

Bensnes, Simon

Working Paper

Time to spare and too much care: Congestion and overtreatment at the maternity ward

Discussion Papers, No. 963

Provided in Cooperation with:

Research Department, Statistics Norway, Oslo

Suggested Citation: Bensnes, Simon (2021) : Time to spare and too much care: Congestion and overtreatment at the maternity ward, Discussion Papers, No. 963, Statistics Norway, Research Department, Oslo

This Version is available at:

<https://hdl.handle.net/10419/250130>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Time to spare and too much care

Congestion and overtreatment at the maternity ward

TALL

SOM FORTELLER

DISCUSSION PAPERS

963

Simon Bensnes

Simon Bensnes

Time to spare and too much care
Congestion and overtreatment at the maternity ward

Abstract:

Identifying the causal effect of resource use on health outcomes is generally complicated by endogenous supply and demand adjustments. This paper tackles these issues in the setting of the maternity ward using the number of women in local areas with the same due date as an instrument for congestion. I find that congestion leads to both fewer and less invasive interventions and better health outcomes, indicating medical overtreatment during slower periods. I also show that absent instrumentation I find similar results and similar signs of bias as the related literature on congestion in maternity wards.

Keywords: Healthcare, crowdedness, congestion, maternity

JEL classification: I12, I14, I18

Acknowledgements: I am grateful for comments from Victoria Marone, Sturla Løkken, Ingrid Huitfeldt, Kjetil Telle and seminar participants at Statistics Norway and ASHEcon 2021. I gratefully acknowledge financial support from the Norwegian Research Council (grant no. 256678)

Address: PB 2633 St. Hanshaugen 0131 Norway, Statistics Norway, Research Department. E-mail: sdb@ssb.no

Discussion Papers

comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

The Discussion Papers series presents results from ongoing research projects and other research and analysis by SSB staff. The views and conclusions in this document are those of the authors.

© Statistics Norway
Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
<http://www.ssb.no/en/forskning/discussion-papers>
<http://ideas.repec.org/s/ssb/disvap.html>

ISSN 1892-753X (electronic)

Sammendrag

Det er generelt komplisert å etablere en kausal effekt mellom ressursbruk og utfall i helsesektoren på grunn av endogene tilpasninger på både tilbuds- og etterspørselssiden. For eksempel gjennom at sykere pasienter kan motta flere og dyrere behandlinger. Denne artikkelen håndterer denne utfordringen gjennom å fokusere på førstegangsfødende ved Norske fødeavdelinger kombinert med å instrumentere for antall kvinner som blir skrevet inn på hver dag. Instrumentet blir konstruert ved å først identifisere hvilken fødeavdeling som vanligvis benyttes innenfor hver enkelt av norske grunnkretser, som brukes til å klassifisere opptaksområder for hver fødeavdeling. Deretter definerer jeg instrumentet som antall kvinner med samme termindato innenfor disse opptaksområdene. Resultatene viser at kvinner som blir skrevet inn på fødeavdelingen på travlere dager opplever færre og mindre inngripende behandlinger, og barna som blir født av disse kvinnene har bedre APGAR-skåre, har lavere reinnleggelsesrater de første 7 og 30 dagene etter fødsel, og er sjeldnere innom sykehuset de første leveårene. Det er ingen indikasjoner på at mors helse forverres i tiden rett etter fødsel målt ved lengde på opphold eller komplikasjoner. På lengre sikt finner jeg ingen indikasjoner på at kvinner som blir skrevet inn i travlere perioder benytter flere helsetjenester. Robustheten av resultatene blir utforsket og utfordret på en rekke måter. De videre analysene viser at store fødeavdelinger blir langt mindre påvirket av travelhet; mødre med hyppigere kontakt med helsevesenet og mødre som er registrert gifte påvirkes mindre enn andre. Samlet viser denne studien at det ikke er noen negativ helseeffekt for kvinner som ankommer fødeavdelingen under spesielt travle perioder. Snarere ser det ut til at barna disse kvinnene føder har en positiv helseeffekt av travelhet. En mulig årsak til dette er overbehandling av fødende ved rolige perioder.

1 Introduction

Healthcare utilization varies within provision systems and is affected by the supply and demand side in about equal measure, both in the US and in Europe (Finkelstein et al., 2016; Godøy and Huitfeldt, 2019). Efforts to uncover the underlying mechanisms driving variations in utilization have produced a substantial literature identifying important factors on the supply side, such as financial incentives and practice styles (e.g. Skinner, 2011). This paper adds to the literature on supply side determinants of healthcare provision by focusing on the role of resource availability through the lens of exogenous short-term variations in congestion.

Observing rising costs in the healthcare sector, a recurring question in economics is whether healthcare is increasingly provided at the “flat of the curve” where the marginal return to health is negligible or even negative. Some argue that such overtreatment may be widespread and potentially a major contributor to rapidly increasing healthcare spending (Wennberg, 2010).¹ Overtreatment can occur because healthcare providers at the point of use (e.g. physicians) often have few economic incentives to reduce costs, such that the marginal cost of treatment equals the marginal value in terms of health improvement: First, there may be financial incentives to provide additional treatment beyond the efficient level, rather than minimizing for given desired patient outcomes. Second, even if a treatment is budget neutral for the provider at the point of use, excess capacity may in itself lead to additional treatment provision (Freedman, 2016). Providers may, for example, choose to provide additional care if they have spare capacity, under the assumption that the marginal medical treatment has a non-negative health return and potentially increases patient utility.

Whether or not overtreatment is an important factor behind rising costs, it is intrinsically difficult to identify its prevalence or even the return to healthcare treatment and spending overall due to endogeneity. On the supply side, providers may increase resources used on unobservably sicker patients, for instance by admitting them on slower days, when staffing

¹It is obvious that patients may value aspects of treatment other than purely health outcomes. If so, the utility maximizing level of care will deviate from the health optimizing level of care. I will refer to “overtreatment” in the sense that treatment provides a non-positive return to patient health.

is relatively higher and more equipment is available. Such supplier behavior could lead to a negative association between spending and health outcomes. On the demand side, healthier patients may sort toward certain suppliers, demand more care or be able to extract more resources from suppliers, which would strengthen any positive correlation between spending and health outcomes.

This paper provides new evidence of the causal effect of resource availability on medical treatment and health outcomes, and shows that overtreatment occurs even without direct financial incentives. To circumvent endogeneity issues, I focus on exogenous variation in short-term congestion at Norwegian maternity wards. I define congestion as the number of women who are admitted to the same ward on the same day. To eliminate threats to identification I instrument for congestion by predicting congestion levels using ward catchment areas and due dates.

An attractive feature of using short-term fluctuations in congestion at maternity wards is that it combines the use of a large and generally healthy population with a policy-relevant source of variation. Assuming that wards cannot perfectly adjust to these shocks, unforeseen congestion leads to a direct reduction in the resources available to each patient.

Results show that short-term variations in resource availability due to congestion have strong effects on healthcare provision during birth. Specifically, women who are admitted on busier days receive fewer and less invasive interventions while giving birth, including the use of analgesics and labor inductions. In addition, I find that newborns are less likely to receive poor APGAR scores and be readmitted the first month after birth.² This latter finding is evidence suggesting that some of the medical interventions induced by reduced congestion levels actually have a negative health return for the child.

Studying the health outcomes of mothers and children up to 3 years after birth, I find little evidence of increased healthcare utilization for mothers who are admitted on more congested days. Assuming that healthcare utilization primarily proxies healthcare needs, this finding suggests that the additional treatment provided due to less congestion has no significant

²APGAR is an acronym for Activity, Pulse, Grimace, Appearance, and Respiration. Each of these are given a score from 0 to 2, summing up to a max of 10.

effect on longer-term health outcomes.

Detailed administrative data also allows me to provide some evidence as to which patients are affected when resources become scarcer. For instance, in line with previous evidence, I find that married women are less affected than non-married women, which may indicate that patients with more bargaining power are less affected as argued by [Facchini \(2019\)](#) who also study resource use in maternity wards.

This paper has four main contributions. First, the paper shows how policy-relevant variations in available healthcare resources play an important role in determining healthcare provision for a general population in a common setting, the maternity ward. Though the principle that inputs (healthcare resources) affect output (healthcare provision) is predictable from a theoretical perspective, the empirical literature has not reached a consensus on the importance of this mechanism. In addition, many past contributions are confined to narrow procedures or patients in life-threatening situations, as opposed to giving birth, which a large share of the population experiences (e.g. [Doyle J. et al., 2015](#); [Woodworth, 2020](#)).

Second, I show that overtreatment, in the sense that the marginal return to health is non-positive, occurs even when physicians have no monetary incentive to increase the amount of care. This finding expands our current understanding of how overtreatment may occur and is informative for shaping policy to reduce its prevalence.

Third, I use an instrumental variable approach, which address important endogeneity challenges. Such issues have generally been difficult to resolve in some of the past literature, which has often has relied on a fixed effects approach controlling for time and provider effects ([Freedman, 2016](#); [Maibom et al., 2020](#)). I show that using a fixed effects framework as an alternative to the IV strategy yields different results in my setting. Some of these differences may be caused by heterogeneous treatment effects, but I find little supporting evidence for that. Rather, I find some evidence of imbalance where parental characteristics are correlated with congestion also conditional on ward-by-date fixed effects. Thus, the findings suggests that the identifying assumptions in the fixed effects framework may fail when studying the effects of healthcare resources.

Last, many past contributions have had little opportunity to explore

treatment provision beyond spending. In contrast, I am able to consider specific interventions at a highly detailed level and measure patient utilization of healthcare services for years after birth.

The remainder of the paper is organized as follows: Section 2 relates and contrasts this paper to the relevant literature, while Section 3 presents the institutional framework and the data acquired. Section 4 presents the instrument along with tests of validity and relevance. In Section 5 I present and challenge the results, while I discuss the findings in Section 6.

2 Relation to literature

The fact that patients with the same health conditions and the same access to healthcare receive different treatment has been widely observed, but is not yet widely understood (Skinner, 2011). While past evidence has shown that patients' non-medical characteristics, preferences and so on are central factors, there are also important determinants on the supply side. For example, Molitor (2018) show that cardiologists who move to hospitals with different practice techniques change their treatment behavior and become more similar to their new colleagues. Molitor (2018) further argues that 60-80% of regional differences in physician behavior is explained by environmental factors.³

A large literature is devoted to estimating the causal effect of healthcare resources on patient outcomes. Using quasi-experimental variation, Doyle J. et al. (2015) find that patients who are referred by ambulance to high-cost hospitals have better health outcomes than otherwise similar patients who are referred to low-cost hospitals. Focusing on emergency patients admitted while away from home and their usual healthcare provider, Doyle (2011) find that patients who are exposed to higher-spending health-

³Economic incentives can also be an important factor on the supply side and has been studied previously, also in the setting of maternity wards (e.g. Chandra et al., 2011; Almond and Doyle, 2011). Economic incentives can also interact with patient characteristics: Johnson and Rehavi (2016) explore how patient knowledge interacts with financial incentives for caesarian sections (C-sections). In particular, they find that women with a medical degree are less likely to have a C-section, and are even less likely to do so in health maintenance organization (HMO) hospitals, where the hospital internalizes the additional costs of overtreatment. However, midwives and physicians in Norwegian maternity wards have no clear financial incentive as wards are 100% publicly financed.

care systems have better outcomes compared to patients admitted to lower-spending healthcare systems. At the same time, [Gowrisankaran et al. \(2017\)](#) find that skilled physicians use more resources, but these physicians also have worse patient outcomes. Given that institutional settings, conditions, patients, and identification strategies vary across papers analyzing the marginal return to healthcare spending, there is no strong consensus as to how widespread wasteful spending is. In particular, the prevalence of wasteful spending may be directly influenced by institution-specific monetary incentives, or lack thereof.

In order to identify the effect of resources on treatment decisions and patient outcomes, it is necessary to identify exogenous shocks to resource availability which is not a function of patient needs. Unforeseen congestion has served this purpose in several past contributions because it is a form of short-term strain on providers which effectively reduces the resources available to each patient. Few parts of the healthcare system offer as good a setting for studying the effects of crowding as the emergency department (ED): The inflow and health status of patients vary substantially, and there are limited options for physicians to reschedule examinations or refer patients to other institutions (e.g. [Evans and Kim, 2006](#); [Gruber et al., 2018](#); [Woodworth, 2020](#)). The findings from this literature vary. [Evans and Kim \(2006\)](#) use variations in staffing at Californian EDs and find little or no evidence of negative effects of congestion. However, [Woodworth \(2020\)](#) use the opening of new EDs as shocks to crowding and find large reductions in mortality when crowding is reduced.

As with EDs, most patients admitted to maternity wards are in need of immediate care and monitoring. The influx of patients and their health status also varies considerably. For these reasons, several papers have studied treatment decisions and outcomes in maternity wards. [Freedman \(2016\)](#) analyze how bed availability in the neonatal intensive care unit affects the likelihood that neonatals are admitted there. For identification, the author relies on short-term variations in the number of available beds at the neonatal intensive care unit within a hospital-year-month. Using data from California and New York, the study shows that for the sickest infants, capacity has no effect on admission probability. However, for healthier babies, capacity has a direct impact on the probability of admission.

Using an identification strategy similar to [Freedman \(2016\)](#), a concurrent study, [Maibom et al. \(2020\)](#), estimates how the number of mothers admitted to the maternity ward affects procedures and health outcomes in Danish hospitals. The authors find that an increase in the number of women admitted reduces the probability that the mother is exposed to medical interventions, but they find no effects on APGAR scores or longer-run health outcomes. [Marks and Choi \(2019\)](#) ask whether hospital spending affect infant health. They propose to use the number of births at hospitals as a measure for crowdedness and an instrument for spending. They conclude that when hospitals are forced to reduce spending on newborns due to congestion, the reduction is enforced in such a way that there is no negative effect on newborn health. [Facchini \(2019\)](#) considers day-to-day variations in the ratio of midwives to patients in an Italian hospital and find that during congested periods, mothers are more likely to undergo an unscheduled c-section. The author argues that this is because c-sections are surgical procedures performed by a surgeon and relatively more c-sections therefore reduce the workload for midwives.

Common to much of the past literature is that identification often rely on residual variation in congestion levels after controlling for provider and time fixed effects. Two central underlying assumptions are that patients' ability to sort to providers is independent of unobservable patient characteristics, and that providers have a negligible opportunity of affecting the timing of admissions and which patients that are admitted. In contrast, identification in this paper relies on an IV strategy which does not have to rely on these assumptions. The application of the IV strategy constitutes a central aspect of this paper's contribution. The following illustrates the challenges presented by the fixed effects in more detail, using the framework in this paper.

First, the fixed effects framework commonly used in the literature rests on the identifying assumption that institution fixed effects remove all cross-sectional bias from patients' allocation to providers. The assumption fails if patient sorting depends on congestion levels at a given time and unobserved patient characteristics. Using the context of this paper as an example, mothers choose a maternity ward for the birth. While I cannot observe the choice of ward directly in my data, 13% of women in the sam-

ple do not give birth at the ward closest to their home. The probability of not giving birth at the closest ward is correlated with various observable sociodemographic variables, and women who give birth at wards further away also experience lower levels of congestion.⁴ Experienced congestion levels are therefore directly correlated with factors which are known to affect outcomes, such as education and maternal age. This is not necessarily an issue, assuming that it is possible to control for all patient characteristics that impact patient sorting. However, this is an untestable assumption, because it is not possible to rule out the possibility that unobserved factors, such as other health conditions or complications during pregnancy, influence where mothers end up giving birth. One could easily imagine that deliveries that are expected to be complicated are more likely to be diverted to a ward with less congestion, outside the catchment area. Were this to be the case, biased estimates would show that lower congestion levels lead to more interventions and more complicated births. The sorting of patients to providers is a general challenge in much of the literature. As long as the choice of provider is endogenous and correlates with unobservables, controlling for providers is not sufficient to remove selection bias.

Second, a similar issue occurs in the time dimension. One can include time effects, for instance date of admission, but as with the provider fixed effects, an implicit assumption is that the timing of all admissions cannot be influenced by either patients or providers based on unobservable characteristics. This assumption fails if providers are able to affect the timing of admission of patients based on characteristics, unobservable by the econometrician, that are correlated with both outcomes and congestion. In the maternity ward setting, midwives and obstetricians have significant leeway in deciding the timing of births through induction, referral to other wards, and delaying admissions. A potential consequence is that mothers scheduled to be admitted for birth induction are not admitted due to an unforeseen influx of mothers in active labor. If fixed effects are relied on for identification, this could represent a challenge to identification because birth induction is per definition less likely for women in active labor, while mothers who are not admitted on that day are much more likely to go

⁴See Table A1 in the Appendix.

through induction at a later stage. Consequently, the estimates may not reflect the effect of congestion on treatment provided to the same population sample, but rather the fact that congestion in itself affects which women are admitted on a specific date. Outcomes other than induction are also likely to be affected by this mechanism, as inductions are usually accompanied by additional complications and medical interventions.

These challenges need not necessarily add up to a severe threat to the validity of the results of any past contributions. There is, however, evidence of imbalance reported in several of the contributions discussed above.⁵ Using an IV strategy, I am able to address all of these issues. Specifically, my instrument for congestion effectively removes all issues stemming from geographical sorting, as discussed in the first point. I solve issues of inter-temporal sorting of mothers by using the due dates of other women rather than admission dates. In particular, the due dates of other women are outside the locus of control for both wards and neighbors, and therefore not endogenous to the admission date, which solves the second issue above. Last, I show that my instrument is uncorrelated with any parental characteristics, which is evidence that the instrument fulfills the exogeneity criterion.

3 Institutions and data

3.1 Institutions

The Norwegian healthcare system is dominated by a single-payer system financed by the national government. It is not possible for individuals to fully opt out of the public healthcare system, but individuals may buy private health insurance in addition.⁶ The public healthcare system is

⁵Maibom et al. (2020) show that parents' income and educational levels is correlated with their measure of crowding and a correlation between transfers and congestion in some samples. Freedman (2016) find that both parental and newborn characteristics are correlated with his measure available neonatal care beds. Marks and Choi (2019) find significant differences for several socio-economic background characteristics for mothers giving birth on days characterized as "slow" and "busy". Note that the authors in all these studies argue that the size of the differences are too small to bear any economic significance, yet, it demonstrated that some imbalances are common in the past literature.

⁶Approx. 10% of the Norwegian population is covered by some kind of private health insurance according to Finans Norge (2019). Of these 90% are covered through their

organized in two tiers. The first tier consists of general practitioners (GPs), who are contracted to a municipality. All inhabitants are initially assigned a GP at birth or immigration, but may change to another physician up to twice a year. Patients seeing a GP are required to pay approx. USD 20-25, adjusted for inflation, per consultation. Additional minor co-payments are required for blood tests etc. When the yearly sum of co-payments exceeds a limit value (USD 250 in 2019), patients are exempt from additional co-payments for the remainder of the calendar year.⁷ In addition to providing diagnostic services and outpatient treatment, GPs also make referrals to specialist healthcare services from public or private providers. As a general rule, patients cannot consult specialists or get services at hospitals with public funding without a referral from a GP.⁸

The second tier of the healthcare system consists of private specialists contracted to the public healthcare system and public hospitals. Patients may buy services from private specialists such as imaging services, etc., but without a referral from their GP, they have to pay the entire cost of the service out-of-pocket. Inpatient services at public hospitals are free of charge at the point of use. This includes all services related to birth. For all inpatient health services, patients are allowed to select the institution at which they wish to receive care. Most of Norway is fairly sparsely populated, with long distances between hospitals. On the whole, then, the choice of institution is therefore not a relevant choice margin for most women when it comes to childbirth. However, in Oslo expectant mothers may choose among several institutions within a small geographical area. In order to choose a ward, expectant mothers can fill out and submit a form to the institution they prefer and may then be accepted or rejected by the institution. During unforeseen congested periods wards are free to refer expectant mothers to other institutions in the early phases of labor as long as this does not potentially compromise the health of the child or the mother. I address the empirical issue of selection of women to institutions in Section 4 below.

employer.

⁷Income for GPs is generated through three channels: (i) co-payments from patients; (ii) reimbursements from the national government through the Control and Payment of Health Reimbursement (KUHR); (iii) a fixed sum per patient on their patient list.

⁸The main exemption is patients who are referred for further treatment from the ED.

During a healthy pregnancy, the GP is the main point of contact for expectant mothers. Expectant mothers may also choose to have regular check-ups, where available, with a certified midwife employed by the municipality. During a normal pregnancy an expectant mother will undergo 8 antenatal check-ups before reaching her due date (helsenorge.no, 2020). These check-ups are exempt from co-payment and include clinical tests, scans, and conversation about health and risk behaviors. Additionally, pregnant women will be offered an ultrasound scan by designated providers free of charge at an approximated gestation age of 18 weeks. This ultrasound scan includes setting a final due date. Importantly, the first check-up includes guidance on how the parents should book a midwifery unit or hospital for delivery.

3.2 Data and sample

Data for this study were collected from four sources. First, the Medical Birth Registry of Norway (MFR) contains health and administrative data on all births in Norway from 1967 onward. From this register I collect several variables on the individual child’s health and procedures and interventions administered during birth.⁹ The register also contains identifiers for the institution in which the mother gave birth, but does not contain the date of the mother’s admission to the institution.¹⁰ Importantly, these data contain the due date for each birth as determined through ultrasound earlier in the pregnancy.

The second source of data is the National Patient Registry of Norway (NPR). This registry records every treatment, admission date and discharge date for all individuals receiving inpatient treatment at a public hospital or other providers since 2008. This registry contains almost all births in Norway in the period.¹¹ These data also contain the day of admission for all mothers as well as medical codes for each procedure administered,

⁹Procedure and complication codes follow several different standards and include the ICD-10/9, ATC, and complications and interventions identified from a checklist from the birth registration form filled out at the ward.

¹⁰The MFR does contain birth date as well as number of days admitted, but these variables are not sufficient to identify the admission date.

¹¹Home births are not necessarily recorded, but nearly all women give birth at a ward or en route to a ward and are then registered at a public ward shortly after delivery.

coded according to the ICD-10 system at the two digit level. I use these data to calculate the number of women admitted to each institution each day. Specifically, for each woman who gives birth, I first find the spell that overlaps the birth date of the child and include any ICD-10 code which refers to the birth (ICD-10 O80-O84). If no spell overlaps the date of birth I use the closest ensuing spell that contains ICD-10 codes related to childbirth, but restrict these to spells starting within 2 days of the birth. Using these two rules, I identify the day of admission for 98% of the sample between 2008 and 2014. I also use these data to identify the most common maternity ward in each neighborhood for each prospective mother.

The third data source is the Control and Payment of Health Reimbursement, a registry over all reimbursements and payments in public healthcare. This registry contains detailed information on each contact any individual has with the public out-patient healthcare system. This includes date, reason for visit and treatment, if any. Though Norway has a small and growing private healthcare sector, nearly all treatment and contacts between patients and the healthcare system go through the public points of service. I use these data to construct measures of healthcare utilization of both mother and child in the years following birth.

The fourth and last source of data are the sociodemographic register data of Statistics Norway. These include detailed data on each individual's education level, municipality and neighborhood of residence, labor market data, earnings and more. First and foremost, these data allow me to include controls in the main analyses and investigate whether the instrument is balanced. Descriptive statistics for all variables included in the analysis are presented in Tables [A2](#) and [A3](#).

In the analysis I restrict the sample to women giving birth for the first time.¹² This is because past birth outcomes may have direct impacts on treatment given during later births. For example, women who have had C-sections before are advised to avoid vaginal births later due to increased risk of serious complications ([Royal College of Obstetricians & Gynaecologists, 2019](#)). I also exclude mothers younger than 16 or older than 40 at the time of birth. This leaves me with 148 022 births across 40 wards. Descriptive

¹²I show results when this restriction is relaxed in Table [A11](#) in the Appendix.

Table 1. Selected descriptive statistics: Parental background and birth statistics.

	Mean	Std. Dev.
Mothers admitted	11.22	7.82
Predicted mothers admitted (instrument)	11.24	8.37
Child female	0.48	
Interventions and complications		
Number of interventions	4.41	2.16
Number of complications	5.48	1.44
APGAR. < 7	0.02	
Induction	0.22	
Analgesic	0.94	
Emergency C-section	0.13	
Child readmitted 30 days	0.05	
Paternal characteristics		
Pensionable income (000's)	447.47	322.21
Age	31.06	6.09
Married	0.18	
Maternal characteristics		
Pensionable income (000's)	314.09	198.97
Age	27.95	4.80
Married	0.18	
Diagnoses related to pregnancy prior to birth	3.06	5.56
Pregnancy related contacts with healthcare providers	13.92	9.51

Note: Pensionable income is CPI adjusted to 2015. Descriptive statistics for all variables are reported in Table A2. N = 148 022.

statistics on selected variables are reported in Table 1.¹³

4 Identification strategy

Establishing a causal chain from congestion to health outcomes is generally challenging, as there are a number of unobserved factors which are likely to influence both patient exposure to congestion on the one hand, and treatments and health outcomes on the other. In the specific case studied in this paper, there are several institutional factors which would warrant concern for selection of mothers to wards and dates based on congestion.¹⁴

¹³See the Appendix for descriptive statistics on all variables used in the analysis.

¹⁴The structure of this section builds on the analogous section in Bhuller et al. (2020).

First, both mothers and wards have some possibility of influencing where the birth will take place. In particular, mothers state their preferred maternity ward, and will give birth there provided that the ward has capacity at the time of admission. As mothers may sort toward different wards on a number of different dimensions, this implies clear selection issues. Second, wards are able to refuse and reroute incoming mothers during times of unforeseen congestion. This kind of referring could bias results both towards and away from zero. On the one hand, wards may be more likely to re-route healthy mothers during congested periods because they are at lower risk of complications during transport. On the other hand, there may be a greater possibility of high-risk mothers being re-routed during congested periods if a ward does not have the capacity to manage a complicated birth. Additionally, which ward mothers are referred to may depend on their risk level. For instance, if pregnancies are considered at-risk at an early stage, mothers living further away from central hospitals may be admitted to these better equipped wards well ahead of the due date to ensure that they will not be in transport at the time of birth or give birth at smaller wards.¹⁵ Wards may also differ as to how they assess incoming mothers and who they refer to other institutions. These assessments may also differ over time and within the individual ward, depending on the characteristics of mothers currently admitted. Depending on the degree of and reasons for rerouting, these mechanisms may bias observational results either upwards or downwards. There are indications that these mechanisms are in play also in past contributions as discussed above.

To demonstrate the presence endogeneity problems in the data used here, I estimate a simple OLS regression of the number of other women admitted at the same ward and day as the focal mother against a set of characteristics linked to the focal mother. The results are reported in Column (1) of Table 2. Even though I include fixed effects for the due date and local ward identifiers, there are strong and significant correlations between several background characteristics and experienced congestion, indicating clear selection patterns. In particular, higher parental education

¹⁵In the sample, 13% of women give birth at an institution other than the one closest to their home.

and maternal age are both positively correlated with congestion levels. The imbalances are reduced, but remain also if I exchange due date and local ward fixed effects with admission date and admitting ward-by-admission year fixed effects. These results are reported in Table A4. This suggests that the simple FE estimates are likely to be biased.

To handle the endogeneity issue, I use an instrumental variable strategy. In essence, I rely on inter-temporal variation in the number of women who have the same due date within a local ward catchment area. Because no individual mother or ward can influence the due dates of mothers in a geographic area, the IV estimate should be free from bias due to endogenous responses to congestion. To construct the instrument, I start by identifying the ward most commonly used in each woman’s neighborhood and year-month. These refer to the statistical unit “grunnkrets” which is the least aggregated geographical statistical unit in Norway, and comprises on average 350 individuals. I then sum up the number of women in these catchment areas by due date. I term this measure “predicted congestion”. Deviations between observed and predicted congestion stem from women being admitted either before or after their due date or being admitted to other wards.

The average number of women admitted each day at a given ward is likely to be correlated with time-invariant ward characteristics. For instance, wards in more densely populated areas are likely to have larger wards and serve a larger patient group. The same wards may also have different routines or leniency towards specific treatments. I therefore condition all estimates on ward fixed effects.¹⁶ I also include due date fixed effects to account for differences over time that are common across wards. Such differences include seasonal variation in birth numbers and the potential effects of vacations and weekends on capacity. Together, these fixed effects net out a large fraction of ward and time effects that may affect both congestion and the treatment of mothers.

The main question asked in this paper is whether congestion in the maternity ward affects health outcomes for mother and child. The model

¹⁶These fixed effects are dummies for the local ward of each woman as opposed to the ward where the birth actually takes place. The reason for this is that the ward they end up giving birth at is not determined randomly.

Table 2. Background variables and experienced vs predicted congestion.

	(1)		(2)	
	Congestion: $\#admitted$		Predicted congestion: $\#\hat{admitted}$	
Father				
Lower secondary	-0.2520***	(0.0456)	-0.0248	(0.0324)
Upper secondary	-0.2572***	(0.0391)	-0.0259	(0.0277)
Master's	0.1143**	(0.0580)	-0.0490	(0.0405)
Age	-0.0017	(0.0030)	0.0011	(0.0022)
Immigrant	0.1206**	(0.0502)	0.0065	(0.0353)
ln(income)	-0.0052	(0.0104)	-0.0009	(0.0073)
Mother				
Lower secondary	0.0188	(0.0417)	0.0137	(0.0301)
Upper secondary	0.0628*	(0.0354)	0.0306	(0.0254)
Master's	0.3358***	(0.0523)	-0.0054	(0.0368)
Age	0.0242***	(0.0043)	-0.0003	(0.0031)
Immigrant	-0.1495***	(0.0491)	0.0009	(0.0349)
ln(income)	0.0000	(0.0095)	-0.0024	(0.0068)
Child female	-0.0100	(0.0268)	0.0098	(0.0192)
Observations	148 022		148 022	
Dep. mean	11.22		11.24	
F-statistic for joint significance	23.93		0.34	
Joint p-value	0.00		0.99	

Note: Column (1) shows the result of a linear regression where the dependent variable is the number of women admitted on the same maternity ward-day as the focal mother. Parental income is transformed using the inverse hyperbolic sine to handle zeros. In addition to the listed variables, the regression includes due date and local ward fixed effects. Column (2) shows the same result for the instrument. In the bottom part of the table I report the F-statistic and p-value for a test of joint significance of the reported variables. Robust standard errors in parentheses.

applied to answer this question is formulated in Equation 1, where y_{ihd} is the outcome (indicator for induced labor, longer-term health outcome etc.) for woman i at ward h admitted on date d .¹⁷ $Admitted_{hd}$ is the number of women admitted to the same ward on the same day as the focal mother is admitted. $\mu_{\hat{h}}$ are fixed effects for the mother's local ward. $due\ date_d$ is the due date fixed effects. X_i are characteristics of the mother, father and child that are determined prior to birth. For both mother and father these include: years of education, income and income squared one year prior to birth, age, and immigration status. For the child, the included

¹⁷When the health of the child is the outcome, i refers to the child.

characteristic is the sex of the child.¹⁸

$$y_{ihd} = \beta \widehat{admitted}_{hd} + \mu_{\widehat{h}} + \eta \widehat{due\ date}_{id} + X'_i \Xi + \varepsilon_{ihd} \quad (1)$$

As demonstrated by Table 2, the number of mothers admitted to a ward on any given day is not random when conditioning on fixed effects. It is thus likely that a simple OLS regression would yield a biased estimate of β . Acknowledging this concern, my identification strategy relies on random variation in the congestion level caused by the number of women with due dates on specific calendar dates within a ward catchment area, $\sum_{i=1}^N \widehat{due\ date}_{hd-i}$. Summing up the number of women with the same due date within each catchment area, I can instrument for the number of women admitted on any day and at any ward, using a first stage defined as:

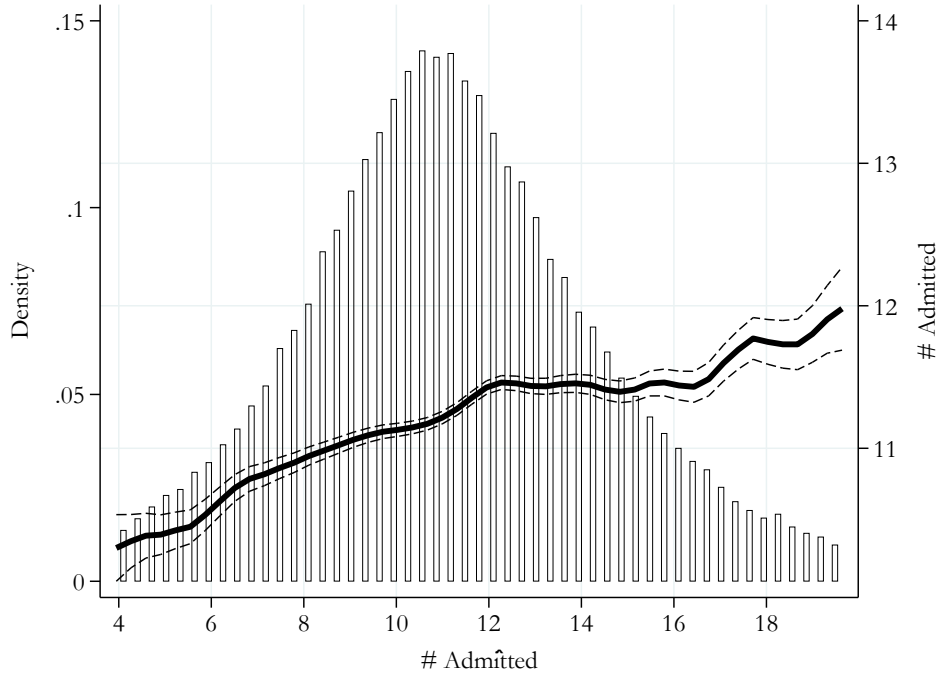
$$\widehat{admitted}_{hd} = \delta \sum_{i=1}^N \widehat{due\ date}_{hd-i} + \mu_{\widehat{h}} + \tau \widehat{due\ date}_{id} + X'_i \Omega + \varepsilon_{ihd} \quad (2)$$

Assuming instrument exogeneity and monotonicity, the estimated 2SLS result can be interpreted as the weighted average of the causal effect of congestion at different levels among the subgroup of women whose exposure to congestion is influenced by the number of women with the same due date in the local area. In other words, the results apply to women who were less likely to be re-routed or select to different wards or admission dates through the mechanisms described above.

4.1 Relevance.

Figure 1 shows the identifying variation in the data and a graphical illustration of the first stage. The histogram shows the distribution of the instrument after it has been residualized against all control variables and the relevant fixed effects included in Equation 1. The figure shows that after a large share of the cross-sectional and common ward inter-temporal

¹⁸One could argue that the length and weight of the child are determined before birth, but as women who are admitted on busier days are less likely to be induced, there may be some effects along these dimensions. I therefore choose not to include these potentially endogenously determined variables in the regression, although they have no significant effect on results.



Note: The histogram shows the residualized variation in the instrument (number of mothers with due date on mother i's admission date at her local area) after conditioning on controls and fixed effects as in Equation 1 (top and bottom 2% excluded). The solid line shows a local linear regression of congestion on the instrument with the 90% CI shown by dashed lines.

Figure 1. Sample variation in predicted congestion

variation in the sample has been subsumed, a policy-relevant amount of variation remains. The mean of the predicted congestion measure (11.24) is very close to the number admitted per institution-day on average (11.22), and the residualized variation in the instrument is quite large: the difference between ± 2 standard deviations in predicted congestion is closer to 17 admissions.¹⁹

The first stage representation plotted by the solid line shows that the congestion level increases quite monotonically in the predicted congestion level.²⁰ There is some flattening and noise in the relationship at higher

¹⁹In Figure A1 I also plot the raw variation in the number of women admitted over one calendar year to an anonymized ward, to provide some descriptive illustration of the raw day-to-day variation of admissions at institution level.

²⁰Note that the monotonicity assumption only requires that the first stage does not change sign, therefore the assumption is not violated even if the strength of the first stage weakens as illustrated in the figure.

Table 3. First stage estimates with various amounts of controls

	(1)	(2)	(3)
	No controls	+ Parental age, immig. stat., gender of child	All controls from Table 2
$\#admitted$	0.0958*** (0.0044)	0.0958*** (0.0044)	0.0959*** (0.0044)
F-stat. on instrument	468.66	469.84	470.97

Note: First stage estimates with variations in included controls. In addition to the listed variables, the regressions include due date and local ward fixed effects. Robust standard errors in parentheses.

predicted congestion levels, which is indicative of more frequent re-routing during particularly congested periods.

Table 3 reports the estimated first stage with various degrees of controls. Starting from the simplest possible first stage, excluding all control variables except due date and local ward fixed effects, Column (1) shows a precisely estimated positive relationship between predicted and actual congestion levels. Adding additional controls does not alter the estimated first stage in any meaningful way, with almost precisely the same F-statistic on the instrument, far exceeding any conventional levels of relevance criteria. The first stage estimate shows that for every additional predicted admission conditional on due date and constant local ward characteristics, the number of women admitted increases by .1 on average.

4.2 Conditional independence.

A necessary condition for the instrumental variable approach to be valid is that the instrument is uncorrelated with other factors affecting the outcome, i.e. that the independence assumption holds. As explained above, this is likely to be the case once fixed effects are included. Nonetheless, it is impossible to provide a definitive answer to the validity of the assumption. It is, however, possible to investigate whether the instrument is balanced across observed characteristics. I do this by regressing the predicted congestion level against a vector of observed characteristics determined well ahead of the due date. These include parental income, education, age and

immigration status as well as the gender of the child. I also include due date and local ward fixed effects as explained above. Results reported in Column (2) of Table 2 show that there is no detectable correlation between the instrument and any characteristic, with no coefficient being statistically different from 0. Note also that the joint explanatory power of all characteristics is very low (p-value = 0.99).²¹ I consider these results to be a strong indicator that the number of women with the same due date within the catchment area of a ward is uncorrelated to other determinants of child and mother outcomes.

To lend further support to this interpretation, Table 3 shows that adding control variables to the first stage has no impact on the estimated relationship between congestion and predicted congestion. This is reassuring, because it renders it less likely that any unobserved characteristics are causing bias to the degree they are correlated with observables.

4.3 *Exclusion.*

In this paper, the exclusion criterion requires that the number of women with the same due date in a limited geographic region does not affect outcomes through any direct channel other than the number of women admitted to the same ward on the same day as the focal mother. One potential violation of this assumption could be that wards foresee congestion ahead of time and therefore adjust the staff on duty. Though this may be the case to some degree, the result of such supply side behavior would most probably attenuate any effects. In other words, if wards perfectly foresee the number of admissions ahead of time based on the due dates of women in the area, the wards could increase staffing to mitigate congestion effects. Were this to be the case it would bias estimated effects towards zero. However, perfect foresight and perfect adjustment to short-term variations is extremely unlikely. First, only 60% of newborns in the sample are born on the admission day, and only 12% are admitted ± 1 day from their due date. Second, midwives and physicians generally have contracts limiting working hours and regulate the number of hours that have to pass between each shift and the number of days it is allowed to work over time, making it

²¹The balancing test is robust to including pregnancy related diagnoses and contact with health care providers during pregnancy as explanatory variables.

difficult to manage staffing to perfectly adapt to admissions.

4.4 *Monotonicity.*

Assuming that the causal effect of congestion is constant across mothers, the instrument need only to satisfy the independence and exclusion criteria. If, however, the effect of congestion is heterogeneous, then the additional assumption of monotonicity is required in order to interpret the IV estimates as an average of the local average treatment effects across various congestion levels.²² In the setting studied in this paper, the implication of the assumption is that no mother would experience a lower congestion level should the predicted congestion level increase. One way to assess the credibility of this assumption is to examine whether there are sub-populations for which the instrument changes sign in the first stage. I pursue this option by first estimating the predicted congestion level for each mother, using only the control variables and fixed effects listed above. Then I divide the sample into quartiles, from lowest estimated congestion propensity to highest. Next, I estimate the first stage for each of these sub-samples, and their reverse, leading to a total of eight separate first stages. The results from this exercise are reported in Table A5. For each of these first stages the instrument remains precise and positive, though gradually weakening toward the most extreme congestion propensities. This latter point is also visually demonstrated in the gradual flattening of the first stage slope in Figure 1.

5 Results

5.1 *Effects on birth outcomes*

I begin by considering outcomes for children and mothers who are measured during or shortly after birth. Reduced form estimates are reported in Table A6 and illustrated in Figures A2 and A3.

The birth records contain a multitude of complications and interventions at a highly detailed level. As a more aggregated measure of treatment

²²There will be a local average treatment effect for mothers who are shifted from being admitted the same day as 7 to 8 mothers, from 8 to 9 and so on. The reported estimates will be an average of these effects.

intensity and as a starting point, I report the estimated effects on the number of interventions in Column (1) in Table 4. The results indicate that the total number of interventions falls by about 10% when the number of admitted mothers increases by 10.

One of the major interventions it is relevant to analyze is the probability of the birth being induced. Inductions are usually performed when the mother is well into post-term pregnancy (defined as gestation age exceeding 14 days beyond the due date), or if there is some other risk to the child's or mother's health ([The Norwegian Medical Association, 2014](#)). The effect of induction on maternal health and perinatal outcomes is unclear, as solid evidence is lacking. However, a recent literature survey on the effects of induction among low-risk mothers found that the procedure has little effect on child health, though there may be a positive effect on APGAR scores and reductions in the probability of an emergency C-section being performed ([Middleton P, 2018](#); [Grobman et al., 2018](#)). When a birth is induced it may take hours or days before the birth transitions into the active phase with contractions. During periods with more admitted mothers, the staff at a maternity ward can avoid intensifying the workload by choosing not to induce additional births. Induced labor is quite common and occurs in about 20% of the births in the sample.

Column (2) in Table 4 shows that every additional mother arriving at the maternity ward the same day as the focal mother reduces the probability of induction by 1 percentage point, and a one standard deviation increase in congestion reduces the probability by more than a third of the sample mean. One way of interpreting this result is that health personnel on the ward take active measures to regulate congestion: by refraining from inducing births during hectic periods they effectively reduce the congestion level in the short-term. If obstetricians and midwives perceive more hectic periods as potentially harmful for patients, this behavior is in line with maximizing the expected health level of patients post-partum. An alternative hypothesis, which cannot be ruled out by the results, is that physicians and midwives provide more interventions as long as there is spare capacity. This would imply that the same amount of resources on a ward, such as time and equipment, is simply divided among a larger number of patients during more congested periods.

During most births, women receive some kind of pain relief. In the analysis sample, 94% receive an analgesic, including epidural.²³ The types and amounts of drugs administered depend on pre-existing conditions, complications and the wishes of the patients (Royal College of Obstetricians & Gynaecologists, 2015). Analgesics comprise a wide range of drugs, from relatively mild to stronger drugs used only in critical situations (Royal College of Obstetricians & Gynaecologists, 2015). The effects of analgesics on the health of both mother and child are subject to debate. Some studies find effects on the fetal heart rate and the probability of emergency C-sections, while others find no adverse effects (Eltzschig et al., 2003). Results, reported in column (3), suggest that there is a 3% reduction in the use of analgesic drugs when congestion increases by one standard deviation.²⁴

According to Norwegian medical guidelines, an emergency C-section can only be performed after a medical assessment deems it the proper treatment due to unforeseen and acute threats to the health of the mother or the child such as pre-eclampsia. However, the wish of the mother should be taken into account during the assessment (Helse Stavanger, 2019). If physicians rely strictly on the medical assessment of the health conditions of mother and child, one would not expect congestion to have any effect on the prevalence of emergency C-sections. The results in Column (4) show that the C-section rate is completely unaffected by the congestion level, suggesting that emergency C-sections are only performed when medically necessary and independently of congestion.

In the final two columns I report results for two indicators of newborn health. The APGAR score is a commonly used measure of children's health outcomes shortly after birth, and is set by a midwife based on a set of predetermined criteria on a checklist. 94% of newborns receive a score of 8 or higher. In past research the score has been shown to negatively correlate with long-term cognitive ability and a score below 7 has long been used as an indicator for poor newborn health (Stuart et al., 2011; Ehrenstein et al., 2009; Casey et al., 2001). Column (5) shows that children born during more congested periods are less likely to receive an APGAR score of less than 7:

²³Analgesics are defined as medicines used with an ATC code starting with N02.

²⁴In Table A7 in the Appendix I also show results for anesthetics, and also for epidural as a specific type of analgesic.

a one standard deviation increase in congestion reduces the probability of a child receiving an APGAR score of less than 7 by about 60%, though the estimate is significant only at the 10% level. The exact size of the effect is therefore rather uncertain, but it suggests that children born during more hectic periods are at least no worse off. One reason for a positive effect on children’s health may be that some of the interventions that are less used during congested periods are marginally harmful for the child. However, these very same interventions may be perceived by the medical staff as marginally health improving, leading to overtreatment of mothers with negative effects on neonatal health.

About 5% of the newborns in the sample are readmitted within the first 30 days after birth. The most common causes are feeding problems, respiratory distress, and jaundice, and increasing readmission rates have been related to insufficient monitoring after birth (Young et al., 2013). In Column (6) I estimate that children born during more congested periods are significantly less likely to be readmitted within in their first month of life. In Table A8 in the Appendix, I show that readmission rates are also lower the first week of life. This evidence suggests that treatment given to women during less congested periods may actually have detrimental effects on children in the earliest part of their life.

I report results for some additional outcomes in Appendix Table A8. These results show that congestion has no impact on length of stay, and women admitted on more congested days are no more likely to suffer from severe lacerations or severe hemorrhages. The frequency of both severe lacerations and severe hemorrhages are considered to be affected by staff through active monitoring and are registered in the medical birth registry. Mothers admitted on more congested days are also no more likely to be readmitted in the first 7 or 30 days after childbirth or be registered with more complications. These results support the conclusion from the main results that the interventions foregone during congested periods have no detectable effects on health.

In the appendix I challenge the main results in several ways. First, I show that estimates are nearly identical when I remove all individual level controls in Table A9, supporting the IV assumptions. Results also remain qualitatively similar when I remove Oslo, the only city with mul-

tiple wards, from the sample in Table A10. In the main sample I exclude women who have previously given birth. I show results for these mothers in Table A11.²⁵ These results show that mothers who have given birth previously are similarly affected in terms of the number of interventions, labor inductions and the use of analgesics, but are less likely to undergo emergency c-sections during congested periods. There is also no detectable effect on APGAR score or readmissions. There are several reasons for why these differences may occur, but these are hard to test in the data. First, having gone through a birth previously gives the staff much more information on risk factors. For instance, successful past vaginal deliveries may make staff more reluctant to perform emergency c-sections, and in particular during congested periods. Second, the mother herself is likely to have other expectations and needs than a first-time mother. Having experience as a parent may reduce the impact of congestion on the probability of later readmissions.

The main results are estimated on a model where congestion is measured as the number of admitted women after netting out local ward and due date fixed effects. There is substantial heterogeneity in terms of the size of wards in the sample, and it is not given that smaller and larger wards are equally affected by one additional admitted patient. I explore this possibility by estimating separate effects for the largest (≥ 90 th percentile) wards and the remaining wards in Table A13.²⁶ These results show that larger wards are able to mitigate most effects of congestion on treatment intensity and there is no evidence that congestion leads to better child health outcomes. For smaller wards, the opposite is true. The most likely explanation for this is that an additional admission at a large ward has a relatively smaller impact on congestion as these have higher admission numbers on average.

Finally, I estimate models where I scale congestion to the admission numbers at the ward level. Specifically, I measure actual and predicted admissions on each day and ward as the centile of the maximum number of admissions observed at the ward in the period covered by the data. Using this measure of congestion instead of the number of admissions gives a

²⁵Note that all mothers, regardless of age, parity and so on, are included when constructing the instrument and the endogenous variable.

²⁶By splitting at the 90th percentile the number of observations in both groups remain roughly similar.

Table 4. Main results.

	(1) N interv.	(2) Ind. labor	(3) Analgesic
$\#admitted$	-0.0418*** (0.0151)	-0.0101*** (0.0032)	-0.0033* (0.0017)
Dep. mean	4.41	0.22	0.94
	(4) Emg. C-section	(5) APGAR < 7	(6) Child readmitted
$\#admitted$	-0.0002 (0.0025)	-0.0016* (0.0009)	-0.0038** (0.0015)
Dep. mean	0.13	0.02	0.05

Note: Results from estimating Equation 1 using 2SLS. Outcomes are defined in column titles. Each coefficient is estimated in a separate regression. Each regression includes controls for parents age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. Robust standard errors in parentheses. F-statistic first stage = 471.

substantially smaller first stage and consequently less precision, but qualitatively similar results as reported in Table A14.

Overall, the main results can be summarized as two general findings: first, mothers who are admitted on busy days receive fewer interventions than mothers on less busy days and have the same number of complications. Second, the children born to the former are healthier, measured by APGAR score and readmission rates, than comparable children born on less busy days. This implies that at least some of the treatments women receive due to being admitted on slow days have no beneficial effect on child health outcomes.

A central question in assessing the validity of these results, is whether the reporting itself is affected by congestion. One could imagine that midwives and physicians prioritize tending to patients rather than reporting complications and treatments during stressful situations, which could cause some of the patterns in the data. This could for example affect the APGAR score: During congested periods midwives may have less time to assess the health of neonatals to find defects or imperfections, and might

assign higher APGAR scores as a result. However, it seems very unlikely that the reporting of major interventions such as induction of labor or use of strong medications should be affected in this manner. Further, readmission rates cannot be directly affected by this kind of reporting behavior. Nonetheless, it must be acknowledged that reporting behavior may affect some outcomes.

5.1.1 Comparison with OLS To illustrate the importance of using an IV strategy, I report the results of an OLS specification in Table 5. Here I replace due date fixed effects with admission date fixed effects, and local ward fixed effects with admitting ward-by-admission year fixed effects. There are clear differences from the IV results. In particular, the OLS results suggest that the number of interventions a woman experiences is unaffected while the induction rate increases. The OLS results also find no effect on APGAR or readmissions. Jointly with the balance test for the OLS specification reported in Table A4 it appears that there is enough spatial or temporal sorting of women to different congestion levels that the OLS results in these data are biased to an extent that results are significantly affected.

An alternative explanation for the differences is that the mothers complying with the instrument differ from those in the overall sample population. That is, the women whose experienced congestion is predictable by the number of women in the local area with due date on the mother's admission date, may not be similar to other women. In other words, the effects estimated of compliers to the instrument cannot be extrapolated to non-compliers that are included in the OLS estimations. To assess whether this may be driving the differences between OLS and IV estimates, I characterize compliers in Table A12 following Bhuller et al. (2020). I start by estimating a propensity for congestion by regressing experienced congestion against all observable characteristics included in the main analysis (excluding the instrument). I then split the sample into quartiles based on this propensity and re-run the first stage separately for each sub-sample. The estimated first stage coefficients are reported in the first rows in Column (1) of Table A12. The next columns report the standard deviation of the estimates along with the population shares and the relative size of the first

Table 5. OLS conditional on admitting ward, admission date, and admitting-ward-by-admission-year

	(1) N interv.	(2) Ind. labor	(3) Analgesic
<i>#admitted</i>	-0.0019 (0.0014)	0.0019*** (0.0003)	-0.0007*** (0.0002)
Dep. mean	4.41	0.22	0.94
	(4) Emg. C-section	(5) APGAR < 7	(6) Child readmitted
<i>#admitted</i>	0.0002 (0.0002)	0.0000 (0.0001)	-0.0002 (0.0002)
Dep. mean	0.13	0.02	0.05

Note: Outcomes defined in column titles. Each regression includes controls for parental age, education, earnings and immigration status as well as the gender of the child and admission date and admitting-ward-by-admission-year fixed effects. Robust standard errors in parentheses.

stage compared to the one estimated on the full sample. Focusing on the latter, it is clear that the size of the first stage estimate relative to that of the full sample is very similar across each sub-sample. The estimates in the next rows show results when the sample is split by each observable characteristic included in the main analysis. Overall, there is little evidence that any sub-population has substantially different compliance, which suggests that heterogeneity does not explain the large differences between the OLS and IV estimates.

Research similar to the current paper have generally found that there are fewer interventions during more congested periods, and a similar conclusion can be drawn from the results presented here (Maibom et al., 2020; Freedman, 2016; Marks and Choi, 2019). However, the comparison between the main results in Table 4 and the OLS results in 5 highlight the potential shortcomings of relying on residualized admission numbers when estimating congestion effects. First, using the similar fixed effects as the past literature, I find similar correlations between parental background characteristics and congestion as reported in Table A4. Then, estimating

the effect of congestion on interventions and health outcomes I find a very limited effect on outcomes which is in line with the conclusion in [Maibom et al. \(2020\)](#) and [Marks and Choi \(2019\)](#). However, when instrumenting for congestion and removing spatial and temporal patient sorting, I find a much clearer and positive effect of congestion on health outcomes, which seems to be cluttered by temporal and spatial sorting in the simpler OLS specification. Therefore, while the results in this paper cannot comment directly on the validity of using residualized admission numbers as an exogenous source for congestion, they caution a strict causal interpretation of consequent results.

5.1.2 Triage and maternal characteristics The fact that women admitted during more congested periods on average undergo fewer interventions may veil heterogeneous effects due to triage: physicians and midwives may reduce the intensity of medical interventions for healthier women, while maintaining the same intensity of care for at-risk pregnancies. I cannot identify which pregnancies are considered at-risk by the healthcare personnel from the data. However, a possible approach to testing for triage is to rely on past observed pregnancy-related contacts with health services and related diagnoses. The underlying assumption is then that past healthcare utilization during pregnancy provides a proxy for mother and child health and risk factors during birth.

In order to identify pregnancy-related contacts with healthcare providers, I sum all contacts the mother had with inpatient and outpatient healthcare providers which can be categorized as pregnancy-related.²⁷ For diagnoses, I sum all pregnancy-related diagnoses made during the same contacts.²⁸ Examples of pregnancy-related diagnoses are pregnancy-related hypertension or pregnancy-induced diabetes. Note that not all contacts lead to a specific diagnosis and some contacts may result in several diagnoses, and if so I count these as a single diagnosis.

To test for triage I re-estimate the baseline model separately for women above and below the median number of contacts and diagnoses. The results

²⁷These are ICPC-2 codes starting with W and ICD-10 codes O10-16; O20-29; O30-36; Z34-35.

²⁸These are ICPC-2 codes A94; W3-5; W28-29; W70-76; W80-81; W84-85, and ICD-10 codes O10-16, O20-29, O30-39, Z35.

for diagnoses (contacts) are reported in Table 6 (A15). Both sets of results indicate that there is limited, if any, triage based on these proxies of at-risk pregnancies with women who are likely to be at higher risk being more affected by congestion. One explanation for this finding could be that mothers with more diagnoses and/or contacts are more exposed to defensive medicine at the baseline. That is, mothers with higher levels of health care utilization prior to birth are more likely to receive treatments with little or no marginal health benefit during birth. Using random variations in capacity at public health clinics, [Harris et al. \(2019\)](#) find that nurses appear to prioritize tasks to maintain the quality of services when time constrained. An implication of their finding for this study could be that midwives reduce their effort towards defensive medicine for at-risk births, but limit effects of congestion on effective interventions. This is also suggested by the fact that interventions are only reduced for mothers who are assumed to be at risk, and this from a higher baseline level, while the same group also experience a positive effect on readmission. In sum, it is therefore not possible to rule out effective triage using the data available in this study.

Existing literature has shown that patient characteristics influence the interventions the patients receive ([Grytten et al., 2013](#); [Johnson and Rehavi, 2016](#)). For instance, [Johnson and Rehavi \(2016\)](#) show that physician mothers are less likely to undergo C-sections than comparable non-physicians, and that most of this effect cannot be explained by sorting. In addition to medical knowledge, one can also imagine that the treatment patients receive in the maternity ward is, at least to some degree, the result of a form of negotiation between the patient and healthcare personnel. For instance, a mother may ask for pain-relief such as an epidural, while the obstetrician is reluctant to fulfill this request based on his evaluation of patient needs. In cases like this, the bargaining power of the mother is likely to be stronger if her partner is present and can put additional pressure on the obstetrician to comply. Though it is not possible to identify the bargaining power of either physicians or patients, I provide results showing that mothers who have a registered partner are less affected by congestion in Table 7.²⁹ The results show that, overall, women who have a partner are

²⁹Note that I cannot observe whether or not the partner is present at any stage during labor. I therefore define “married” as the mother being registered with the father as a

less affected by congestion. These results indicate that patient bargaining power has an impact on interventions, yet other mechanisms may be at play. For instance, midwives and obstetricians may treat mothers differently depending on their marital status, regardless of the partner being present in the maternity ward.

partner in administrative registries prior to labor.

Table 6. Results by pregnancy-related issues diagnosed during pregnancy.

		(1)	(2)	(3)
		N	Ind.	Analgesic
		interv.	labor	
Not diagnosed with pregnancy-related issue	$\#ad\hat{m}itted$	-0.0320 (0.0281)	-0.0089 (0.0055)	-0.0051 (0.0036)
Dep. mean		4.28	0.16	0.94
Diagnosed with pregnancy-related issue	$\#ad\hat{m}itted$	-0.0620** (0.0294)	-0.0150** (0.0064)	-0.0009 (0.0029)
Dep. mean		4.64	0.28	0.95
		(4)	(5)	(6)
		Emg.	APGAR < 7	Child
		C-section		readmitted
Not diagnosed with pregnancy-related issue	$\#ad\hat{m}itted$	-0.0021 (0.0044)	-0.0024 (0.0017)	-0.0012 (0.0026)
Dep. mean		0.10	0.01	0.04
Diagnosed with pregnancy-related issue	$\#ad\hat{m}itted$	0.0025 (0.0050)	-0.0014 (0.0018)	-0.0060** (0.0030)
Dep. mean		0.16	0.02	0.06

Note: Median number of diagnoses set during pregnancy is 2. F-stat for women not diagnosed during pregnancy = 124 (N = 58 004), F-stat for women diagnosed during pregnancy = 143 (N = 68 712). Outcomes defined in column titles. Each estimate is from a separate regression and includes controls for parental age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. See the Appendix for the definition of diagnoses coded as pregnancy related-issues. Robust standard errors in parentheses.

Table 7. Results by marital status

		(1)	(2)	(3)
		N	Ind.	Analgesic
		interv.	labor	
Unmarried	$\#ad\hat{m}itted$	-0.0563*** (0.0170)	-0.0101*** (0.0036)	-0.0035* (0.0020)
Dep. mean		4.40	0.22	0.94
Married	$\#ad\hat{m}itted$	-0.0037 (0.0343)	-0.0101 (0.0073)	-0.0025 (0.0038)
Dep. mean		4.49	0.24	0.94
		(4)	(5)	(6)
		Emg.	APGAR < 7	Child
		C-section		readmitted
Unmarried	$\#ad\hat{m}itted$	-0.0009 (0.0028)	-0.0021** (0.0010)	-0.0039** (0.0017)
Dep. mean		0.13	0.02	0.05
Married	$\#ad\hat{m}itted$	0.0059 (0.0060)	0.0002 (0.0022)	-0.0033 (0.0033)
Dep. mean		0.15	0.02	0.05

Note: F-stat. unmarried = 378 (N = 121 867), F-stat married = 93 (N = 26 130). Outcomes defined in column titles. Each regression includes controls for parental age, education, earnings and immigration status as well as the gender of the child, due date fixed effects and ward fixed effects. Robust standard errors in parentheses.

5.1.3 Longer-term outcomes In this section I focus on longer-term health outcomes for mother and child.³⁰ To identify health outcomes, I measure various forms of healthcare used for mother and child when the child is 1, 2 and 3 years old. Outpatient treatments are measured as visits to a registered GP, while inpatient treatment is defined as visits to public hospitals. Effects are reported by age in panels A-C in Table 8. Starting with treatments for the child, it appears that the negative effect of congestion on readmission rates the first 30 days is mirrored in the hospitalization rates during the first 2 years of life. Thus, children born during more congested periods are no less healthy than those born during less congested periods.

Postpartum depression has received significant public attention and has also been the focus of recent contributions in the economics literature (Persson and Rossin-Slater, 2019). Following birth, a non-negligible share of women go through a period with depression which may hinder optimal care for the child, with potential negative effects on child development. Considering that the women admitted on hectic days receive less treatment, they may be more likely to experience birth as a negative event, which may increase their probability of developing postpartum depression. On the other hand, one might imagine that women who give birth with fewer interventions, such as induction and emergency C-section, have shorter restitution periods and therefore experience birth as a more positive event. I therefore estimate the utilization of psychologists as well as overall outpatient and inpatient interventions for the mother. The results are reported in columns (3) - (5) in Table 8. Overall, the results show that mothers giving birth during more congested periods have the same utilization rates of all these services as other mothers. Because the utilization rates of medical services do not necessarily reflect mothers' health status, I also include one additional measure of maternal health outcomes: sick leave. Sick leave for longer periods (> 3 days) is prescribed by GPs or physicians at hospitals and triggers sickness pay from the government. Workers can take shorter periods of sick leave without any loss of income. The results show no evidence that sick leave is affected by congestion except a small effect three years after birth which is significant at the 10%

³⁰There are no effects on 30-day and 1-year mortality rates for children. Results available upon request.

Table 8. Longer-term outcomes.

Panel A	Child age 1					
	(1) GP, child	(2) Hosp. child	(3) GP, mother	(4) Hosp. mother	(5) Psych. mother	(6) Sick leave, days
$\hat{\#admitted}$	-0.0012 (0.0028)	-0.0444** (0.0202)	-0.0088 (0.0336)	0.0092 (0.0176)	-0.0021 (0.0097)	-0.1455 (0.3342)
Dep. mean	0.10	1.68	4.69	1.26	0.05	11.97
Panel B	Child age 2					
	(1) GP, child	(2) Hosp. child	(3) GP, mother	(4) Hosp. mother	(5) Psych. mother	(6) Sick leave, days
$\hat{\#admitted}$	-0.0027 (0.0051)	-0.0386** (0.0181)	-0.0018 (0.0466)	0.0208 (0.0251)	0.0104 (0.0124)	0.5594 (0.4041)
Dep. mean	0.32	0.98	6.33	1.89	0.09	24.18
Panel C	Child age 3					
	(1) GP, child	(2) Hosp. child	(3) GP, mother	(4) Hosp. mother	(5) Psych. mother	(6) Sick leave, days
$\hat{\#admitted}$	0.0062 (0.0048)	0.0100 (0.0175)	-0.0633 (0.0488)	0.0027 (0.0278)	0.0127 (0.0142)	0.7938* (0.4571)
Dep. mean	0.28	0.94	6.93	2.35	0.10	31.51

Note: Each coefficient is estimated in a separate regression. The outcome “GP” measures the number of primary care physician visits registered for the individual child or mother in each year of life of the child. The outcome “Hosp.” is constructed similarly but measure hospital visits. The outcome “Psych.” is the number of registered visits with certified therapists for the mother. Each regression includes controls for parents’ age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. Robust standard errors in parentheses.

level only.

6 Discussion

The reasons for differential utilization of healthcare across individuals with equal access remains an unanswered question. An expanding literature considers the extent to which supply side incentives, personnel and resource scarcity are part of these causes. In the maternity ward setting, past studies have shown that women are more likely to have an unplanned C-section if they are in labor at night than during the day due to physicians leisure requirements (Costa-Ramón et al., 2018), and that congestion impacts the

provision of interventions (Maibom et al., 2020; Freedman, 2016; Marks and Choi, 2019; Facchini, 2019). This paper uses an instrumental variable approach to resolve threats to identification present in some past contributions. I show that when the relative resources available to women in labor are reduced due to more women giving birth in the same ward, they experience fewer and less invasive interventions. I cannot detect any negative health effects in the long-term for either mother or child. Rather, children born during more congested periods seem to have higher APGAR scores and are less likely to be readmitted, which suggests positive health effects for the neonatals, and overtreatment

Comparing results from the IV-specification with an OLS estimate that include fixed effects for admission date and admitting ward-by-admission year I find meaningful differences and I also find evidence of violations of the identifying assumptions in the OLS approach. I emphasize that these findings do not imply that there is significant bias in the discussed literature and do not speak directly to the findings and methods of past contributions, but caution against the general validity of estimates on congestion effects in health care relying on conditional independence after controlling for fixed effects.

The results in this paper indicate that there is some degree of overtreatment, or at least treatment with zero or negligible effects, of women during less congested periods. The implication of this finding is that when resources are available, there will be additional interventions, even when the intervention is unnecessary or even potentially harmful. Overall, it appears that a policy which reduces spare capacity from the current level in these data will have no negative effects on mother or child.

References

- Almond, D. and Doyle, J. J. (2011). After midnight: A regression discontinuity design in length of postpartum hospital stays. *American Economic Journal: Economic Policy*, 3(3):1–34.
- Bhuller, M., Dahl, G. B., Løken, K. V., and Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, 128(4):1269–1324.
- Casey, B. M., McIntire, D. D., and Leveno, K. J. (2001). The continuing value of the apgar score for the assessment of newborn infants. *New England Journal of Medicine*, 344(7):467–471. PMID: 11172187.
- Chandra, A., Cutler, D., and Song, Z. (2011). Who ordered that? the economics of treatment choices in medical care. In *Handbook of health economics*, volume 2, pages 397–432. Elsevier.
- 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.
- Doyle, J. J. (2011). Returns to local-area health care spending: evidence from health shocks to patients far from home. *American Economic Journal: Applied Economics*, 3(3):221–43.
- Doyle J., J. J., Graves, J. A., Gruber, J., and Kleiner, S. A. (2015). Measuring returns to hospital care: Evidence from ambulance referral patterns. *Journal of Political Economy*, 123(1):170–214.
- Ehrenstein, V., Pedersen, L., Grijota, M., Nielsen, G. L., Rothman, K. J., and Sørensen, H. T. (2009). Association of apgar score at five minutes with long-term neurologic disability and cognitive function in a prevalence study of danish conscripts. *BMC pregnancy and childbirth*, 9(1):1–7.
- Eltzschig, H. K., Lieberman, E. S., and Camann, W. R. (2003). Regional anesthesia and analgesia for labor and delivery. *New England Journal of Medicine*, 348(4):319–332. PMID: 12540646.

- Evans, W. N. and Kim, B. (2006). Patient outcomes when hospitals experience a surge in admissions. *Journal of Health Economics*, 25(2):365 – 388.
- Facchini, G. (2019). Low staffing in the maternity ward: Keep calm and call the surgeon. *Working paper*.
- Finans Norge (2019). Statistikk - behandlingsforsikring. <https://www.finansnorge.no/statistikk/skadeforsikring/helseforsikring/behandlingsforsikring/> [Accessed: 14.11.2019].
- Finkelstein, A., Gentzkow, M., and Williams, H. (2016). Sources of geographic variation in health care: Evidence from patient migration. *The quarterly journal of economics*, 131(4):1681–1726.
- 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.
- Godøy, A. and Huitfeldt, I. (2019). Regional variation in health care utilization and mortality. Technical report.
- Gowrisankaran, G., Joiner, K. A., and Léger, P.-T. (2017). Physician practice style and healthcare costs: Evidence from emergency departments. Technical report, National Bureau of Economic Research.
- Grobman, W. A., Rice, M. M., Reddy, U. M., Tita, A. T., Silver, R. M., Mallett, G., Hill, K., Thom, E. A., El-Sayed, Y. Y., Perez-Delboy, A., et al. (2018). Labor induction versus expectant management in low-risk nulliparous women. *New England Journal of Medicine*, 379(6):513–523.
- Gruber, J., Hoe, T. P., and Stoye, G. (2018). Saving lives by tying hands: The unexpected effects of constraining health care providers. Working Paper 24445, National Bureau of Economic Research.
- 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.
- Helse Stavanger (2019). Keisersnitt. [helse-stavanger.no/behandler/keisersnitt](https://www.helse-stavanger.no/behandler/keisersnitt) [Accessed: 14.11.2019].
- helsenorge.no (2020). Check-ups during pregnancy. <https://www.helsenorge.no/en/pregnancy-and-maternity-care-in-norway/check-ups-during-pregnancy/> [Accessed: 15.10.2020].
- 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.
- Maibom, J., Sivertsen, H. H., Simonsen, M., and Wust, M. (2020). Maternity ward crowding, procedure use and child health. *Journal of Health Economics, Fortcoming*.
- Marks, M. and Choi, M. K. (2019). Baby boomlets and baby health: Hospital crowdedness, hospital spending, and infant health. *American Journal of Health Economics*, 5(3):376–406.
- Middleton P, Shepherd E, C. C. (2018). Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database of Systematic Reviews*, 5.
- Molitor, D. (2018). The evolution of physician practice styles: evidence from cardiologist migration. *American Economic Journal: Economic Policy*, 10(1):326–56.
- Persson, P. and Rossin-Slater, M. (2019). When dad can stay home: Fathers’ workplace flexibility and maternal health. Technical report, National Bureau of Economic Research.
- Royal College of Obstetricians & Gynaecologists (2015). The management of third- and fourth-degree perineal tears. <https://www.rcog.org.uk/globalassets/documents/guidelines/gtg-29.pdf> [Accessed: 17.12.2019].

- Royal College of Obstetricians & Gynaecologists (2019). Birth after previous caesarean birth, green-top guideline no. 45. https://www.rcog.org.uk/globalassets/documents/guidelines/gtg_45.pdf [Accessed: 25.11.2019].
- Skinner, J. (2011). Chapter two - causes and consequences of regional variations in health care. In Pauly, M. V., McGuire, T. G., and Barros, P. P., editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 45 – 93. Elsevier.
- Stuart, A., Olausson, P. O., and Källen, K. (2011). Apgar scores at 5 minutes after birth in relation to school performance at 16 years of age. *Obstetrics & Gynecology*, 118(2):201–208.
- The Norwegian Medical Association (2014). Veileder i fødselshjelp 2014. <https://beta.legeforeningen.no/contentassets/9c5011955cc644a595cec42c01934c92/veileder-i-fodselshjelp-2014.pdf> [Accessed: 17.12.2019].
- Wennberg, J. E. (2010). *Tracking medicine: a researcher's quest to understand health care*. Oxford University Press.
- Woodworth, L. (2020). Swamped: Emergency department crowding and patient mortality. *Journal of Health Economics*, 70:102279.
- Young, P. C., Korgenski, K., and Buchi, K. F. (2013). Early readmission of newborns in a large health care system. *Pediatrics*, 131(5):e1538–e1544.

A Appendix

A.1 *Additional results*

Table A1. Mothers' use of alternative wards and background characteristics

	Birth in less common ward	#admitted
Father lower secondary schooling	-0.0153*** (0.0029)	-0.2866*** (0.0452)
Father upper secondary schooling	-0.0102*** (0.0024)	-0.2802*** (0.0386)
Father Master's degree	-0.0002 (0.0033)	0.1139** (0.0566)
Father's age	0.0000 (0.0002)	-0.0016 (0.0030)
Father is immigrant	0.0003 (0.0030)	0.1213** (0.0495)
ln(father's income)	-0.0060*** (0.0007)	-0.0188* (0.0103)
Mother lower secondary schooling	-0.0095*** (0.0027)	-0.0027 (0.0415)
Mother upper secondary schooling	-0.0005 (0.0022)	0.0616* (0.0352)
Mother Master's degree	-0.0060** (0.0030)	0.3221*** (0.0511)
Mother's age	-0.0010*** (0.0003)	0.0220*** (0.0043)
Mother is immigrant	-0.0169*** (0.0029)	-0.1878*** (0.0484)
ln(mother's income)	-0.0066*** (0.0006)	-0.0148 (0.0094)
Child female	-0.0006 (0.0017)	-0.0114 (0.0265)
		-2.2623*** (0.0732)
Observations	148 022	148 022
F-statistic	24.3	11.22
Dep. mean	0.13	100.55

Note: Column (1) shows estimated results from a regression of a dummy indicating that the mother was admitted to a hospital other than the most commonly used in her area. Column (2) shows the estimated results from a regression of the number of women admitted the same day as the focal mother against background characteristics as well a dummy indicating that the mother was admitted to a hospital other than the most commonly used in her area. In addition to listed variables, the regression includes due date and local ward fixed effects. Robust standard errors in parentheses.

Table A2. Descriptive statistics: Parental background and birth statistics.

	Mean	Std. Dev.	Obs
Mothers admitted	11.22	7.82	148 022
Predicted mothers admitted (instrument)	11.24	8.37	148 022
Weight (gr)	3447.51	536.36	148 022
Length (cm)	50.00	2.46	144 333
Child female	0.48	0.50	148 045
Interventions and complications			
Number of interventions	4.41	2.16	148 022
Number of complications	5.48	1.44	148 022
APGAR. < 7	0.02	0.13	148 022
Induction	0.22	0.42	148 022
Analgesic	0.94	0.24	148 022
Emergency C-section	0.13	0.34	148 022
Child readmitted 30 days	0.05	0.22	148 022
Paternal characteristics			
Pensionable income	447.47	322.21	148 022
Age	31.06	6.09	148 022
Immigrant	0.18	0.38	148 022
Lower secondary schooling	0.27	0.44	148 022
Upper secondary schooling	0.39	0.49	148 022
Bachelor's degree	0.22	0.41	148 022
Master's degree	0.12	0.33	148 022
Married	0.18	0.38	148 022
Maternal characteristics			
Pensionable income	314.09	198.97	148 022
Age	27.95	4.80	148 022
Immigrant	0.19	0.39	148 022
Lower secondary schooling	0.24	0.43	148 022
Upper secondary schooling	0.27	0.44	148 022
Bachelor's degree	0.36	0.48	148 022
Master's degree	0.13	0.34	148 022
Married	0.18	0.38	148 022
Diagnoses related to pregnancy prior to birth	3.57	5.85	126 760
Pregnancy-related contacts prior to birth	16.26	8.22	126 760

Note: There are fewer observations for healthcare utilization prior to birth because there are no inpatient health data in 2007

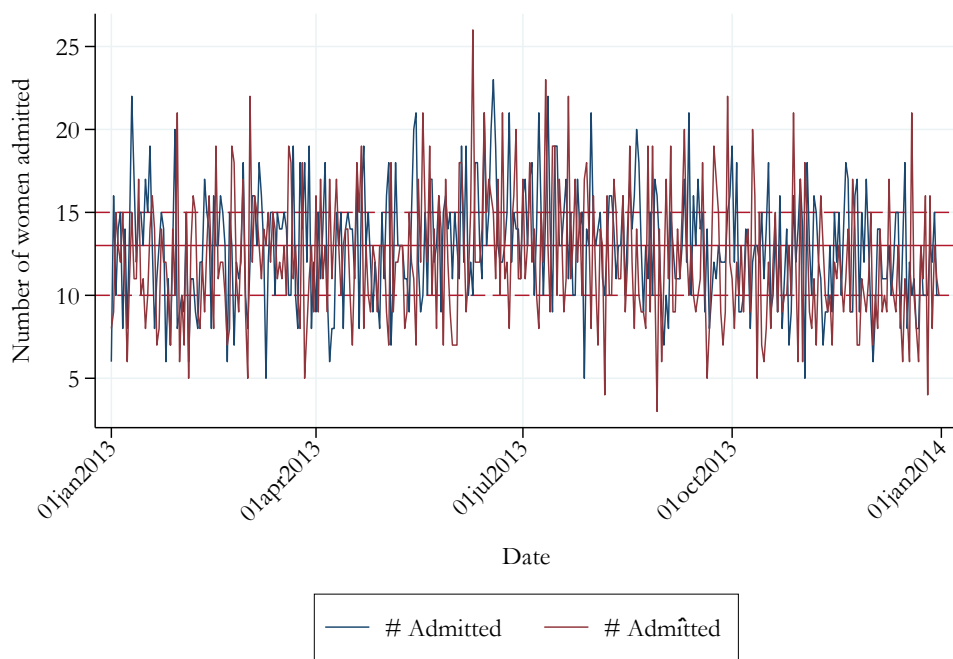
Table A3. Descriptive statistics: Health outcomes by child age.

	Mean	Sd	Obs
Age child 1 year			
GP visits, child	0.10	0.39	148 022
Hospital visits, child	1.68	2.88	148 022
GP visits, mother	4.69	5.24	148 022
Hospital visits, mother	1.26	2.46	148 022
Visits to mental healthcare providers, mother	0.05	1.13	148 022
Sick leave, mother	11.98	44.85	148 022
Age child 2 years			
GP visits, child	0.32	0.71	148 022
Hospital visits, child	0.98	2.38	148 022
GP visits, mother	6.33	6.99	148 022
Hospital visits, mother	1.89	3.47	148 022
Visits to mental healthcare providers, mother	0.09	1.45	148 022
Sick leave, mother	24.19	56.11	148 022
Age child 3 years			
GP visits, child	0.28	0.65	148 022
Hospital visits, child	0.94	2.38	148 022
GP visits, mother	6.93	7.44	148 022
Hospital visits, mother	2.36	3.83	148 022
Visits to mental healthcare providers, mother	0.10	1.60	148 022
Sick leave, mother	31.51	63.15	148 022

Table A4. Background variables and experienced congestion controlling for year-by-admitting-hospital.

	(1)		(2)	
	Bivariate		Joint	
Father				
Lower secondary	-0.0005*	(0.0003)	-0.0297	(0.0316)
Upper secondary	0.0002	(0.0003)	-0.0206	(0.0270)
Master's	-0.0002	(0.0003)	-0.0530	(0.0393)
Age	0.0025	(0.0042)	-0.0004	(0.0021)
Immigrant	-0.0004	(0.0003)	0.0106	(0.0345)
ln(income)	0.0033***	(0.0011)	0.0170**	(0.0072)
Mother				
Lower secondary	-0.0004	(0.0003)	0.0243	(0.0295)
Upper secondary	0.0003	(0.0003)	0.0310	(0.0249)
Master's	0.0002	(0.0003)	0.0273	(0.0360)
Age	0.0039	(0.0032)	0.0022	(0.0030)
Immigrant	-0.0006*	(0.0003)	-0.0423	(0.0341)
ln(income)	0.0022*	(0.0012)	0.0020	(0.0066)
Child female	0.0000	(0.0004)	0.0007	(0.0188)
Observations	148 022		148 022	
Dep. mean	11.22		11.22	
F-statistic for joint significance			1.21	
Joint p-value			0.26	

Note: Column (1) shows the results of separate linear regressions where the dependent variable is the number of women admitted on the same maternity ward-day as the focal mother and the left hand side variable is listed. In addition to the listed variables, the regression includes admission-date-by-admitting-hospital FE. Column (2) shows the resulting estimation when all listed variables are jointly included. In the bottom part of the table I report the F-statistic and p-value for a test of joint significance of the reported variables. Robust standard errors in parentheses.



Horizontal lines indicate 25th and 75h percentile and mean for the hospital-year

Figure A1. Raw variation in number of mothers admitted to an anonymized hospital in 2013.

Table A5. Test of monotonicity assumption

	(1)	(2)
	<i>#admitted</i>	<i>#admitted</i>
	Base sample	Reverse sample
Congestion propensity quartile	First	Not first
<i>#admitted</i>	0.0426*** (0.0087)	0.0820*** (0.0047)
N	36986	111010
F-stat	23.83	303.00
	Second	Not second
<i>#admitted</i>	0.0372*** (0.0070)	0.0869*** (0.0052)
N	36992	111012
F-stat	28.47	282.71
	Third	Not third
<i>#admitted</i>	0.0447*** (0.0068)	0.1086*** (0.0055)
N	36982	111009
F-stat	42.78	391.22
	Fourth	Not fourth
<i>#admitted</i>	0.0415*** (0.0080)	0.0740*** (0.0043)
N	36989	111014
F-stat	27.23	298.89
Due date	Yes	Yes
Hospital	Yes	Yes

Note: First stage estimates. Samples split by congestion propensity score estimated by running an OLS specification of congestion against all controls and fixed effects included in the main specification, excluding the instrument. Each coefficient is estimated in a separate regression. Robust standard errors in parentheses.

Table A6. Reduced form

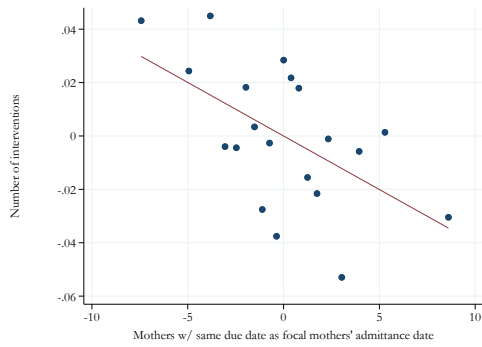
	(1) N interv.	(2) Ind. labor	(3) Analgesic
$\sum_{i=1}^N due\ date_{hd-i}$	-0.0040*** (0.0014)	-0.0010*** (0.0003)	-0.0003* (0.0002)
Dep. mean	4.41	0.22	0.94
	(4) Emg. C-section	(5) APGAR < 7	(6) Child readmitted
$\sum_{i=1}^N due\ date_{hd-i}$	-0.0000 (0.0002)	-0.0002* (0.0001)	-0.0004** (0.0001)
Dep. mean	0.13	0.02	0.05

Note: Outcomes defined in column titles. Each regression includes controls for parental age, education, earnings and immigration status as well as the gender of the child, the due date and local ward fixed effects. Robust standard errors in parentheses.

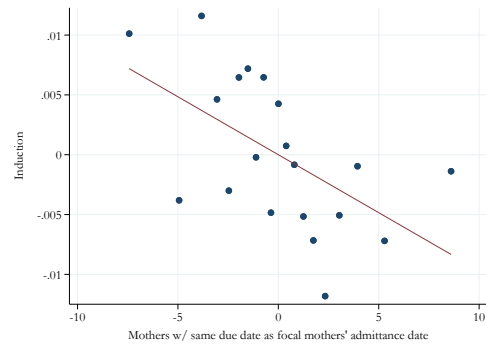
Table A7. Use of pain relievers.

	(1) Analgesic	(2) Anesthesia	(3) Epidural
$\#admitted$	-0.0033* (0.0017)	-0.0005 (0.0019)	-0.0042 (0.0037)
Dep. mean	0.94	0.12	0.47

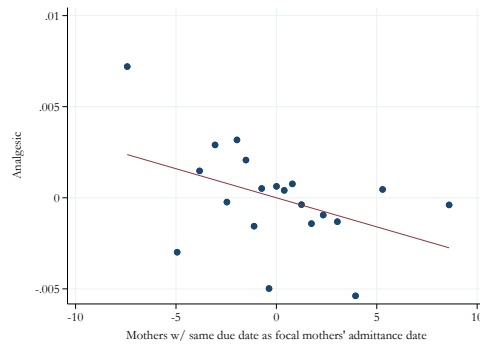
Note: Each coefficient is estimated in a separate regression. Each regression includes controls for parents age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. Robust standard errors in parentheses.



(a) Number of interventions.



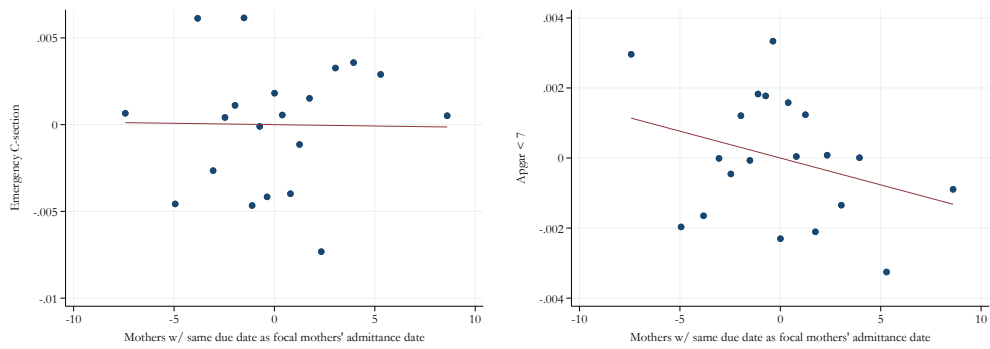
(b) Induced labor.



(c) Any analgesic.

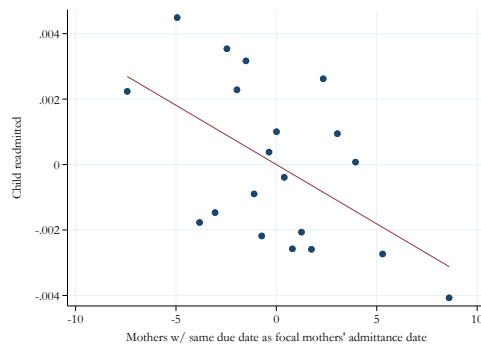
Figure A2. Plots of reduced form.

Note: Each figure reports the binned residuals of outcomes and the number of mothers admitted. The variables are residualized by means of the controls used in Equation 1. See note under Table 4 for details.



(a) Emergency C-section.

(b) Apgar < 7.



(c) Child readmitted.

Figure A3. Plots of reduced form.

Note: Each figure reports the binned residuals of outcomes and the number of mothers admitted. The variables are residualized by means of the controls used in Equation 1. See note under Table 4 for details.

Table A8. Additional Outcomes.

	(1)	(2)	(3)	
	Hospital days	Severe laceration	Severe postpart. hemorrhage	
$\#ad\hat{m}itted$	0.0122 (0.0128)	0.0019 (0.0013)	0.0006 (0.0011)	
Dep. mean	3.24	0.03	0.02	
	(4)	(5)	(6)	(7)
	Mother read. 7 days	Mother read. 30 days	Child read. 7 days	N compl.
$\#ad\hat{m}itted$	-0.0002 (0.0010)	-0.0009 (0.0012)	-0.0027** (0.0013)	-0.0003 (0.0106)
Dep. mean	0.02	0.03	0.04	5.48

Note: Outcomes listed in column titles. “read.” is short for “readmitted”. “compl.” is short for “complications”. “postpart” is short for “postpartum”. Each coefficient is estimated in a separate regression. Each regression includes controls for parents age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. Robust standard errors in parentheses.

Table A9. Main model without controls.

	(1)	(2)	(3)
	N interv.	Ind. labor	Analgesic
$\#ad\hat{m}itted$	-0.0423*** (0.0152)	-0.0101*** (0.0032)	-0.0034* (0.0017)
Dep. mean	4.41	0.22	0.94
	(4)	(5)	(6)
	Emg. C-section	APGAR < 7	Child readmitted
$\#ad\hat{m}itted$	-0.0001 (0.0026)	-0.0016* (0.0009)	-0.0038** (0.0015)
Dep. mean	0.13	0.02	0.05

Note: Outcomes listed in column titles. Each coefficient is estimated in a separate regression. Each regression includes controls for parents age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. Robust standard errors in parentheses. F-statistic first stage = 469.

Table A10. Main model without mothers living in Oslo.

	(1) N interv.	(2) Ind. labor	(3) Analgesic
$\#ad\hat{m}itted$	-0.0846*** (0.0253)	-0.0151*** (0.0053)	-0.0059** (0.0029)
Dep. mean	4.48	0.22	0.94
	(4) Emg. C-section	(5) APGAR < 7	(6) Child readmitted
$\#ad\hat{m}itted$	0.0045 (0.0042)	-0.0039** (0.0016)	-0.0054* (0.0029)
Dep. mean	0.13	0.02	0.06

Note: Each coefficient is estimated in a separate regression. Each regression includes controls for parents age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. Robust standard errors in parentheses. F-statistic first stage 306.

Table A11. Main model with higher parity children.

	(1) N interv.	(2) Ind. labor	(3) Analgesic
$\#ad\hat{m}itted$	-0.0438*** (0.0121)	-0.0108*** (0.0028)	-0.0079*** (0.0029)
Dep. mean	3.39	0.18	0.78
	(4) Emg. C-section	(5) APGAR < 7	(6) Child readmitted
$\#ad\hat{m}itted$	-0.0041** (0.0017)	-0.0000 (0.0007)	-0.0008 (0.0014)
Dep. mean	0.06	0.01	0.05

Note: Each coefficient is estimated in a separate regression. Each regression includes controls for parents age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. Robust standard errors in parentheses. F-statistic first stage = 558.

Table A12. Compliers

	(1)	(2)	(3)	(4)
	First stage	se	Population share	FS / FS all
Full sample	.0959	.0044	1	1
<i>Predicted congestion quartile</i>				
1st	.0426	.0087	.25	.44
2nd	.0372	.007	.25	.39
3rd	.0447	.0068	.25	.47
4th	.0415	.008	.25	.43
<i>Mother's age</i>				
<= median (28)	.0768	.0064	.47	.8
> median (28)	.1011	.006	.53	1.05
<i>Father's age</i>				
<= median (30)	.0927	.0068	.43	.97
> median (30)	.0948	.0058	.57	.99
<i>Mother's income</i>				
<= median	.0876	.0063	.5	.91
> median	.0964	.0062	.5	1.01
<i>Father's income</i>				
<= median	.0832	.0063	.5	.87
> median	.1043	.0061	.5	1.09
<i>Mother's education</i>				
< Bachelor's degree	.0919	.0062	.51	.96
>= Bachelor's degree	.0921	.0062	.49	.96
<i>Father's education</i>				
< Bachelor's degree	.0852	.0054	.66	.89
>= Bachelor's degree	.0972	.0075	.34	1.01
<i>Mother's immigration status</i>				
Not immigrant	.0973	.005	.81	1.01
Immigrant	.0847	.0099	.19	.88
<i>Father's immigration status</i>				
Not immigrant	.0959	.0049	.82	1
Immigrant	.0931	.0103	.18	.97
<i>Child's gender</i>				
Male	.0957	.0062	.52	1
Female	.0967	.0064	.48	1.01

Note: Characterization of compliers. Sample split by predicted congestion based on observable characteristics, and separately in mutually exclusive groupings (denoted in italics) for each observable characteristic used in the main sample. The first column shows the estimated first stage for each sub-sample with standard errors in Column (2). Population shares for each sub-sample are reported in Column (3) and the relative size of the first stage for each sub-sample is reported in Column (4).

Table A13. Results by ward size.

		(1)	(2)	(3)
		N	Ind.	
		interv.	labor	Analgesic
< 90th percentile ward size	$\#admitted$	-0.0912*** (0.0260)	-0.0132** (0.0056)	-0.0102*** (0.0035)
Dep. mean		4.22	0.22	0.93
≥ 90th percentile ward size	$\#admitted$	-0.0501* (0.0267)	-0.0114** (0.0055)	-0.0006 (0.0028)
Dep. mean		4.69	0.22	0.95
		(4)	(5)	(6)
		Emg. C-section	APGAR < 7	Child readmitted
< 90th percentile ward size	$\#admitted$	0.0021 (0.0046)	-0.0051*** (0.0017)	-0.0081** (0.0032)
Dep. mean		0.14	0.02	0.06
≥90th percentile ward size	$\#admitted$	-0.0008 (0.0043)	-0.0006 (0.0016)	-0.0015 (0.0023)
Dep. mean		0.13	0.02	0.04

Note: F-stat for mothers at smaller wards = 306 (N = 85 802), F-stat for mothers at larger wards = 134 (N = 62 203). Outcomes defined in column titles. Each regression includes controls for parental age, education, earnings and immigration status as well as the gender of the child, the due date and local ward fixed effects. Robust standard errors in parentheses.

Table A14. Results using alternative measures of congestion.

	(1) N interv.	(2) Ind. labor	(3) Analgesic
<i>Centile</i> ($\#admitted$)	-4.3484*** (1.3961)	-0.8598*** (0.2978)	-0.1008 (0.1553)
Dep. mean	4.41	0.22	0.94
	(4) Emg. C-section	(5) APGAR < 7	(6) Child readmitted
<i>Centile</i> ($\#admitted$)	-0.0386 (0.2254)	-0.1530* (0.0865)	-0.1512 (0.1537)
Dep. mean	0.13	0.02	0.05

Note: F-statistic running centile = 60. Each coefficient is estimated in a separate regression. Each regression includes controls for parents age, education, earnings and immigration status as well as the gender of the child and due date and local ward fixed effects. Robust standard errors in parentheses.

Table A15. Results by pregnancy-related contacts prior to birth.

		(1)	(2)	(3)
		N	Ind.	Analgesic
		interv.	labor	
< median	<i>#admitted</i>	-0.0314	-0.0135**	-0.0041
	<i>#contacts</i>	(0.0275)	(0.0058)	(0.0035)
Dep. mean		4.31	0.19	0.93
≥ median	<i>#admitted</i>	-0.0709**	-0.0121*	-0.0024
	<i>#contacts</i>	(0.0309)	(0.0065)	(0.0030)
Dep. mean		4.63	0.26	0.95
		(4)	(5)	(6)
		Emg.	APGAR < 7	Child
		C-section		readmitted
< median	<i>#admitted</i>	0.0004	-0.0012	-0.0033
	<i>#contacts</i>	(0.0046)	(0.0017)	(0.0027)
Dep. mean		0.12	0.02	0.05
≥ median	<i>#admitted</i>	-0.0017	-0.0025	-0.0054*
	<i>#contacts</i>	(0.0051)	(0.0019)	(0.0031)
Dep. mean		0.15	0.02	0.06

Note: F-stat for mothers with fewer contacts = 131 (N = 61 067), F-stat for mothers with more contacts = 131 (N = 65 649). Outcomes defined in column titles. Each regression includes controls for parental age, education, earnings and immigration status as well as the gender of the child, the due date and local ward fixed effects. See the Appendix for the definition of pregnancy-related contacts. Robust standard errors in parentheses.

A.2 Data details

The medical data used in this paper come from three separate sources: The Medical Birth Registry of Norway (MFR), National Patient Registry of Norway (NPR), and the Control and Payment of Health Reimbursement (KUHR). Each of these registries follows separate coding procedures. I detail the coding systems and codes used below.

The Medical Birth Registry of Norway (MFR) All births in Norway have been required to be reported to the MFR since 1967. For each birth, midwives or physicians fill out a specific form on which they check off boxes for various complications, treatments, interventions and birth outcomes. Over time, the level of detail on the forms has increased and reports also contain reporting on several interventions and complications through international coding systems. In recent years the MFR has used a combination of different coding standards: mainly ICD-10, ATC, and a check-list system, but also NCSP and NCMP. Categories on the check list are loosely based on medical treatments and diagnoses also found in other coding systems such as ICD-10, but do not completely overlap. Examples of cross-list items included are “Anesthesia”, and “perineal rupture grade 1 or 2” both of which can be answered “yes” or “no”. When possible I have relied on ICD-10 codes and ATC codes, but for most variables I rely on the check-list system. I detail below how each of the medical outcomes is defined for each registry.

The National Patient Registry of Norway (NPR) The NPR contains all records of inpatient treatments in Norwegian public hospitals since 2008. Each episode is recorded with institution, start and end dates and procedural codes in the ICD-10 standard. I use the registry to identify admission and discharge dates for birth-related hospitalization spells, institutions (ward/hospital), and treatments related to pregnancies. As ICD-10 codes are only available at the 2-digit level, I refer to the MFR for data on interventions during childbirth.

The Control and Payment of Health Reimbursements (KUHR) Each treatment provided at a public outpatient institution is reported to the KUHR

for reimbursement. Public outpatient institutions include all GP services, public psychiatric services, etc. The data contain dates, service provider type, and treatment provided. The data additionally contain information on the reason for the visit in many cases. The ICPC-2 is the common coding system for reporting in the KUHR.

Table A16. Construction and sources of health-related variables.

Variable	Variable type	Code type(s)	Codes used	Source(s)
Induction	Dummy	Check-list		MFR
Emergency C-section	Dummy	Check-list		MFR
Analgesic	Dummy	ATC and check-list	N02	MFR
N interventions	Integer	Check-list		MFR
APGAR 5 min. < 7	Dummy	Constructed from check-list		MFR
GP visits	Integer			KUHR
Hospital visits	Integer			NPR
Psychiatric visits	Integer			KUHR
Child readmitted within 30 days	dummy			NPR
Diagnoses related to pregnancy prior to birth	integer	ICPC-2 ICD-10	A94, W3-5, W28-29, W70-76, W80-81, W84-85 O10-16, O20-29, O30-39, Z35	KUHR NPR
Pregnancy-related contacts with healthcare providers	integer	ICPC-2 ICD-10	W Z34-35, O10-16, O20-29, O30-36	KUHR NPR