GABRIEL FACCHINI

LOW STAFFING IN THE MATERNITY WARD: KEEP CALM AND CALL THE SURGEON[1]

Abstract. This paper investigates the relationship between workload and choice of treatment in a large but understudied segment of the healthcare sector – maternity wards. Using detailed microdata on childbirth, I exploit quasi-random assignment of patients attempting to have a natural delivery to different ratios of patients-to-midwives and compare their likelihood of changing delivery method. I find that women who face a ratio higher than 1.33 are 34% more likely to give birth by cesarean section (C-sections). This effect is larger for patients who were already admitted with a higher risk of C-section, implying that provision of proper and timely care matters more for this type of patients. Because C-sections are faster than vaginal deliveries – in which the patient follows the course of labor –, the medical team may find it appealing to do more C-sections when time constrained. Using civil status as a proxy for bargaining power -assuming single women are on average more likely to be alone, I find that only single patients are subjected to unnecessary surgery. This provides evidence that high midwives' workload is yet another factor which triggers physician-induced-demand for C-sections.
Keywords. Cesarean section, capacity utilization, workload, midwives, physician-induced-demand

[1] Department of Economics, European University Institute, Florence, Italy. Email: gabriel.facchini@eui.eu. I would like to extend my gratitude to Andrea Ichino and Jérôme Adda for their invaluable support and advice throughout this research project. I also benefited greatly from discussions with Juan Dolado, Daniela Iorio, Matilde Machado and Matthew Neidell. 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. I also thank seminar participants at the EUI Micro-Econometrics Working Group. 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 patients' health outcomes, experts currently point towards the reduction of waste as the best way to go.[2] Among the several sources of waste, two widely cited ones are the lack of adoption of known best practices (e.g. effective preventive care) and overtreatment, that is, the carrying out of treatments that cannot possibly improve the patient's health (e.g. cases of physician induced demand). These two sources of waste are particularly salient in maternity ward settings.

The role of midwives -as opposed to physicians- in assisting birth speaks to the first point. Whereas relevant public health authorities have recently recognized that midwife-led care during labor is safer for low-risk pregnancies,[3] the media and midwifery colleges have long spoken of a "shortage of mid- wives",[4] which was also acknowledged by the World Health Organization (WHO) in 2009.[5] At the same time, cesarean sections (C-sections) rank high among greatly overused interventions[6], and governments and clinicians have expressed concern about its potential negative impact on patients' health.[7] Indeed, C-sections not only cost more than vaginal deliveries, but they also imply higher risks for both mother and infant[8] and, according to a growing medical literature, are associated to lower long-term outcomes of children's health.[9] In addition, because vaginal delivery after a C-section (VBAC) is very unlikely,[10] one C-section sets a path dependency for more C-sections in future births. There is also evidence that women who follow a C-section are

[2] See, for example, Berwick and Hackbarth (2012).
[3] For example, the National Institute for Health and Care Excellence (NICE) updated its guidelines in this direction in 2014.
[4] 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".
[5] Büscher *et al.* 2009.
[6] While the international healthcare community considers an ideal rate of C-sections to be between 10-15%, 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).
[7] WHO Statement on Caesarean Section Rates, WHO (2015).
[8] See Deneux-Tharaux *et al.* 2006; Gregory *et al.* 2012; Curtin *et al.* 2015.
[9] 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).
[10] VBAC rate is only 8.3% in the US, and 12% in Italy.

more likely to have less children,[11] something that is particularly alarming in developed countries with already low fertility rates.

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 high workload -with midwives devoting less time to each patient, therefore rising the probability of complications that lead to surgery- or because physicians may find it optimal to induce some patients towards a C-section independently of their health status. Since a C-section takes less time than a vaginal birth – no need to wait for the appropriate dilation of the cervix –, midwives' workload can be reduced by shifting patients to surgeries.

This study causally tests whether patients follow a different delivery method depending on the effective staff level in the maternity ward at the moment of admission. It exploits a simple natural experiment: the majority of patients follow the natural course of birth and only go to the hospital once labor has already started and/or their water has broken (unlike, for example, scheduled cesarean sections). The effective staff level (e.g. the staff per patient ratio) observed by these patients at admission is orthogonal to their demographic and health characteristics (and to their ex-ante probability of delivering by C-section). The effective staff level at admission changes with the number of patients who arrived before and the number of midwives present in the delivery room, two variables that are unknown for the incoming patient.

The data for this project comes from a census of births from a large public hospital in Italy for the period 2011-2014. Three features of this dataset make it well suited for tackling the issue at hand. First, birth certificates have precise information on delivery method, allowing the identification of scheduled and unscheduled patients. Second, using patient's ID, each certificate was merged with hospital administrative data containing the exact time of admission and discharge. I use this information to compute the actual number of patients in the delivery room at each point in time. Finally, this is complemented with data on the number of midwives scheduled by month, day of the week and shift.

Results suggests that there is a non-linear relationship between effective midwifery staff and de- livery method: a newly admitted patient who faces a ratio of patients-to-midwives higher than 1.33 is 34% more likely to give birth by C-section. This means that, for first-time mothers, about 1.2 p.p. (or 5.7%) of all C-sections (both scheduled and unscheduled) are the consequence of low midwifery staffing.

[11] Norberg and Pantano 2016.

The second part of the analysis looks at possible mechanisms behind this 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 one should see patients with marginally lower health being 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 babies had an extreme weight at birth, and a high-health type, with all the remaining patients. Indeed, the gap between the probability of having a C-section between a low-health and a high-health patient widens with a higher workload.

Another factor that can explain the rise in C-sections alongside with workload is the presence of physician induced demand (PID). Because C-sections are faster than vaginal births, when faced with time constraints, physicians may decide to put some patients through surgery -without a medical necessity for it-, reducing the midwives' workload. Within the agency discrimination framework, physicians will choose to practice an unnecessary surgery on patients with lower bargaining power. This study tests for the presence of agency by comparing single women and non-single women, assuming that single patients are – on average – more likely to be alone in the delivery room. In those cases, the physician will need less effort in convincing the patient to have a C-section. Indeed, the data shows that the gap in the probability of delivering by C-section between these two groups is statistically significant only for high ratios of patients-to-midwives. On the other hand, I find that married and low-risk patients are between 24% and 35% more likely of not attaining skin-to-skin contact with their newborn when the number of patients per midwife is high. This provides more evidence that, by performing more C-sections, physicians are avoiding some bad outcomes.

This paper contributes to several strands of literature. First, it adds to existing work on the effect of staff ratios on health outcomes. Previous studies find none or very small effects when using census discharge data (Evans and Kim 2006; Cook *et al.* 2012), and a negative impact of crowding on health when focusing on patients in the Emergency Department (ED) (de Araujo *et al.*, 2013). This difference between areas makes sense given the particular time constrains of patients in the ED. The maternity wards lay somewhere in between these two. However, there is no study looking at the effect of staff ratio in maternity wards using a casual approach. The one that comes closest to this is Balakrishnan and Soderstrom (2000), using data from 225,473 maternity admissions at 30 hospitals in the state of Washington. They identify crowded days using a percentile cut-off from the distribution of patients' admissions for each hospital-year combination and the rate of C-sections as outcome. They find a positive and significant correlation between the two, only for those pregnancies that are classified as at-risk of C-section. A shortcoming of this paper is that they cannot differentiate between scheduled and unscheduled patients in their data, rising concerns about causal relationships. It could be the case that days with more patients are those with more planned C-sections, without necessarily having any effect on patients' health outcomes. I contribute to this literature by causally estimating the effect of low staffing ratios on delivery method.

Second, there is a vast number of empirical studies that look at different causes for the exceedingly high levels of C-sections. Starting from the paper by Gruber and Owings (1996) where they use physician's income drop 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 (Keeler and Brodie 1993; Lawthers *et al.* 1992; Currie and MacLeod 2008; Dranove and Watanabe 2009), and physician's scheduling convenience (Lefèvre 2014).[12] I provide of scheduled and unscheduled patients. Second, using patient's ID, each certificate was merged with hospital administrative data containing the exact time of admission and discharge. I use this information to compute the actual number of patients in the delivery room at each point in time. Finally, this is complemented with data on the number of midwives scheduled by month, day of the week and shift.

Third, this study also relates to the literature that empirically tests possible mechanisms behind PID. Two recent papers use information asymmetry variations in the maternity ward set up. Grytten *et al.* (2011) compare expert and non-expert patients and conclude that a model of statistical discrimination (expert patients are better at communicating with the physician) explain their results better than one of agency discrimination (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. They do so by comparing physician patients with non-physician patients, in settings with and without financial incentives to perform C-sections. I add to this body of work by using a different approach to test for bargaining power: whether the mother is alone in the delivery room.

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

[12] For an extensive review of this literature see Allin *et al.* 2015.

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.[13] 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 seen necessary.

The remaining patients, those attempting to follow the natural course of labor and vaginal delivery, are the 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 into a labor and delivery room and assigned a gynecologist and a midwife. If everything goes as plan and the patient can have a vaginal delivery, the midwife will be the one helping her throughout 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 having a C-section instead.

More importantly, the actual presence of some of these medical conditions depend heavily on the subjective opinion of the gynecologist.[14] The presence of this gray area -or asymmetry of information on when is a C-section necessary gives the gynecologist more room to suggest the patient to follow a surgery, 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 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 a vaginal delivery.[15]

[13] Most inducements are performed on pregnancies that have past their due date and still haven't started labor.
[14] Two of these more 'subjective' conditions are dystocia (abnormally slow labor) and fetal distress.
[15] For a deeper discussion on the Italian Health System see Francese *et al.* 2014.

3. EMPIRICAL METHODOLOGY

3.1 A natural experiment

An ideal experiment to test for an effect of low-staffing in the maternity ward 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 vaginal delivery, and uses the exogenous variability in the number of patients and midwives present at admission to causally identify the impact of low staffing 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 this sample of patients, their pre-admission probability of developing a complication and needing 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.[16]

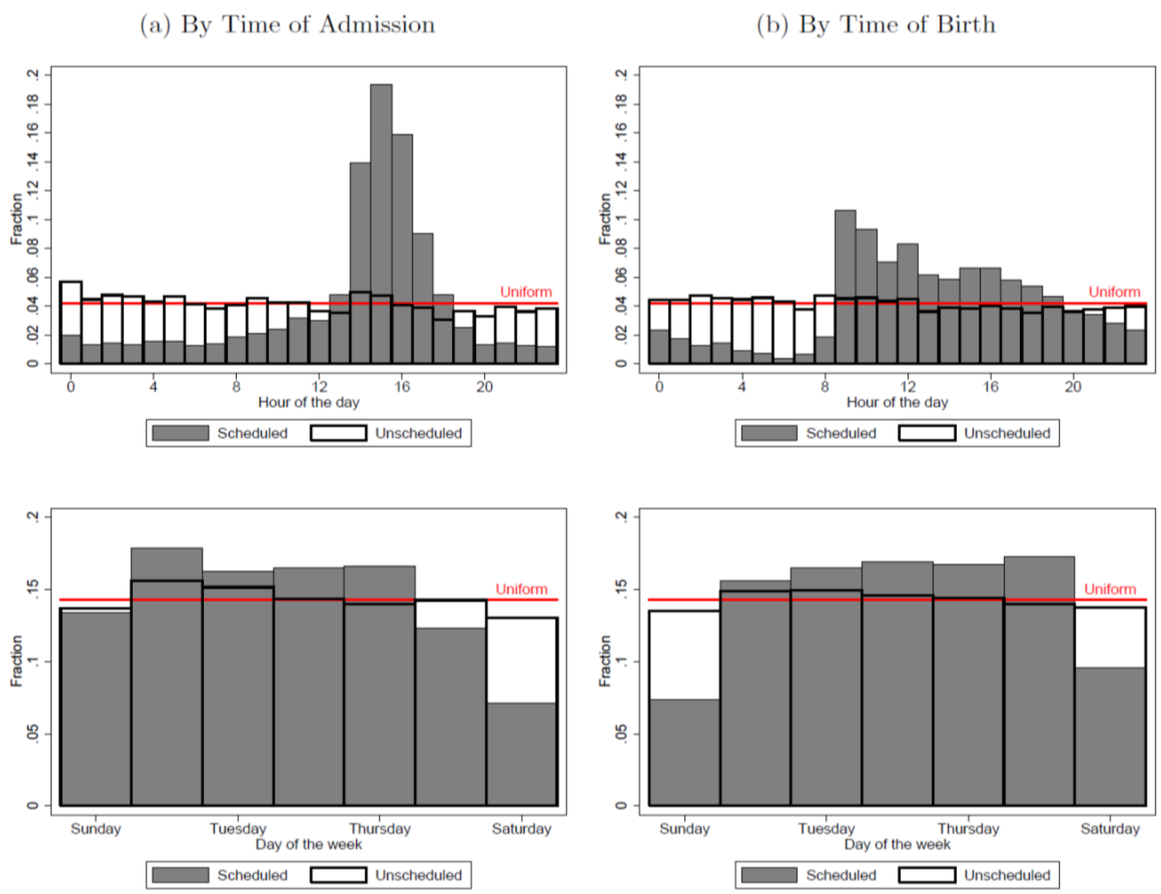
The left column of Figure 3.1 shows the distribution of admissions by hour of the day and day of the week. The right column does the same for births. Both are estimated for scheduled and unscheduled patients for comparison. We can immediately see that admissions of scheduled patients are concentrated 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 across the day. When looking at the distribution by days of the week, again 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.

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

3.2 Data

Previous studies looking at newborns' health tend to use anonymous birth certificates since they are publicly available for many countries and for long periods of time. However, these datasets 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), for each hospital.

Figure 3.1. Distribution of admission and birthes



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;[17] and (ii) hospital admissions.[18] Birth

[17] Certificato di assistenza al parto (CEDAP).
[18] Scheda di Dimissione Ospedaliera (SDO).

certificates constitute a census of all births that took place in the hospital in this period. It contains information on mother characteristics (e.g. com- munity 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 merged together.

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

However, the richness of this dataset comes at a cost: because the information available corresponds only to one hospital in a four-year period the sample size is relatively small. Furthermore, due to the path dependence of treatment in second and higher order births, this study focuses on first-time mothers. There were approximately 5,240 singleton births at this hospital in the sample period. From this, about 870 observations are plural births and/or delivered by urgent C-section which will not be considered in the analysis because of their particular characteristics and handling within the hospital. Then further restricting the sample to non-induced planned-vaginal deliveries the number of observations goes down to around 2,685. Finally, after dropping observations with missing time of admission, maternal age, education, birth order, weight and prenatal visits, the number of observations in the working sample is about 2,600. 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 12% had an in-labor C-section. Patients are on average 31 years old, only 36% has a university degree, and 44% are single. There are few cases with bad outcomes: only 4.6% have a 5-minute APGAR score below 9, and about 5% are born prematurely or weighting less than 2,500 grams. Columns 2 and 3 report statistics for patients with a low and high ex-ante risk of C-section respectively. Columns 4 and 5 do the same by civil status. By construction, patients from the high-risk are more likely to give birth by C-section, use of a neonatal intensive care unit, and have an APGAR score below 9. They are also more likely to be single and less likely to have a university degree. Finally, single patients are less likely to have a university degree and more likely to delivery by C-section, although other outcomes are similar to the married subsample.

Table 3.1. Descriptive statistics

	All	Low-risk	High-risk	Married	Single
Outcomes	88.1	88.8	85.1	89.5	86.6
% vaginal birth					
% in-labor C-section	11.9	11.2	14.9	10.5	13.4
Other interventions/health outcomes					
% operative birth	13.3	13.6	11.9	13.5	12.6
Average length-of-stay (hours)	76.0	75.3	79.0	76.7	75.9
% need of NICU	7.3	4.8	17.8	5.9	8.1
% lack of skin-to-skin contact	19.3	16.0	33.5	18.0	20.0
% non-exclusive breastfeeding	36.0	33.9	46.7	35.5	36.5
% APGAR score below 9	4.6	3.3	10.5	3.8	5.4
Mother's characteristics					
Average age	31.1	31.2	30.6	31.2	30.8
% with university degree	35.9	36.3	33.9	41.3	30.2
% single	44.2	43.5	47.0	0.0	100.0
Pregnancy's characteristics					
% born before 37 weeks of gestation	5.3	2.7	16.6	5.2	5.2
% with at least 1 ER visit	11.5	0.0	60.6	10.4	13.3
Newborn's characteristics					
% male	51.0	50.2	54.3	51.9	50.4
Average weight at birth	3,235	3,271	3,085	3,234	3,234
% low birthweight (<2,500 grams)	4.9	0.0	26.1	4.9	4.9
% high birthweight (>4,000 grams)	3.9	0.0	20.4	4.4	3.7
Observations	2,613	2,118	495	1,300	1,028
	All	**Low-risk**	**High-risk**	**Married**	**Single**

Note: Statistics for main sample of unscheduled first-time mothers, from 2011-2014. High-risk are patients who, at admission, have a higher probability of needing a C-section. Those are defined as patients with newborns with extreme birthweight and patients with an emergency department visit during pregnancy. Low-risk are those without any of those characteristics.

3.3 An exogenous measure of midwives' workload

A good measure of effective staff contains information on both number of patients and personnel. For the maternity wards setting of this paper I use the ratio between the number of patients and the number of midwives in the delivery room.[19] The richness of

[19] 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. Furthermore, Appendix B repeats the main analysis using solely the number of patients as the covariate of interest, and results are qualitatively the same. In light of these results,

the data in hand allows 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. But 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 patients stay on average 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 obvious options are at the time of admission and at the time of delivery. The last one has the advantage 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.[20]

On the second issue, it is important to clarify 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). One could think that, because the time of the admitted patient is random, there is no risk of endogeneity here. 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. Hence this study includes in the numerator all patients but those already scheduled to give birth by C-section.[21] Instead, the number of scheduled C-sections is included in the regression as control (see econometric specification below). More specifically, in this paper the workload observed by patient i at admission time t is define as

$$R_{it} = \frac{PVB_{it}}{MW_{it}} \tag{1}$$

in the main paper I will use the ratio of patients-to-midwives since it provides an advantage with regard to external validity (findings become less dependent on the size of the hospital studied).

[20] In the following section I perform several robustness check measuring staff levels at different points in time during a patient's stay, and discuss the results.

[21] 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 will need a midwife to help them.

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 mean number of midwives and patients (with its standard deviation) in the delivery room by day of the week and shift of admission. The number of midwives is higher during the morning shift (5), and lower at nights and Sundays (3). On the other hand, the average number of patients is virtually the same across days of the week and shifts, with a slightly lower level on Sundays.[22]

Table 3.2. Number of midwives and patients by day of the week and shift

Day	Shift	Midwives (mean)	Midwives (sd)	Patients§ (mean)	Patients§ (sd)
Weekdays	Morning (7am - 1pm)	5	0	7.31	2.81
	Afternoon (1pm - 7pm	4	0	7.48	2.89
	Night (7pm - 7am)	3	0	7.32	2.86
Saturdays	Morning (7am - 1pm)	4	0	7.53	2.63
	Afternoon (1pm - 7pm	4	0	7.41	2.70
	Night (7pm - 7am)	3	0	7.26	2.71
Sundays	Morning (7am - 1pm)	3	0	7.09	2.73
	Afternoon (1pm - 7pm)	3	0	7.08	2.76
	Night (7pm - 7am)	3	0	6.94	2.68

§ Number of patients waiting who attempt to have a vaginal birth.

Table 3.3 shows the distribution of the ratio of patients to midwives for the whole sample and then disaggregated by shift of admission. The ratio is unimodal and slightly skewed to the right.[23] At the median, there are 2 patients for every midwife in the delivery room. The 25th and 75th percentiles are 30% (below) and 34% (above) the median, respectively. Note that shifts later in the day have higher values of the ratio, meaning, more crowding. 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).[24] The bottom rows of the Table 3.3 show the cutoff values for the lowest and highest quintiles (and by construction for the three middle quintiles altogether). The lowest

[22] The difference with Sunday is due to the fact that there are less induced births.
[23] See Figure D.1 for a graphic representation of the density distribution of the ratio by shift.
[24] In Figure D.2 one can see how the average ratio of patients to midwives by hour of admission shows a discrete jump up with each change in shift due to one less midwife being present.

quintile will be considered a case with no crowding, with a mean of 1 patient per midwife. The middle quintiles have a mean ratio of 1.9, somehow crowded. The highest ratio, with a mean of 3.2 patients per midwife, will be referred to as highly crowded or chaos.

Table 3.3. Descriptive statistics for ratio of patients to midwives by shift of admission

	All	Morning (7am - 1pm)	Afternoon (1pm - 7pm)	Night (7pm - 7am)
p1	0.60	0.40	0.50	0.67
p5	0.80	0.75	0.75	1.00
p25	1.40	1.20	1.50	1.67
p50	2.00	1.50	1.75	2.33
p75	2.67	2.00	2.33	3.00
p95	3.67	3.00	3.25	4.00
p99	4.50	4.00	4.50	4.67
mean	2.06	1.60	1.91	2.36
sd	0.86	0.70	0.74	0.88
<20th Percentile	1.33			
>80th Percentile	2.67			
Obs.	2,613	636	641	1,336

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:[25]

$$y_{it} = \alpha + \beta R_{it} + \theta X_{it} + \gamma_t cs_{it} + dow \times shift_t + year_t + month_t + \epsilon_{it} \quad (2)$$

where y_{it} is a dummy variable indicating whether birth i admitted at time t had an in-labor C-section, and R_{it} is the ratio of patients-to-midwives observed at admission as explained above. X_{it} contains individual-level control variables of mother and pregnancy characteristics.[26] To further control for supply side changes in physicians' availability I

[25] A probit model was also estimated assuming a normal distribution of the error term and results virtually the same (See Table D.3).

[26] 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

include the number of scheduled C-sections that took place while the indexed patient was in the delivery room (*cs*). Since most supply side changes in the maternity ward take place between shifts and days, in the most demanding specification I also add fixed effects for day-of-the-week (*dow*) times *shift*.[27] To control for seasonal and secular variation in outcomes, I also include monthly and yearly dummy variables. β is the coefficient of interest. As discussed above, if physicians are more likely to perform a C-section when the ratio of patients to midwives is high, then β should be positive.

Two models are estimated for the probability of delivering by C-section. First, I use the ratio of patients to midwives added linearly to the model. Because there can be nonlinear effects between staffing and delivery method, for the second model I split the sample in three categories based on the ratio of patients-to-midwives: low, medium, and high (or chaos). All those observations with a ratio below the 20th percentile are in the first group. These are cases of no crowding, or very low ratio of patients to midwives. The second group includes those observations between the 20th and 80th percentiles, and are categorized are cases with some crowding. Finally, the last group consists of all those above the 80th percentile. These are situations of very high ratios of patients to midwives. The cut offs for these groups are reported in the bottom of Table 3.3. In these models, the lowest quintile (low staffing is considered the reference group.[28] Table D.1 shows the coefficients of a regression of each of the pre-treatment controls on the ratio of patients-to-midwives. The lack of statistical significance for all cases provides support to the exogeneity assumption of my measure of staffing. Furthermore, for the non-linear specification, Table D.2 shows that the mean of the pre-treatment characteristics are not statistically different across the three groups of staffing (low, medium and high). Again, this emphasizes the strength of the quasi-natural experiment.

The last part of the analysis aims at understanding the mechanisms through which physicians decide to recommend some patients to change delivery method. Two hypotheses are tested. First, it could be the case that high values of the ratio of patients-to-midwives 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 that need more care) will be the most affected. At higher ratios, the probability of C-section should rise

is born with low weight (less than 2,500 grams), and a dummy for whether the mother had at least one emergency checkup during pregnancy.

27 This means that all the variation in this specification comes from within same day of the week and shift. For example, I would be comparing a mother who arrived on a Tuesday afternoon shift and finds many patients waiting with another woman arriving a different Tuesday afternoon but who observes few patients waiting.

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

faster for this group than for other patients -all else constant- due to their pre-treatment lower health. These patients with a higher risk are identified as those with extreme birthweight (below 2,500 grams or above 4,000 grams) or with at least one emergency visit to the hospital during pregnancy.

The second hypothesis has to do with agency discrimination. When resources are constrained, e.g. high ratio, physicians may see optimal to shift some patients to the operative theater and perform 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. This paper uses the patient's civil status as a proxy to whether the she is alone in the delivery room.[29] The assumption here is that, on average, single women are more likely to be alone in the delivery room.[30] 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 harder to analyze the pros and cons of each alternative-.

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

$$y_{it} = \alpha + \beta_1 R_{it} + \beta_2 R_{it} \times D_{it} + \beta_3 D_{it} + \theta X_{it} + \gamma_t cs_{it} + dow \times shift_t + year_t + month_t + \epsilon_{it} \tag{3}$$

where D_{it} is either one of two variables: an indicator for whether the patient has a high-risk of C-section, or whether she is single. The remaining variables are defined as in Eq. (2), adding civil status as a control. I expect high-risk and single patients to be more affected by a high ratio of patients, hence, a positive β_2 in both cases.

4. Results

Table 4.1 presents the results of estimating Eq. (2). Starting from a regression with only the covariate of interest and fixed effects for year, month and day of the week in the first column, each remaining column sequentially adds more controls. The second column adds controls for mother and pregnancy characteristics, the third adds the number of scheduled

[29] This variable is constructed only with married and single women. For the sake of clarity, all women outside these two categories (divorced, separated and widows) are not considered.

[30] 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).

C-sections taking place during patient's labor, the fourth column includes hour of admission fixed effects, while the last one instead uses shift-of- admission interacted with day-of-the-week fixed effects. This last model is the preferred one since it accounts for possible supply changes shift and day that may take place in the ward (apart from midwives). To save space, only the coefficients of treatment are included, but results for other co- variates are comparable to previous studies.[31] The numbers in parentheses in the table are standard errors. The average value of each dependent variable is included at the bottom of each panel to help understand whether coefficients are economically important. For all remaining estimations in this paper I will use the specification model in column (5).

Panel (A) of the table reports results for the Ratio of patients to midwives as a continuous variable, and Panel (B) reports results using a dummy variable for different levels of workload in order to test for non-linearities. First thing to notice is that coefficients across columns (models) only change in the third decimal. This is a good sign of exogeneity of the ratio of patients-to-midwives. Although the coefficient for the linear specification is not statistically significant, in the second panel the probability of having a C-section is about 4 p.p. (34%) higher for those who face a ratio of patients-to-midwives in the middle of the distribution compared to those in the reference group. For those patients arriving when the ratio of patients-to-midwives is very high (last quintile), there is not statistically significant effect on the probability of C-section. This may be due to some capacity constraints on the operative theater when workload is at its highest levels.

This effect would imply a 5.7% (or 1.2 p.p.) rise in total C-sections (scheduled and unscheduled), which is economically important and 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 p.p., all else equal. 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 A.7 presents results of the effect by whether the patient arrived in a weekday or weekend, and by shift of admission. The estimations are very imprecise due to the few number of observations in each cell, and render all differences insignificant. Nevertheless, point estimates are slightly higher in weekends, as well as for admissions during the morning shift. Table A.8 shows results for a robustness check where I measure effective staff level at different points in time between a patient's admission and delivery. The effect

[31] See full regressions in Appendix D, Table D.7 and Table D.8.

of congestion disappears the further away from admission it is measured, which can be a result of the endogeneity issue mentioned before: physicians can adjust the timing of births. Finally, Table A.9 presents results for a placebo test where workload is measured 24 hours after admission (instead at admission as before). As expected, for all different specifications, the placebo is always statistically and clinically insignificant.

Table 4.1. Effect of effective staffing on the Probability of C-section

	(1)	(2)	(3)	(4)	(5)
Panel (A)					
Ratio patients	0.0019	0.0038	0.0046	0.0007	0.0004
to midwives	(0.0074)	(0.0074)	(0.0074)	(0.0083)	(0.0085)
Panel (B)					
20-80th Percentile	0.0432***	0.0441***	0.0452***	0.0410***	0.0408***
	(0.0149)	(0.0148)	(0.0148)	(0.0154)	(0.0156)
>80th Percentile	0.0193	0.0238	0.0261	0.0187	0.0197
	(0.0196)	(0.0195)	(0.0196)	(0.0210)	(0.0213)
Observations	2,613	2,613	2,613	2,613	2,613
Mean dep.	0.119	0.119	0.119	0.119	0.119
Time FE	✓	✓	✓	✓	✓
Controls		✓	✓	✓	✓
Other patients			✓	✓	✓
Hour FE				✓	
Shift*dayofweek FE					✓

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
Reported coefficients are average marginal effects.

4.1 How do physincians choose which patients to send to the operative theater?

This part of the study digs deeper into the mechanisms behind the effect of staffing on the rate of C-sections. As mentioned before, two hypotheses are tested. First, low-staffing means there is less midwifery-time available for each patient, which may result in more patients needing C-section due to the lack of proper care. This effect should be higher for those patients who were admitted with an already higher risk of C-section. Secondly, physicians and midwives may actively decide to perform a C-section on some patients in moments of low-staffing to reduce the number of patients in the delivery room. In this case I expect patients with lower bargaining power -which I proxy by civil status- being more treated than others.

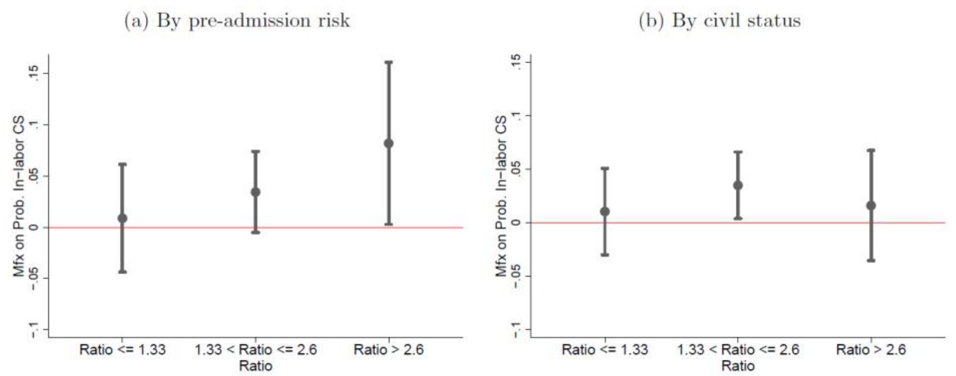
Table 4.2 reports the average marginal effects obtained for each group from estimating Eq. (3). As expected, a higher number of patients per midwife rises the probability of C-section more for single patients but not for married ones. Points estimates suggest that high-risk patients are more affected by workload than low-risk patients, although the only statistically significant coefficient is for this group. However, estimates are very imprecise.

Table 4.2. Effect of effective staffing on the Probability of C-section - by groups

	Low-risk	High-risk	Married	Single
20-80th Percentile	0.0330*	0.0587	0.0270	0.0516**
	(0.0175)	(0.0368)	(0.0202)	(0.0247)
>80th Percentile	0.0023	0.0755	0.0125	0.0182
	(0.0229)	(0.0557)	(0.0265)	(0.0332)
Observations	2,118	495	1,300	1,028
Mean	0.11	0.15	0.11	0.13

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$. Reported coefficients are average marginal effects from a regression of the probability of C-section on the interaction of treatment, a variable for being high risk and a variable for being single. The number of observations when using marital status is slightly smaller because the variable is missing for 11% of the working sample.

Figure 4.1. Difference in the effect of staffing on the probability of C-section by type of patient



Note: Dots are the average marginal effect of whether the patient is high-risk (a) or single (b). Bars are 90% confidence intervals.

Another way to look at it is by comparing the average marginal effects of being high-risk and single, across the different levels of the ratio of patients-to-midwives. This can be seen in Figures 4.1a and 4.1a respectively. Note that the effect of staffing, in both cases, is not statistically significant when the ratio is low. For the comparison based on ex-ante risk, the point estimate for the difference in the probability of C-section between the two groups gets higher with workload -albeit not statistically significant-. This is reasonable since ex-ante high-risk patients for whom the marginal benefit from midwives' attention is higher.

Instead, for the case of married vs. single mothers, the difference is statistically significant only for those in the middle of the distribution, but goes down again when workload is high. At high levels of workload, it is more likely that capacity constraints in the operative theater emerge as well. These "extra" C-sections only based on midwives' workload and not due to patients' health-status should go down during the busiest times.

4.2 The effect on other interventions and morbidity outcomes

The estimates above demonstrate that, when the ratio of patients-to-midwives is high, physicians send some patients to the operative theater to have a C-section. These patients are typically patients with a higher-risk of needing a C-section, or single women. However, are physicians using their high bargaining power to transfer some patients so midwives can provide give better care for the remaining patients? To test this, I estimate Eq. (3) again but now the outcome variable is one of the five indicators of morbidity and interventions mentioned before. If a high ratio lowers the quality of care, then those type of patients who are not likely to be sent to the operative theater would be the ones more affected by this.

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).[32]

The restricted-use version of the birth certificates in hand contains, however, some other measures of health and registers of medical interventions that are associated with

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

health outcomes. The measures that occur in at least 1% of births are: having an operative birth[33], length-of-stay after birth (LOS), whether the newborn was transferred to a neonatal intensive care unit (NICU), no skin-to-skin contact, lack of exclusive breastfeeding, and whether the newborn had an APGAR score below 9.[34] A higher probability of needing NICU, having an operative birth[35] or a longer time in the hospital during crowded times can be signals of lower quality. Similarly, if human resources are scarce, physicians 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 breast-feeding.

While it is clear why a higher probability of going to NICU having a low APGAR score, or staying longer in the hospital are not desirable, there are also compelling arguments regarding the importance of the remaining set of outcomes. In a systematic review, Ip *et al.* (2007) finds that breastfeeding is associated with both decreased risk for many early-life diseases and conditions as well as with health benefits to women.[36] 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 amongst preterm babies in hospital (Lawn *et al.* 2010). In the case of operative births, even though it is still widely used, this delivery method is becoming less popular due to some evidence showing it increases maternal morbidity and can cause significant fetal morbidity (Ali and Norwitz 2009; Murphy *et al.* 2011; Towner *et al.* 1999).

Table 4.3 displays the average marginal effects for each of the four groups of women (high and low risk, married and single), and for the five outcomes above mentioned. Estimates are quite imprecise given the small sample size and the rarity of these morbidities. However, there is a statistically significant, large and positive effect of the high ratios of patients-to-midwives on the probability of not achieving skin-to-skin contact with the infant. Furthermore, this effect is only present for married patients, who are not more likely to get surgery when workload rises. These patients are between 24% and 35% more likely to not attain skin-to-skin contact with their newborn when the number of patients

---

[33] 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.

[34] 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.

[35] A higher likelihood for operative birth has been linked to scarce or absent midwifery care and the presence of obstetrician or physicians instead (Hatem *et al.* 2008).

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

per midwife is higher. This provides further evidence of the fact that, by shifting delivery method for some patients, physicians are avoiding some bad outcomes to occur.

Table 4.3. Effect of effective staffing on other health outcomes

	Low-risk						High-risk					
	Op. birth	LOS⋆	NICU§	No s2s‡	NEB†	Apgar<9	Op. birth	LOS⋆	NICU§	No s2s‡	NEB†	Apgar<9
20-80th Percentile	-0.0164	-0.0246	-0.0134	0.0298	-0.0325	-0.0096	-0.0616	-0.0085	-0.0013	0.0133	-0.0392	0.0070
	(0.0194)	(0.0184)	(0.0123)	(0.0216)	(0.0296)	(0.0104)	(0.0389)	(0.0414)	(0.0304)	(0.0497)	(0.0629)	(0.0298)
>80th Percentile	0.0257	-0.0380	-0.0064	0.0426	-0.0179	-0.0011	-0.0213	-0.0168	-0.0357*	0.0329	-0.0592	-0.0075
	(0.0264)	(0.0234)	(0.0160)	(0.0287)	(0.0384)	(0.0144)	(0.0270)	(0.0274)	(0.0195)	(0.0310)	(0.0422)	(0.0181)
Observations	2,613	2,521	2,609	2,297	2,044	2,613	2,613	2,521	2,609	2,297	2,044	2,613
Mean dep.	0.133	4.274	0.0728	0.193	0.360	0.0463	0.133	4.274	0.0728	0.193	0.360	0.0463

	Married						Single					
	Op. birth	LOS⋆	NICU§	No s2s‡	NEB†	Apgar<9	Op. birth	LOS⋆	NICU§	No s2s‡	NEB†	Apgar<9
20-80th Percentile	-0.0339	-0.0061	0.0124	0.0449*	-0.0183	-0.0033	-0.0213	-0.0168	-0.0357*	0.0329	-0.0592	-0.0075
	(0.0241)	(0.0222)	(0.0140)	(0.0265)	(0.0367)	(0.0124)	(0.0270)	(0.0274)	(0.0195)	(0.0310)	(0.0422)	(0.0181)
>80th Percentile	0.0096	-0.0194	0.0111	0.0699**	-0.0517	0.0117	-0.0119	-0.0159	-0.0086	0.0416	0.0210	-0.0142
	(0.0329)	(0.0291)	(0.0171)	(0.0345)	(0.0472)	(0.0178)	(0.0359)	(0.0342)	(0.0272)	(0.0432)	(0.0559)	(0.0235)
Observations	2,328	2,264	2,324	2,057	1,826	2,328	2,328	2,264	2,324	2,057	1,826	2,328
Mean dep.	0.133	4.274	0.0728	0.193	0.360	0.0463	0.133	4.274	0.0728	0.193	0.360	0.0463

Reported coefficients are average marginal effects. Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$
⋆LOS: Length-of-stay after birth (in log-hours); §NICU: Neonatal Intensive Care Unit; ‡No s2s: No skin-to-skin contact; †NEB: Non-exclusive breastfeeding.

4.3 Other possible channels?

Beyond the mechanisms mentioned in the previous section, there are -at least- two more channels that can explain the rise in C-sections along with the rise of the ratio of patients-to-midwives. The first and most obvious option is that patients who are admitted in low and high staffing times are different. Nevertheless, 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 the day and week.

The other possible explanation is that those type of patients who get these 'extra' C-sections actually prefer this delivery method. However, because the focus is exclusively on in-labor C-sections, the above estimates correspond to women who have already agreed on attempting labor in the process to attempt a vaginal delivery. Hence the effect is more likely to arise from decisions made in the delivery room regarding when to stop labor and change treatment, than from maternal preferences for C-sections. Nevertheless, because 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 send to surgery.

4.4 Can these 'extra' C-sections be avoided?

Results above suggest that physicians do more surgeries when staffing is low. First-time mothers facing a ratio of patients-to-midwives between 1.33 and 2.66 are 4 p.p. (or 34%) more likely to have an in-labor C-section. A policy to eliminate overcrowding from maternity wards would have a very significant effect on the already high levels of C-sections seen in Italy. How to do that is not clear.

Considering only the hospital used in the analysis, in the absence of crowding, the "extra costs" for the public health system is of about €17,700 a year.[37] This is of course not enough to hire the necessary number of midwives to avoid low-staffing situations. Of course, this analysis is not complete since one should include other costs, like the drop in skin-to-skin contact when staffing is low, or the other non-financial costs of C-sections mentioned in the introduction of this study.

[37] Back of the envelope calculations suggest that there are about 86 'extra' C-sections in the 4 years in the sample due to crowding. 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 (€820 time the number of extra C-sections (107) divided by the number of years (4) gives €17,700.

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). 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.

5. Conclusions

In this paper I use a natural experiment set up -that patient characteristics is orthogonal to the level of staffing at the hospital at the moment of admission- and detailed data on births to estimate the impact of staffing on physician's treatment decisions. More specifically, I investigate whether different levels of midwifery effective staffing (patients-to-midwives) influence the probability that a patient will be sent to have a cesarean section. The contribution is threefold. First it proposes an innovative empirical approach that allows me to estimate physician's responses to exogenous shocks to effective staffing. Second, it provides suggestive evidence that physicians do not choose at random which patients to over-treat, but may instead use their bargaining power. Lastly, it brings to light yet another cause for the high C-section rates we see today: low effective staffing.

Focusing exclusively on patients attempting labor and vaginal delivery, this study finds that first- time mothers who -at admission- face a ratio of patients-to-midwives higher than 1.3 are about 34% (or 4 p.p.) more likely to change delivery method. There are two type of patients who are more affected by this. First, patients who upon admission have an already higher risk of C-section are more likely to develop complications due to limited care when few midwives are available. Secondly, single women, due to their lower bargaining power. I provide evidence that physicians may decide to induce some patients towards having a C-section to speed up the delivery and release the pressure on midwives in the delivery room. In summary, the evidence provided here suggests that physicians' way to deal with an exogenous shock in demand (patients) is to induce some patients towards an intervention that is faster, maximizing the aggregate health in the maternity ward.

My estimates imply that total number of C-sections for first-time mothers could be reduced by about 5.7% (1.2 p.p.) if situations of low-staffing are avoided. This would be a very important achievement given the already overly high rates of C-sections observed in developed countries. Nevertheless, it is not clear that public healthcare systems can quickly afford to tackle this issue.

REFERENCES

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

Ali U.A. and Norwitz E.R. (2009), *Vacuum-assisted vaginal delivery*, «Reviews in Obstetrics and Gynecology», 2, 1, p. 5

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

Balakrishnan R. and Soderstrom N.S. (2000), *The cost of system congestion: Evidence from the healthcare sector*, «Journal of Management Accounting Research», 12, 1, pp. 97-114

Berwick D.M. and Hackbarth A.D. (2012), *Eliminating waste in us health care*, «Jama», 307, 14, pp. 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

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

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, pp. 340-348

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

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, p. 1

de Araujo P., Khraiche M. and Tukan A. (2013), *Does overcrowding and health insurance type impact patient outcomes in emergency departments?*, «Health Economics Review», 3, 1, pp. 1-7

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), pp. 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», 12, 1, pp. 69-94.

Evans W.N. and Kim B. (2006), *Patient outcomes when hospitals experience a surge in admissions*, «Journal of Health Economics», 25, 2, pp. 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, pp. 262-277

Gregory K.D., Jackson S., Korst L. and Fridman M. (2012), *Cesarean versus vaginal delivery: Whose risks? Whose benefits?*, «American Journal of Perinatology», 29, 1, pp. 7-18

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

Gruber J. and Owings M. (1996), *Physician financial incentives and cesarean section delivery*, «The Rand Journal of Economics», 27, 1, 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, pp. 163-180

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

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

Johnson E.M. and Rehavi M.M. (2016), *Physicians treating physicians: Information and incentives in childbirth*, «American Economic Journal: Economic Policy», 8, 1, pp. 115-141, *http://www.aeaweb.org/articles.php?doi=10.1257/pol.20140160*

Keeler E.B. and Brodie M. (1993), *Economic incentives in the choice between vaginal delivery and cesarean section*, «The Milbank Quarterly», pp. 365-404

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

Lawthers A.G., Laird N.M., Lipsitz S., Hebert L., Brennan T.A. and Localio A.R. (1992), *Physicians' perceptions of the risk of being sued*, «Journal of Health Politics, Policy and Law», 17, 3, pp. 463-482

Lefèvre M. (2014), *Physician induced demand for c-sections: does the convenience incentive matter?*, Health, Econometrics and Data Group (HEDG) Working Papers, 14/08, pp. 1-22

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

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, 3

Murphy D.J., Macleod M., Bahl R. and Strachan B. (2011), *A cohort study of maternal and neonatal morbidity in relation to use of sequential instruments at operative vaginal delivery*, «European Journal of Obstetrics & Gynecology and Reproductive Biology», 156, 1, pp. 41-45

Norberg K. and Pantano J. (2016), *Cesarean sections and subsequent fertility*, «Journal of Population Economics», 29, 1, pp. 5-37

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

Towner D., Castro M.A., Eby-Wilkens E. and Gilbert W.M. (1999), *Effect of mode of delivery in nulliparous women on neonatal intracranial injury*, «New England Journal of Medicine», 341, 23, pp. 1709-1714

APPENDIX A: THE WORKING SAMPLE AND SCHEDULED PATIENTS

The working sample used in the main paper is restricted to only those unscheduled patients who at- tempt 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

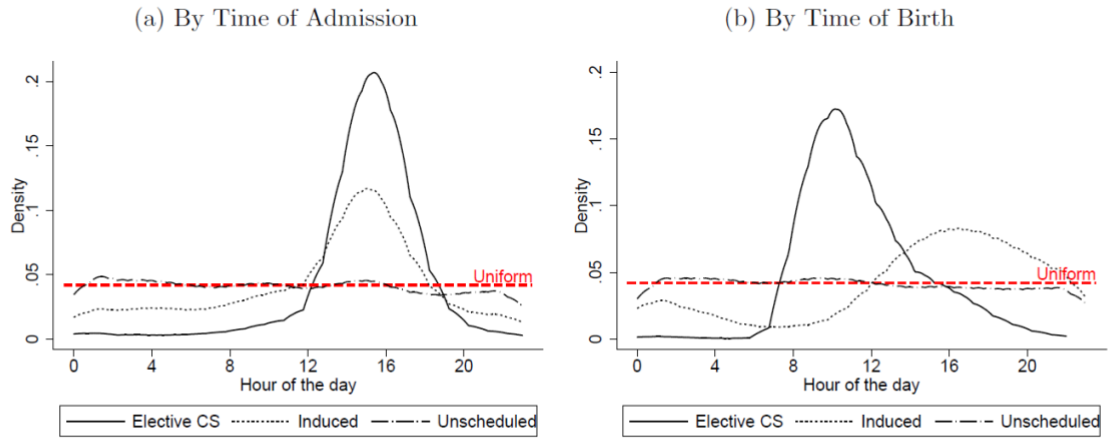
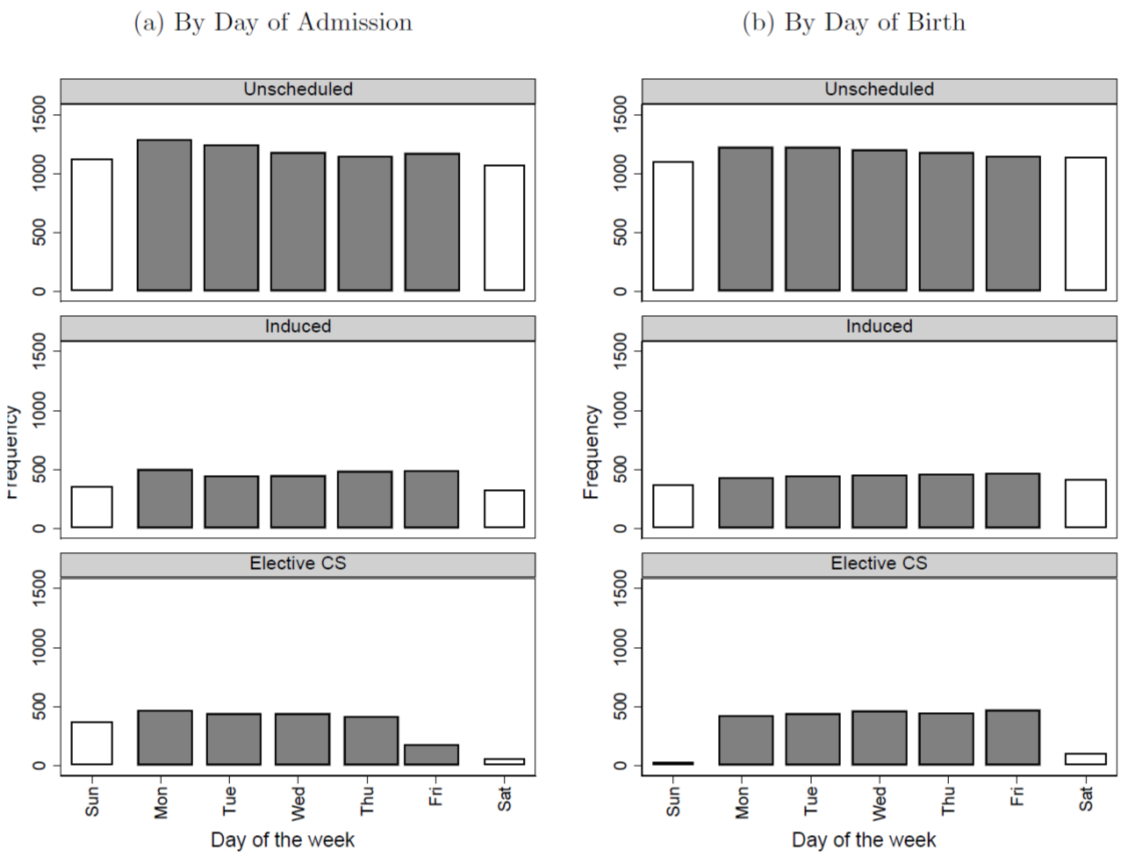


Figure A.2. Frequency of admissions and births by day



APPENDIX B: A MEASURE OF WORKLOAD WITHOUT ADJUSTING FOR SUPPLY SIDE FACTORS

The measure of workload used in the main paper is the ratio of patients-to-midwives, hence it takes into account both demand and supply side effects. Specifying the covariate of interest as a ratio may put some constraints on the estimated coefficient. This appendix repeats the main estimations but using instead the number of unscheduled patients waiting to give birth (without adjusting for the number of midwives).

Figures B.1 shows a histogram of the number of unscheduled patients observed by each patient at admission. The mode is 3, and the mean is slightly above at 3.34. As in the main paper, I divide this variable in quintiles to test for non-linearities in its effect on the probability of C-section. Table B.1 describes the number of observations and limits for each quintile.

Figure B.1. Histogram unscheduled patients

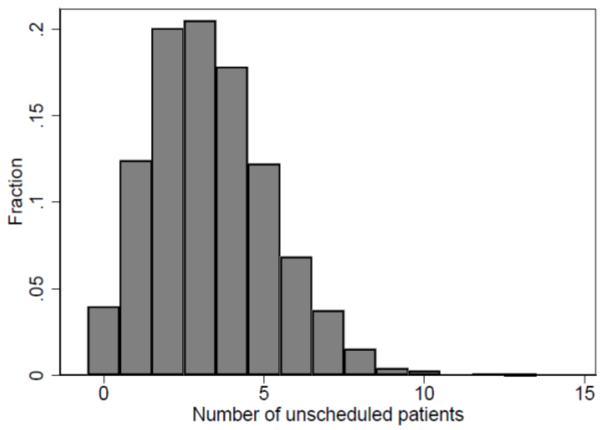


Table B.1: Descriptive statistics of quintiles

Quintile	N	Min	Max
1	953	0	2
2	536	3	3
3	466	4	4
4	320	5	5
5	338	6	13

Finally, Table B.2 presents the results from running the preferred model using the number of unscheduled patients as regressor. Similar to the findings in the main paper, there seems to be a non-linear relationship between workload and the probability of C-section. This effect starts to rise already in the second quintile and slowly declines in the fourth and fifth quintiles. This provides more assurance to the results using the ratio of patients-to-midwives.

Table B.2. Average marginal effect on probability of C-section

	Linear	Non-linear
Number of unscheduled patients	0.0044	
	(0.0036)	
2nd Quintile		0.0309*
		(0.0177)
3rd Quintile		0.0503**
		(0.0200)
4th Quintile		0.0384*
		(0.0223)
5th Quintile		0.0266
		(0.0212)
Observations	2,613	2,613
Mean dep.	0.119	0.119

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

APPENDIX C: 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 (1) to (4) in Table C.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 C.1. Alternative model specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Ratio	0.0004	0.0524**	0.1195**	0.1627		
	(0.0085)	(0.0240)	(0.0535)	(0.1283)		
Ratio Square		-0.0104**	-0.0369**	-0.0633		
		(0.0044)	(0.0182)	(0.0719)		
Ratio Cubic			0.0030*	0.0091		
			(0.0018)	(0.0157)		
Ratio Quadratic				-0.0005		
				(0.0011)		
2nd Quintile					0.0554***	
					(0.0199)	
3rd Quintile					0.0329*	
					(0.0196)	
4th Quintile					0.0299	
					(0.0211)	
5th Quintile					0.0167	
					(0.0216)	
20-80th Percentile						0.0408***
						(0.0156)
>80th Percentile						0.0197
						(0.0213)
Observations	2,613	2,613	2,613	2,613	2,613	2,613
AIC	1544.59	1542.58	1543.18	1545.06	1541.89	1539.65

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

APPENDIX D: OTHER GRAPHS AND TABLES

Figure D1-2. Distribution of admissions and births

Figure D.1: Density distribution of Ratio

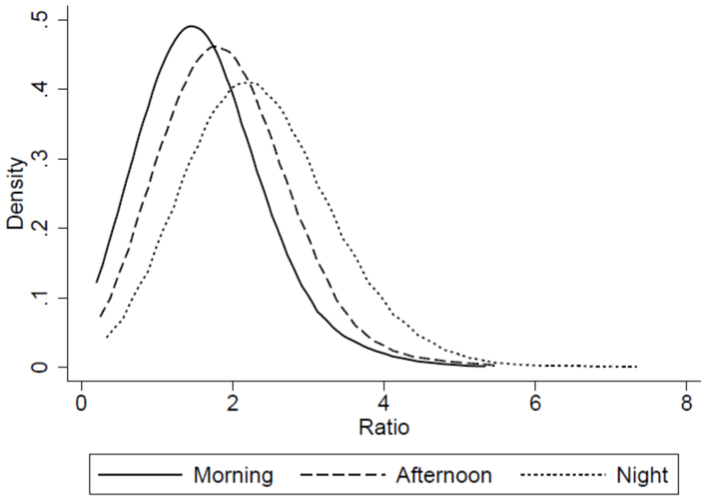
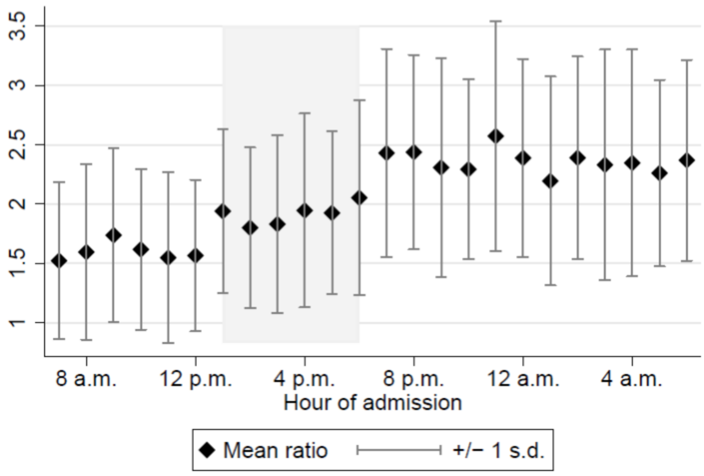


Figure D.2: Ratio by time of admission



Note: Shaded area marks the afternoon shift (1 pm – 7 pm)

Table D.1. Regression of pre-treatment characteristics on Ratio of patients-to-midwives

Dependent variable	Coef. of Ratio
Mother's characteristics	
With university degree	-0.0141
	(0.0129)
Above 36 yo	-0.0033
	(0.0117)
Pregnancy's characteristics	
Preterm (before 37th week)	-0.0080
	(0.0057)
At least 1 ER visit	-0.0005
	(0.0076)
Newborn's characteristics	
Male	-0.0045
	(0.0132)
Low weight at birth	-0.0054
	(0.0055)
Observations	2,613

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table D.2. Pre-treatment variables balanced across treatments and control

	Level of Ratio				
	<20th Percentile	20-80th Percentile	>80th Percentile	(1) vs. (2)	(1) vs. (3)
Mother's characteristics					
% of mothers with university degree	0.386	0.358	0.340	0.271	0.112
	(0.022)	(0.013)	(0.018)		
% older than 36 yo	0.302	0.288	0.266	0.567	0.182
	(0.021)	(0.012)	(0.017)		
Pregnancy's characteristics					
% of births before 37 weeks of gestation	0.057	0.058	0.040	0.942	0.191
	(0.011)	(0.006)	(0.008)		
% of pregnancies with at least 1 ER visit	0.110	0.123	0.100	0.431	0.597
	(0.014)	(0.009)	(0.012)		
Newborn's characteristics					
% of male newborns	0.508	0.517	0.497	0.755	0.704
	(0.023)	(0.013)	(0.019)		
% of low-weight newborns (<2,500 grams)	0.055	0.051	0.042	0.746	0.305
	(0.010)	(0.006)	(0.008)		

Standard errors in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$

Table D.3. Probability of C-section using a Linear Probability and Probit Model

	LPM	Probit
Panel (A):		
Ratio patients to midwives	0.0007	0.0012
	(0.0083)	(0.0081)
Panel (B):		
20-80th Percentile	0.0410***	0.0437***
	(0.0154)	(0.0166)
>80th Percentile	0.0187	0.0222
	(0.0210)	(0.0221)
Observations	2,613	2,613
Mean dep.	0.119	0.119

Robust standard errors in parentheses
$***p < 0.01$, $**p < 0.05$, $*p < 0.1$

Table D.4. LPM of C-section by day and staff shift

	All	Weekday	Weekend	Morning	Afternoon	Night
20-80th Percentile	0.0410***	0.0377**	0.0608*	0.0467*	0.0341	0.0242
	(0.0154)	(0.0176)	(0.0325)	(0.0270)	(0.0300)	(0.0245)
>80th Percentile	0.0187	0.0195	0.0251	0.0054	0.0477	0.0014
	(0.0210)	(0.0245)	(0.0429)	(0.0574)	(0.0523)	(0.0282)
Observations	2,613	1,883	730	636	641	1,336
Mean dep.	0.119	0.114	0.132	0.104	0.137	0.118

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table D.5. Effect of staffing for different windows of time since admission

	1 hour	2 hours	3 hours	4 hours	5 hours	6 hours
20-80th Percentile	0.0408***	0.0490***	0.0469***	0.0399**	0.0308*	0.0241
	(0.0156)	(0.0156)	(0.0163)	(0.0162)	(0.0164)	(0.0164)
>80th Percentile	0.0197	0.0213	0.0320	0.0136	0.0100	0.0077
	(0.0213)	(0.0209)	(0.0222)	(0.0220)	(0.0217)	(0.0216)
Observations	2,613	2,613	2,613	2,613	2,613	2,613

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table D.6. Effect ofstaffing at admission and 24hs after

	Workload at admission		Workload 24hs after admission	
Panel (A)				
Ratio patients to midwives	0.0019	0.0004	0.0020	-0.0010
	(0.0074)	(0.0085)	(0.0073)	(0.0081)
Panel (B)				
20-80th Percentile	0.0432***	0.0408***	-0.0019	-0.0065
	(0.0149)	(0.0156)	(0.0160)	(0.0161)
>80th Percentile	0.0193	0.0197	0.0011	-0.0091
	(0.0196)	(0.0213)	(0.0202)	(0.0213)
Time FE	✓	✓	✓	✓
Controls		✓		✓
Other patients		✓		✓
Shift*DOW FE		✓		✓

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table D.7. LPM using a continuous measure of staffing

	(1)	(2)	(3)	(4)
Ratio patients	0.0019	0.0038	0.0046	0.0007
to midwives	(0.0074)	(0.0074)	(0.0074)	(0.0083)
University degree		-0.0178	-0.0180	-0.0167
		(0.0134)	(0.0134)	(0.0135)
Old (>34 years old)		0.0489***	0.0492***	0.0510***
		(0.0153)	(0.0153)	(0.0153)
Male		0.0351***	0.0352***	0.0348***
		(0.0125)	(0.0125)	(0.0126)
Preterm (<37 wofg)		0.0079	0.0089	0.0055
		(0.0326)	(0.0326)	(0.0329)
Low weight at birth (< 2.5kg)		0.0284	0.0287	0.0318
		(0.0350)	(0.0350)	(0.0348)
ER visit		0.0062	0.0066	0.0116
		(0.0201)	(0.0201)	(0.0202)
>6 hours waiting		-0.0052	-0.0031	-0.0028
		(0.0148)	(0.0155)	(0.0156)
>14.2 hours waiting		0.0458***	0.0516**	0.0496**
		(0.0160)	(0.0204)	(0.0210)
Observations	2,613	2,613	2,613	2,613
Yea, Month, Dow FE	✓	✓	✓	✓
Controls		✓	✓	✓
# Scheduled CS			✓	✓
Hour FE				✓

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$

Table D.8. LPM using a categorical variable for staffing

	(5)	(6)	(7)	(8)
20-80th Percentile	0.0530***	0.0545***	0.0555***	0.0555***
	(0.0155)	(0.0154)	(0.0155)	(0.0166)
>80th Percentile	0.0413**	0.0467**	0.0485**	0.0477**
	(0.0187)	(0.0187)	(0.0188)	(0.0214)
University degree		-0.0172	-0.0175	-0.0162
		(0.0133)	(0.0133)	(0.0135)
Old (>34 years old)		0.0490***	0.0494***	0.0507***
		(0.0153)	(0.0153)	(0.0153)
Male		0.0350***	0.0351***	0.0346***
		(0.0125)	(0.0125)	(0.0126)
Preterm (<37 wofg)		0.0070	0.0079	0.0049
		(0.0324)	(0.0324)	(0.0327)
Low weight at birth (< 2.5kg)		0.0294	0.0296	0.0328
		(0.0350)	(0.0350)	(0.0347)
ER visit		0.0051	0.0055	0.0103
		(0.0201)	(0.0201)	(0.0201)
>6 hours waiting		-0.0045	-0.0023	-0.0017
		(0.0148)	(0.0155)	(0.0156)
>14.2 hours waiting		0.0471***	0.0532***	0.0523**
		(0.0159)	(0.0204)	(0.0209)
Observations	2,613	2,613	2,613	2,613
Yea, Month, Dow FE	✓	✓	✓	✓
Controls		✓	✓	✓
# Scheduled CS			✓	✓
Hour FE				✓

Robust standard errors in parentheses. $***p < 0.01, **p < 0.05, *p < 0.1$