolution of (1) variable between different DRG versions must reasonably be capturing the change in how full is coverage for two delivery methods by DRG tariffs.

Although several previous studies (*Gruber et al.* 1999, *Allin et al.* 2015, etc) examined the ratio (or price differentials) of tariffs for C-section and normal delivery as a measure of the financial incentive, this approach will not be as informative in the DRG refinement setting in the presence of multiple tariff groups. First, it does not take into account the heterogeneity of cases in terms of initial reimbursement (and thus the extent to which the DRG payment covers actual hospital costs) and instead takes average fees in a given year geographical area and year. Second, as discussed in Section 2,

Figure 3: Financial incentives for scheduled C-section ($r_{2012} - \overline{r_{2010}}$, horozantal axis) and changes in scheduled C-section rates after 2012 DRG reform (difference of scheduled C-section rates between 2012 and 2011, vertical axis), for treatment groups



Legend : size corresponds to the number of cases in the treatment group

in France tariffs in ex-OQN and ex-DG sectors are not directly comparable since they include different components. Thus, their evolution relative to the baseline period, rather than their static level at a given period, will be a more consistent way to approximate financial incentives generated by the reform.

As shown in Figure 2, after the DRG refinement several groups of cases underwent substantial changes in relative reimbursement rates, while many cases were almost unaffected or almost unaffected by the reform. Thus, unaffected observations form an equivalent of a control group found in DID specifications were a treatment variable is binary.

In addition, it is worthwhile to note that on average for-profit and non-profit institutions were affected slightly differently by the reform. As shown in Table 1, the mean value of (1) was positive (negative) for for-profit (non-profit) hospitals, indicating that C-section became relatively more (less) profitable to perform in these institutions.

Difference-in-differences (DID) specifications are a common tool to estimate the causal effect of policy changes. Since the outcome variable, scheduled C-section indicator, is binary, we refer to our model as linear probability DID.

To test the effect of the price incentive, we proceed to estimate

$$CSsch_{iht} = \alpha + X'_{iht}\beta_1 + \beta_2 \Delta price_{iht} \cdot POST_t + d'_{iht}\xi + \gamma_t + \theta_h + \epsilon_{iht} \quad (2)$$

where $CSsch_{iht}$ is a binary outcome for a scheduled C-section for patient i, healthcare institution h, and year t. Term α is constant, X_{iht} is a column vector of controls including age, secondary ante-partum (i.e. occurring before labor begins) diagnoses [16] and other observed clinical characteristics, gynecologist availability index and pseudo-postal code average income, $\Delta price_{iht}$ is the price incentive defined as (1), $POST_t$ is a post-reform period dummy taking the value 1 for years 2012 and 2013. [17] 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 and hospital fixed effects γ_t and θ_h respectively.

Even though, in the general case, the linear probability model is biased and suffers from heteroskedasticity problems, it gives correct average treatment effects, provided that the data generating process is specified correctly.

The main coefficient of interest is β_2. In these specifications, it has the interpretation of the number of percents by which the probability of a scheduled C-section increases if the DRG relative tariff ratio for scheduled C-section increases by one unit relative to the baseline period level. In other terms, if normal delivery tariffs are kept fixed, 100% increase in C-section tariff relative to the DRG schedule of the baseline year would on average lead to $\beta_2 \cdot \overline{r_{2010}}\%$ increase of scheduled C-section probability.

For DID estimates to be valid, treatment groups must satisfy the common trend assumption. This assumption may be violated if time-varying unobserved characteristics affected groups differently. To lend this assumption more credibility, we analyze only groups that fall within the margin of 15%

[16] Inclusion of ante-partum controls is optional since treatment groups are in part determined, in a non-linear way, by them. However, their inclusion is susceptible of decreasing standard errors of estimates.

[17] The inclusion of $POST_t$ variable serves to set to zero all values of $\Delta price_{iht}$ for observations before March 2012. In general, reimbursement rates underwent a negligible change in 2011 after an inflation and cost-driven re-adjustment of tariffs that were unlikely to influence the medical practice. The impact of this change is further tested as a robustness checks in Section 6

change relative to the average scheduled C-section rate of 7.53%, registered in the pre-reform years of 2010 and 2011, which is equivalent to 0.56% deviation from the pre-reform scheduled C-section rate in the post-reform. This makes us eliminate 8.1% of cases from the sample.

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.

The models are tested for all hospitals combined, and separately for private for-profit (ex-OQN) and for public and private non-profit (ex-DG) care institutions.

5. Main results

The results of various model specifications are presented in Table 4. The introductory logistic model (1) is run on the same covariates as the main specification presented in Section 4, excluding price incentive measures and treatment group dummies. The results of (1) are in line with general knowledge about this type of medical interventions. Namely, older women are in a greater risk of complications during vaginal delivery. Scheduled C-sections are usually conducted before labor is likely to start and, in the same time, are less likely to be performed on very pre-term mothers. In addition, previous C-sections are associated to an increased risk of normal delivery failure, although they are not considered per se as a counter-indication to a normal delivery by the French Health Authority (HAS).[18]

[18] It is worthwhile to note that French Healthcare Authority (HAS) periodically monitors the medical pertinence for C-sections. It concluded that in order to avoid evening and nighttime deliveries, French obstetricians tend to schedule C-sections too early. Moreover, keeping healthcare practitioner individual record of C-section may lead to changes in medical practices. Following this intervention, C-section rates dropped from 22,2% in 2010 to 17% in 2013 in hospitals concerned by monitoring. *(Vanlerenberghe, 2015)*

Table 4: Impact of tariff changes on the probability of scheduled C-section, 2010-2013

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Difference of price ratios (relative to version 11b)	—	-0.0011 (0.0011)	-0.0017 (0.0011)	0.0048 (0.0044)	-0.0012 (0.0011)	-0.002* (0.001)	0.004 (0.004)
Age	0.04*** (0.0006)	0.0023*** (<0.0001)	0.0024*** (<0.0001)	0.0023*** (0.0001)	0.0023*** (<0.0001)	0.002*** (<0.0001)	0.002*** (0.0001)
Gestational age 22-35 weeks	-0.72*** (0.02)	-0.018*** (0.002)	-0.0131*** (0.0024)	-0.0466*** (0.0061)	-0.017*** (0.002)	-0.01*** (0.002)	-0.04*** (0.006)
Gestational age 40-44 weeks	-1.35*** (0.008)	-0.0431*** (0.0003)	-0.0442*** (0.0003)	-0.0368*** (0.0005)	-0.042*** (0.0002)	-0.04*** (0.00001)	-0.03*** (0.0005)
Act performed on Sunday or holiday	-6.47*** (0.01)	-0.0401*** (0.0005)	-0.0386*** (0.0005)	-0.0443*** (0.001)	-0.045*** (0.0005)	-0.04*** (0.0006)	-0.042*** (0.001)
Number of previous C-sections	0.64*** (0.003)	0.0992*** (0.0005)	0.0958*** (0.0006)	0.0975*** (0.001)	0.097*** (0.0005)	0.093*** (0.0005)	0.094*** (0.001)
Gynecologist concentration index	-0.01*** (0.001)	-0.0008*** (<0.0001)	-0.0009*** (<0.0001)	-0.0006*** (<0.0001)	-0.0004*** (<0.0001)	-0.0004*** (<0.0001)	-0.0004*** (0.0001)
Average postal code income per person	0.0001 (<0.0001)	-0.0001*** (<0.0001)	-0.0001*** (<0.0001)	0.0001** (<0.0001)	-0.0001*** (<0.0001)	-0.0001*** (<0.0001)	0.0001*** (<0.0001)
Antepartum diagnosis controls	Y	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y	Y
Hospital fixed effects	Y	N	N	N	Y	Y	Y
Num. of observations	2,612,472	2,369,472	1,784,737	584,735	2,369,472	1,784,737	584,735
R-squared	—	0.438	0.41	0.536	0.444	0.416	0.541
Pseudo R-squared	0.15	—	—	—	—	—	—

*Legend: ***- 1 % sign., **-5 % sign., *- 10% sign. Coefficient std. errors in parentheses.* (1) Logit model for all hospitals, whole sample, with year, antepartum diagn. and hospital controls; (2) and (5) DID, linear probability for all hospitals with specified controls; (3) and (6) DID, linear probability for public and private not-for-profit hospitals; (4) and (7) DID, linear probability, for private for-profit hospitals

The main models of interest are (2)-(7). Models (2)-(4) estimate the impact of financial incentives using all the controls specified in Section 1 without hospital fixed effects. Compared to model (1), the signs of most coefficients remain unchanged. The main coefficient of interest β_2 representing the change in financial incentives compared to the baseline period is insignificant, suggesting that the financial incentives caused by the DRG reform did not affect obstetric providers' practice. We will explore models with alternative formulations of price incentive in Section 6.

The introduction of hospital fixed effects into the model in specifications (5)-(7) led to a very moderate increase in the share of explained variation, but did not alter the main results. Coefficients for $r_t - \overline{r_{2010}}$ remain insignificant in the model tested on all hospitals (5) and private clinics (7), while β_2 becomes minimally significant at 10% for private clinics. However, given that $\overline{r_{2010}}$ is comprised with the interval $(0.83; 2.6)$, the magnitude of the effect of 100% C-section tariff increase, $\beta_2 \cdot \overline{r_{2010}}$, is negligible.

It is worthwhile to note that, in the specifications that we tested, the coefficient for the price incentive tends to be positive for models tested on private hospitals, and negative for models run on all hospitals and non-profit hospitals.

Finally, similarly to *Gruber et al.* (1999) and *Milcent and Rochut* (2009), we find that gynecologist availability decreases the chance of a scheduled C-section being chosen as a delivery option. It is consistent with their explanation that demand inducement that occurs in the settings of obstetrician-patient information asymmetry is one of the channels augmenting C-section rates.

6. Robustness checks

6.1. Alternative price incentive measures

The measure $r_k - \overline{r_{2010}}$ is not the only way how financial incentives can be represented. Alternative measures include :

1. $\left(r_t - \overline{r_{2010}}\right) / \overline{r_{2010}}$, or change of price ratios, relative to the hypothetical level of price ratios in 2010.
2. $\Delta_t - \overline{\Delta_{2010}}$, where Δ_k stands for inflation-adjusted price differential $CS_k - ND_k$ in year k, in prices of 2010. Thus, it reflects the change in the price differential Δ_k compared its hypothetical level in 2010
3. $\left(\Delta_t - \overline{\Delta_{2010}}\right) / \overline{\Delta_{2010}}$, which is the same measure as 2, calculated relative to the hypothetical level of the price differential in 2010.

By definition, all these price incentive measures have value 0 in 2010 because of the absence of change (i.e. DRG versions and, consequently, r_t (or Δ_t) coincide).

To estimate the impact of the price incentive, we tested the same specifications as in Section 4, the only difference being the price incentive measure used. The results for β_2 coefficient for the specifications including year and

Table 5: Coefficients for other price incentive measures, by hospital type

Measures of financial incentive	Hospital type All	Non-profit	For-profit
	-0.0009	-0.0024	0.0064
$\left(r_t - \overline{r_{2010}}\right)/\overline{r_{2010}}$	(0.002)	(0.0025)	(0.0056)
	(0.62)	(0.23)	(0.21)
	-0.0008*	-0.001**	0.0046
$\Delta_t - \overline{\Delta_{2010}}$	(0.0005)	(0.0005)	(0.0038)
	(0.09)	(0.03)	(0.23)
	-0.0003	-0.0006	0.0016
$\left(\Delta_t - \overline{\Delta_{2010}}\right)/\overline{\Delta_{2010}}$	(0.0004)	(0.0004)	(0.0013)
	(0.48)	(0.22)	(0.23)

*Legend: ***- 1 % sign., **-5% sign., *- 10% sign. Coefficient std. errors and p-values are given in parentheses under the coefficient. DID, linear probability model with year and hospital fixed effects as in models (5)-(7) in Table 4*

hospital fixed effects (as in models (5)-(7) in Section 5) are presented in Table 5.

The results are in line with the estimates obtained in Section 5. The main coefficient of interest β_2 is statistically significant only for measure 2 at 10% and 5% significance level in regressions of all institution types and public hospitals respectively. However, combined with previous results and given the insignificance of this coefficient for measures 1 and 3, we cannot interpret it as robust evidence for the presence of the effect of tariff incentives on obstetrical practice.

6.2. Considerable incentive changes

The changes in tariffs resulting from DRG refinement concerned, without exception, all patient groups. However, as noted earlier, some of them were more affected by the reform than others.

Under the hypothesis that only considerable DRG tariff incentive changes would be able to change the medical practice, or even be noticed by a obstetrical care providers, we impose additional criteria on the sample. First, in the post-reform period we include only those observations for which the change in the relative price relative to its level in 2010 exceeded 15% , that is $\left(r_t - \overline{r_{2010}}\right)/\overline{r_{2010}} > 0.15$, $t = \{2012, 2013\}$. Second, to match post-reform

Table 6: Impact of tariff changes on the incidence of scheduled C-section in 2010-2013, cases with changes in relative price > 15%

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Difference of price ratios (relative to version 11b)	-0.0084***	-0.0028	-0.0042*	-0.0139	-0.003	-0.004*	0.004
	(0.0023)	(0.0021)	(0.0023)	(0.0104)	(0.002)	(0.0022)	(0.004)
Age	0.0051***	0.003***	-0.0038***	0.0012***	0.0032***	0.0036***	0.002***
	(0.0001)	(0.0001)	(0.0001)	(0.0004)	(<0.0001)	(0.0001)	(0.00005)
Gestational age 22-35 weeks	-0.0919***	-0.0543***	-0.0601***	-0.057***	-0.054***	-0.058***	-0.04***
	(0.0094)	(0.0085)	(0.0122)	(0.002)	(<0.0001)	(0.0121)	(0.006)
Gestational age 40-44 weeks	-0.1179***	-0.0846***	-0.085***	-0.068***	-0.083***	-0.0842***	-0.03***
	(0.0014)	(0.0014)	(0.0014)	(0.0007)	(<0.0001)	(0.0013)	(0.0005)
Act performed on Sunday or holiday	-0.0926***	-0.0739***	-0.0785***	-0.059***	-0.0843***	-0.086***	-0.042***
	(0.0028)	(0.0025)	(0.002)	(0.0089)	(0.0025)	(0.0027)	(0.001)
Number of previous C-sections	0.2729***	0.0957***	0.0941***	0.1064***	0.0921***	0.091***	0.094***
	(0.0018)	(0.0021)	(0.0023)	(0.0062)	(0.002)	(0.0023)	(0.001)
Gynecologist concentration index	-0.0015***	-0.0009***	-0.001***	-0.001*	-0.0005***	-0.0006***	-0.0004***
	(0.0001)	(0.0002)	(0.0002)	(0.0006)	(0.0002)	(0.0002)	(<0.0001)
Postal code income per person	$-6.8 \cdot 10^{-5}$	-0.0002	-0.0001	-0.0015***	$-8.8 \cdot 10^{-5}$	$-4 \cdot 10^{-5}$	-0.0009*
	(<0.0001)	(0.0001)	(0.0001)	(0.0005)	(<0.0001)	(<0.0001)	(0.0006)
Antepartum diagnosis controls	N	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y	Y
Hospital fixed effects	N	N	N	N	Y	Y	Y
Num. of observations	188,438	188,438	177,084	11,354	188,438	177,084	11,354
R-squared	0.199	0.357	0.359	0.355	0.368	0.368	0.374

*Legend: ***- 1 % sign., **-5 % sign., *- 10% sign. Coeff, std. errors in parentheses.* (1), (2), (5) DID, linear probability for all hospitals, with specified controls; (3), (6) DID, linear probability for public and private not-for-profit hospitals; (4), (7) DID, linear probability, for private for-profit hospitals

observations with comparable pre-reform ones, we include only those cases in pre-reform period for which $\left(\overline{r_{2012}} - r_t\right)/r_t > 0.15$, $t = \{2010, 2011\}$, where r_{2012} is a hypothetical price ratio post-reform DRG tariff level in 2012.[19]

These exclusion criteria reduced the size of sample by 13.8 times from 2,369,472 to 188,438 observations. The estimation results for this sample are presented in Table 6. Models (2), (3) and (4) run on patients treated all, not-

[19]Using 2012 and 2013 DRG tariffs gives the same result because in 2013 tariffs were only adjusted by the inflation rate that disappears while calculating the expression

for-profit and private- for-profit hospitals respectively do not provide evidence for the role of financial incentives in influencing the clinical decision in favor of a scheduled C-section. As in Section 5, the inclusion of hospital fixed effects in models (5)-(7) did not change the results, the main coefficient of interest β_2 remaining insignificant for models (5) and (7), minimally significant in the model (6) tested on patients treated in not-for-profit hospitals.

6.3. Effects of price incentives, by year

In the models tested so far, the effect of changes in price incentives has been evaluated as homogeneous for the post-reform period. However, it can be tested separately for different smaller periods, since the intensity of the effect can potentially change over time. For example, if financial incentives matter in the decision, it may take time for healthcare practitioners to understand the monetary consequences of the reform, resulting in more sizable effects in later periods.

In addition, as mentioned in Section 5, in 2011 there was a negligible adjustment of tariffs that did not completely mirror the inflation rate. This change could not plausibly be a significant factor influencing the delivery mode decision, and very likely had passed unnoticed by both hospital administrators and care provider. In the models presented earlier, this change was ignored because of the interaction of a price incentive variable with the post-reform indicator. However, as a robustness check, we explicitly test the effect of this tariff change along with changes in the post-reform years 2012 and 2013.

To make it possible, we include interactions of a specified price incentive measure with a set of year dummies. Thus, the model tested can be specified as:

$$CSsch_{iht} = \alpha + X'_{iht}\beta_1 + \sum_{n=0}^{2} \beta_{2+n}\Delta price_{iht} \cdot Year_{2011+n} + d'_{iht}\xi + \gamma_t + \theta_h + \epsilon_{iht} \quad (3)$$

where $\beta_2, \beta_3, \beta_4$ are the coefficients of interest for a price incentive measure interacted with year dummy variables $Year_t$. The models also include year and hospital fixed effects.

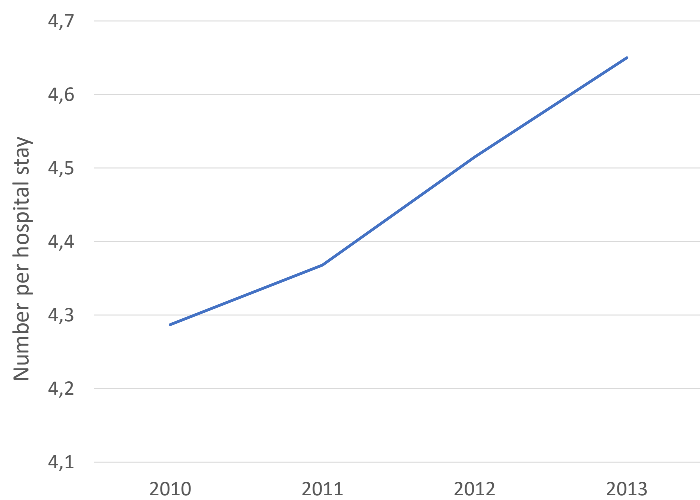
Table 7: Coefficients for price incentives, effects by year

	2011	2012	2013
Price inventive measure			
$r_k - \overline{r_{2010}}$			
	0.02	-0.003	-0.0017
All hospitals	(0.0306)	(0.0014)	(0.0015)
	(0.51)	(0.83)	(0.26)
	0.013	-0.0007	-0.003**
Not-for profit hospitals	(0.03)	(0.001)	(0.0015)
	(0.66)	(0.61)	(0.04)
	3.12*	0.0018	0.0074
Private clinics	(1.78)	(0.0052)	(0.0056)
	(0.079)	(0.72)	(0.18)
Price inventive measure			
$\Delta_k - \overline{\Delta_{2010}}$ *(in thousand euros)*			
	-0.0005	0.0014	0.0002
All hospitals	(0.0005)	(0.0012)	(0.0013)
	(0.36)	(0.22)	(0.85)
	-0.0008	0.0016	0.0006
Not-for profit hospitals	(0.0006)	(0.0013)	(0.0014)
	(0.17)	(0.21)	(0.66)
	0.0134**	0.003	0.0034
Private clinics	(0.0067)	(0.0042)	(0.0046)
	(0.04)	(0.46)	(0.45)

*Legend: ***- 1 % sign., **-5% sign., *- 10% sign. Coefficient std. errors and p-values are given in parentheses under the coefficient. DID, linear probability models including year and hospital fixed effects*

The estimation results for the coefficients of interest are summarized in Table 7. As in earlier specifications, the results do not provide evidence for the presence of the impact of financial invectives on the probability of scheduling a C-section. A significantly negative coefficient appears in 2013 in the model run on not-for-profit hospitals using the measure $r_k - \overline{r_{2010}}$, which nonetheless is not present when $\Delta_k - \overline{\Delta_{2010}}$ is tested. In addition, the coefficient for β_2 reflecting the impact of minor tariff changes in 2011 is significant at 5% and 10% significance levels in specifications run on for-profit

Figure 4: Average number of encoded comorbidities per child delivery case in 2010-2013



hospitals for measures $r_k - \overline{r_{2010}}$ and $\Delta_k - \overline{\Delta_{2010}}$ respectively. However, the magnitude of effect produced by these coefficients is negligible. In addition, its unstable character is indicative of a multicollinearity problem if effects for 2011 are included in the model.

7. Discussion and policy implications

In this study we have shown the absence of a link between financial incentives and the choices between a scheduled C-section and a normal delivery. The results demonstrate that neither public nor private healthcare providers in France significantly responded to changes in reimbursement rates caused by a DRG refinement.

The results of the study are most likely driven by three channels. First, the results can be explained by the fact that in France healthcare providers' contacts are drawn up in a way that makes them insensitive to hospital-level financial objectives. To put it differently, it is difficult for healthcare institutions to enforce their financial policies on individual providers may enjoy a sufficient level of autonomy and legal guarantees. In addition, healthcare management can potentially face catastrophic reputation costs if attempts to exert influence on medical practices become public. Arguably, this situation is even less likely to occur in bigger maternity wards where policy coordination between administration and individual healthcare providers is generally considered to be a more serious issue.

Secondly, it is important to emphasize that the cost of understanding the financial consequences of DRG reforms can be fairly high for both health care management and individual practitioners. This may render their perception about the arising stimuli imprecise. Anecdotal evidence also suggests that many administrators managing healthcare institutions were unaware of the DRG reform.

Lastly, at the level of healthcare institutions, the reform eased the financial pressure on bigger maternity wards since the reform resulted in their overall revenue increases. In other words, DRG refinement benefited mostly those healthcare providers that had an increased share of high-cost patients presenting with complicated medical comorbidities, targeted by the policy intervention. Thus, after the reform, bigger maternity wards did not have strong incentives to react to changes in terms relative reimbursement rates because of their improved financial standing. On the other hand, overall revenues of small obstetric care providers generally remained stable or even decreased, making this channel improbable for them.

The evidence presented in this paper suggests that the main objective of the DRG reform - decreasing financial risk for obstetric healthcare providers - was feasibly met without significant impact on medical practices in question. However, the overall impact of the DRG refinement reform in terms of public wealth is not clear. Although unobservable transnational and administrative costs related to encoding medical procedures likely increased as a consequence of the DRG reform, the exact magnitude of this effect is not clear nor can be observed.

Our main contribution to the previous research is twofold. First, our methodology introduced a considerable improvement to measuring price incentives. Unlike previous studies, we take into account the dependence between severity levels for both modes of delivery at the patient level, which allows us to exploit more fully the resulting variation in tariff incentives. In the specific case of DRG refinement, ignoring this relationship can potentially lead to a significant measurement error and attenuation bias. Second, this study fills the relative lack in research on the impact of DRG refinement on medical practice and is the first study analyzing this impact in the area of obstetrics.

Changes in reimbursement rates could have made "upcoding" a possible strategy for healthcare providers, generating more profits without affecting the volume or quality of services. Such an effect was observed by *Dafny* (2005), who finds that price changes resulting from abolition of age criteria in DRGs (DRG aggregation) in the USA in 1988 and subsequent recalibration of tariffs led to an increase of the share of top-coded of patients, while evidence that this price shock affected services volume or intensity was weak. In our sample, the average number of encoded comorbidities per child delivery was higher in the post reform period than before it (see Figure 4). However, the trend of increasing number of encoded comorbidities existed before the 2012 reform, arguably making this channel unlikely to be triggered by the reform.

Another concern is the fact that DID strategy allows eliminating only time-invariant component in group unobserved heterogeneity. Such an event could be the publication of good practice manual by HAS in 2012, which was coincidental with the reform in DRGs. However, despite the importance of this manual, it contained no strong indications or counter-indications for scheduled C-sections, instead listing in an non-exhaustive fashion around 7 common medical conditions for which a scheduled C-section can be recommended. Hence, the extent to which it could affect the actual obstetrician's practice appears to be restricted to cases where a patient presented with only relatively common comorbidities, the final decision still remaining at obstetrician's and, in some cases, patient's discretion.

The efficiency of the estimates could have increased if additional explanatory varibles had been included. For example, our database lacks information about patients' occupation which can affect provider's choice between delivery options if patient is perceived by a physician as lacking relevant obstetrical knowledge (*Johnson and Rehavi*, 2016). Similar to other related studies mentioned in the introduction, we cannot directly address the problem of patient self-selection into private/public healthcare providers based on unobserved characteristics, such as patient's risk-aversion and idiosyncratic preferences. In France under certain conditions women can have a freedom of choice between a C-section and a normal delivery. This freedom is generally larger in private hospitals, where women can have more influence on the treatment plan they receive. Therefore, if there is a preference for a C-section among these women, private hospitals will have a higher C-section rate.

8. Conclusion

Our study estimated the impact of DRG refinement and subsequent changes in reimbursement rates for child delivery modes on the choice between scheduled C-section and normal delivery. This policy shock occurred in 2012 in an attempt to reduce financial risk for obstetric care providers by introducing additional parameters into the price formula, such as multiparity, multiple birth, gestational age and new comorbidities, bringing the payment system closer to the cost-plus approach. Thus, patients who were identical in terms of observed characteristics would have had different reimbursement rates before and after the reform, changing the relative profitability between child delivery options. To exploit this exogenous variation in DRG tariffs, we adopted different approaches to measuring price incentives and tested difference-in-difference models. The results suggest that the reform did not have a significant impact on healthcare providers' obstetrical practice. Models testing only bigger changes in incentives and specifications with alternative measures of financial incentives lead to the same conclusion. Moreover, there is no evidence that effects varied across years and for different types of care institutions.

The results are likely due to three main channels. First, the transmission of hospital-level financial stimuli to individual care providers appears to be limited. Secondly, the costs of understanding of changes in DRG tariff stimuli can be sufficiently for hospital administrators. Lastly, bigger obstetric care providers benefited from overall increases reimbursement rates due to recalibration of tariffs for the most severely coded patients targeted by the reform. As a result of an improved financial position, they faced less pressure to raise profits by changing the medical practice.

This study contributes to the existing literature on the impact of financial incentives provoked by DRG refinements. The study concentrates on the choice between two treatment options, one of which is often an elective surgery and the other is a non-surgical procedure. To the best of our knowledge, the consequences of price refinement have never been studied specifically in the realm of obstetrics. Methodologically, our contribution consists in proposing an approach to modeling changes in DRG financial incentives at the patient level. In contrast to earlier studies that considered *average* fees for delivery modes (as in Allin et al. (2015)) to calculate the corresponding fee incentives, we construct these measures based on observed and predicted DRG tariff groups for both child delivery modes according to

DRG classifications before and after the refinement reform. The resulting estimates exploit tariff variations in a more complete fashion, thus alleviating a potential source of attenuation bias due to measurement error.

The approach we used to estimate the effect of change in relative profitability on child delivery decision can potentially be improved by modeling the probability for each severity level in each delivery outcome. This would allow relaxing the assumption of a perfect knowledge of main comorbidities, replacing it instead with estimated probabilities for each severity level for both delivery modes. Thus, under this specification a care provider would be making her/his decision based on *expectations* of hospital-level financial rewards conditional on ante-partum diagnoses observed at the time of decision.

Another 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 trend for merging or closing small maternity wards, leading to an emergence of larger institutions belonging to different financial groups. The way these industrial changes interacted with price incentives is an unanswered question.

Appendix 1

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 a gynecologist between 0am and 6am, 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, this dataset does not allow to identify a patient, and 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 me to find a corresponding medical activities contained in the previous dataset. It is worthwhile to note that since in France each version of GHM comes into effect on the 1st of March of each year, 'RSA' dataset classifies diagnoses based on two versions of GHM, 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 GHM, 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 GHM version 11d in 2012, not only diagnosis roots underwent a major revision, but also severity levels.

Appendix 2

Table 8: Correlation matrix of main variables

	Sch. C-section	$r_k - \overline{r_{2010}} = 0$	$\Delta_t - \overline{\Delta_{2010}}$	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											
$r_k - \overline{r_{2010}} = 0$	0.001	1										
$\Delta_t - \overline{\Delta_{2010}}$	-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	Gyn. availability
Average pseudo postal code income	1	
Gyn. availability	0.171	1

Table 9: Obstetrical DRGs before and after 2012 reform

DRG before reform (GHM 11b, 11c)			DRG after reform (GHM 11d, 11e)		
GHM code	Num. of severity levels	Name of procedure	GHM code	Num. of severity levels	Name of procedure
14C02	3	Cesarean section	14C06	4	Cesarean section, with child dead
			14C07	4	Ceserean section, multiple pregnancy
			14C08	4	Ceserean 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 10: 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 11: Effect of gestational age on severity level of a Cesarean 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 12: 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

References

[1] 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.*

[2] Barlet, Muriel and Coldefy, Magali and Collin, Clémentine and Lucas-Gabrielli, Véronique (2012). Local Potential Accessibility (LPA): A New Measure of Accessibility to Private General Practitioners. *Questions d' économie de la santé*, 174, p1-8, `http://www.irdes.fr/EspaceAnglais/Publications/IrdesPublications/QES174.pdf`

[3] Brown, H. (1996). Physician demand for leisure: implications for cesarean section rates. *Journal of Health Economics*, 15(2), pp.233-242.

[4] Diagnosis-Related Groups in Europe : Moving towards transparency, efficiency and quality in hospitals. Edited by Busse R., Geissler A., Quentin W., Wiley M. *Open University Press* (2011)

[5] Clemens, J., Gottlieb J. (2014). Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health? *American Economic review*, 104 (4), pp. 1320-49.

[6] Coudin, E., Pla, A., Samson, A. L. (2015). GP responses to price regulation: evidence from a French nationwide reform. *Health economics*, 24(9), pp.1118-1130.

[7] Currie J.W., MacLeod B. (2008) First Do No Harm? Tort Reform and Birth Outcomes. *The Quarterly Journal of Economics*, 123, (2), pp. 795–830.

[8] Dafny, L.(2005). How Do Hospitals Respond to Price Changes? *The American Economic Review*, 95(5), pp. 1525-1547.

[9] Delattre, E., Dormont, B. (2003). Fixed fees and physician-induced demand: A panel data study on French physicians. *Health Economics*, 12(9), pp.741-754

[10] Delattre, E., Dormont, B. (2005). La régulation de la médecine ambulatoire en France: quel effet sur le comportement des médecins libéraux?. *Dossiers solidarité et santé*, (1), pp.135-161.

[11] Dubay, L., Kaestner, R., Waidmann. T. (1999). The Impact of Malpractice Fears on Caesarean Section Rates. *Journal of Health Economics*, 18 (4) pp.491-522.

[12] Ellis R. (1998) Creaming, skimping and dumping: provider competition on the intensive and extensive margins, *Journal of Health Economics*, Vol. 17, pp. 537—555.

[13] Ellis R., McGuire T. (1996) Hospital response to prospective payment: Moral hazard, selection, and practice-style effects, *Journal of Health Economics*, Vol. 15, pp. 257–277.

[14] Gilman B. (2000) Hospital response to DRG refinements : the impact of multiple reimbursement incentives on inpatient length of stay. *Health Econ Policy Law*, Vol. 1, Issue 1, pp.79-90.

[15] Goddard M. , Hauck K., Smith P.C. (2006) Priority setting in health - a political economy perspective *Health Economics*, online version

[16] Grant, D. (2009) Physician financial incentives and cesarean delivery: new conclusions from the healthcare cost and utilization project. *Journal of health economics*, 28(1), pp.244-250.

[17] Gruber J., Owings M. (1996) Physician financial incentives and cesarean section delivery, *RAND Journal of Economics*, 27 (1), pp. 99-123.

[18] Gruber, J., Kim J., and Mayzlin D. (1999). Physician Fees and Procedure Intensity: The Case of Cesarean Delivery *Journal of Health Economics*, 18 (4), pp. 473-490.

[19] Grytten, J., Monkerud, L., Sorensen, R. (2012). Adoption of diagnostic technology and variation in caesarean section rates: a test of the practice style hypothesis in Norway. *Health services research*, 47(6), pp.2169-2189.

[20] Grytten, J., Skau I., Sorensen R. (2011). Do Expert Patients Get Better Treatment than Others? Agency Discrimination and Statistical Discrimination in Obstetrics. *Journal of Health Economics*, 30 (1), pp. 163-180.

[21] Grytten, J., Skau, I., Sorensen, R. (2013). Do mothers decide? The impact of preferences in healthcare. *Journal of Human Resources*, 48(1), pp. 142-168.

[22] Hafsteinsdottir E., Siciliani L. (2010). DRG prospective payment systems: refine or not refine? *Health Economics*, 19, pp. 1226-1239

[23] Janulevicuite J., Askildsen J., Kaarboe O., Siciliani L.(2016) How Do Hospitals Respond To Price Changes? Evidence from Norway. *Health Economics*, 25, pp. 620–636.

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

[25] Lo J. (2003) Patients' attitudes vs. physicians' determination: implications for cesarean sections. *Social Science and Medicine*, 57(1), pp.91-96.

[26] Lo J. (2008) Financial incentives do not always work - an example of cesarean sections in Taiwan. *Health policy*, 88(1), pp.121-129.

[27] Milcent C., Rochut J. (2009) Tarification hospitalière et pratique mèdicale La pratique de la césarienne en France *Revue èconomique* , 60, pp. 489-506.

[28] Niskanen, W.A. (1971) Bureaucracy and Representative Government (Chicago: Aldine, Atherton).

[29] Niskanen, W.A. (1975) Bureaucrats and Politicians Journal of Law and Economics, Vol. 18, Issue 3, pp.617-43.

[30] Papanicolas I., McGuire A. (2015). Do financial incentives trump clinical guidance? Hip Replacement in England and Scotland, *Journal of Health Economics*, 44 (2015), pp.25-36

[31] Rochaix, L., Jacobzone, S. (1997). L'hypothèse de demande induite: un bilan économique. *Economie prévision*, 129(3-4), pp.25-36.

[32] Shurtz I. (2013) The impact of medical errors on physician behavior: Evidence from malpractice litigation. *Journal of Health Economics*, 32 (2), pp. 331-340.

[33] Siciliani, L.. (2006) Selection of treatment under prospective payment systems in the hospital sector, *Journal of Health Economics*, 25, pp.479-499.

[34] 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.

[35] Verzulli R., Fiorentini G., Lippi Bruni M., Ugolini C. (2016) Price changes in regulated healthcare markets: do public hospitals respond and how? *Health Economics*, online version

[36] 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), pp. 234-242.