Low Staffing in the Maternity Ward: Keep Calm and Call the Surgeon

Gabriel Facchini*

May, 2019

For the most recent version, visit http://gabrielfacchini.weebly.com/research

Abstract

This paper examines how workload causally affects the provision of care in a large but understudied segment of the healthcare sector – maternity wards. I use detailed patient-level administrative data on childbirth, and exploit quasi-random assignment of unscheduled patients to different staffing ratios. Indeed, nature distributes births and associated problems uniformly across time and this is still the case for unscheduled patients attempting to have a vaginal birth. I find that patients who at admission observe a higher ratio of patients-to-midwives are more likely to get a C-section. I provide suggestive evidence of two mechanisms at play behind this increase. First, the raise in workload lowers the quality of care and this leads to more patients needing a C-section. Consistent with this, patients who were admitted with an ex-ante higher risk of C-section are more affected by workload. Second, the raise in C-sections is driven by single women, while married women are largely unaffected by workload. Because C-sections are faster than vaginal deliveries, and single women are -on average- more likely to be alone in the delivery room, the medical team may find optimal to do more C-sections on single patients when time-constrained. This study provides evidence that midwives' workload is yet another factor which triggers physician-induced-demand for C-sections.

JEL Classification: D82, H42, H51, I18, J13, J16, J22

Keywords: cesarean section, capacity utilization, workload, midwives, physician induced demand, bargaining power

*Department of Applied Economics, Universitat Autònoma de Barcelona, Spain. Email: gabriel.facchini@uab.cat.
I would like to extend my gratitude to Andrea Ichino and Jérôme Adda for their invaluable support and advice throughout this research project. Thanks also to Juan Dolado, Matilde Machado, Joseph Doyle, Libertad González, Daniela Iorio and Matthew Neidell for their help and feedback on this project. I am extremely grateful to Carlo Dani, Simone Pratesi, Federico Mecacci, Lucia Pasquini, Franca Rusconi and Luigi Gagliardi for their clinical expertise, and to Tommaso Lanis and Francesca Superbi for their assistance in accessing the data. All errors remain my own.

1 Introduction

Over the last decades health care systems in developed countries have been under constant pressure to reduce costs, despite facing an increasing demand for health care services. In order to avoid a trade-off between cutting down on costs and a negative impact on patient's health outcomes, experts currently point towards the reduction of waste as the best way to go (Berwick and Hackbarth, 2012). Among the several sources of waste, two widely cited ones are: (i) failures of care delivery -the lack of adoption of known best practices-, and (ii) overtreatment -the carrying out of treatments that cannot possibly improve patients' health. These two sources of waste are particularly salient in maternity ward settings.

The presence of midwives -as opposed to physicians- in assisting birth speaks to the first point. Even though there is consensus about the crucial role of midwives in guaranteeing access to quality services and better health outcomes for both mother and child (Bender et al., 2014), the WHO has already expressed concern about a worldwide "shortage of midwives" (Büscher et al., 2009).[1]

Meanwhile, cesarean sections (C-sections) are the most commonly performed operating room procedures in the US (McDermott et al., 2017) and rank highly among greatly overused interventions. While the international healthcare community considers an ideal rate of C-sections to be between 10-15% (World Health Organization, 1985), country average rates in Europe vary from as low as 15.6% in The Netherlands to as high as 36.8% in Italy (OECD data 2012). This large variation in C-section rates has lead governments and clinicians to express their concern about its possible overuse and potential negative impact on patients health (World Health Organization, 2015). Indeed, C-sections not only cost more than vaginal deliveries, but they also imply higher risks for both the mother and the infant (Deneux-Tharaux et al., 2006; Gregory et al., 2012; Curtin et al., 2015; Costa-Ramón et al., 2018) and, according to a growing medical literature, are associated to lower long-term outcomes of children's health.[2] In addition, because vaginal delivery after a C-section (VBAC) is very unlikely, one C-section sets a path dependency for more C-sections in future births.[3] There is also evidence that women who follow a C-section are more likely to have less children (Norberg and Pantano, 2016) -something that is particularly alarming in developed countries with already low fertility rates- and to develop postnatal depression (Tonei, 2019).

In light of these concerns, a natural question is whether a situation of low staffing can result in more unnecessary C-sections being performed. This can happen either as a direct consequence of

[1] This has also been mentioned by midwifery colleges. In a 2015 report, The Royal College of Midwives estimates that the UK "...needs 2,600 more midwives to be able to cope with the number of births the country is experiencing...". The Federal Association of Midwives of Spain (FAME) has as main objective to address the shortage of midwives in the health care system. The president of the Italian Midwifery Association recently stated that "...there is a shortage of midwives. Too few to guaranty the proper level of care that other European Countries have".

[2] Infants born by C-section are not exposed to the maternal bacteria of the birth canal and as a consequence have different intestinal bacteria, which can affect their immune system and other important processes. For a meta-analysis of this literature see (Blustein and Liu, 2015). There is also an emerging economic literature looking at the causal link between C-sections and children's health. They find that babies born by C-section are more likely to visit the ED within their first year of life (Card et al., 2018) and more likely to be hospitalized for asthma syndromes up to their adolescence (Jachetta, 2015) with respect to babies born vaginally.

[3] VBAC rate is only 8.3% in the US, and 12% in Italy.

high workload -with midwives devoting less time to each patient, therefore rising the probability of complications that lead to surgery- or because physicians find it optimal to induce some patients towards a C-section independently of their health status. There are two differences between C-sections and vaginal births that make a shift in delivery method a potential way of reducing midwives' workload. First, C-sections take less time than vaginal births since there is no need to wait for the appropriate dilation of the cervix. Second, C-sections are less intensive in midwives' time and more intensive in physicians' time.[4]

To identify the causal effect of workload on the probability of getting a C-section, I exploit a simple natural experiment that varies the level of workload each patient is exposed to in an arguably random manner. This is the case for all unscheduled patients, those that follow the natural course of birth and only go to the hospital once labor has already started and/or their water has broken. Because their precise day and time of admission to the hospital is unknown, the maternity's workload level upon being admitted is orthogonal to the patient's demographic and health characteristics (and to their ex-ante probability of delivering by C-section). In addition, my measure of workload varies with both the number of patients who arrived before and the number of midwives on duty in the delivery room, two variables that are unknown to the incoming patient. I provide evidence that the time of admission is uniformly distributed across hours of the day and days of the week only for these unscheduled patients. Moreover, I show that workload is not correlated with commonly known predictors of C-sections like birth-weight, gender, being a first-time mother, age, and ER visits during pregnancy.

The data for this paper comes from a census of births from a large public hospital in Italy for the period 2011-2014. Three features of this data set make it well suited for tackling the issue at hand. First, Italian birth certificates have precise information on delivery method, allowing the identification of patients attempting a vaginal delivery and those who have a scheduled C-section.[5] This information is crucial to exclude scheduled patients from the analysis sample since their time of admission to the hospital has been arranged in advance with the physician. Second, using patient's ID, each certificate was linked with time stamps containing the exact time of admission and discharge. I use this information to compute the actual number of patients waiting in the delivery room at each point in time. Finally, this is complemented with data on the number of midwives scheduled for duty by month, day of the week and shift. This allows me to adjust demand by supply-side factors. My measure of workload is then calculated as the ratio between the observed number of unscheduled patients waiting to give birth at the time the indexed patient is admitted and the number of midwives scheduled to be in the delivery room.

The analysis is divided in two parts. First, I show that there is a non-linear relationship between midwives' workload and delivery method: patients admitted when the ratio of patients-to-midwives is above the 20th percentile are 1.9-2.7 percentage points more likely to give birth by C-section. This

[4]Although recovery time is longer after a C-section relative to a vaginal birth, post-birth care is administered by another staff-team outside the delivery room in the post-natal unit.

[5]Even though this may seem as very straight forward information, vital statistics in the US didn't report any information on 'trial of labor' until 2004, making it impossible to separate scheduled from unscheduled patients before that date.

means that about 16 percent of unscheduled C-sections are a consequence of low midwifery staffing. Moreover, this effect is only present for patients who are admitted during weekdays (when C-section staff resources are higher) and for first-time mothers. There are two reasons for this heterogeneous effect by birth order. First, women who already had a previous birth and are attempting a vaginal delivery have almost surely had a vaginal delivery the time before (path dependence). Second, the time between the onset of labor and birth is significantly longer for first pregnancies, which makes them more midwife-time intensive. In addition, I fail to find any significant effects of workload on measures of infant and maternal morbidity. When looking at measures of treatment intensity, I find that workload doesn't affect the probability of having an instrumented birth, meaning the rise in C-section mirrors the drop in natural vaginal births.

The second part of the analysis looks at potential mechanisms behind the change in delivery method. One possibility is that, in situations with a high ratio of patients-to-midwives, the time dedicated to each patient is lower and the quality of care inappropriate, eventually resulting in the need for C-section. If that is true, then patients with marginally lower health should be more affected. In order to test this hypothesis, two types of patients are compared: a low-health type, formed by those patients who had an emergency visit during their pregnancy or whose newborn had an extreme weight at birth, and a high-health type, with all remaining patients. I find that the gap in the probability of having a C-section for these two groups widens with rising workload.

Another factor that can explain the increase in C-sections with workload is the presence of physician induced demand (PID). Because C-sections are faster than vaginal births, when time constrained, physicians may decide to put some patients through surgery and reduce the midwives' workload. Within the agency discrimination framework, physicians are more likely to practice an unnecessary surgery on patients with lower bargaining power. I tests for the presence of agency using patient's civil status, that is, comparing single and married patients. The assumption is that single patients are -on average- more likely to be alone in the delivery room, reducing the physician's cost of inducing a C-section. Indeed, the probability of delivering by C-section between these two groups is virtually the same for low-workload levels. However, single patients observe a raise in the probability of C-section along the raise in workload.

1.1 Related literature

This paper contributes to a large literature looking at the effects of workload (measured as patient-to-staff ratios) on patient health outcomes. Most prior studies use exogenous variation in stemming from new legislation raising the mandatory minimum staffing ratios (Tong, 2011; Cook et al., 2012; Lin, 2014; Matsudaira, 2014; Chen and Grabowski, 2015). One exception is Evans and Kim (2006), who use variation in the number of admissions in the two days after as an exogenous shock to the effective staff level. These papers find mixed results in terms of the effect of workload on quality of care. My contribution to this literature is twofold. First, previous studies using legislative changes exploit an aggregate, permanent and positive variation in capacity (Harris et al., 2019), while I use *temporary* and *stochastic* changes in the staffing ratios that vary at the *patient* level. Second, to the best of my knowledge, I provide the first causal estimates of the effect of workload on health outcomes in the maternity ward setting.

In contrast, there is little empirical work on the effects of workload on the provision of health care (i.e. the channel through which workload may affect health outcomes). There are, however, several recent and concurrent papers tackling this issue using health-care provider administrative data. Harris et al. (2019) look at exogenous changes in nurse capacity in five public clinics in Tennessee and find that providers value sufficient time spent with patients over seeing more patients. Alkalay et al. (2018) use data from eleven primary care clinics in Israel and exploit the absence of colleagues as a source of exogenous variation in physician workload, and find that referrals to specialists and lab tests go down with shorter visits. Neprash (2016) finds that primary care physicians perform fewer procedures and record fewer diagnoses for appointments that start later due to physicians being behind schedule. Finally, Freedman et al. (2018) use unexpected schedule changes as variation in primary care physician's time pressure and find that higher pressure reduces the number of diagnoses recorded and increases both scheduled and unscheduled follow-up care. My paper differs from these papers in three important ways. First, all these papers focus on primary care, an environment where time pressure is among the lowest in health care. Meanwhile, a maternity-ward's demand is mainly driven by unscheduled patients that need assistance within a relatively short span of time.[6] Second, with the exception of Freedman et al. (2018), these papers use shocks to provider availability and time (supply side) while I exploit stochastic variation in patients arrivals driven by nature (demand side). Finally, with the exception of Alkalay et al. (2018), all these papers use data from the US, a context where financial incentives have been shown to significantly affect health-care provision (Clemens and Gottlieb, 2014; Ho and Pakes, 2014).[7] On the contrary, physicians in Italy perceive a fix salary, and non-financial incentives may play a bigger role in the supply-driven variation in treatment.[8]

A large related literature studies how physicians may induce a patient's demand against the physician's interpretation of the best interests of the patient (Johnson, 2014). Many of these papers have focused on the maternity ward set-up and the decision of delivery method due to its discretionary nature. One strand of this literature assumes the presence of information asymmetry and exploits variation in a wide range of physicians' incentives to induce. Starting from Gruber and Owings (1996) where they use a fall in physician's as a trigger for more C-sections, to other incentives like relative prices between C-sections and vaginal deliveries (Gruber et al., 1999; Alexander et al., 2013; Allin et al., 2015), defensive medicine (Currie and MacLeod, 2008; Dranove and Watanabe,

[6]In a recent working paper, de Elejalde and Giolito (2019) use a policy change in co-payment in Chile that incentivizes patients to move to private hospitals and observe that this increase in the number of patients at private hospitals is linked to a rise in the probability of having a C-section. Their work differs crucially from mine in that they use an 'expected' measure of workload at the 'hospital' level, and show that hospitals increase the number of scheduled C-sections in anticipation of a positive demand shock.

[7]Although this is not exclusive to the US. Brekke et al. (2017) find that physicians in Norway respond to fee changes.

[8]Recent work suggest that non-financial incentives may also be important for the US market. Using survey data linked to fee-for-service Medicare expenditures, Cutler et al. (2013) find that "physicians' responsiveness to financial factors play a relatively small role in explaining equilibrium variations in utilization patterns in our context".

2009; Bertoli and Grembi, 2019), and physician's scheduling convenience (Lefèvre, 2014).[9] Unlike previous studies, my findings suggest that physicians may have incentives to perform medically unnecessary C-sections for reasons other than personal gain. Indeed, if physicians cannot adjust the price nor the quantity, they may resort to lowering the quality of treatment. In the maternity ward setting, a patient's unnecessary C-section can alleviate midwives' workload and allow them to better attend other patients.

Another strand of the literature on PID uses heterogeneity in information asymmetry to explain the variation in C-section levels. Two recent papers compare the treatment received by expert and non-expert patients in the maternity ward set up. Grytten et al. (2011) observe that, in an institutional context with incentives to reduce C-sections, non-expert patients get less C-sections than expert patients. They conclude that a model of statistical discrimination (expert patients are better at communicating with the physician) explains their results better than one of agency discrimination (the physician influences the diagnosis and treatment for non-expert patients). On the contrary, Johnson and Rehavi (2016) find evidence that physicians are more likely to exploit the information asymmetry when it is profitable (agency discrimination). They do so by comparing physician patients with non-physician patients, in settings with and without financial incentives to perform C-sections. My paper differs from these two papers in that I refrain from exploiting variation in information asymmetry. Instead, I use variation in the presence of another person in the delivery room caring for the patient's interest (husband) and raising the physician's cost of inducing an unnecessary treatment.

The remainder of this paper is organized as follows: Section 2 describes the clinical and institutional setting. Section 3 discusses the identification strategy and describes the data. Section 4 reports the results, and Section 5 concludes.

2 Clinical and Institutional Setting

Maternity wards receive two types of patients: scheduled and unscheduled. The former includes patients admitted for an elective C-section and those who will be induced.[10] For patients following an elective C-sections the date of delivery is set in advance, and there is no possibility for changing delivery method (unless the mother goes into labor before). These pregnancies typically present some health condition that constitute a risk for the mother and/or the baby if delivered vaginally. Similarly, induced patients already know in advance the date they will be induced but, although they will attempt a vaginal delivery, the physician may still decide to change delivery method on the way if considered necessary.

The remaining patients, those attempting to follow the natural course of labor and vaginal delivery, are the main focus of this study. For these patients the process starts with frequent contractions and/or because they believe their water has broken (spontaneous onset of labor). Once the mother arrives to the hospital she is evaluated and if in active labor, she is admitted

[9] For an extensive review of the literature on PID in the maternity ward setting see Allin et al. (2015).

[10] Most inducements are performed on pregnancies that have past their due date and still haven't started labor.

into the labor and delivery room and assigned a gynecologist and a midwife. If everything goes as planned and the patient is able to have a vaginal delivery, the midwife will be the one helping her through out the whole process. Nevertheless, during labor there are several medical conditions that can emerge and complicate a vaginal birth, putting in danger the health of the infant and/or the mother. Under these circumstances, the midwife and gynecologist may decide to recommend to have a C-section instead.

More importantly, the presence of some of these medical conditions depends heavily on the subjective opinion of the gynecologist.[11] This gray area -or asymmetry of information- on when a C-section is necessary gives the physician more room to recommend surgery to the patient, even when not medically needed.

The maternity unit analyzed in this paper is part of one large teaching hospital in Italy. The staff working in the delivery room are paid a fixed salary, meaning they have no personal financial incentive to recommend any particular treatment. On the other hand, hospitals are reimbursed depending on a DRG (Diagnosed-related group) tariff system, which in general gives a higher reward for a C-section than for a vaginal delivery.[12]

3 Empirical Methodology

3.1 A natural experiment

An ideal experiment to test for an effect of workload on patients' delivery method would imply assigning parturient women randomly between two different hospital types: a first one with already a large number of patients and a second type, identical to the first, but with few patients and hence ready to focus entirely on the coming patient. For obvious reasons this is not possible to implement in practice.

This paper focuses on patients who attempt to have a vaginal delivery, and uses plausibly exogenous variability in the number of patients and midwives present at admission to causally identify the impact of workload on delivery method. For the majority of births, the time of arrival is unknown to the hospital beforehand. In the same way, the level of capacity utilization of the maternity ward in a given point in time is unknown for future patients until they reach the hospital. For these patients, their pre-admission probability of developing a complication and delivering by C-section is orthogonal to the level of crowding at the hospital.

The study sample includes all births that, up to the point of arriving to the hospital, followed the "natural" course of pregnancy and labor. This means leaving out all scheduled deliveries where the physician decided, together with the patient, the date when the birth should take place. This type of patients are those who had an elective C-section or who were pharmaceutically induced to start labor.[13]

[11] Two of these more 'subjective' conditions are dystocia (abnormally slow labor) and fetal distress

[12] For a deeper discussion on the Italian Health System please see Francese et al. (2014).

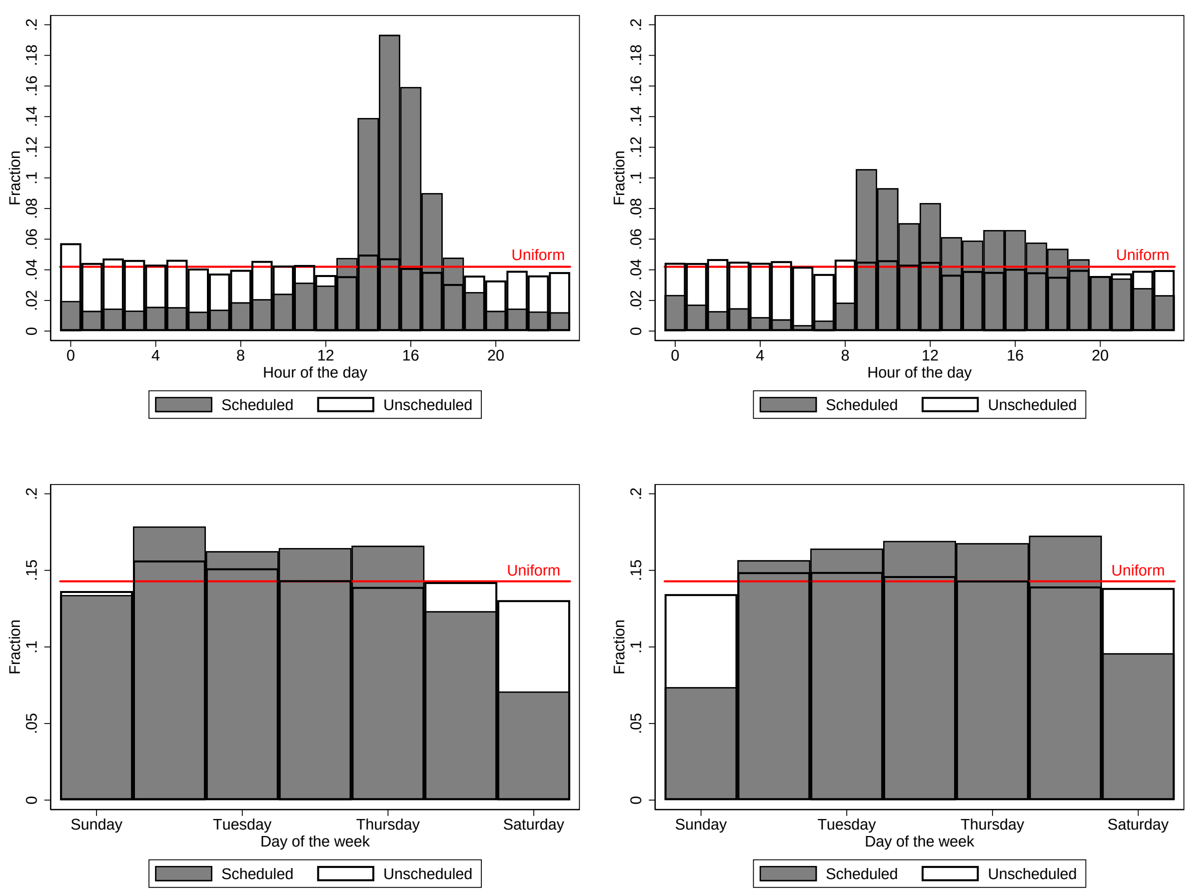
[13] For more evidence supporting the criteria for selecting the working sample see Appendix A.

The left column of Figure 3.1 shows the distribution of admissions by hour of the day and day of the week. The right column does the same for births. Both are estimated for scheduled and unscheduled patients for comparison. We can immediately see that admissions of scheduled patients are concentrated in the afternoon, while births start at 9 a.m. and become less and less frequent as the day goes by. Instead, both admissions and births for unscheduled patients are very close to a uniform distribution within a day. Similarly, when looking at the distribution by days of the week, unscheduled patients are randomly distributed while scheduled patients are less common to be admitted on Saturdays, and less likely to have surgery on Sundays and Saturdays. In Appendix A, I repeat the exercise disaggregating scheduled patients in scheduled C-sections and induced deliveries. Results are qualitatively the same, with both categories showing non-uniform distributions across days and hours.

Figure 3.1: Distribution of admissions and births.

(a) Admission

(b) Birth



Notes: Scheduled patients include elective C-sections and patients who were induced into labor. Unscheduled patients are those who started labor and attempt a vaginal delivery.

3.2 Data

Previous studies looking at newborns' health typically use anonymous birth certificates which are publicly available for many countries and for long periods of time. However, these data sets commonly lack information on key variables needed for a rigorous study of staffing levels, namely, the exact date and time of admission of patients (demand side) and the number of staff available (supply side) in each delivery unit.

This study utilizes data from the Maternity Department of the Azienda Ospedaliero Universitaria Careggi (AOUC) for the years 2011 through 2014. This is the biggest hospital in the Province of Florence with more than 3,000 deliveries per year. The primary databases used are two: (i) birth certificates (Certificato di assistenza al parto); and (ii) hospital admissions (Scheda di Dimissione Ospedaliera). Birth certificates constitute a census of all births that took place in the hospital in this period. They contain information on mother characteristics (e.g. community of residence, education, civil status, age, previous deliveries, etc.), pregnancy characteristics (e.g. weeks of gestation, controls, assisted reproduction, etc.) and birth characteristics (e.g. time of birth, type of labor, attendant, place, weight of the baby etc.). The administrative hospital admission data provides information on the time of admission and time of discharge for each patient. Using unique mother-pregnancy identifiers, both databases can be linked.

The aforementioned data on patients is complemented with information on the level of staff scheduled to be present in the delivery room each month, day of the week and shift combination. Note that this is not the effective level of staff present at each point in time but the staffing rule of the delivery unit. Anecdotal evidence suggests that deviations from planned levels are rare, even because the hospital calls in someone else when an employee misses a shift.

There were approximately 11,359 singleton births at this hospital in the sample period. After excluding all scheduled patients the sample goes down to 6,479 births. Finally, after dropping observations with missing time of admission, maternal age, education, birth order, weight and prenatal visits, and patients arriving the first and last 15 days in the sample, the number of observations in the working sample is about 6,142. The models described below are fitted to this sample.

Table 3.1 summarizes the variables used in the analysis. The first column corresponds to the whole sample. Most of the patients who attempt a vaginal delivery succeeded. Only about 10.4% had an unscheduled C-section. Patients are on average 32.5 years old, 31% have a university degree, and 64% are married. There are few cases with bad outcomes: only 4.7% have a 5-minute APGAR score below 9, about 6% are born prematurely, and 5% are born weighting less(more) than 2,500(4,000) grams. Columns two and three report statistics for patients with a low (healthy) and high (unhealthy) ex-ante risk of C-section respectively. Columns four and five do the same by civil status. By construction, healthy patients are less likely to give birth by C-section, to use the intensive care unit, and to have an APGAR score below 9. They are also more likely to be first-time mothers and more likely to have a university degree. Finally, married patients are more likely to have a university degree and less likely to be first-time mothers. Married patients have a slightly lower probability of bad health outcomes across the board.

Table 3.1: Descriptive statistics: workload

	(1) All	(2) Healthy	(3) Unhealthy	(4) Married	(5) Single
Outcomes					
% vaginal births	89.6	90.8	85.3	90.8	88.0
% unscheduled C-sections	10.4	9.2	14.7	9.2	12.0
% operative births	8.8	9.0	7.9	8.2	9.8
% transfered to neonatal ICU	7.2	4.0	19.4	6.2	8.3
% lack of skin-to-skin contact	18.0	14.3	32.2	16.8	19.1
% non-exclusive breastfeeding	32.2	30.5	40.0	30.3	34.9
% Apgar score<9	4.7	2.9	11.6	4.1	5.5
Mother characteristics					
(mean) age	32.5	32.6	32.4	32.8	31.9
% university degree	30.9	31.7	28.1	34.0	26.5
% married	63.6	63.9	62.8	100.0	0.0
% first-time mothers	42.1	43.0	38.3	37.1	51.2
Pregnancy characteristics					
% preterm (<37 wofg)	6.0	2.7	18.8	5.7	6.3
% with at-least 1 ER visit	12.6	0.0	60.3	12.5	13.3
Infant characteristics					
% male	51.6	50.9	54.3	52.1	51.3
(mean) weight in grams	3,265.6	3,302.2	3,127.2	3,276.9	3,249.1
% low birthweight (< 2,500gr)	5.1	0.0	24.2	4.9	5.2
% high birthweight (> 4,000gr)	5.0	0.0	24.0	5.2	4.7
Observations	6142	4856	1286	3477	1986

Notes: RPM is the ratio of patients to midwives at admission. Patients is the number of patients waiting at admission who attempt to have a vaginal birth. Morning shift: 7am to 1pm. Afternoon shift: 1pm to 7pm. Night shift: 7pm to 7am. At the bottom, the values of the 20th and 80th percentile of the RPM distribution.

3.3 An exogenous measure of midwives' workload

A good measure of workload contains information on both the number of patients and hospital's staff. For the maternity ward setting I use the ratio of patients to midwives (RPM) in the delivery room.[14] The richness of the data in hand allows me to construct a very precise measure of the number of parturient women in the maternity ward at any point in time and to differentiate between those waiting to give birth and those in postpartum. There are yet two decisions to be taken regarding the moment at which this ratio is calculated and the type of patients to include in the numerator. On the former, because the median patient stays 7 hours in the delivery room between admission and birth, it is not obvious at what time to measure the level of staffing. The two most straightforward options are at the time of admission and at the time of delivery. The last one has the advantage

[14] One drawback of this measure is that it constraints the coefficient of interest due to the simultaneous variations in numerator and denominator. The fact that my preferred model specification uses fixed effects by shift and day-of-the-week means that all the variation used for the estimation comes solely from fluctuations in the numerator, alleviating this issue.

of measuring staff when needed the most, meaning, when the mother needs help to give birth. The problem with this option is that, given that physicians can rush a delivery (e.g. by doing a C-section), the level of staffing at time of birth can be endogenously determined. On the other hand, even though the level of staffing at time of admission can be relatively less relevant, it is indeed an exogenous shock. For these reasons I will use the ratio of patients to midwives calculated at the time of admission of each patient.[15]

On the second issue, it is important to decide which patients are included in this measure of staffing. The first option would be to include all patients (regardless of whether they are scheduled or induced). Nevertheless, since the outcome of interest is the probability of C-section, counting elective C-sections in the measure of staffing would make it biased. To see this, note that when there are more elective C-sections there are also more gynecologists ready to perform them. Incorporating elective C-sections in the numerator would not only include a demand side but also a change in the supply of physicians who can perform C-sections. For this reason, I include in the numerator all patients but those already scheduled to give birth by C-section.[16]

More specifically, the workload observed by a patient admitted at time t is defined as

$$\mathrm{RPM}_t = \frac{\mathrm{PVB}_t}{\mathrm{MW}_t} \tag{1}$$

where PVB is the number of patients waiting to attempt a vaginal birth, and MW is the number of midwives scheduled to be present in the delivery room.

Table 3.2 shows the distribution for the RPM and the corresponding values for number of patients and midwives. The ratio is unimodal and slightly skewed to the right (see also Figure 3.2a). At the median, there are 2.3 patients for every midwife in the delivery room. The 25th and 75th percentiles are 26% (below) and 30% (above) the median, respectively. Finally, Figure 3.2b shows the mean and one standard deviation intervals of the ratio by hour of admission. This is the variation I use in my most demanding specification including hour-of-admission fixed effects. Note that shifts later in the day have higher values of the ratio, meaning, higher workload. Indeed, there are two discrete jumps in the ratio at 1 p.m. and at 7 p.m.. Remember that the distribution of patients is rather uniform across the day, hence this upward shift in the ratio comes exclusively from a lower supply (less midwives present).

[15] In the following section I perform several robustness check measuring staff levels at different points in time during a patients stay and discuss the results.

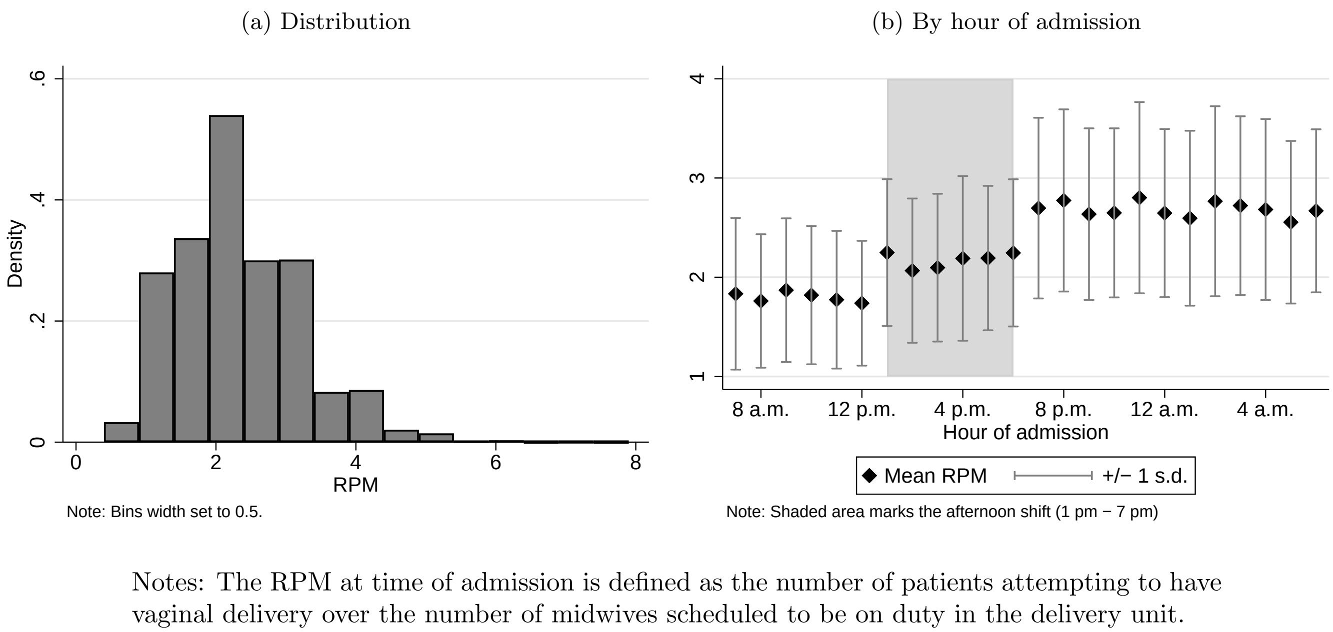
[16] Note that this is not the same sample as the study sample because it also includes induced deliveries. Those are not at risk of contaminating the measure because they will still attempt a vaginal delivery, and need a midwife to help them.

Table 3.2: Descriptive statistics: workload

	Ratio of patients to midwives	Patients waiting	Midwives scheduled
p1	0.8	2.0	3.0
p5	1.0	3.0	3.0
p25	1.7	5.0	3.0
p50	2.3	7.0	3.0
p75	3.0	9.0	4.0
p95	4.0	12.0	5.0
p99	4.8	14.0	5.0
mean	2.3	7.1	3.6
sd	0.9	2.7	0.8
20th Percentile	1.6		
80th Percentile	3		

Notes: At the bottom, the values of the 20th and 80th percentile of the ratio of patients-to-midwives. There are 5 midwives in the morning shift on weekdays, 4 midwives in the afternoon shifts on weekdays and the morning and afternoon shifts on Saturdays, and 3 midwives in the night shift on weekdays and Saturdays and all Sunday shifts.

Figure 3.2: Distribution of the ratio of patients to midwives



Notes: The RPM at time of admission is defined as the number of patients attempting to have vaginal delivery over the number of midwives scheduled to be on duty in the delivery unit.

3.4 Econometric specification

The first part of the analysis estimates OLS regressions of a binary indicator for C-section on the treatment variable along with demographic and clinical controls. A simple reduced-form linear probability model of the following type is used:[17]

$$\mathrm{y}_{it} = \alpha + \beta\, \mathrm{RPM}_t + \theta\, \mathrm{X}_i + \theta_t + \epsilon_{it} \tag{2}$$

where y_i is a dummy variable indicating whether birth i had an unscheduled C-section, and RPM_t is the ratio of patients-to-midwives observed at admission time t -constructed as described in Section 3.3. X_i contains individual-level control variables of mother and pregnancy characteristics[18]. To control for seasonal and secular variation in outcomes, I include additive fixed effects for shift, day of the week, month, and year of admission. Since most supply side changes in the maternity ward take place between shifts and days of the week, I perform robustness checks using fixed effects for day-of-the-week times shift, and using hour of admission fixed effects. Results are virtually the same. β is the coefficient of interest. As discussed above, if physicians are more likely to perform a C-section when the RPM is high, then we expect β to be positive.

Two models are estimated for the probability of delivering by C-section. First, I use the RPM added linearly to the model. Because there can be non linear effects between workload and delivery method, for the second model I split the sample in three categories based on the distribution of the RPM: low, medium, and high. All those observations with a ratio below the 20^{th} percentile are in the first group. These are cases of *low* workload. The second group includes those observations between the 20^{th} and 80^{th} percentiles, and are categorized as cases with *medium* workload. Finally, the last group consists of all those above the 80^{th} percentile, and I will refer to them as *high* workload cases. The cut offs for these groups are reported in the bottom of Table 3.2. In these models, the lowest quintile (low workload) is considered the reference group.[19]

The key identifying assumption of my empirical model is that changes in workload are uncorrelated with unobserved patient characteristics. While this is not directly testable, we provide suggestive evidence in Table 3.3. Panel A shows the coefficients of a regression of each of the pre-treatment controls on the RPM, while Panel B uses the non-linear specification of RPM. I find that none of these coefficients are statistically significant, which empathizes the strength of the natural experiment used.

The second part of the analysis aims at understanding the mechanisms through which workload affects the choice of delivery method. Two hypothesis are tested. First, it could be the case that

[17]A probit model was also estimated assuming a normal distribution of the error term and results virtually the same (See Table C.1).

[18]These include: a dummy for whether the mother is above 34 years old, a dummy for whether the mother has a university degree, a dummy for whether this is her first pregnancy, a dummy for whether the infant is a male, a dummy for whether is a pre-term birth (below 37 weeks of gestation), a dummy for whether the baby is born with low weight (less than 2,500 grams), and a dummy for whether the mother had at least one emergency check up during pregnancy.

[19]See Appendix B for a more detailed discussion on model selection, where models of different polynomial degrees and categorical definitions of workload are tested.

Table 3.3: Balance table

	(1) Age<30	(2) University	(3) Married	(4) Primipara	(5) Preterm	(6) ER visit	(7) Male	(8) Low bw	(9) High bw
Panel (A)									
RPM	-0.003	0.003	-0.003	0.004	-0.001	-0.005	-0.009	-0.000	0.005
	(0.008)	(0.008)	(0.008)	(0.008)	(0.004)	(0.005)	(0.008)	(0.004)	(0.004)
Panel (B)									
Med. RPM	-0.026	-0.011	-0.008	0.007	-0.001	0.007	-0.017	-0.006	-0.004
	(0.017)	(0.017)	(0.018)	(0.018)	(0.009)	(0.012)	(0.018)	(0.008)	(0.008)
High RPM	-0.005	0.009	-0.001	0.000	-0.004	-0.020	-0.017	0.001	0.008
	(0.023)	(0.023)	(0.024)	(0.024)	(0.011)	(0.015)	(0.024)	(0.011)	(0.011)
Observations	6142	6142	5463	6142	6142	6142	6142	6142	6142
Shift FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column regresses a different maternal or pregnancy control on a measure of workload. All models contain fixed effects for year, month and day of the week of admission. Panel (A) shows coefficients for regressions using as treatment a linear form of workload (RPM). Panel (B) shows coefficients for regressions using as treatment two dummy variables for workload being between the 20th and 80th percentile of the RPM, or above. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

high workload results in less midwifery time available for each patient. Under this scenario, patients who were admitted with an already higher risk of C-section (and presumably need more care) will be the most affected. At higher ratios, the probability of C-section should rise faster for this group than for other patients -all else constant- due to their pre-treatment lower health. These patients are identified as those with extreme birthweight (below 2,500 grams or above 4,000 grams) or with at least one emergency visit to a hospital during pregnancy.

The second hypotheses has to do with agency discrimination. When resources are constrained, e.g. high RPM, physicians may see optimal to shift some patients to the operative theater to get a C-section. This would reduce the workload on midwives by reducing the number of patients waiting in the delivery room. Because patients are heterogenous, physicians will find it easier to offer this treatment to some patients than others. I use patient's civil status as a proxy for bargaining power assuming that, on average, single women are more likely to be alone in the delivery room.[20] In those cases, the physician only needs to convince one person about the change in procedure -not to mention the patient is in labor and in a lot of pain, which makes it harder to analyze the pros and cons of each alternative.

To test whether physicians' treatment covaries with the patients' characteristics above mentioned, I estimate the following regression:

$$\mathrm{y}_{it} = \alpha + \beta_1\,\mathrm{RPM}_t + \beta_2\,\mathrm{RPM}_t \times \mathrm{D}_i + \beta_3\,\mathrm{D}_i + \theta\,\mathrm{X}_i + \theta_t + \epsilon_{it} \tag{3}$$

where D_i is either one of two variables: an indicator for whether the patient is healthy (low ex-ante probability of C-section), or whether she is married. The remaining variables are defined as in equation (2), adding civil status as a control. I expect healthy and married patients to be less affected by a high RPM, hence, a negative β_2 in both cases.

4 Results

4.1 Provider response to workload

Table 4.1 presents the results of equation (2). Columns one and two show the estimates of β for the linear model with and without shift fixed effects respectively, while columns three and four do the same for the non-linear model that includes dummy variables for different levels of workload. All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4.

Results for the RPM added as a linear variable are not statistically significant. However, there is evidence of a non-linear effect of workload on delivery method. Relative to the low-workload case, the probability of C-section raises 1.9 percentage points with medium workload, and an extra 0.8 percentage points with high workload. This implies an 18% and 26% raise in unscheduled C-sections with respect the mean (10.4%). A similar pattern emerges when using quadratic polynomial for RPM instead of dummies, with a positive linear term and a negative and smaller quadratic term

[20] For a single woman in Tuscany, the odds of being alone in the delivery room are 1.25 times larger than the odds for a married woman being alone (ARS Toscana, 2013).

Table 4.1: Effect of workload on the probability of C-section

	(1)	(2)	(3)	(4)
RPM	0.007	0.008		
	(0.005)	(0.005)		
Med. RPM			0.019**	0.019*
			(0.010)	(0.010)
High RPM			0.025*	0.027*
			(0.013)	(0.014)
Observations	6142	6142	6142	6142
Mean dep.	0.104	0.104	0.104	0.104
Shift FE		Yes		Yes

Notes: All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4. Medium RPM is a dummy variable for patients who observe a RPM between the 20^{th} and 80^{th} percentile, while High RPM is a dummy for patients who observe an RPM above the 80^{th} percentile. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

(Table B.1 in Appendix B). This suggests that the probability of C-section follows raises at a decreasing rate with workload.

Taken together, the effects of medium and high levels of workload would imply, on average, a 2.1 percentage points (16%) increase in unscheduled C-sections. This is reasonable when compared with previous studies looking at all C-sections and changes in monetary compensation. Allin et al. (2015) find that doubling the compensation for a C-section relative to a vaginal delivery increases the likelihood that a physician opts for the former by just more than 5 percentage points, all else equal. Similarly, Gruber et al. (1999) suggests that cesarean delivery rates would rise by 3.9% in response to each $100 increase in the compensation received for a C-section, all else equal.

Table 4.2 performs a battery of robustness checks of the econometric model used. Column one is my benchmark estimates from Table 4.1, where I use shift fixed effects and cluster standard errors by date of admission. Column two uses hour of admission fixed effects. This a more restrictive model where we compare patients who arrived at the same hour of the day but observed different workload levels. Estimates are virtually the same. Column three, instead, uses shift times day-of-the-week fixed effects. This model exhaust all possible combinations of shift and days that drive changes in resources. Again, results are quite similar, and if anything, slightly larger. In column four I cluster standard errors by week of admission instead of date, but again, it doesn't make a difference. Finally, column five presents results for a placebo test where workload is measured 24 hours after admission (as opposed to at the time of admission). As expected, the estimates for both coefficients are statistically and clinically insignificant.

Table 4.2: Robustness checks

	(1) Benchmark	(2) Hour	(3) Shift x Day of the week	(4) Week	(5) Placebo 24hs after
Med. RPM	0.019*	0.019*	0.021*	0.019*	0.005
	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)
High RPM	0.027*	0.027*	0.029*	0.027**	0.010
	(0.014)	(0.015)	(0.015)	(0.014)	(0.013)
Observations	6142	6142	6142	6142	6142
Fixed Effects	Shift	Hour	Shift x DoW	Shift	Shift
Cluster s.e.	Date	Date	Date	Week	Date

Notes: All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4. Column 1 is the benchmark regression. Column 2 uses hour fixed effects instead of shift fixed effects. Column 3 uses interacted shift and day-of-the-week fixed effects. Column 4 clusters standard errors at the level of the week of admission instead of date. Column 5 uses a measure of RPM 24 hours after admission of the indexed patient. Medium RPM is a dummy variable for patients who observe a RPM between the 20^{th} and 80^{th} percentile, while High RPM is a dummy for patients who observe an RPM above the 80^{th} percentile. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

4.1.1 Additional Treatment Margins

The evidence above suggests that workload shifts delivery method from vaginal birth to C-section. However, workload can affect other treatment decisions in childbirth. Two such decisions are performing an operative vaginal birth and/or an episiotomy.[21] A higher likelihood in these procedures has been linked to scarce or absent midwifery care and the presence of obstetricians or physicians instead (Hatem et al., 2008). Finally, a third margin is the use of anesthesia, and in particular, the epidural. The use of analgesics can be considered a more discretionary decision, and one is intensive in midwives' and anesthetist's time. Under time pressure, physicians may decide to reduce the use of this input.

Table 4.3 presents estimates of equation (2) using indicators for operative birth, episiotomy, any analgesic (including epidural), and epidural as dependent variables. Most coefficients are not statistically significant. There is, however, a 3.2 p.p. raise in the likelihood of having an episiotomy for high levels of workload. This may reflect a drop on quality of care when resources are the most constrained. In addition, although not statistically significant, the point estimates for receiving analgesics is always negative. This is consistent with the idea that these are a discretionary and time-intensive treatment.

[21] Operative vaginal delivery refers to a delivery in which the physician uses forceps or a vacuum device to assist the mother in transitioning the fetus to extra-uterine life. An episiotomy is a surgical cut performed at the opening of the vagina during childbirth to help a difficult delivery.

Table 4.3: Effect of workload on ancillary procedures

	(1) Operative birth	(2) Episiotomy	(3) Any analgesic	(4) Epidural
Med. RPM	-0.004	0.011	-0.027	-0.017
	(0.010)	(0.013)	(0.017)	(0.021)
High RPM	0.006	0.032*	-0.033	-0.016
	(0.014)	(0.018)	(0.023)	(0.028)
Observations	6142	6141	6142	3086
Mean dep.	0.088	0.186	0.367	0.731
Shift FE	Yes	Yes	Yes	Yes

Notes: All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4. Medium RPM is a dummy variable for patients who observe a RPM between the 20^{th} and 80^{th} percentile, while High RPM is a dummy for patients who observe an RPM above the 80^{th} percentile. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

4.1.2 Heterogeneity

Table 4.4 presents estimates dividing the sample based on birth order and day of admission. I find that workload affects delivery method only for first-time mothers. There are two possible explanations for this heterogeneous effect by birth order. First, there is a path-dependence in delivery method. This means that non-first-time mothers are positively selected (they have achieved vaginal birth) and they already know the process (lower asymmetry of information with the physician). Second, the time between the onset of labor and birth is significantly longer for first pregnancies. This means that the costs in terms of midwife time of not performing a C-section is higher for first-time mothers.

In addition, I find that only patients admitted on weekdays are more likely to get a C-section for medium levels of workload (a raise of 2.8 p.p.). This is consistent with the hypothesis of physician-induced-demand. Because scheduled C-sections are performed on weekdays, there are more resources available to perform them during those days. Previous work has shown that the availability of medical resources may positively affect their rate of utilization (Freedman, 2016). On the contrary, for high workload, point estimates are virtually the same for weekday or weekend admissions once we control for shift fixed effects. This supports the idea that high workload levels lead to lower quality of care and the need for C-sections. Unlike the case of medium workload, here the decision to change delivery method is not a discretionary one. I will explore both of these channels further in the next section.

Table 4.4: Additional estimates

	Primipara		Not-primipara		Weekday		Weekend	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Med. RPM	0.054***	0.055***	-0.004	-0.005	0.023**	0.028**	0.003	-0.002
	(0.016)	(0.017)	(0.012)	(0.013)	(0.010)	(0.012)	(0.025)	(0.025)
High RPM	0.030	0.034	0.024	0.024	0.018	0.029*	0.034	0.028
	(0.021)	(0.023)	(0.017)	(0.019)	(0.014)	(0.017)	(0.032)	(0.031)
Observations	2583	2583	3559	3559	4500	4500	1642	1642
Fixed Effects		Shift		Shift		Shift		Shift

Notes: All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4 -with the exception of columns 5 to 8 which don't include day of the week fixed effects-. Medium RPM is a dummy variable for patients who observe a RPM between the 20^{th} and 80^{th} percentile, while High RPM is a dummy for patients who observe an RPM above the 80^{th} percentile. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

4.2 Mechanisms at play

The estimates above strongly suggest that higher workload leads to more C-sections. In this section I directly test whether this effect is due to the worsening of care or physician induced demand. Under the first hypothesis, patients who are of relatively weaker health should be affected more by workload. On the other hand, if physicians decide to perform an unnecessary C-section, they will find it easier for patients with a low-bargaining power. Table 4.5 displays estimates of the coefficients in equation (3). Columns one and two use a model where the covariate interacted with workload is a dummy variable for being married, while columns three and four instead use a dummy for whether the patient is healthy (those who had not visit the ER during pregnancy and whose newborn doesn't have an extreme birth-weight).

As expected, the workload coefficients are positive and statistically significant. Single mothers admitted with medium or high levels of workload have C-section rates that are 4.1 and 4.8 percentage points higher than those admitted with low workload, respectively. The coefficients on workload (β_1) and the coefficients on the interaction between workload and married-patient (β_2) are close in magnitude and of opposite sign -although the interaction with high RPM is not statistically significant. Thus, unlike single patients, married women appear to be unaffected by workload. Furthermore, the coefficient on being married (β_3) is not statistically different from zero, suggesting that for low levels of workload there is not difference in the likelihood of getting a C-section between single and married patients.[22]

The pattern is similar when comparing the effect of workload on healthy and unhealthy patients. The workload coefficients are positive, and statistically significant for high RPM. Unhealthy

[22] Using data from the state of New Jersey (US), Currie and MacLeod (2017) observe that, conditional on C-section risk, single women are more likely to have C-sections, as are African-American and Hispanic women, less educated women, older mothers, and mothers of first born children. They do not comment on what might be behind this gap in probabilities by demographic groups, and do not separate low and high workload levels.

mothers admitted with a high RPM have C-section rates that are 7.8 percentage points higher than those admitted with low RPM. Again, the coefficient on high RPM (β_1) and the coefficient on the interaction between high RPM and healthy-patient (β_2) are close in magnitude and of opposite sign. This means that healthy patients seem to be less affected by high workload.

Table 4.5: Heterogeneity effect of workload

	Married		Healthy	
	(1)	(2)	(3)	(4)
Med. RPM	0.044**	0.041**	0.022	0.023
	(0.017)	(0.018)	(0.023)	(0.024)
High RPM	0.051**	0.048**	0.075**	0.078**
	(0.024)	(0.024)	(0.035)	(0.035)
Covariate	0.005	0.005	-0.035	-0.033
	(0.018)	(0.018)	(0.021)	(0.021)
Med. RPM x Cov.	-0.038*	-0.038*	-0.005	-0.006
	(0.021)	(0.021)	(0.025)	(0.025)
High RPM x Cov.	-0.043	-0.043	-0.061*	-0.062*
	(0.028)	(0.028)	(0.036)	(0.036)
Observations	5463	5463	6142	6142
Shift FE		Yes		Yes

Notes: Columns 1 and 2 report coefficients for workload interacted with a dummy variable for being married (base category is being single). Columns 3 and 4 report coefficients for workload interacted with a dummy variable for being healthy (base category is being unhealthy, defined as those with newborns with either an extreme birth-weight or patients with an emergency department visit during pregnancy.). All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4 -with the exception of those used in constructing the healthy dummy-. Medium RPM is a dummy variable for patients who observe a RPM between the 20^{th} and 80^{th} percentile, while High RPM is a dummy for patients who observe an RPM above the 80^{th} percentile. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

4.2.1 Other possible channels?

So far the evidence suggests that both physician-induced-demand and deterioration of care play in role in the link between workload and the likelihood of C-section. Beyond these mechanisms, there are at least two more channels that can explain it. The first and most obvious is that patients who are admitted in low and high workload levels are different. However, all tests performed in this study and previous research support the idea that, for those patients attempting a vaginal delivery, their time of arrival to the hospital is randomly distributed across time and independent of hospital conditions. Of course we cannot totally rule out that some unobserved patient characteristics make them sort into high or low levels of admission and simultaneously vary their likelihood of needing a C-section.

Another possible explanation is that those type of patients who get these 'extra' C-sections actually have a preference for this delivery method. For example, Grytten et al. (2013) find that immigrant-mothers' mode of delivery in Norway is affected by the rate of C-section in their home country -a proxy for preferences-, with a stronger effect for scheduled C-sections. However, because here I focus exclusively on unscheduled C-sections, my estimates are obtained from a sample of patients who have expressed a revealed preference for vaginal delivery by attempting labor. Hence the effect is more likely to arise from decisions made in the delivery room regarding when to stop labor and change treatment rather than from maternal preferences for C-sections. Nevertheless, because the data comes from a public hospital, patients may be denied an elective C-section -even when preferred- if there is no medical reason for it. Hence it is not possible to totally rule out that some demographic groups may be more inclined towards having a C-section and physicians internalize this when deciding which patient is sent to surgery.

4.3 The effect of workload on infant's health

The estimates above demonstrate that patients arriving in low and high workload levels receive different treatment in birth. However, one may be also interested in the consequences of high workload on infant's health after birth. In order to test this, I re-estimate equation (2) using indicators for infant's morbidity as dependent variables.

In the economics literature the most commonly studied health outcomes for births are: weight, fetal mortality and maternal mortality. Nevertheless both maternal and fetal deaths are extremely rare events (4 per 100,000 births and 2.7 per 1,000 births respectively for Italy). In the case of weight-at-birth, because treatment here is defined at the moment of admission to the hospital, it is considered a pre-defined outcome (not affected by treatment).[23]

The restricted-use version of the birth certificates in hand contains, however, other measures of morbidity. Those that occur in at least 1% of births are: the lack of skin-to-skin contact, the lack of exclusive breastfeeding, whether the newborn had an APGAR score below 9, and whether the newborn was transferred to a neonatal intensive care unit (NICU).[24] A higher probability of needing NICU or a low Apgar score can be signals of lower quality. Similarly, if human resources are scarce, midwives may decide to skip some steps of the service considered important but not essential. For example, they may decide that helping the newly mother achieve skin-to-skin contact with her newborn is not as important as helping another woman in labor to deliver. The same reasoning applies for not giving exclusive breastfeeding.

While it is clear why a higher probability of going to NICU or having a low APGAR score are

[23] In fact, weight at birth is one of the variables used to assess the balancing of the sample between treatment and control groups.

[24] The Apgar score is a method used to quickly summarize the health of newborn children. The Apgar scale is determined by evaluating the newborn baby on five simple criteria on a scale from zero to two, then summing up the five values thus obtained. The resulting Apgar score ranges from zero to 10.

not desirable, there are also compelling arguments regarding the importance of the remaining set of outcomes. In a systematic review, Ip et al. (2007) find that breastfeeding is associated with both decreased risk for many early-life diseases and conditions as well as with health benefits to women[25]. At the same time, skin-to-skin contact has been shown to increase the probability and length of exclusive breastfeeding (Moore et al., 2007), as well as substantially reducing neonatal mortality among preterm babies in hospitals (Lawn et al., 2010).

Table 4.6 displays estimates of the effect of workload on each one of these four outcomes. Like previous studies, I fail to find an effect of workload on health outcomes. None of the coefficients are statistically different from zero. However, these results should be taken with caution. Some estimates are quite imprecise given the small sample size and the rarity of some of these conditions. In addition, there may be other conditions that are affected by workload but unobserved in the data at hand.

Table 4.6: Effect of workload on infant health outcomes

	No skin-to-skin contact	Non-exclusive breastfeeding	Apgar score<9	Neonatal ICU
Med. RPM	0.013	-0.013	-0.001	-0.009
	(0.014)	(0.018)	(0.007)	(0.008)
High RPM	0.014	-0.017	0.007	0.001
	(0.017)	(0.025)	(0.009)	(0.011)
Observations	5415	4799	6142	6131
Mean dep.	0.180	0.322	0.047	0.072
Shift FE	Yes	Yes	Yes	Yes

Notes: All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4. Medium RPM is a dummy variable for patients who observe a RPM between the 20^{th} and 80^{th} percentile, while High RPM is a dummy for patients who observe an RPM above the 80^{th} percentile. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

5 Conclusions

This paper presents a novel approach to understand whether workload variations affect the quality of care. My identification strategy relies on exploiting plausibly exogenous variation in the ratio of patients-to-midwives (RPM) among unscheduled patients who attempt a vaginal delivery. I find that the likelihood of getting a C-section is higher for patients who observe a RPM above the 20th percentile. I fail to find any effect of workload on other margins of treatment or on newborn's health -although the morbidity estimates not very precise.

I provide suggestive evidence of two mechanisms behind the raise in C-sections with workload.

[25] "Breastfeeding and Maternal and Infant Health Outcomes in Developed Countries", AHRQ Publication No. 07-E007, April 2007.

First, I find that patients who have an ex-ante higher risk of needing a C-section are more likely to have one when workload is high. This supports the hypothesis that a raise in workload lowers the quality of care, and relatively weaker patients are the most affected by it. Second, I find that, while single and married women have the same probability of C-section when workload is low, the likelihood raises with workload for single women only. I interpret this within an agency discrimination model, where single women are more likely to be alone in the delivery room and are easier to induce.

My estimates imply that approximately 16 percent of unscheduled C-sections are a consequence of high workload. Eliminating low-staffing instances from maternity wards would have a very significant effect on the already high levels of C-sections seen in developed countries (Italy included). Considering only the difference in costs between a vaginal and C-section delivery, these “extra C-sections” cost about €45,567 per year for the hospital under analysis.[26] Although this is not enough to hire the necessary number of midwives to assure an RPM below the 20th percentile, other costs beyond the hospital's financial burden should be considered (e.g. patients satisfaction, short and long-term effects of C-sections on mothers and infant's health, etc.). Another possible policy is to concentrate maternity wards in fewer but bigger units and benefit from the economies of scale emerging. The larger the population a hospital serves, the lower the coefficient of variation of demand, and hence the higher the occupancy rate (Long and Feldstein, 1967).[27]

Finally, this study opens up some interesting avenues for future research on the long-term effects of C-sections using workload as an instrumental variable. For instance, I will also examine whether being born by C-section affects infants' health later in life and whether having a C-section affects a woman's fertility decisions.

[26]Back of the envelope calculations suggest that there are about 222 “extra” C-sections in the 4 years in the sample due to workload. According to the prices on acute interventions published by the Italian Ministry of Health, a vaginal delivery without complication is rated at €1,272, while a C-section costs €2,092. Hence the difference (€820times the number of extra C-sections (122) divided by the number of years (4) gives €45,567.

[27]For the hospital in case this may not really be a suitable alternative since it is already a large maternity ward and the only on its city.

References

Alexander, D. et al. (2013), Does physician compensation impact procedure choice and patient health?, Technical report.

Alkalay, A., Eizenberg, A., Lahad, A., Shurtz, I. et al. (2018), Physician workload and treatment choice: the case of primary care, Technical report, CEPR Discussion Papers.

Allin, S., Baker, M., Isabelle, M. and Stabile, M. (2015), Physician incentives and the rise in c-sections: Evidence from canada, Technical report, National Bureau of Economic Research.

Bender, P. T. H., Homer, C., Matthews, Z., Nove, A., Sochas, L., Campbell, J. and deBernis, L. (2014), The state of the world's midwifery: a universal pathway, a woman's right to health.
URL: *https://eprints.soton.ac.uk/391219/*

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

Berwick, D. M. and Hackbarth, A. D. (2012), 'Eliminating waste in us health care', *Jama* **307**(14), 1513–1516.

Blustein, J. and Liu, J. (2015), 'Time to consider the risks of caesarean delivery for long term child health', *BMJ: British Medical Journal* **350**.

Brekke, K. R., Holmås, T. H., Monstad, K. and Straume, O. R. (2017), 'Do treatment decisions depend on physicians' financial incentives?', *Journal of Public Economics* **155**, 74–92.

Büscher, A., Sivertsen, B., White, J. et al. (2009), 'Nurses and midwives: a force for health', *WHO Europe, Copenhagen* .

Card, D., Fenizia, A. and Silver, D. (2018), The health effects of cesarean delivery for low-risk first births, Technical report, National Bureau of Economic Research.

Chen, M. M. and Grabowski, D. C. (2015), 'Intended and unintended consequences of minimum staffing standards for nursing homes', *Health Economics* **24**(7), 822–839.

Clemens, J. and Gottlieb, J. D. (2014), 'Do physicians' financial incentives affect medical treatment and patient health?', *American Economic Review* **104**(4), 1320–49.

Cook, A., Gaynor, M., Stephens Jr, M. and Taylor, L. (2012), 'The effect of a hospital nurse staffing mandate on patient health outcomes: Evidence from california's minimum staffing regulation', *Journal of Health Economics* **31**(2), 340–348.

Costa-Ramón, A. M., Rodríguez-González, A., Serra-Burriel, M. and Campillo-Artero, C. (2018), 'It's about time: Cesarean sections and neonatal health', *Journal of health economics* **59**, 46–59.

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

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

Curtin, S., Gregory, K., Korst, L. and Uddin, S. (2015), 'Maternal morbidity for vaginal and cesarean deliveries, according to previous cesarean history: New data from the birth certificate, 2013.', *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* **64**(4), 1.

Cutler, D., Skinner, J., Stern, A. D. and Wennberg, D. (2013), Physician beliefs and patient preferences: a new look at regional variation in health care spending, Technical report, National Bureau of Economic Research.

de Elejalde, R. and Giolito, E. (2019), More hospital choices, more c-sections: Evidence from chile, Technical report, Institute for the Study of Labor (IZA).

Deneux-Tharaux, C., Carmona, E., Bouvier-Colle, M.-H. and Bréart, G. (2006), 'Postpartum maternal mortality and cesarean delivery', *Obstetrics & Gynecology* **108**(3, Part 1), 541–548.

Dranove, D. and Watanabe, Y. (2009), 'Influence and deterrence: How obstetricians respond to litigation against themselves and their colleagues', *American Law and Economics Review* .

Evans, W. N. and Kim, B. (2006), 'Patient outcomes when hospitals experience a surge in admissions', *Journal of Health Economics* **25**(2), 365–388.

Francese, M., Piacenza, M., Romanelli, M. and Turati, G. (2014), 'Understanding inappropriateness in health spending: The role of regional policies and institutions in caesarean deliveries', *Regional Science and Urban Economics* **49**, 262–277.

Freedman, S. (2016), 'Capacity and utilization in health care: the effect of empty beds on neonatal intensive care admission', *American Economic Journal: Economic Policy* **8**(2), 154–85.

Freedman, S., Golberstein, E., Huang, T.-Y., Satin, D. and Barrie Smith, L. (2018), Docs with their eyes on the clock? the effect of time pressures on primary care productivity, Technical report, Working paper.

Gregory, K. D., Jackson, S., Korst, L. and Fridman, M. (2012), 'Cesarean versus vaginal delivery: whose risks? whose benefits?', *American Journal of Perinatology* **29**(01), 07–18.

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

Gruber, J. and Owings, M. (1996), 'Physician financial incentives and cesarean section delivery', *The Rand Journal of Economics* pp. 99–123.

Grytten, J., Skau, I. and Sørensen, R. (2011), 'Do expert patients get better treatment than others? agency discrimination and statistical discrimination in obstetrics', *Journal of Health Economics* **30**(1), 163–180.

Grytten, J., Skau, I. and Sørensen, R. (2013), 'Do mothers decide? the impact of preferences in healthcare', *Journal of Human Resources* **48**(1), 142–168.

Harris, M. C., Liu, Y. and McCarthy, I. (2019), Capacity constraints and the provision of public services: The case of workers in public health clinics, Technical report, National Bureau of Economic Research.

Hatem, M., Sandall, J., Devane, D., Soltani, H. and Gates, S. (2008), 'Midwife-led versus other models of care for childbearing women', *The Cochrane Library* .

Ho, K. and Pakes, A. (2014), 'Physician payment reform and hospital referrals', *American Economic Review* **104**(5), 200–205.

Ip, S., Chung, M., Raman, G., Chew, P., Magula, N., DeVine, D., Trikalinos, T. and Lau, J. (2007), 'Breastfeeding and maternal and infant health outcomes in developed countries.', *Evidence report/technology assessment* pp. 1–186.

Jachetta, C. (2015), Cesarean sections and later child health outcomes. unpublished.

Johnson, E. M. (2014), 'Physician-induced demand'.

Johnson, E. M. and Rehavi, M. M. (2016), 'Physicians treating physicians: Information and incentives in childbirth', *American Economic Journal: Economic Policy* **8**(1), 115–41.
URL: *http://www.aeaweb.org/articles.php?doi=10.1257/pol.20140160*

Lawn, J. E., Mwansa-Kambafwile, J., Horta, B. L., Barros, F. C. and Cousens, S. (2010), 'Kangaroo mother care to prevent neonatal deaths due to preterm birth complications', *International Journal of Epidemiology* **39**(suppl 1), i144–i154.

Lefèvre, M. (2014), 'Physician induced demand for c-sections: does the convenience incentive matter?'.

Lin, H. (2014), 'Revisiting the relationship between nurse staffing and quality of care in nursing homes: An instrumental variables approach', *Journal of health economics* **37**, 13–24.

Long, M. F. and Feldstein, P. J. (1967), 'Economics of hospital systems: peak loads and regional coordination', *The American Economic Review* **57**(2), 119–129.

Matsudaira, J. D. (2014), 'Government regulation and the quality of healthcare evidence from minimum staffing legislation for nursing homes', *Journal of Human resources* **49**(1), 32–72.

McDermott, K., Freeman, W. and Elixhauser, A. (2017), 'Overview of operating room procedures during inpatient stays in us hospitals, 2014', *HCUP Statistical Brief* **233**, 1–18.

Moore, E. R., Anderson, G. C., Bergman, N. et al. (2007), 'Early skin-to-skin contact for mothers and their healthy newborn infants', *Cochrane Database Syst Rev* **3**.

Neprash, H. T. (2016), Better late than never? physician response to schedule disruptions, Technical report, Working paper.

Norberg, K. and Pantano, J. (2016), 'Cesarean sections and subsequent fertility', *Journal of Population Economics* **29**(1), 5–37.

Osservatorio di Epidemiologia dell'Agenzia Regionale di Sanit (ARS) della Toscana (2013), Nascere in toscana. anni 2008-2011, Technical report.
URL: *http://www.ars.toscana.it/files/pubblicazioni/Volumi/2013/72_cap_2013.pdf*

Tonei, V. (2019), 'Mothers mental health after childbirth: Does the delivery method matter?', *Journal of health economics* **63**, 182–196.

Tong, P. K. (2011), 'The effects of california minimum nurse staffing laws on nurse labor and patient mortality in skilled nursing facilities', *Health Economics* **20**(7), 802–816.

World Health Organization (1985), 'Appropriate technology for birth', *Lancet* **2**, 436–437.

World Health Organization (2015), WHO statement on caesarean section rates, Technical report.
URL: *http://www.who.int/reproductivehealth/publications/maternal_perinatal_health/cs-statement/en/*

A The working sample and scheduled patients

The working sample used in the main paper is restricted to only those unscheduled patients who attempt to have a vaginal delivery after going through labor, and leaves out scheduled patients. Scheduled patients can be further divided in two groups: (i) elective C-sections, and (ii) pharmacologically-induced patients. This appendix shows evidence of how the latter group's transition through the maternity ward resembles more that of elective C-section rather than the one of unscheduled patients, and hence should not be included in the working sample.

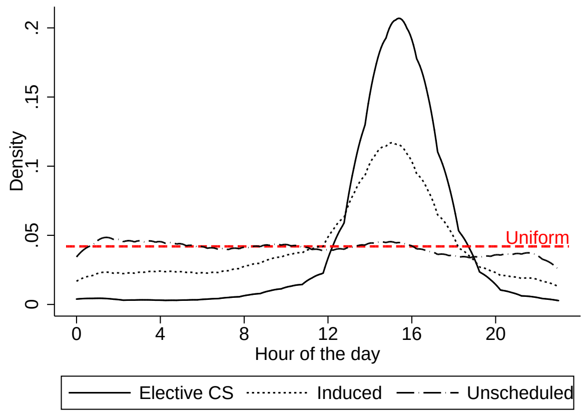
One important caveat of the data is that one cannot disentangle scheduled from unscheduled patients among those who were pharmacologically induced. However, anecdotal evidence from the ward's staff suggest that most of them are scheduled (e.g. overdue pregnancy). Furthermore, a descriptive analysis of the data seems to corroborate that. Figures A.1 and A.2 present the distribution of patients across hours and days as performed in section 1.3.1 of the main paper except that now scheduled patients are further divided between elective C-sections and induced. Starting from Figure A.1, it shows that there is a pick in admissions for both elective C-sections and induced patients during the afternoon shift, and then again a pick in time of birth (although the pick is later in the day for induced patients relative to the elective C-sections). Nevertheless, the picks are less pronounced for induced patients, suggesting that some of them may be arriving at random hours of the day like unscheduled patients do.

Even though the distribution by hours of induced patients seem to follow that of elective C-sections, their distribution by day of the week instead is closer to that of unscheduled patients. Even though admissions are slightly lower during weekends, births are evenly distributed across all days of the week. This is probably due to the fact that, as long as everything goes well, these patients are taken care of by midwives (not physicians).

The evidence provided in this appendix supports the idea of excluding both elective C-sections and pharmacologically induced patients from the working sample, but to include the latter group in the treatment variable given that they are primordially seen by midwives.

Figure A.1: Distribution of admissions and births by hour.

(a) By Time of Admission



(b) By Time of Birth

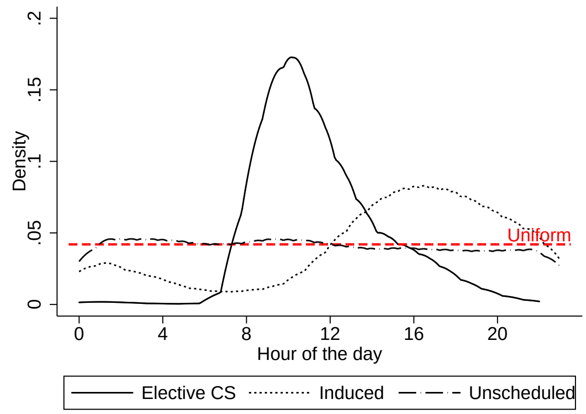
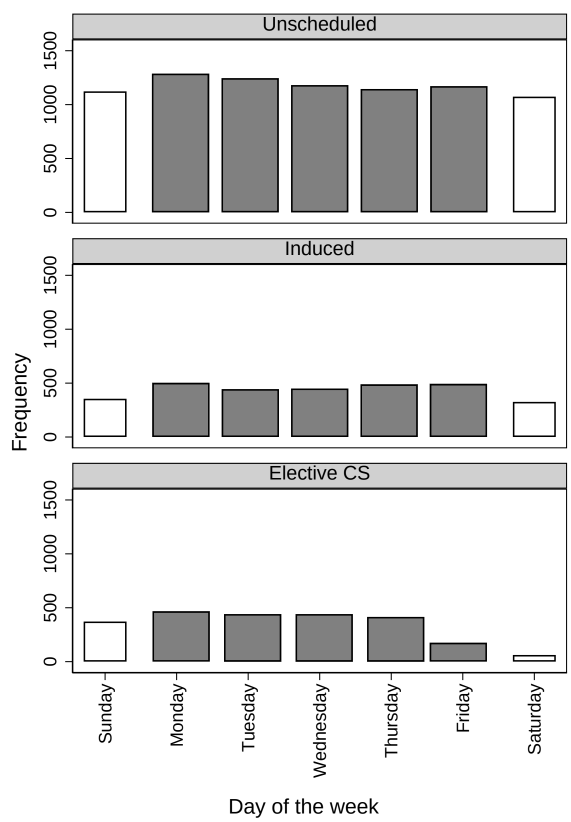
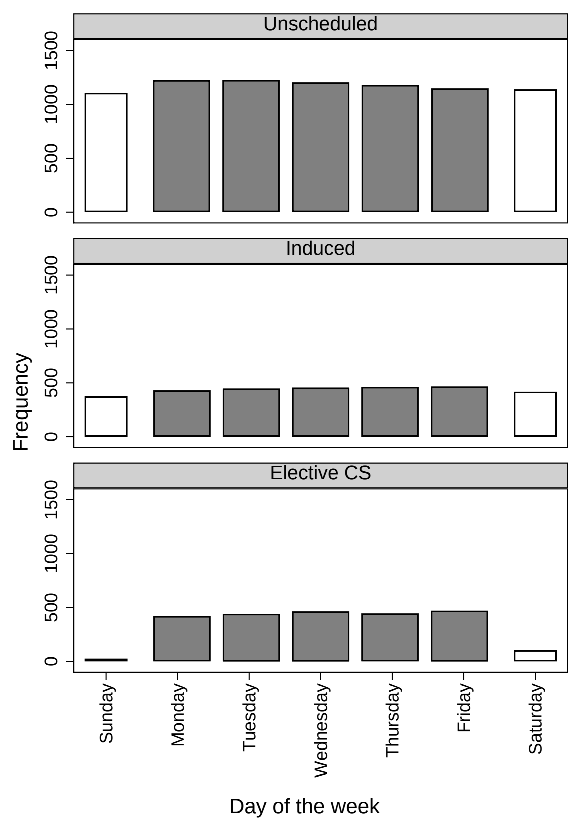


Figure A.2: Frequency of admissions and births by day.

(a) By Day of Admission



(b) By Day of Birth



B Robustness to alternative models

In the main paper two functional forms are tested for the effect of workload on the probability of C-section: a linear specification, and a non-linear one using a categorical variable constructed from the 20th and 80th percentiles. This appendix elaborates further on the model selection and tests other specifications. Columns one to four in Table B.1 present the coefficients for different polynomial degrees of the ratio of patients-to-midwives, with the Akaike Information Criteria (AIC) reported at the bottom. It seems that, within these polynomial functional forms, the data at hand is better represented by either a squared or cubic polynomial, given their statistical significance and their low AIC.

Column five presents results using a categorical variable with the quintiles of the distribution of the ratio of patients-to-midwives (where the reference group is the first quintile). This specification gives the model more flexibility to fit the data, at the cost of estimating more coefficients. Results suggest that there is a sudden rise in the probability of C-section for patients who see a ratio of patients-to-midwives in the second quintile, which then falls slowly until the fifth quintile where is no longer statistically distinguishable from the reference group. This decay in the probability of C-section for higher workloads may be associated with capacity constraints on the operative theater (beds, number of gynecologists, etc.).

Given the previous , I created a variable with three categories where the 3 middle quintiles of the ratio of patients-to-midwives have been coded together in one group (<20th percentile, 20-80th percentile, >80th percentile). This specification has the advantage of capturing the higher level of C-sections that occurs in the middle of the workload distribution, while diminishing the number of coefficients to be estimated and augmenting precision. Results are presented in the sixth column.

Table B.1: Effect of workload on the probability of C-section

	(1)	(2)	(3)	(4)	(5)
RPM	0.008	0.034**	0.037		
	(0.005)	(0.016)	(0.040)		
RPM × RPM		-0.005*	-0.006		
		(0.003)	(0.013)		
RPM × RPM × RPM			0.000		
			(0.001)		
20-40th RPM				0.019	
				(0.012)	
40-60th RPM				0.017	
				(0.014)	
60-80th RPM				0.022*	
				(0.013)	
>80th RPM				0.027*	
				(0.014)	
Med. RPM					0.019*
					(0.010)
High RPM					0.027*
					(0.014)
Observations	6142	6142	6142	6142	6142
Mean dep.	0.104	0.104	0.104	0.104	0.104
Shift FE	Yes	Yes	Yes	Yes	Yes
AIC	2766	2766	2768	2770	2767

Notes: All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4. Medium RPM is a dummy variable for patients who observe a RPM between the 20^{th} and 80^{th} percentile, while High RPM is a dummy for patients who observe an RPM above the 80^{th} percentile. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

C Other Graphs and Tables

Table C.1: Robustness: LPM vs Probit

	LPM		Probit	
	(1)	(2)	(3)	(4)
RPM	0.008		0.008	
	(0.005)		(0.005)	
Med. RPM		0.019^{*}		0.018^{*}
		(0.010)		(0.010)
High RPM		0.027^{*}		0.026^{*}
		(0.014)		(0.014)
Observations	6142	6142	6142	6142
Shift FE	Yes	Yes	Yes	Yes

Notes: Columns 1 to 3 use a linear regression (Linear Probability Model), while columns 4 to 6 use a Probit model. Reported coefficients are average marginal effects. All models contain the mother and pregnancy controls and time fixed effects mentioned in Section 3.4. Medium RPM is a dummy variable for patients who observe a RPM between the 20^{th} and 80^{th} percentile, while High RPM is a dummy for patients who observe an RPM above the 80^{th} percentile. Standard errors, clustered by date of admission, are in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$