NBER WORKING PAPER SERIES

THE HEALTH IMPACTS OF HOSPITAL DELIVERY PRACTICES

David Card
Alessandra Fenizia
David Silver

Working Paper 25986
http://www.nber.org/papers/w25986

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2019

We thank Pablo Muñoz and Alice Wu for exceptional research assistance, Pat Kline, Marit Rehavi, Chris Walters, and Matt Notowidigdo for many detailed comments, and seminar participants at UC Berkeley, UC Merced, UCLA, UCSB, Princeton and NBER Summer Institute (Children) for helpful comments and suggestions. We are also extremely grateful to Laura Giuliano for conversations that led to this project. Our use of the data in this project is covered under UC Berkeley Federalwide Assurance #00006252. Silver gratefully acknowledges funding from the National Institute on Aging, Grant Number T32-AG000186. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Health Impacts of Hospital Delivery Practices
David Card, Alessandra Fenizia, and David Silver
NBER Working Paper No. 25986
June 2019
JEL No. D22,I11,I18,J13

ABSTRACT

Hospital treatment practices vary widely, often with little connection to the medical needs of patients. We assess the impact of these differences in the context of childbirth. We focus on low-risk first births, where c-section rates vary enormously across hospitals, and where policymakers have focused much of their attention in calls for reducing unnecessary c-sections. We find that proximity to hospitals with high c-section rates leads to more cesarean deliveries, fewer vaginal births after prolonged labor, and higher average Apgar scores. Infants born in these hospitals are less likely to be readmitted in the year after birth, but more likely to visit the emergency department for a respiratory-related problem. They also have lower mortality rates, driven by a reduction in the joint probability of prolonged labor and subsequent death. A stylized cost benefit analysis suggests that re-allocating births to high c-section hospitals could lead to net social benefits.

David Card
Department of Economics
549 Evans Hall, #3880
University of California, Berkeley
Berkeley, CA 94720-3880
and NBER
card@econ.berkeley.edu

Alessandra Fenizia
Department of Economics
549 Evans Hall, #3880
University of California, Berkeley
Berkeley, CA 94720-3880
afenizia@econ.berkeley.edu

David Silver
Department of Economics
Julis Romo Rabinowitz Building
Princeton University
Princeton, NJ 08544
david.silver@princeton.edu

Hospitals vary widely in clinical practices and the utilization of intensive treatments.[1] Whether and how these differences translate into health outcomes is a key policy concern (e.g. Baicker et al., 2012). The traditional view is that supplier-induced treatments have little impact on patient outcomes (e.g. Fisher et al., 2003a, 2003b), perhaps because poor hospital management contributes to wasteful practices (Bloom et al., 2015). A recent study by Doyle et al. (2015), however, points to significant health benefits of being routed to a high-cost hospital.

In the context of childbirth, rates of cesarean delivery – the most common surgical procedure in the United States – have been documented to vary dramatically across hospitals in ways uncorrelated with medical need (Kozhimannil et al., 2013; Kozhimannil et al. 2014). The usual finding from observational comparisons is that cesarean delivery is associated with *worse* outcomes for infants and mothers.[2] This evidence has led the American College of Obstetricians and Gynecologists, the US Department of Health and Human Services, and the Joint Commission to issue a series of clinical guidelines with the goal of reducing cesarean delivery rates, with either implicit or explicit focus on hospitals with high c-section rates (ACOG, 2006; Spong et al., 2012; ACOG/SMFM, 2014; Joint Commission, 2016).[3]

These guidelines focus on cesarean delivery by first-time mothers with no clear indications of medical need, a group we term "low-risk first births" (LRFBs).[4] In our California-based data, as in the nation as a whole, about one quarter of LRFBs are delivered by c-section, with significant variation across hospitals even within narrow geographic areas. We interpret this variation as in part reflecting

[1] See Romley et al. (2011) and Doyle et al. (2015) for analyses of hospital-specific spending and outcomes, Rysavy (2015) for a study of variation in treatment of pre-term infants, and Barnato et al. (2005) for an analysis of hospital-specific variation in the context of racial disparities in treatment of AMI. Skinner (2012) provides a review of the related literature on regional variation in intensity of care or spending; recent contributions in this vein include Cutler et al. (2019) and Finkelstein et al. (2016).

[2] See e.g., Clark and Silver (2011), Gregory et al. (2012), Goer et al. (2012), and Hyde et al. (2012).

[3] In 2020 the Joint Commission will begin publicly reporting hospitals with consistently high cesarean delivery rates among "low-risk women having their first birth with a term, singleton baby in a vertex position" (Joint Commission, 2018).

[4] Specifically, following Osterman and Martin (2014), we define LRFBs as term (37+ weeks gestation), singleton, vertex first births, and further exclude mothers under 18 or over 35, mothers with BMI above the 90th percentile, and mothers with eclampsia, pre-eclampsia, growth restrictions, or >20 pre-natal visits.

hospital-specific practice standards over how long labor is allowed to progress before patients are referred for c-section.[5]

Among LRFBs, the tradeoff between earlier c-section and longer labor is widely debated.[6] On one hand, prolonged labor may amplify fetal stress and cause health problems for children, as well as increasing the risk of hemorrhage and lacerations for mothers. On the other, cesarean deliveries typically cost more than vaginal births (Podulka et al., 2011; Sakala et al., 2013); they often preclude future vaginal births; and they are correlated with worse health outcomes, including poorer respiratory health of infants (Hyde et al., 2012), higher risk of maternal infection, and prolonged recovery for mothers.[7] Nearly all the evidence on the health effects of delivery mode is observational, however, making causal interpretations difficult. Moreover, a systematic accounting of costs and benefits of earlier c-section versus longer labor is lacking.

In this paper, we attempt to fill this gap by assessing the leading costs and benefits of alternative hospital delivery practices.[8] Systematic sorting of patients across hospital types is a key empirical challenge. Our research design uses the fact that mothers tend to deliver at the nearest hospital (e.g., Phibbs et al., 1993; Currie and MacLeod, 2017). Building on McClellan, McNeil, and Newhouse (1994), we classify hospitals in each Hospital Referral Region (HRR) into two groups based on their risk-adjusted average cesarean rate for LRFBs and use the relative distance from a mother's home zip code to the nearest high c-section (H) hospital versus the nearest low c-section (L) hospital as an instrumental variable for delivery at an H hospital.

---

[5] The leading motive reported for c-sections among LRFBs is the failure of labor to progress (ACOG/SMFM, 2014). Progress of labor is commonly judged against Friedman's curve (Friedman, 1955).
[6] See for example the commentary by Cohen et al. (2018), criticizing the new guidelines for labor management published by American College of Obstetricians and Gynecologists.
[7] In Appendix A we present an abbreviated summary of recent studies on health risks associated with c-section.
[8] Two prospective RCTs of "active management of labor" interventions to reduce c-section rates (López-Zeno et al., 1992; Frigoletto et al. 1995) reached different conclusions about whether such programs had an effect on c-section rates. More recently, Gimovsky and Berghella (2016) implemented a small (N=78) RCT to extend labor for women with a prolonged second stage, which substantially reduced c-section rates. These studies were under-powered for studying subsequent health effects on mothers and infants.

We implement this design using a California data set that combines hospital discharge records for mothers and newborns, birth certificate information, inpatient and outpatient records for infants in the year after birth, and parallel records for mothers in the year before birth. These data provide indicators of infant health for a large sample of LRFBs, as well as detailed information on maternal characteristics and predetermined risk factors.

A concern with inferences based on relative distance is that patient health may be correlated with proximity to high- or low-intensity hospitals (Hadley and Cunningham, 2004). To address this, we focus on *within*-Hospital Service Area (HSA) comparisons and control for the fraction of government-insured mothers in a mother's zip code. Conditional on HSA, the local share of government-insured mothers, and distance to *any* hospital, we show that relative distance to a high c-section hospital is uncorrelated with a wide set of maternal characteristics and risk factors, including education, race, insurance status, prenatal care, maternal smoking, and infant birth weight. We also document the insensitivity of all our key results to including any, all, or none of these additional factors in our models.

Relative distance to a high versus low c-section hospital strongly affects the probability of delivery at a high c-section hospital and the probability of cesarean delivery. The increase in cesarean deliveries is largely offset by reductions in vaginal births occurring a day or more after admission to the hospital, as would be expected if physicians at high c-section hospitals tend to refer patients for c-section earlier in the labor process.

We find that *hospital compliers*, who shift their place of delivery from low to high c-section hospitals based on relative distance, have lower education than other mothers and are more likely to be covered by Medicaid or other government insurance. Mothers who deliver by c-section when shifted to a high c-section hospital – *hospital and procedure compliers* – have even lower education and are even more likely to be covered by government insurance, suggesting that practice style differences have larger impacts on relatively disadvantaged patients.[9]

Turning to impacts at birth, our instrumental-variables estimates show that delivery at an H hospital is associated with significantly higher 5-minute Apgar scores (with an effect size of around 0.11σ's per birth). Much of the gain comes from a reduction in the fraction of late vaginal births (≥1 day in hospital) with low Apgar scores, consistent with existing studies of the negative effects of prolonged labor (e.g., Altman et al., 2015). We also find slightly shorter average hospital stays for the newborn, though the effect is statistically insignificant.

For mothers, we find that delivering at H hospitals leads to substantial reductions in the rates of perineal lacerations and other traumas. We also document that delivery at H hospitals leads to a shorter average time between admission and birth but a longer average time from birth to discharge, with no significant change in the overall length of stay, similar to the results for infants.

In the year after birth we find that infants delivered at H hospitals have a higher probability of an ED visit. Consistent with a growing literature documenting adverse respiratory development after c-section (Hyde et al., 2012), the majority of these additional visits are attributable to respiratory-related diagnoses. Offsetting these effects, however, we find a lower probability of readmission to hospital, concentrated in the neonatal period (first 28 days).

An important issue for interpreting these impacts is whether they might be driven in part by *other* practice differences between H and L hospitals – a so-called "correlated treatments" problem (McClellan et al., 1994). To address practice differences that affect ED use, we classify hospitals by the probability that infants born there have an ED visit in the year after birth and form a second instrumental variable based on the relative distance to high ED use hospitals. We then fit models that include two

[9] In California over our sample period the MediCal system paid the same amount to hospitals for vaginal and cesarean delivery, so this finding is not driven by financial incentives to perform more cesareans on government-insured patients. See Alexander (2015) for a discussion of delivery mode choice and financial incentives under Medicaid in a national sample of births.

endogenous hospital characteristics: high versus low cesarean delivery rate, and high versus low ED use rate. We follow a parallel approach for inpatient readmissions, classifying hospitals as high versus low-readmission, forming a third instrumental variable based on the relative distance to high readmission hospitals, and fitting models that control for both delivery practices and infant readmission differences.

We find that hospital-specific factors affecting the relative fraction of cesarean deliveries are very weakly correlated with factors affecting ED use or re-admission rates of infants born in that hospital. As a result, controlling for a second channel (or for all three channels at once) has little effect on our conclusions about the impacts of delivering at a high c-section hospital on ED visits or inpatient readmissions.

We then turn to consider the effects of delivery practices on infant deaths. Delivery at a high c-section hospital is associated with a relatively large reduction in infant mortality -- on the order of 2.4 fewer deaths per 1,000 deliveries -- with *p*-values around 0.02. This effect is highly robust to controlling for maternal demographics, infant risk factors, and any independent effect of delivery at high inpatient readmission hospitals. Moreover, about two-thirds of the overall reduction in death is attributable to a decline in the joint event of late delivery (i.e., delivery ≥1 day after mothers' admission) and death. Taken together with our findings on Apgar scores and readmissions in the neonate period, we believe the evidence points to a significant risk of death after prolonged labor that is partially mitigated by practices at H hospitals.

As in other settings, the applicability of our findings depends on whether the treatment effects associated with delivery at a high c-section hospital are different for compliers than for other infants. Using a series of correlated random coefficient (CRC) models (Garen, 1984; Heckman and Vytlacil, 1998), we find evidence of larger health impacts on infants whose mothers have unobserved preferences for low-c-section hospitals. One possible interpretation of this finding is that mothers who prefer vaginal births are less likely to switch to high c-section hospitals because of distance, but the treatment effects for their infants are larger if they do switch. In any case, the estimated *average* treatment effects in our CRC models are very close to the *local average* treatment effects from our simpler IV models.

We conclude by putting together our estimates in a partial cost-benefit analysis that accounts for the lifetime costs of respiratory disease, the delivery and health costs of subsequent births to mothers who receive a primary c-section, and the impacts on mortality. We estimate that the net benefit of delivering an extra 1,000 births at a high c-section hospital is between $7 and $8 million. Even assuming that the mortality effect is only one-half as large as our main estimate we find a net benefit of around $1 million. While we cannot necessarily generalize these findings to other settings, we take some comfort in the fact that California's aggregate c-section rates are quite similar to the US as a whole in our period of study.[10] At a minimum, our estimates suggest that policies aimed at reducing primary c-section rates by extending the time until surgical intervention may have substantial offsetting costs.

Our paper contributes to an important literature estimating the effects of hospital practices on patient health outcomes. We build directly on the work of McClellan, McNeil, and Newhouse (1994), Cutler (2007), and Chandra and Staiger (2007), who use distance-based designs to study treatment of heart attacks. In recent work, Doyle et al. (2015) leverage quasi-random assignment of ambulance companies to study the health outcomes of elderly Medicare patients with non-deferrable conditions routed to high-cost hospitals. Our paper extends this literature to another domain – childbirth – where we know far less about the impacts of hospital practices, but where, as in the majority of hospital-based care, ambulance-based designs are infeasible.[11] We believe our empirical strategy serves as a viable alternative for evaluating the impacts of hospital practices in this and other non-emergency care settings.

Second, we contribute to a developing literature on the relationship between management and productivity in healthcare (e.g. Bloom et al., 2015). Our results suggest that the clinical management of

[10] California's overall c-section rate of 33.2% in 2011 was comparable to the national rate of 32.8% (Martin et al., 2013a).
[11] Only 16% of emergency department visits arrive by ambulance, regardless of whether they are admitted to the hospital (authors calculations based on 2016 National Hospital Ambulatory Medical Care Survey). In the Traditional Medicare hospitalizations studies in Doyle et al. (2015), only one quarter of non-deferrable hospitalizations were brought to the hospital's emergency department by ambulance.

the timing of interventions (in our case, the timing of c-sections within the process of labor) is an important factor in their effectiveness: earlier interventions improve survival outcomes in what would have become difficult labors. In a simple context where c-sections are performed only if the mother's length of labor passes some predetermined threshold, our results suggest that hospitals that set earlier thresholds will consequently have higher c-section rates and better outcomes among marginal births.

Finally, we contribute to a growing economics literature on the consequences of treatment choices at delivery. Jensen and Wüst (2013) show that breech births benefit from c-section delivery. Jachetta (2014) uses malpractice premiums to instrument for area-level c-section rates and finds that cesarean delivery leads to higher incidence of asthma. Halla et al. (2018) document a fertility effect of cesarean delivery for Austrian mothers. Costa-Ramón et al. (2018) use arrival time at the hospital as an exogenous determinant of c-section rates and find that cesarean delivery leads to lower Apgar scores. We quantify some of the main infant health effects of alternative delivery practices while paying close attention to the issue of instrument validity and the possibility of multi-dimensional practice style differences across hospitals.

The paper proceeds as follows. Section II provides an overview of childbirth and hospital delivery practices and lays out our econometric framework. Section III describes our data sources, sampling scheme, and the construction of our instruments. Section IV discusses first-stage results and complier characteristics. Section V provides the main results of our analysis of health outcomes. Section VI summarizes our results in a simple cost-benefit framework. Section VII concludes.

II. An Overview of C-Section and Our Modeling Approach

II.a. Hospital Practice Styles and Delivery Outcomes

Figure I provides a stylized overview of the pathways leading to c-section. The left branch shows the pathway for mothers with a planned (or scheduled) c-section. This group includes women who have had a previous c-section, those with breech pregnancy or multiple fetuses, and those with risk factors like obesity or eclampsia (Declercq et al, 2006; Zhang et al. 2010). Their c-sections occur with no attempted labor, often several days before normal term.

The right branch shows the pathway for mothers who reach normal term with no scheduled intervention. Typically, a mother-to-be shows signs of labor and is admitted to hospital where her progress is monitored and pain relief and labor-augmenting medications are administered.[12] Barring other factors, a decision to perform c-section at hospital h is reached when labor time exceeds the threshold T_h (which can vary with maternal characteristics and other information), resulting in an unscheduled or intrapartum c-section. Hospital practices vary widely over how long to allow labor to proceed (Zhang et al., 2010; Kozhimannil et al, 2013, 2014), leading to wide variation in the probability of intrapartum c-section.

Given these two very different pathways we focus on "low-risk" first births, eliminating twins, breech presentations, births to mothers younger than 18 or over 35, and five other risk factors.[13] We classify hospitals as having high or low cesarean rates for LRFBs relative to other hospitals in the same HRR, and use a distance-based design to identify the causal effects of delivering at a high c-section hospital. Our interpretation is that risk-adjusted differences in c-section rates for LRFBs reflect differences in the hospital-specific threshold T_h.[14] Several factors could play a role in this variation, including financial incentives (e.g. Gruber and Owings, 1996; Gruber et al. 1999; Alexander, 2015), malpractice pressures or experiences (e.g. Dubay et al. 1999; Baicker et al., 2006; Currie and MacLeod, 2008; Shurtz, 2013), and differences in medical training (e.g. Epstein and Nicholson, 2009). Rather than try to identify these factors, however, we take a data-driven approach and simply classify hospitals as H

---

[12] Declercq et al. (2006) report that that 76% of all U.S. mothers had epidural anesthesia during labor. Many practitioners believe this slows labor and makes c-section more likely, though the evidence is controversial – see Howell (2000) and Klein (2006). Reporting rates of epidural anesthesia appear to vary widely across hospitals in our sample and we do not attempt to control for this factor.

[13] We do not attempt to eliminate scheduled c-sections since the classification depends on indicators of labor on the mother's discharge record which are known to be under-reported (Henry et al., 1995). See the discussion below.

[14] We present supporting evidence on the timing of birth relative to the day of arrival of the mother below.

or L. This is similar to the approach taken by Finkelstein et al. (2016) in classifying HRRs based on average spending by Medicare-insured patients.

II.b. Econometric Framework

Next, we briefly summarize our econometric framework. This consists of simple linear models for the choice of a high c-section hospital by mother *i* (denoted by H_i), cesarean delivery (C_i), and a health outcome for the infant (y_i):

$$H_i = \delta_0 + \delta_1 Z_i + \delta_x X_i + u_i \quad (1)$$

$$C_i = \pi_0 + \pi_1 Z_i + \pi_x X_i + \eta_i \quad (2)$$

$$y_i = \tau_0 + \tau_1 Z_i + \tau_x X_i + \xi_i \,. \quad (3)$$

In each case the explanatory variables are Z_i, the relative distance from the mother's home to a low versus high c-section hospital, and X_i, a vector of controls. Equations (1) and (2) are first-stage models for choice of hospital type and c-section delivery. Equation (3) is a reduced-form model for the effect of relative distance on health.

In this setup there are two possible IV estimators. The first is formed by dividing the reduced form effect of relative distance (τ_1) by the first stage effect on the type of hospital (δ_1). This estimator scales the reduced form effect *per additional delivery* at H hospitals. The second IV estimator divides τ_1 by the first stage effect of relative distance on the probability of c-section (π_1), scaling the reduced form effect *per additional c-section*.

To clarify the interpretation of these estimators, consider a binary version of relative distance, Z_i^B, which indicates whether a mother's home is closer to a high c-section hospital or not.[15] Let H_{0i} and H_{1i} represent indicators for whether mother *i* would choose a high c-section hospital when $Z_i^B = 0$ or $Z_i^B = 1$, and let C_{0i} and C_{1i} represent indicators for whether she would deliver by c-section. The potential responses to changes in Z_i^B are represented by the 4-tuple $(H_{0i}, H_{1i}), (C_{0i}, C_{1i})$. For example, mothers with $(H_{0i}, H_{1i}), (C_{0i}, C_{1i})=$ (0,1),(0,0) are those who switch hospital types in response to distance but who always deliver vaginally (i.e., H-compliers/V always-takers, using the standard LATE nomenclature).

We make three standard assumptions. First, we assume there are no H-defiers (i.e., $H_{1i} \geq H_{0i}$). Second, we assume that distance has no direct effect on delivery mode, so $C_{1i} = C_{0i}$ if $H_{1i} = H_{0i}$.[16] Finally, we assume there are no H-complier/C-defiers, ruling out rank-reversals in treatment intensity for H compliers.[17] Under these assumptions there are three groups of H-complying mothers: (1) those who switch from vaginal to cesarean delivery when closer to an H hospital (*H&C compliers*); (2) those who always deliver by cesarean (*H complier & C always-taker*); (3) those who always deliver vaginally (*H complier & V always-taker*).

Let $\rho_1(x), \rho_2(x), \rho_3(x)$ represent the population shares of these three groups for patients with $X_i = x$. Then, for a given *x*-group the first-stage effect of relative distance on the probability of

[15] In our analysis below we focus on a continuous measure of relative distance. However, we get similar IV estimates using an indicator for being closer to a high c-section hospital.
[16] This rules out the possibility, for example, that being closer to the hospital affects the stage of labor at arrival, which in turn affects the probability of c-section. To address concerns about the role of travel time we include a measure of the mother's distance to the nearest hospital of any type in all our models.
[17] Rank invariance is routinely assumed in the analysis of quantile treatment effects (e.g., Chernozhukov and Hansen, 2005). Currie and MacLeod (2017) consider a setting where rank-reversal could be important. They document substantial heterogeneity in physician diagnostic ability, which could lead to rank reversals if better diagnosticians are concentrated at certain hospitals. Importantly, our low-risk first birth sample excludes most of the higher-risk births considered by Currie and MacLeod (2017).

delivery at an H hospital (i.e., δ_1) identifies $\rho_1(x)+\rho_2(x)+\rho_3(x)$, and the first-stage effect on the probability of c-section (i.e., π_1) identifies $\rho_1(x)$. More generally, assuming that X consists of a set of dummies for mutually exclusive subgroups, the first-stage coefficients for the overall sample identify the *average* fractions of H compliers and H&C compliers across subgroups (see Appendix B).

Next, let $Y_i(h,c,z,x)$ represent a potential health outcome that would be observed for birth i conditional on hospital type *h,* delivery mode *c,* relative distance *z*, and covariates *x*. We assume:

$$Y_i(h,c,z,x) = Y_i(h,c,x)$$

i.e., that health does not depend on relative distance conditional on delivery mode, hospital type, and *x*. The treatment effects for the three subgroups of H-compliers whose hospital choice changes with Z_i^B are:

$$\mu_1(x) = E[Y_i(1,1,x) - Y_i(0,0,x) \mid H \,\&\, C \textit{ complier}]$$

$$\mu_2(x) = E[Y_i(0,1,x) - Y_i(1,1,x) \mid H \textit{ complier} \,\&\, C - AT]$$

$$\mu_3(x) = E[Y_i(0,0,x) - Y_i(1,0,x) \mid H \textit{ complier} \,\&\, V - AT].$$

Note that μ_1, the treatment effect for H&C compliers, combines the effects of a switch from vaginal to cesarean delivery and other effects of delivering at an H hospital, such as a shorter labor. For the other complier groups delivery mode is fixed and the health effects depend only on other practice differences at H hospitals (e.g., a shorter labor for C-always takers).

The reduced-form health effect for a given *x*-group is an estimate of $\rho_1(x)\mu_1(x)+\rho_2(x)\mu_2(x)+\rho_3(x)\mu_3(x)$ which combines the probabilities of each complier group with their treatment effects. Scaling by the first-stage effect on the probability of delivery at an H hospital yields an estimate of the average treatment effect per H complier:

$$(\rho_1(x)\mu_1(x)+\rho_2(x)\mu_2(x)+\rho_3(x)\mu_3(x))/(\rho_1(x)+\rho_2(x)+\rho_3(x)). \qquad (4)$$

Scaling by the first stage effect on the probability of C-section, on the other hand, yields an estimate of $\mu_1(x)+(\rho_2(x)\mu_2(x)+\rho_3(x)\mu_3(x))/\rho_1(x)$. If there are no health effects on the C or V always-takers, this is an estimate of $\mu_1(x)$. Given the strong likelihood that practice differences affect the health of the always-taker groups, however, we focus on estimates that scale by the fraction of births moved from L to H hospitals and interpret the IV estimate as a weighted average of effects for the 3 complier groups.

III. Data Sources, Sample Overview, Relative Distance Instrument

III.a. Data Sources

We use a linked cohort data set created by the California Office of Statewide Health Planning and Development (OSHPD) that combines patient discharge (PD) records, emergency department (ED) records, ambulatory surgery (AS) records, and vital statistics (VS) records for all in-hospital births between 2007 and 2011.[18] Specifically, discharge records for the birth stay of the infant and mother are linked with birth certificate data, PD/ED/AS and VS records over the following year, and PD/ED/AS records for mothers in the year prior to birth. The resulting data set includes VS-derived information on the mother (e.g., education, race, weight, and prenatal care) and infant (gestation, birthweight, Apgar score), as well as PD-derived information on diagnoses at delivery. Pre-birth PD/ED/AS records provide additional indicators of maternal health (such as the number of ED visits in the year prior to birth). The post-partum PD/ED/AS and VS records provide our main health outcomes (hospital and ED visits and infant death in the year after birth).

A limitation of this data set is the absence of information on visits to physician offices, community clinics, and similar facilities. Thus, we miss health problems for infants or mothers that are treated in these settings rather than at licensed hospitals or AS centers. And despite the relatively rich

[18] This is known as PDD/ED/AS/Linked Birth Cohort data and is available to researchers through OSHPD. See Appendix C for more information on the characteristics of the data and the derivation of our samples.

information from the birth stay records and the birth certificate, it also lacks direct clinical information on factors like whether a cesarean occurred after a trial of labor.[19] The offsetting benefit is that we have a relatively large sample size, allowing us to detect plausible-sized effects with an IV research design.

III.b. Sample Overview

Table I provides an overview of the characteristics of all 2.7 million births in California during our 5-year sample window (column 1) and of all low-risk first births (column 2). We define LRFBs as singleton non-breech first births delivered at 37+ weeks of gestation, corresponding to the two lowest risk groups in Robson's (2001) classification. We further eliminate mothers under 18 or over 35 and those with any of 5 other risk factors: eclampsia, pre-eclampsia, growth restrictions, BMI >90th percentile, or >20 prenatal visits. We do ***not*** condition on other risk factors (including birthweight), allowing us to test for orthogonality of our distance-based instrument with indicators of infant health.

The entries in column 1 show that about 50% of all California mothers are Hispanic, one-half have no more than high school education, and one-half have their delivery paid by government insurance (mainly Medi-Cal, the state's version of Medicaid). All three rates fall to around 40% among LRFB mothers. LRFB mothers are also younger and lighter. Overall about one third of all California births and one quarter of LRFBs were delivered by c-section during our sample period, similar to the national averages reported by Osterman and Martin (2014).

III.c. Construction of Relative Distance Instrument

Our distance-based design relies on a prior classification of hospitals. Since c-section rates vary across regions of California, we elected to define high and low c-section hospitals *within* Health Referral

[19] Martin et al. (2013b) evaluate the quality of the medical and health data recorded on birth certificates and conclude that while some information (e.g. parity) is relatively accurate, other information (e.g., fetal intolerance of labor) is poorly recorded.

Regions (HRRs).[20] As detailed in Appendix C, we fit a logit model for cesarean delivery on our LRFB sample, including hospital dummies and a set of risk factors. We classify a hospital as "high c-section" (H) for mothers in a given HRR if its risk-adjusted c-section rate (i.e., the hospital effect in the logit) is above the patient-weighted mean for all hospitals in that HRR. Otherwise it is classified as "low c-section" (L).[21] We note that if a hospital serves mothers from two adjacent HRRs it may be classified as H for mothers in one HRR and L for those in the other, depending on the c-section rates at other hospitals in the two HRRs.

Appendix Table I shows that the average cesarean delivery rate of LRFBs is 29% at H hospitals and 22% at L hospitals. About two-thirds of the difference is attributable to a higher rate of c-sections that are performed in cases where there are no indications of a trial of labor (i.e., unscheduled c-sections). H hospitals are also more likely to be for-profit (18% vs. 9%) and less likely to have a NICU unit (74% vs. 86%). Nevertheless, the two types of hospitals have similar average numbers of deliveries per year (3,695 versus 3,635). Even L hospitals perform an average of 800 cesarean deliveries per year on LRFBs, so the staff at these hospitals have wide experience with the procedure.

We then calculate the distance from the centroid of a patient's home zip code to the centroid of the zip code of the nearest H hospital (d_{Hi}) and the nearest L hospital (d_{Li}). We define the relative distance $Z_i \equiv d_{Li} - d_{Hi}$ and a simple binary indicator for being closer to an H hospital $Z_i^B = 1[Z_i \geq 0]$. We also define the distance to nearest hospital $d_i^{MIN} = \min[d_{Li}, d_{Hi}]$.

The third column of Table I presents the characteristics of the subsample of LRFBs that have non-missing patient zip code information, non-missing values for all the control variables in our main

[20] Hospital referral regions (HRRs) represent regional health care markets for tertiary medical care, defined by the Dartmouth Atlas (see Dartmouth Atlas, undated). There are 25 HRRs in our sample of LRFBs. If we classified hospitals on a statewide basis, we would have many more high c-section hospitals in Southern and Central California and many more low c-section hospitals in Northern California.

[21] One concern is that the estimated hospital logit coefficients models for c-section simply reflect unobservable case-mix differences. To probe whether our classification algorithm is biased by case-mix differences, we reclassified hospitals with various sets of risk-adjusters. Reassuringly, the resulting classifications are all nearly perfectly correlated with our preferred classification, suggesting that case-mix differences, although likely present, are unlikely to significantly affect our results.

specifications, and have $d_{Hi} \leq 20$ miles, $d_{Li} \leq 20$ miles, and less than 20 miles between the mother's home zip code and actual hospital she delivered in. These restrictions eliminate about 20% of LRFBs, leaving a final analysis sample of 491,604 births. A comparison of columns 2 and 3 suggests that this sample is quite similar to the overall LRFB sample.

Figure II illustrates the strong relationship between relative distance and hospital choice for LRFB mothers. Here we plot the fraction of mothers in each zip code who deliver at an H hospital against the value of Z_i for mothers in that zip code. The data suggest a nearly symmetric S-curve relationship, tending toward a minimum of about 10% when $Z_i < -15$ and a maximum of about 90% when $Z_i > 15$.

III.d. Evaluating Instrument Validity

A concern with a distance-based IV strategy is that relative distance may be correlated with underlying determinants of patient health (e.g., Hadley and Cunningham, 2004; Garabedian et al., 2014). To assess this concern, we estimated a series of OLS models for a set of observed maternal characteristics and infant risk factors, looking for evidence of a correlation with relative distance. Specifically, for each risk factor we fit a model of the form:

$$R_i = \psi_0 + \psi_1 Z_i + \psi_X X_i^0 + \zeta_i \qquad (5)$$

where X_i^0 is a set of basic control variables we include in all our outcome models. This includes HSA effects, year effects, distance to the nearest hospital (d_i^{MIN}), and a simple measure of neighborhood economic status based on the fraction of *all* mothers (including first- and higher-parity births) from the same zip code who were covered by government insurance at their delivery. The latter variable is included to control for the possibility that lower-income families are sorted within HSAs and are more likely to live near hospitals with higher (or lower) risk-adjusted c-section rates.

Results for 31 different maternal characteristics and risk factors are summarized in Table II. On the mother's side we examine age, education, race, insurance status, height, weight, BMI, use of hospital or ED in the year prior to birth, timing and number of prenatal care visits, diabetes, herpes, asthma, and smoking. For the infant we examine gestation, birth weight, and incidence of low birth weight. We also examine average income in the mother's zip code and 4 other characteristics of new mothers in that zip code (calculated leaving out the mother in question): mean education, the mean fraction with less than high school education, and the mean fractions black and Hispanic.

Column 1 reports estimates of the ψ_1 coefficients from this exercise, with standard errors clustered at the mother's zip code in column 2. Controlling for HSA and the fraction of mothers in the same zip code with government insurance, mothers living closer to H or L hospitals are statistically indistinguishable from one another. Only 2 of the 31 estimated effects of relative distance are near statistical significance: reported maternal herpes (*t*-statistic = -1.73) and maternal asthma (*t*-statistic = 2.00). To summarize the overall pattern of the differences in risk factors we fit logit models for 3 outcomes -- an infant ED visit in the year after birth, an infant readmission to hospital, and infant death -- using all 31 predictors in the table. We find no evidence that infants whose mothers live nearer to H hospitals have higher or lower predicted risks of these outcomes. Finally, we perform a joint F-test of the null hypothesis that the coefficients on all 31 listed characteristics are jointly zero in a regression of relative distance on those characteristics and our baseline controls. The *F*-statistic from this test is 1.041 – close to its expected value under the null, with a *p*-value of 0.406.[22]

For comparative purposes columns 3 and 4 show the estimated coefficients from models of the relationship between the risk factors and *actual delivery* at a high c-section hospital (i.e., we replace Z_i in equation (5) with H_i). Consistent with the motivation for an instrumental-variables approach in

[22] Consistent with the concerns raised by Hadley and Cunningham (2004), there is some evidence of sorting at broader geographies. The F-test **without** conditioning on mother's HSA rejects that these 31 characteristics are orthogonal to relative distance ($F(31,1249) = 2.969$, $p < 0.001$), highlighting the importance of comparing mothers who are served by the same hospital market for the validity of our distance-based design.

the first place, we observe significant patterns of sorting across hospital types. Conditional on our basic controls, mothers who deliver at H hospitals are more likely to have government insurance; they have higher rates of visiting the hospital in the year before delivery; they have slightly *lower* reported rates of diabetes, herpes, and asthma; and their infants are more likely to be categorized as low birthweight (<2500g). On balance, it appears that infants delivered at H hospitals have lower expected health: predicted ED visits, inpatient readmissions and death rates are all higher for these infants. Thus, we might expect simple observational comparisons to show that H-delivery leads to worse infant health.

In contrast, the results in column 1 show that our approach of looking within HSAs and isolating choices based on the relative distance to H versus L hospitals overcomes this sorting. In the analysis below, however, we perform several additional checks to ascertain that our IV estimates remain stable as we add various controls.

IV. First-Stage Relationships and Complier Characteristics

With this background, we turn to our first-stage models and a characterization of the compliers affected by relative distance. The first two rows of Table III present estimates of the first-stage effects of relative distance on the probabilities of giving birth in an H hospital and delivering by cesarean. We show 4 sets of estimates. The first two columns use the continuous version of relative distance (Z_i), with only our basic controls (X_i^0) or a full set of controls that includes all the variables in Table II (in some cases expanded to a set of dummies or a polynomial function). The third and fourth columns show analogous estimates using the binary version of our instrument (Z_i^B).

The estimate of 1.586 in the first row of the first column means that a mother living 10 miles closer to an H-hospital is 15.86 percentage points (ppt's) more likely to deliver at an H hospital (controlling for X_i^0), while the 1.014 estimate in the third column means that a mother who is simply closer to an H hospital is 10.14 ppt's more likely to deliver at such a hospital. Both estimates are highly significant (t=10.28 and 8.28). As expected given the results in Table II, the addition of controls for maternal demographics and risk factors has almost no effect on the estimates of these effects, but does lead to a small gain in precision.

Relative distance also has a strong effect on the probability of c-section delivery. A mother living 10 miles closer to an H hospital has a 1.86 ppt higher probability of c-section; a mother who is simply closer to an H-hospital than an L hospital has a 1.09 ppt higher probability of c-section. Again, these effects are virtually the same when we add the additional controls.

The third and fourth rows of Table III show the effects of relative distance on cesarean deliveries that occur with or without indications of labor.[23] These estimates imply that 70-75% of the extra c-sections attributable to living closer to an H hospital are unscheduled. Since indicators of labor are known to be under-reported, however, this fraction is a *lower bound* on the share of unscheduled procedures among those affected by the instrument. Indeed, the 16% under-reporting rate found by Henry et al. (1995) suggests that at least one-half of the "scheduled" c-sections attributed to distance are actually intrapartum procedures where the hospital record failed to report indications of labor.[24]

The next rows of Table III show the estimated effects of relative distance on the probabilities of c-section and vaginal births on the day the mother arrived at the hospital versus 1 or more days later. (We do not have information on the hour of arrival or birth). Most of the extra c-sections for mothers who are closer to an H hospital occur on the day of arrival, while most of the reduction in vaginal

[23] Specifically, we follow the existing literature (e.g., Gregory et al., 2002; Johnson and Rehavi, 2016) and classify unscheduled or scheduled c-sections based on the presence or absence of at least one of a set of ICD-9-CM diagnosis codes devised by Henry et al. (1995) that indicate dystocia or fetal distress during labor.

[24] Henry et al. (1995)'s data show that of 831 primary c-sections for non-breech births that were clinically coded as having trial of labor, 701 had indications of labor on the discharge record, implying a 15.6% under-reporting rate. If we assume that distance only affects unscheduled c-section rates, we would expect to find that *scheduled* c-sections account for 15.6% of the overall first stage effect. Martin et al. (2013b) report that "trial of labor" reported on the birth certificate under-reports actual trial of labor for cesarean delivered births by 12-25%.

deliveries is for births after a day or more in hospital. As expected, we find that practice styles at H hospitals tend to cut short longer labors, shifting later vaginal births to earlier cesarean deliveries.

Next, we show estimates of the effect of relative distance on cesarean deliveries at H and L hospitals, respectively. Being closer to an H hospital leads to relatively large rise in the probability of delivery by c-section at an H hospital, offset by a reduction in c-sections at L hospitals. Under our assumptions on compliance, the latter effect provides an estimate of the share of H-complier/C-ATs, who switch hospital types in response to relative distance but have a cesarean delivery regardless of where they present.

Finally, in the bottom panel of the table we show the implied breakdowns of the overall H-complier population into its three constituent subgroups To calculate these fractions using the continuous distance instrument we use the changes in probabilities associated with a 7-mile reduction in relative distance to an H hospital (which has about the same effect on the probability of H-delivery as simply being closer to an H-hospital). Both the continuous and binary versions of the instrument imply that H&C compliers represent about 11% of the overall H complier population, while C-ATs represent about 20%, and V-ATs represent 69%.

To further explore the effects of relative distance on hospital and delivery mode choice, we used the discrete version of our instrument to estimate the mean characteristics of the overall group of H-compliers and the subgroup of H&C compliers. The results are summarized in Table IV. A comparison of the demographic characteristics of all LRFB mothers (column 1) and the H-compliers (column 2) shows that the compliers are more likely to be Hispanic and to have at most a high school education. They are also more likely to have government insurance, and to have visited an ED in the year prior to birth.[25] Consistent with evidence from other settings – e.g., Beckert et al.'s (2012) study of hip

[25] Although not reported in the table we also looked at the fractions with Kaiser insurance, and find very low rates among H compliers and H&C compliers. We have estimated our main models excluding Kaiser insurees and find that the resulting IV estimates are very similar to those from our larger sample. This is as expected given there are so few Kaiser insurees among the compliers.

replacement patients – these comparisons suggest that families whose hospital choices are most affected by distance tend to be less advantaged.

The H&C compliers (column 3) are even more highly selected. For example, only 14% have a college degree; they are also more likely to live in zip codes with lower incomes and higher fractions of government-insured mothers, and to have been users of the ED prior to the birth. The implication is that differences in hospital practices have a larger impact on the delivery modes of lower-SES mothers, even conditional on choosing a hospital based on relative distance. This is the complement of the findings reported by Johnson and Rehavi (2016), who conclude that the delivery modes of physician mothers are *less* responsive to the financial incentives faced by their doctors.[26]

Panel D of Table IV shows the fractions of compliers whose infants have above-median predicted probabilities of an ED visit or inpatient readmission in the year after birth.[27] Infants of H&C compliers have a particularly high risk of an ED visit, consistent with the fact that their mothers tend to be less-educated women with high rates of government insurance who were themselves relatively heavy users of the ED. They also have high predicted readmission rates, suggesting that these infants may be particularly vulnerable to problems at delivery. The last row of Panel D shows the predicted probabilities of c-section by group. Our focus on LRFBs limits variation in these pre

V. Impacts of High C-Section Delivery Practices on Infant and Maternal Health

V.a. Outcomes at Delivery

With this background, we now turn to the health effects of hospital-specific delivery practices. We begin in Table V with outcomes realized at delivery. For infants, we examine the 5-minute Apgar score,

---

[26] An earlier study Grytten et al. (2011) finds that in Norway, where c-section rates are among the lowest in the OECD, physician mothers are ***more*** likely to have c-section. They attribute this to enhanced agency of these mothers in the hospital setting.

[27] We estimate logit models for the probabilities, including the full set of controls but excluding HSA dummies. We emphasize that higher risk of an ED visit represents a combination of worse health and a higher probability that the family takes the infant to an ED rather than a doctor office or clinic.

incidence of a birth injury, admission to the NICU, use of ventilation, and length of hospital stay.[28] For mothers, we focus on labor-related injuries and the length of the hospital stay. For each outcome, we show the mean value among all LRFBs (in the first column), the OLS coefficients from regressions of the outcome on H_i or C_i and our richest set of controls (in the second and third columns), the estimated reduced-form effect based on the continuous instrument Z_i (in the fourth column), and the IV estimate of the effect of delivery practices at H hospitals (in the final column).

To keep the table manageable, we show only the results from models that include our richest set of maternal characteristics and risk factors. Specifications that include the basic control set X_i^0 are very similar and are available on request. We also report only IV estimates that scale the health effects per extra delivery at an H hospital. Estimates that scale the effects per extra cesarean delivery are roughly 9 times larger and have essentially the same *t*-statistics.

Apgar scores are widely used as indicators of newborn health (see Casey et al., 2001).[29] They range from 0 to 10, with most infants scoring above 8 and a mean close to 9. In OLS models, delivery at an H hospital has a modest positive effect on the 5-minute Apgar, while c-section delivery has a negative effect. The reduced-form effect of relative distance on the mean 5-minute Apgar score is positive and significant ($t \approx 3$), with an associated IV estimate that implies a relatively large positive effect of shifting deliveries to H hospitals – roughly an effect size of 0.11σs' per additional H delivery.

To help understand this effect we examined the joint potential outcome distribution for delivery mode, timing of birth relative to mother's arrival at hospital (same day or 1+ days later) and high- versus low Apgar (<9 or ≥9) for the compliers who deliver at L and H hospitals. The results, summarized in Appendix Table II, show that being closer to an H hospital leads to a 2.1 ppt increase in

[28] We use Vital Statistics reports of Apgar scores, birth injuries, NICU admissions, and ventilation.

[29] The Apgar is based on 5 components (breathing, heart rate, muscle tone, reflexes, and skin color) each of which is scored 0 1 or 2. See Finster and Wood (2005) for a brief history and discussion of the test. We have also looked at the 1-minute Apgar – the results are similar but slightly attenuated.

the probability of an Apgar score of 9 or higher (t=2.1). Sixty percent of this shift (1.2 ppts) is attributable to a reduction in the fraction of infants who are born vaginally after a delayed delivery and have 5-minute Apgar <9, while another 30% (0.7 ppts) is attributable to a reduction in the fraction of infants who are born by cesarean after a delayed delivery and have Apgar <9. These results suggest that the practice of ending labor earlier at H hospitals has a net positive effect on the infants of both the C&H compliers and C always-takers; these results are also complementary to findings from observational studies which suggest that longer labor has a negative effect on Apgar scores of infants who are delivered vaginally (e.g., Altman et al., 2015).

Looking next at the incidence of birth injuries, the OLS estimates are both negative, with a relatively large negative partial correlation between c-section and the risk of injuries. The reduced-form effect is also negative, though relatively imprecise, leading to an IV estimate that implies a reduction in birth injuries of about 1 per 100 deliveries moved to an H hospital. [30] We emphasize, however, that we cannot rule out a zero effect.

With respect to NICU admissions, OLS models show a negative association with delivery at an H hospital and a positive association with c-section delivery. The reduced-form effect of relative distance is negative and implies an IV effect per delivery that is comparable to the OLS estimate, with t≈2. A potential concern here is that both the OLS and IV estimates may be confounded by the lower fraction of H hospitals with a NICU unit (74% vs. 86% at L hospitals). One way to deal with this is to measure NICU admissions *including transfers* from the birth hospital to another facility, while controlling for the presence of a NICU unit at the birth hospital. Such a model has to be interpreted cautiously because the decision to deliver at a hospital with or without a NICU is potentially endogenous. With that in mind, however, we find that incorporating transfers and controlling for a

[30] Alexander et al. (2006) study fetal injuries after c-section and note that the highest rates of injuries are fetuses born after an unsuccessful trial of forceps or vacuum delivery. To the extent that doctors at H hospitals use these procedures less, and opt for c-section earlier, these injuries will be avoided.

NICU at the birth hospital reduces the estimated effects of H delivery, leading to an IV estimate that is indistinguishable from zero, though still negative.

Next, we look at the incidence of assisted ventilation immediately following delivery.[31] OLS model shows that ventilation is slightly less likely for deliveries at H hospitals but is more likely for those delivered by c-section. The latter effect is consistent with a large literature suggesting that breathing problems are more likely for babies delivered by cesarean (see Appendix A).[32] The reduced-form effect of proximity to an H hospital is positive and relatively large in magnitude (with a t-statistic of 2.4), leading to an IV estimate of +2.7 ppts higher risk of ventilation per delivery – nearly twice the baseline rate among LRFBs of 1.5%.

We have also examined effects on the incidence of longer-term assisted ventilation, which has a much lower average rate (0.12%). The estimated IV effect of delivery at an H hospital on this outcome is small and statistically insignificant (+0.15%, with a standard error of 0.12%), though we cannot rule out a sizeable effect relative to the baseline rate.

Next, we examine three measures of the length of hospital stay (LOS) for newborns: (1) the raw LOS (calculated as number of days from birth to discharge); (2) a censored version that top-codes LOS at 6 days; (3) an indicator for LOS of 5 or more days. The mean LOS for LRFBs is 2.3 days. OLS models show a modest positive effect of H-delivery and a large positive effect of cesarean delivery on mean LOS. In contrast, the IV estimates are negative, though only the estimate for the censored LOS measure is significantly different from the corresponding OLS estimate. These results suggest that practice differences at high c-section hospitals do not lead to longer average infant stays for the compliers, despite the large observational difference in LOS for cesarean versus vaginal births.

[31] This is defined as "infant given manual breaths for any duration with bag and mask or bag and endotracheal tube within the first several minutes from birth" (NCHS, 2016). Martin et al. (2013b) find that assisted ventilation at delivery is reported with some error (with a roughly 30% false negative rate in the one state in their analysis with a large enough sample).

[32] The 1.5% incidence rate we measure in our sample is comparable to the rate of 1.8% measured by Angus et al. (2001) using 1994 discharge data for California and New York.

Our interpretation of the small IV-estimated impact on LOS is that this represents a mixture of effects for the three complier sub-groups (as in equation 4). Some H&C compliers would have had a long labor at an L hospital: for these infants the switch to earlier cesarean delivery at an H hospital could *lower* their LOS. Likewise, some H-complier/C-ATs would have an earlier c-section at an H hospital, avoiding fetal stress and lowering their post-birth LOS. Offsetting these effects, however, some H&C compliers could require more recovery time after cesarean delivery at an H hospital than they would need after vaginal delivery at an L hospital. On average, the net effect across all the groups appears to be close to zero.

Proceeding to maternal outcomes, we begin with two measures of injuries during labor: trauma to the perineum and vulva, and 2nd degree or higher perineal laceration (PL). Trauma includes the more serious PLs, as well as vulvar/perineal hematomas and anal sphincter tears. The incidence of trauma injuries is relatively high among first-birth mothers (46% on average); about two-thirds of these are second-degree or higher PL's. As shown by the OLS estimates, the rate of birth injuries is substantially lower at H hospitals and among cesarean deliveries. In fact, the rates of injuries are under 1% for c-section births. The reduced form effects of relative distance on rates of both injuries are large and negative, leading to implied reductions in the risks of these injuries on the order of 0.1 per delivery at a high c-section hospital, or about 0.9 per complying c-section.

To help interpret the IV estimates we estimated the potential outcomes for incidence of PL among the H&C complier population.[33] These calculations assume that delivery at an H hospital has no effect on PL rates for V-ATs (which seems plausible) and for C-ATs (which also seems plausible and is consistent with the negligible rate of PL among c-section deliveries). The results imply that the rate of PL is over 90% for H&C compliers who deliver at L hospitals but falls to essentially 0 for those who deliver at H hospitals. The high rate for H&C compliers who deliver at L hospitals is indicative of long

[33] We implement this approach using the binary version of relative distance.

and difficult labors when these mothers deliver vaginally, highlighting the tradeoff between performing intrapartum c-sections and waiting for labor to progress.

The final three rows of Table V investigate effects on maternal length of stay. OLS models show that mothers who deliver at H hospitals have about a 0.09-day longer LOS, while those who deliver by c-section have 1.4-day longer stay in the hospital (similar to the OLS effects on infant LOS). We then split maternal LOS into two components: (1) the number of days from admission to the birth;[34] and (2) the number of days from birth until maternal discharge, i.e. the post-birth LOS. As expected, OLS models show that mothers who deliver at H hospitals have a shorter period from hospital admission to birth, offset by a longer post-birth stay, whereas mothers who deliver by c-section have longer pre- and post-birth stays.

Turning to the IV estimates, we find that the effect of delivering at an H hospital on the pre-birth stay is even more negative than the OLS estimate (-0.065 vs. -0.056) whereas the effect on the post-birth stay is less positive (0.052 vs. 0.142). As a result, our IV estimates imply that delivery at an H hospital has a small (statistically insignificant) negative effect on overall maternal LOS, similar to our findings for infant LOS.

Our interpretation is that H hospitals tend to cut short the pre-birth stay for C-AT's and C&H compliers, with mixed effects on length of the post-birth stay for the C&H compliers. Some H&C compliers would have significant injuries if they delivered at an L hospital, so the switch to cesarean delivery could have relatively small effects on the length of post-birth stay.[35] Other H&C compliers require an extra day or two in hospital if they deliver by cesarean at an H hospital rather than vaginally at an L hospital. Taking account of the shorter average time to delivery and the modest increase in post-

[34] Unfortunately, we do not observe exact time of admission or birth. We only know the days elapsed between her admission to the hospital and the birth date of the baby. According to Declercq et al. (2006), the mean time in labor for first-time mothers is around 11 hours. Consistent with this, the delay from admission to birth is 0 or 1 day in 95% of cases.

[35] We have also investigated other indicators of prolonged labor, including a code on the birth certificate, which yields qualitatively similar results. However, average reported rates for prolonged labor on the birth certificate vary widely across hospitals (from 0 to 16%) so we are reluctant to attach much weight to this variable and do not use it elsewhere in the paper.

birth recovery time, practices at H hospitals are associated with about the same overall length of stay as those at L hospitals.[36]

b. Post-Delivery Admission Outcomes - Infants

We now turn to our main measures of the health effects of alternative delivery practices, based on hospital and ED/ASC visits in the year after birth. For infants, in the upper panel of Table VI, we show results for six outcomes: (1) any ED/ASC visit or inpatient readmission in the year after birth; (2) any ED visit; (3) any ED visit for acute respiratory conditions; (4) any readmission in the first month after birth (the neonatal period); (5) any readmission in the first month excluding immediate transfers from the initial birth stay; and (6) any inpatient stay in the year after birth.

The estimated OLS coefficients imply that delivery at an H hospital is associated with a small (and insignificant) increase in the risk of any type of visit, a larger increase in the probability of an ED visit, driven by visits for acute respiratory conditions, and a mixed pattern of effects for the inpatient stay measures. Interestingly, the OLS models show modest negative effects of c-section delivery on all three inpatient measures but positive effects on ED visits.

The reduced-form effects of relative distance on both measures of ED use are positive and statistically significant (with *t* statistics around 2.6). Scaling by the first stage, the IV estimates imply that delivery at an H hospital leads to an 8-ppt increase in the risk of visiting the ED in the year after birth per delivery, 60% of which is attributable to visits for acute respiratory conditions. The effect on respiratory visits represents a 35% increase over the mean rate, confirming the prevailing view in the literature that cesarean delivery is associated with an elevated risk of respiratory-related problems (see Hyde et al. 2012 and Appendix A).

[36] We have investigated a few other outcomes at birth, namely unplanned hysterectomy and asphyxia of the neonate. In our sample, the mean rate of unplanned hysterectomy is only 1 in 10,000; we find no evidence of an effect of c-section, but the precision of our estimates is low. Asphyxia is also rare (3.2 per 1000), and we find very weak evidence (t=0.5) that c-sections reduce asphyxia. There is somewhat stronger evidence that c-section reduces asphyxia-related infant deaths.

In contrast to the effects on ED visits, there is negative impact on hospital readmissions, with the IV point estimate implying 2.4 fewer readmissions per 100 deliveries in the first *month* of life. The estimated effect for neonatal readmissions is robust to exclusion of transfers immediately after birth and is about the same size as the estimated effect on the probability of any inpatient stay in the first *year* of life.

More insight into the timing of the effect on re-hospitalizations is provided in Figure III, which plots the cumulative fraction of all LRFBs with a readmission at various time horizons after birth, and the estimated IV effect of delivery at an H hospital on the probability of readmission up to each point. As would be expected, the probability of readmission rises steadily in the year after birth. In contrast, the IV effect emerges in the first month and then stabilizes, indicative of a health gap immediately after birth.

Although we noted in the discussion of Table II that relative distance to an H hospital (our main instrumental variable) is orthogonal to a wide set of maternal characteristics and risk factors, it is informative to see how the estimated effects of relative distance on infant ED visits and readmissions are affected by including or excluding these characteristics.[37] The results of such an exercise are summarized in Figures IV.a and IV.b, where we show the reduced form effects from a series of models, starting with a specification that includes only our basic control set (X_i^0), then selectively including 12 different sets of characteristics. (Figure IV.b also presents results for infant death – we defer the discussion of these to later in the paper). For reference at the bottom of the figure we show the estimated reduced form effects when we include all the controls together, as we do in Table VI. Consistent with the results in Table II, we find that adding any subset of controls has a negligible effect on the

[37] See Altonji, Elder and Taber (2005) and Oster (2017) for discussions.

magnitude (or precision) of the reduced form effects of relative distance on either measure of ED visits or on the reduced form effect on inpatient readmissions.[38]

To further probe the robustness of the reduced form impacts we conducted a second exercise where we estimated the *range* of possible reduced form impacts that could be obtained by adding combinations of the groups of additional controls to our basic control set. Specifically, for a given *k=1...11* we randomly selected *k* of the 12 possible groups of controls, then re-estimated the reduced form models. By repeatedly sampling we obtain a distribution of potential reduced form estimates for each *k*. The panels in Figure V show the minimum, maximum, mean, and median reduced form estimate for all ED visits (panel a), ED visits for acute respiratory conditions (panel b), inpatient readmissions (panel c), and infant death (panel d). In all cases we see that even selectively choosing subsets of control variables the range of possible reduced form health effects is quite narrow.

c. Post-Delivery Admission Outcomes - Mothers

Returning to Table VI, the lower panel presents a parallel set of models for maternal ED visits and readmissions in the year after birth. OLS models show that delivery at an H hospital is associated with small increases in the probability of visiting a hospital, ED, or ASC in the following year, and very small effects on ED visits or inpatient readmissions. Unlike the models for infants, however, the reduced-form effects of relative distance on these outcomes are relatively small in magnitude. Indeed, the IV point estimates are 5-6 times smaller than the corresponding estimates for infants. We conclude that the short/medium term effects of delivery at an H hospital on mothers are quite small.[39]

[38] The first stage effect of relative distance on the probability of delivering at an H hospital is extremely stable across specifications that add any or all the extra controls.

[39] We note that our samples do not have sufficient power to investigate issues such as the possibility that a higher cesarean delivery rate leads to elevated risk of placenta previa in subsequent births.

d. Allowing for Other Hospital Practice Differences

Our results so far suggest that delivery practices at H hospitals lead to two offsetting effects on infant health: a lower probability of hospital readmissions; and an increase in ED visits, particularly for acute respiratory conditions. A potential confound for interpreting both findings is that the impacts could be driven in part by *other* practice differences between H and L hospitals that affect readmissions and ED visits in the year after birth. For example, hospitals vary widely in the average waiting times at their emergency departments (Ding et al., 2010). If H hospitals tend to have less-congested EDs, mothers who deliver at a nearby H hospital may be more likely to use the ED, leading to a positive correlation between delivery at an H hospital and post-partum ED visits.

A simple way to address these concerns is to classify hospitals by the (risk-adjusted) rate that infants born in that hospital have an ED visit in the year after birth, or are readmitted, and then add controls for these additional characteristics in our models. Since the choice of delivery hospital is endogenous, we follow the same steps we used to develop our primary instrumental variable (Z_i) and form two new instruments based on relative distance to a high- versus low-ED hospitals, and high- versus low-readmission hospitals. We then fit models for ED visits that include *two* endogenous characteristics of the delivery hospital -- high/low cesarean rate and high/low ED visit rate -- and estimate it using the two relevant instrumental variables. Similarly, we fit models for inpatient readmissions that include indicators for high/low cesarean rate and high/low readmission rate at the delivery hospital.

The results of this exercise are shown in Table VII (for ED visits) and Table VIII (for readmissions). The first four columns of each table report the first stage models for the two endogenous variables. For reference, we show models that include only a single instrumental variable for each endogenous characteristic, as well as models that include both instruments. In both tables we find that the first-stage systems are approximately diagonal, with the choice of a high c-section hospital driven by relative distance to an H hospital, and choice of either a high-ED hospital or a high-readmission hospital driven by relative distance to that type of hospital. This diagonality reflects the fact that the correlation between high cesarean delivery (H) status and either high-ED status or high readmission status is very close to zero.

Columns 5-8 of Table VII report reduced-form coefficients and IV models for the probability of an ED visit for any reason in the year after birth, while columns 9-12 report a parallel set of models for the probability of a visit for an acute respiratory condition. Comparing the reduced form models that include only relative distance to an H hospital (columns 5 and 9) with models that also include relative distance to a high-ED use hospital (columns 6 and 10), we see that the magnitude of the estimated reduced form effect of relative distance to a high c-section hospital is essentially unaffected by the addition of the second channel, though relative distance to a high-ED use hospital clearly affects ED use. The invariance of the reduced form coefficients, together with the near-diagonality of the first stage system, means that the IV models (columns 7-8 and 11-12) show very similar impacts of delivery at a high-cesarean hospital, regardless of whether we include a second channel for other hospital practices that affect ED use.

The results in Table VIII for inpatient readmissions in the neonatal period (columns 5-8) or the first year of life (columns 9-12) are quite similar. For these outcomes, however, we find that the addition of the second instrument leads to a slight attenuation (10-20%) of the estimated reduced form impacts of proximity to a high c-section hospital. As a result, the IV models show a slightly smaller effect of delivery at a high c-section hospital on post-partum inpatient admissions. For readmissions in the neonatal period the estimated IV effect falls from 0.024 readmissions per delivery at an H hospital to 0.020 (about 1/2 of a standard error) but remains highly significant (t=2.8). For admissions over the entire year the estimated IV effect falls from 0.024 to 0.017 (also about 1/2 of a standard error) but is only marginally significant in the more general model (t=1.4).

Based on the estimates in Table VII and Table VIII we conclude that there are systematic differences across hospitals that affect the rates that newborns come back for ED visits or are readmitted. Nevertheless, hospitals with relatively high cesarean delivery rates are no more likely than other hospitals to have high ED use rates, or high readmission rates. Consequently, when we control for a second channel in our models for ED use or readmission, we find very similar impacts of hospital delivery practices. We have also estimated 3-channel models for ED visits and readmissions that include all three characteristics of the delivery hospital (high or low c-section; high or low infant ED use; and high or low infant readmission) and use relative distances to the nearest high versus low hospital in each category as instruments. These models also give rise to estimates of the effects of delivery practice differences that are very similar to those in Table VI.[40]

e. Impacts on Infant Death

In our final step, we turn to the effects of hospital delivery practices on infant death in the year after birth. A concern for this analysis is that the infant mortality rate for LRFBs is extremely low (1.2 per 1,000 births, versus 5.5 for all births), making it hard to detect even large proportional changes in the death rate. To partially address this, we develop a proximate measure based on an "adverse event" in the neonatal period, which we define as either death or 6+ days in the hospital. This has a much higher rate (67 per thousand births), mostly driven by the hospitalization measure.

To set the stage for this analysis, Appendix Table III shows the characteristics of all LRFBs and those that resulted in an infant death within a year of birth. Infants that die are about 7 times more likely to be low birth weight (<2,500 grams) than those that survive; they also have relatively low Apgar scores. Indeed, over 30% have a 5-minute Apgar score below 7, indicative of poor health at birth.

[40] In a previous version of this paper, we estimated effects of delivering at H hospitals among otherwise low-risk breech births, whose c-section rate is nearly 100%, but who still tend to choose nearby hospitals. Among this breech sample, point estimates of the effect of H delivery were very close to 0 for our main outcomes, but with somewhat large standard errors. The breech results are consistent with the findings in this draft. Because of the imprecision and concerns about comparability, we omit the analysis of breech births here.

Infants that ultimately die are more likely to be transferred to a NICU unit and more likely to be readmitted as an in-patient. They are also more likely to have been delivered by cesarean, though the latter correlation is driven by *scheduled* procedures: the mean death rate is 1.1/1000 for vaginal births, 0.9 for unscheduled c-sections, but 2.9 for scheduled cesareans.[41]

Table IX shows OLS, reduced-form, and IV results for four outcomes: (1) an adverse event in the neonatal period; (2) death within the first year; (3) the combined outcome of a late delivery (after the mother has been in hospital more than 1 day) and an adverse event in the neonatal period; (4) the combined outcome of a late delivery and death in the first year. All the models include the full set of controls used throughout earlier tables, plus relative proximity to a high-readmission hospital (the second instrumental variable used in Table VIII) to control for other practice differences across hospitals that may affect inpatient readmissions or death.

The OLS models in the first two columns show that delivery at an H hospital has a negative correlation with all four outcomes, whereas c-section delivery has a positive correlation with all four. The reduced-form models show a relatively large negative effect of proximity to an H hospital on the probability of an adverse event in the neonate period, and on the joint occurrence of late delivery and an adverse event. Both effects are significant, echoing our finding in Table VI that relative proximity to an H hospital is associated with a significant reduction in the probability of readmission in the first month of life. The associated IV estimates show that delivery at an H hospital reduces the probability of an adverse neonatal event by 26/1000, of which 16/1000 (62%) is attributable to a reduction in the joint occurrence of a long labor and a subsequent adverse event in the neonatal period.[42]

Consistent with the reductions in the probability of an adverse event, the reduced form and IV estimates also show significant negative impacts on infant mortality. The estimated IV coefficient

---

[41] MacDorman et al. (2008) compare scheduled c-sections to vaginal births for low-risk mothers and find higher neonatal mortality for the scheduled c-sections. They frame this as an "intention to treat" analysis.

[42] On average 47.8% of all deliveries occur 1 or more days after the mother's arrival at hospital, whereas the reduced form effect on the co-occurrence of a long labor and an adverse postnatal outcome accounts for 62.4% of the overall reduced form effect on an adverse outcome.

implies a life-saving effect of 2.36 infant deaths per 1,000 births at an H versus L hospital, with a 95% confidence interval of 0.4 to 4.3 deaths prevented per 1,000 births. Again, a large share of this effect (0.162/0.236=69%) is attributable to a reduction in deaths that occur for infants delivered after the mother was in hospital for one or more days.

The IV point estimate of the effect of H delivery on infant death (2.4/1000) is large relative to the mean death rate of 1.2/1000. To help understand this magnitude we estimated the mean potential death rate of the hospital complier group when delivering at L hospitals. The estimated rate is 2.2 deaths per thousand (with a standard error 1.0 deaths), which is substantially higher than the overall mean for LRFBs, suggesting that the infants of distance-complying mothers are frailer than the overall population.[43]

We probe the robustness of the estimated reduced form effect of relative distance on infant mortality in Figures IV.b and V.d.[44] As with the reduced form effects on ED use and readmission, we find that the reduced form impacts on death are quite stable and are largely unaffected by addition of control variables such as birth weight that are highly correlated with the risk of death. The one set of covariates that noticeably affects the estimated reduced form death effect is characteristics of the home zip code (mean household income, mean maternal education and dropout rate, and mean fractions of black and Hispanic mothers). Adding these 5 variables causes the reduced form effect to fall in magnitude by about 0.02, or about 5%.

Further insight into the relationship between relative distance and infant mortality is provided by the simple bin scatter plot in Figure VI. For reference, we also show a bin scatter of the first stage

[43] We use the discrete version of our instrument to estimate this. We obtained a similar estimate using the weighting procedure suggested by Abadie (2003).

[44] For comparability with the other coefficients reported in Figures IV and V, in the reduced form models for death we exclude the relative distance to a high inpatient hospital variable that is included in all models in Table VI. Sensitivity results including this extra variable are quite similar but in all cases the magnitude of the reduced from death effect is slightly larger (typically around -0.0038 versus -0.0036).

relationship. Both relationships appear relatively linear. The reduced form scatter shows clear evidence of a systematic negative relationship between infant mortality and delivery at an H hospital.

A possible objection to our linear modeling framework is that the risk of death is so low that a linear probability model is inappropriate. To assess this, we re-estimated the reduced-form model using both probit and logistic regressions. In both cases, relative distance has a significant negative effect on the risk of death (with p-values of 0.032 for the logit and 0.031 for the probit, versus 0.022 in the OLS reduced form model). Moreover, the implied average marginal effects of a 10-mile reduction in the relative distance to an H hospital are quite similar to the effects implied by our OLS reduced form: -0.033 for the logit and -0.032 for the probit, versus -0.036 for the OLS model. We also conducted a similar analysis using a binary version of the instrument -- which yields a very similar IV estimate of the effect of death -- and found that computing the reduced form effect using OLS or logit gives nearly the same effect. We conclude that relative proximity to an H hospital has a comparably-sized negative effect on the probability of death regardless of functional form.

How do our estimates relate to the existing literature? With respect to the health impacts of earlier c-section, Tolcher et al. (2014) present a meta-analysis of the related literature on the effects of the delay between the time a decision is made to perform c-section and delivery. They find an inconclusive link, though some studies – e.g. Thomas et al. (2004) – find that that extended delay is associated with worse outcomes. Rennie and Rosenbloom (2011) review animal and human studies on the timing of delivery and the risk of hypoxic ischemic encephalopathy (brain injury due to asphyxia) and conclude there is strong evidence of a link, though the measured incidence of this condition in our data is extremely low (0.2 per 1,000 births). Finally, it is worth noting that there is extensive litigation in the U.S. arising from claims that c-section was performed "too late" or was not performed when it was indicated, resulting in injuries or death of the infant. Our reading is that courts have often agreed with the plaintiffs, despite the lack of a clear scientific consensus.

f. Heterogeneity in the Health Effects of Delivery at High C-Section Hospital

An important issue for the interpretation and extrapolation of our findings is the degree of heterogeneity in the treatment effects associated with delivery at a high c-section hospital.[45] To address this, we extend our instrumental variables setup using a simple control function approach that allows for a random effect in the impact of H delivery (Garen, 1984; Heckman and Vytlacil, 1998; Wooldridge, 2015). Specifically suppose that the causal model relating health outcome y_i to patient characteristics X_i and type of hospital H_i is:

$$y_i = \beta_0 + \beta_{1i} H_i + \beta_x X_i + \varepsilon_i ,$$

where β_{1i} is a random coefficient and ε_i is a structural error incorporating the unobserved determinants of health. We assume that:

$$E[\beta_{1i} \mid H_i, X_i, Z_i] = \beta_1 + \lambda_u u_i + \lambda_x (X_i - \overline{X})$$

$$E[\varepsilon_i \mid H_i, X_i, Z_i] = \theta_u u_i ,$$

where u_i is the error in the first stage equation (1) for H_i. Here β_1 represents the average treatment effect (ATE) of delivery at an H hospital and $\lambda_u u_i$ represents a potential self-selection effect that arises if mothers with a stronger preference for H hospitals have larger or smaller treatment effects from delivering there. Similarly, the term $\lambda_x (X_i - \overline{X})$ represents potential heterogeneity in the treatment effect with respect to (predetermined) maternal and infant characteristics. Finally, the term $\theta_u u_i$ captures any correlation between latent health and the unobserved component of preferences for an H hospital.

---

[45] There is a large and growing literature on heterogeneous treatment effects: see Imbens and Wooldridge (2009) for a general discussion and Cornelissen et al. (2016) for a recent survey emphasizing heterogeneity in marginal treatment effects.

As shown by Heckman and Vytlacil (1998) and Wooldridge (2015) this model can be estimated in two steps by first estimating the first stage model for hospital type, obtaining the residual $\hat{u}_i$, and then estimating a second step model:

$$y_i = \beta_0 + \beta_1 H_i + \beta_x X_i + \lambda_u H_i \hat{u}_i + \lambda_x H_i (X_i - \overline{X}) + \theta_u \hat{u}_i + \varepsilon_i'.$$

This model includes the estimated first stage residual $\hat{u}_i$, an interaction between $\hat{u}_i$ and H_i, and interactions between H_i and the other covariates. Excluding the interaction terms leads to an estimate for β_1 that is numerically equivalent to the standard IV estimate. Adding the interaction terms allows for heterogeneity in the effect of H delivery that can be correlated with either observed characteristics or unobserved preferences. To account for the fact that the first-stage residual is a generated regressor, we conduct inference on the second-step parameters via a block bootstrap, clustered as usual by mother's zip code.

Table X presents estimated control function models for three key infant health outcomes: the five-minute Apgar score; an indicator for an "adverse event" in the neonatal period (6+ days in hospital or death); and infant death. For each outcome we present a benchmark model with no interactions (yielding the IV coefficients already shown in Table V and Table IX), a second model that adds the interaction between $\hat{u}_i$ and H_i, and a third model that also adds interactions with two key observable markers of infant health – birthweight and gestation.

Looking across the three sets of models in Table X we see three interesting patterns. Most importantly, estimates of the ATEs of H-delivery from models that allow for self-selection and heterogeneous treatment effects with respect to birthweight and gestation are very close to the LATEs from our baseline IV procedure. Second, across all three outcomes we see consistent evidence of negative "Roy sorting". The impacts on Apgar scores, adverse neonatal events, and death are all larger in absolute value for infants whose mothers have a stronger preference for L hospitals. A third pattern is that there is relatively little heterogeneity in the effects of H delivery across infants of different birth weights and gestations.

Our finding of negative Roy sorting is reminiscent of the finding by Chandra and Staiger (2017) that hospitals with greater comparative advantage in more intensive AMI treatments are less likely to use them. An important difference, however, is that we are identifying the relationship between a *patient's* preferences for different hospitals and the effect of having a *policy* of earlier c-section. One interpretation of a negative estimate for λ_u is that mothers who prefer L hospitals are at greater risk for long and difficult labors that are ended earlier at H hospitals, with potentially larger than average health benefits for their infants.

VI. Summary and Discussion

To summarize our findings, in Table XI we present a simple cost-benefit calculation of the effects of allocating more low-risk first births to deliver at high c-section hospitals. For the sake of this calculation, we draw on our estimates but also bring in some outside estimates (e.g. the value of a statistical life).

Our findings suggest that the primary benefits of reallocating births to high c-section hospitals are reduced hospital readmissions for the infant and reduced infant mortality. The estimates in Table VI imply that moving 1,000 deliveries from L to H hospitals would lead to about 23.4 fewer readmissions in the first year of life. Assuming an average hospital stay has a cost of about $10,800 (Wier et al., 2011), this amounts to a saving of about $250,000. The estimates in Table IX imply that the same switch in place of delivery would save about 2.36 infant lives. A typical value of a statistical life (VSL) from our sample period is on the order of $6 million. If we use the estimate for US-DOT (2008) of $5.8 million, we can infer that the mortality-reduction benefit is around $13,700,000 per 1,000 births.

On the other hand, the main costs we need to account for involve direct medical costs of delivery, the morbidity costs associated with extra infant ED visits (and potentially lifelong asthma --

see Keag et al., 2018), and finally the cascading effects on future births. Given that most of the additional c-sections performed in high c-section hospitals would be long vaginal deliveries in L hospitals, and our findings on length of stay, we assume the associated cost difference is $1300 (Podulka et al., 2011). We note that this estimate is on the lower end of published cost differences. Assuming instead a $10,000 cost difference between all cesarean and vaginal deliveries would add $2 million to the cost side of our cost-benefit calculation.

Average ED visit costs are on the order of $2200. This cost accrues to 80 infants per 1,000 delivered at high c-section hospitals. Next, we add in the lifelong costs of asthma using the present discounted value of $1,039 per year healthcare expenditures over 80 years, assuming a discount rate of 3%, which comes out to $31,379. Finally, to account for the fact that mothers delivering by cesarean section at first birth are almost surely going to have c-section deliveries on all future births, and vice versa, we scale all these costs by the US average completed fertility of 2.1 (Johnson and Li, 2007), with the exception of direct medical costs where we assume a cost difference in future births of $2300 = $5700 (cesarean delivery without complication) - $3400 (vaginal delivery without complication). Summing these costs up yields the cost of reallocating 1,000 extra births to H hospitals of $6,072,000. This is less than one-half of the estimated benefits arising from reduced readmissions and deaths, suggesting that there could be a large benefit to reallocating deliveries to high c-section hospitals.[46]

The main driver of this net benefit is the reduction in infant mortality. As we have noted, our point estimate of the mortality effect has a relatively wide 95% confidence interval. If we assumed that the true mortality effect was one-half as large as our point estimate (equivalent to using an estimate 1.4 standard errors below the point estimate, or at the 8.4 percentile of the confidence interval) the net benefit would fall to about $1,000,000.

---

[46] One potential cost we ignore is the effect of cesarean delivery on the risk of abnormal placentation in future births (see Appendix A, Table A2; and the review by Clark and Silver, 2011). We also ignore potential costs on reduced future fertility. In an earlier version of this paper we examined future fertility and found no significant effect, though our design lacks power to detect small effects. Offsetting these costs we ignore any potential benefit of reduced vaginal trauma on future risk of pelvic floor disorder (e.g., Gyhagen et al., 2015 and the review by Keag et al., 2018).

While these calculations are quite simplistic, they point to the possibility that, despite their direct and indirect costs, high c-section practice styles have other benefits that may offset these costs. Of course, there may still be ways to both reduce c-section rates and improve outcomes. Our findings suggest that the simplest approach of waiting longer for a vaginal birth to occur comes with risks, even among relatively healthy first births.

Lastly, we note that our cost-benefit calculations are context-specific. The net benefits of intensive delivery practices may well vary according to the prevailing intervention rates. We note here that in unreported results available upon request, we found no such evidence of interactions with prevailing HRR-level c-section rates, despite considerable variation in c-section rates across HRRs in California.

VII. Conclusion

Our analysis suggests that high c-section hospitals' delivery practices have important benefits for low-risk first births. These benefits derive from the truncation of long, difficult labors that risk causing serious harm when allowed to proceed, as many would at low c-section hospitals. We find some indications of negative Roy sorting – mothers who have an unobserved preference for low c-section hospitals tend to have infants who would benefit from the delivery practices at high c-section hospitals – but no clear evidence that returns to intensive treatment practices vary across regions with different baseline c-section rates. While we do not provide clear guidance for optimal c-section utilization, our evidence casts some doubt on the opinion that current delivery practices among first-time mothers represent a "c-section epidemic" (Morris, 2013).

Of course, it is important to note some important limitations of our analysis. First, and most obviously, our analysis of first births does not speak to delivery practices in second and higher-order births, where rates of vaginal birth after c-section (VBAC) dropped from a high of nearly 30% in 1996 to around 10% since 2004, propelled by ACOG guidelines (ACOG/SMFM, 2014). Rising c-section rates among first births compound the effects of these changes on the overall c-section rate, albeit in a limited way since only 15% of all births occur to mothers with a prior c-section. Second, our data, while expansive, tend to capture relatively serious health outcomes. We do not capture subtler hypothesized effects of delivery mode on outcomes such as weight gain or cognitive functioning that may be reported in primary care visit data and education records later in life. Further analysis along these dimensions would help provide a more complete picture of the costs and benefits of delivery practices.

Despite these caveats, our analysis provides a first step towards understanding the potential ramifications of proposals to reduce low-risk primary cesarean delivery rates. Methodologically, we provide some guidance on ways to evaluate and extend distance-based research designs for analyzing hospital practices. Our findings suggest that designs based on *local* differences in hospital types and patient location could prove useful for achieving improved balance in similar designs, thus facilitating research in non-emergency settings where ambulance-based designs (such as in Doyle et al., 2015) are infeasible. Finally, we show how concerns over “correlated treatments” can be addressed using a straightforward multi-channel instrumental variables approach. These issues arise frequently in evaluations in healthcare (e.g. McClellan et al., 1994; Doyle et al., 2015), education (e.g. Walters 2015), welfare programs (e.g. Doyle, 2007), crime (e.g. Kling, 2006), and public assistance (e.g. Maestas et al. 2013). An approach similar to ours could be applied in a variety of settings where data on multiple characteristics of the unit of randomization – the hospital in our case – are available.

UC BERKELEY & NBER
UC BERKELEY
PRINCETON UNIVERSITY

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113(2), 231-263.

Alexander, D. (2015). Does physician pay affect procedure choice and patient health? Evidence from Medicaid c-section use. Working paper.

Alexander, J. M., et al. (2006). Fetal injury associated with cesarean delivery. *Obstetrics and Gynecology*, 108(4), 885-890.

Altman, M., Sandström, et al. (2015). Prolonged second stage of labor is associated with low Apgar score. *European Journal of Epidemiology*, 30(11), 1209-1215.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, 113(1), 151-184.

American College of Obstetricians and Gynecologists (ACOG) (2006). ACOG Committee Opinion No. 340. Mode of term singleton breech delivery. *Obstetrics and Gynecology*, 108(1), 235.

-- (2014). Safe prevention of the primary cesarean delivery. *American Journal of Obstetrics and Gynecology* 123(1), 693-711.

Angus, D. C., et al. (2001). Epidemiology of Neonatal Respiratory Failure in the United States: Projections from California and New York. *American Journal of Respiratory and Critical Care Medicine*, 164(7), 1154-1160.

Baicker, K., Buckles, K. S., and Chandra, A. (2006). Geographic variation in the appropriate use of cesarean delivery. *Health Affairs*, 25(5), w355-w367.

Baicker, K., A. Chandra and J. Skinner (2012). Saving Money or Just Saving Lives? Improving the Productivity of U.S. Health Care Spending. *Annual Review of Economics*, 4, 33-56.

Barnato, A. E., et al. (2005). Hospital Level Racial Disparities in Acute Myocardial Infarction Treatment and Outcomes. *Medical Care*, 43(4), 308-319.

Beckert, W., Christensen, M., and Collyer, K. (2012). Choice of NHS-funded hospital services in England. *The Economic Journal*, 122(560), 400-417.

Bloom, N., Propper, C., Seiler, S., & Van Reenen, J. (2015). The impact of competition on management quality: evidence from public hospitals. *The Review of Economic Studies*, *82*(2), 457-489.

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.

Centers for Disease Control and Prevention (2009). Asthma's impact on the nation: Data from the CDC National Asthma Control Program. Atlanta, GA: CDC. Accessed at https://www.cdc.gov/asthma/impacts_nation/asthmafactsheet.pdf

Chandra, A. and Staiger, D. O. (2007). Productivity spillovers in healthcare: evidence from the treatment of heart attacks. *Journal of Political Economy*, 115(1), 103-140.

Chandra, A. and Staiger, D. O. (2017). Identifying sources of inefficiency in health care. NBER Working Paper #24015. Cambridge, MA: NBER.

Chernozhukov, V., and Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1), 245-261.

Clark, E. A., and Silver, R. M. (2011). Long-term maternal morbidity associated with repeat cesarean delivery. *American Journal of Obstetrics and Gynecology*, 205(6), S2-S10.

Cohen, W. R., Sumersille, M., and Friedman, E.A. (2018). Management of labor: Are the new guidelines justified? *Journal of Midwifery and Women's Health*, 63(1), 10-13.

Cornelissen, T., Dustmann, C., Raute, A., and Schoenberg, U. (2016). From LATE to MTE: Alternative approaches for the evaluation of policy interventions. *Labour Economics*, 41, 47-60.

Costa-Ramón, A. M., Rodriguez-Gonzalez, A., Serra-Burruel, M. and Campillo-Arero, C. (2018). It's about time: Cesarean sections and neonatal health. *Journal of Health Economics*, 59, 46-59.

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

Currie, J., and MacLeod, W. B. (2017). Diagnosing Expertise: Human Capital, Decision Making, and Performance among Physicians. *Journal of Labor Economics*, 35(1), 1-43.

Cutler, D. (2007). The lifetime costs and benefits of medical technology. *Journal of Health Economics*, 26(6), 1081-1100.

Cutler, D., Skinner, J. S., Stern, A. D., & Wennberg, D. (2019). Physician beliefs and patient preferences: A new look at regional variation in health care spending. *American Economic Journal: Economic Policy*, *11*(1), 192-221.

Dartmouth Atlas, undated. List of Hospital Referral Regions. Available at: http://www.dartmouthatlas.org/tools/faq/researchmethods.aspx.

Declercq, E. R., Sakala, C., Corry, M. P., and Applebaum, S. (2006). Listening to Mothers II: Report of the Second National US Survey of Women's Childbearing Experiences. New York: Childbirth Connections.

Ding, R., Mccarthy, M.L., Desmond, J.S., Lee, J.S., Aronsky, D., and Zeger, S.L. (2010). Characterizing waiting room time, treatment time, and boarding time in the emergency department using quantile regression. *Academic Emergency Medicine*, 17(8): 813-823.

Doyle Jr, 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.

Dubay, L., Kaestner, R., and Waidman, T. (1999). The impact of malpractice fears on cesarean section rates. *Journal of Health Economics*, 18(4), 491-522.

Epstein, A. J., & Nicholson, S. (2009). The formation and evolution of physician treatment styles: An application to cesarean sections. *Journal of Health Economics*, *28*(6), 1126-1140.

Finkelstein, A., Gentzkow, M., and Williams, H. (2016). Sources of Geographic Variation in Health Care: Evidence From Patient Migration. *Quarterly Journal of Economics*, 131(4), 1681-1726.

Finster, M., and Wood, M. (2005). The Apgar score has survived the test of time. *The Journal of the American Society of Anesthesiologists*, 102(4), 855-857.

Fisher, E., Wennberg, D.E., Stukel, T.A., Gottleib, D.J., Lucas, F.L. and Pinder, E.L. (2003a). The Implications Of Regional Variations In Medicare Spending. Part 1: The Content, Quality, And Accessibility Of Care. *Annals of Internal Medicine*, 138(4), 273–289.

Fisher, E., Wennberg, D.E., Stukel, T.A., Gottleib, D.J., Lucas, F.L. and Pinder, E.L. (2003b). The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine*, 138(4), 288–298.

Friedman, E.A. (1995) Primigravid labor: A graphicostatistical analysis. *Obstetrics and Gynecology*, 6(6), 567-589.

Frigoletto, F. D., Lieberman, E., Lang, J. M., Cohen, A., Barss, V., Ringer, S., and Datta, S. (1995). A Clinical Trial of Active Management of Labor. *The New England Journal of Medicine*, 333(12), 745-750.

Garabedian, L. F., Chu, P., Toh, S., Zaslavsky, A. M., and Soumerai, S. B. (2014). Potential Bias of Instrumental Variable Analyses for Observational Comparative Effectiveness Research. *Annals of Internal Medicine*, 161(2), 131-138.

Garen, John (1984). The returns to schooling: A selectivity bias approach with a continuous choice variable. *Econometrica*, 52(5), 1199-1218.

Gimovsky, A. C., and Berghella, V. (2016). Randomized controlled trial of prolonged second stage: extending the time limit vs usual guidelines. *American Journal of Obstetrics and Gynecology*, 3(214), 361-e1.

Goer, H., Romano, A. and Sakala, C. (2012). Vaginal or cesarean birth: What is at stake for women and babies? New York: Childbirth Connections.

Gregory, K. D., Korst, L. M., Gornbein, J. A., and Platt, L. D. (2002). Using Administrative Data to Identify Indications for Elective Primary Cesarean Delivery. *Health Services Research*, 37(5), 1387-1401.

Gregory, K. D., Jackson, S., Korst, L., and Fridman, M. (2012). Cesarean versus Vaginal Delivery: Whose Risks? Whose Benefits? *American Journal of Perinatology*, 29(01), 07-18.

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

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

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

Gyhagen, M., Akervall, S. and Milsom, I. (2015). Clustering of pelvic floor disorders 20 years after one vaginal or one cesarean birth. *International Urogynecological Journal*, 26, 1114-1121.

Hadley, J. and Peter P. Cunningham (2004). Availability of safety net providers and access to care of uninsured persons. *Health Services Research*, 39(5), 1527-1546.

Halla, M., Mayr, H., Pruckner, G. J., and Garcia-Gomez, P. (2018). Cutting fertility? The effect of cesarean deliveries on subsequent fertility and maternal labor supply. Johannes Kepler University Linz Unpublished Working Paper.

Heckman, J.J. and Vytlacil, E. (1998). Instrumental variables models for the correlated random coefficients model: Estimating the average return to schooling when the return is correlated with schooling. *Journal of Human Resources*, 33(4), 974-987.

Henry, O. A., Gregory, K. D., Hobel, C. J., and Platt, L. D. (1995). Using ICD-9 codes to identify indications for primary and repeat cesarean sections: Agreement with clinical records. *American Journal of Public Health*, 85(8), 1143.

Howell, C. J. (2000). Epidural versus non-epidural analgesia for pain relief in labour. *The Cochrane Database of Systematic Reviews*, (2), CD000331.

Hyde, M. J., Mostyn, A., Modi, N., and Kemp, P. R. (2012). The health implications of birth by caesarean section. *Biological Reviews*, 87(1), 229-243.

Imbens, G.W. and Wooldridge, J.M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5-86.

Jachetta, C. (2014) Cesarean Sections and Later Child Health Outcomes. Unpublished manuscript. University of Illinois Urbana Champaign Department of Economics. Downloaded from http://www.christinejachetta.com/

Jensen, V. M. and Wüst, M. (2015). Can cesarean section improve child and maternal health? The case of breech babies. *Journal of Health Economics*, 39, 289-302.

Johnson, H.P. and Li, Q. (2007). Birth rates in California. California Counts 9(2). San Francisco, CA: Public Policy Institute of California.

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.

Joint Commission (2016). Specifications manual for Joint Commission National Quality Measures (v2016A1): measure information form. Perinatal Care PC-02: Cesarean Birth.

Joint Commission (2018). The Joint Commission Will Begin Publicly Reporting Cesarean Section Rates by July 2020. Retrieved from https://www.jointcommission.org/assets/1/18/Peformance_Measure_PC-02_JCP1218.pdf

Keag, O. E., Norman, J. E., and Stock, S. J. (2018). Long-term risks and benefits associated with cesarean delivery for mother, baby, and subsequent pregnancies: Systematic review and meta-analysis. *PLOS Medicine* 15(1), e1002494.

Klein, M. C. (2006). Does epidural analgesia increase rate of cesarean section? *Canadian Family Physician*, 52(4), 419.

Kling, J. (2006). Incarceration length, employment and earnings. *American Economic Review* 96(3), 863–876.

Kozhimannil, K. B., Law, M. R., and Virnig, B. A. (2013). Cesarean delivery rates vary tenfold among US hospitals; reducing variation may address quality and cost issues. *Health Affairs*, 32(3), 527-535.

Kozhimannil, K. B., Arcaya, M.C. and Subramanian, S.V. (2014). Maternal clinical diagnoses and hospital variation in the risk of cesarean delivery: Analyses of a national US hospital discharge database. *PLOS Medicine*, 11(10), e1001745.

López-Zeno, J. A., Peaceman, A. M., Adashek, J. A., and Socol, M. L. (1992). A controlled trial of a program for the active management of labor. *New England Journal of Medicine*, 326(7), 450-454.

MacDorman, M. F., Declercq, E., Menacker, F., and Malloy, M. H. (2008). Neonatal mortality for primary cesarean and vaginal births to low-risk women: Application of an "intention-to-treat" model. *Birth*, 35(1), 3-8.

Maestas, N., Mullen, K. J., & Strand, A. (2013). Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt. *American Economic Review*, *103*(5), 1797-1829.

Martin, J. A., Hamilton, B. E., Ventura, S. J., Osterman, M. J. K., and Mathews, T. J. (2013a). Births: final data for 2011. National Vital Statistics Reports 62(1)

Martin, J.A., Wilson, E.C., Osterman, M.J.K., Saadi, E.W., Sutton, S.R., and Hamilton, B.K. (2013b). Assessing the quality of medical and health data from the 2003 birth certificate revision: Results from two states. National Vital Statistics Reports 62(2).

McClellan, M., McNeil, B. J., and Newhouse, J. P. (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *Journal of the American Medical Association*, 272(11), 859.

Morris, T. (2013). *Cut It Out: The C-Section Epidemic in America*. NYU Press.

Mueller-Smith, M. (2015). The criminal and labor market impacts of incarceration. Working paper.

National Center for Health Statistics (NCHS) (2016). Guide to Completing the facility worksheets for the certificate of live birth and report of fetal death. National Vital Statistics System (2003 revision). Atlanta: Center for Disease Control and Prevention.

Oster, E. (2017). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business and Economic Statistics*, 1-18.

Osterman, M.J.K, and Martin, J. A. (2014). Trends in low-risk cesarean delivery in the United States, 1990-2013. National Vital Statistics Reports 63(6).

Phibbs, C. S., Mark, D. H., Luft, H. S., Peltzman-Rennie, D. J., Garnick, D. W., Lichtenberg, E., & McPhee, S. J. (1993). Choice of hospital for delivery: a comparison of high-risk and low-risk women. *Health Services Research*, *28*(2), 201.

Podulka, J., Stranges, E., and Steiner, C. (2011). Hospitalizations related to childbirth, 2008. HCUP statistical brief, 110.

Rennie, J. and Rosenbloom L. (2011). How long have we got to get the baby out? A review of the effects of acute and profound intrapartum hypoxia and ischaemia. *The Obstetrician and Gynecologist*, 13, 169-174.

Robson, M.S. 2001. Classification of cesarean sections. *Fetal and Maternal Medicine Review*, 12(1), 23-39.

Romley, J. A., A. B. Jena, and D. P. Goldman (2011). Hospital Spending and Inpatient Mortality: Evidence from California. *Annals of Internal Medicine*, 154(3), 160-167.

Rysavy, M.A. et al. (2015). Between-hospital variation in treatment and outcomes in extremely preterm infants. *New England Journal of Medicine*, 372(19), 1801-1811.

Sakala, Carol, Suzanne F. Delbanco and Harold D. Miller (2013). The cost of having a baby in the United States. Truven Health Analytics Marketscan Study.

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

Skinner, J. (2012). Causes and Consequences of Regional Variations in Health Care. *Handbook of Health Economics*, 2, 45-93.

Spong, C. Y., Berghella, V., Wenstrom, K. D., Mercer, B. M., and Saade, G. R. (2012). Preventing the first cesarean delivery: Summary of a joint Eunice Kennedy Shriver National Institute of Child Health and Human Development, Society for Maternal-Fetal Medicine, and American College of Obstetricians and Gynecologists Workshop. *Obstetrics and Gynecology*, 120(5), 1181-1193.

Thomas, J., Paranjothy, S., and James, D. (2004). National cross sectional survey to determine whether the decision to delivery interval is critical in emergency caesarean section. *British Medical Journal*, 328(7441), 665-671.

Tolcher, M. C., Johnson, R. L., El-Nashar, S. A., and West, C. P. (2014). Decision-to-incision yime and neonatal outcomes: A systematic review and meta-analysis. *Obstetric Anesthesia Digest*, 35(1), 536-548.

U.S. Department of Transportation (2008). Revised departmental guidance: treatment of the value of preventing fatalities and injuries in preparing economic analyses. Memorandum to Secretarial Officers, Modal Administrators re: Treatment of the Economic Value of a Statistical Life in Departmental Analyses.

Walters, C. R. (2015). Inputs in the production of early childhood human capital: Evidence from Head Start. *American Economic Journal: Applied Economics*, *7*(4), 76-102.

Wier, L. M., Pfuntner, A., Maeda, J., Stranges, E., Ryan, K., Jagadish, P., Collins Sharp, B., and Elixhauser, A. (2011). HCUP facts and figures: statistics on hospital-based care in the United States, 2009. Rockville, MD: Agency for Healthcare Research and Quality (2011).

Wooldridge, Jeffrey M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, 50(2), 420-445.

Zhang, J., et al. (2010). Contemporary cesarean delivery practice in the United States. *American Journal of Obstetrics and Gynecology*, 203(4), 326.e1-326.e10.

Figure I: Pathways to C-Section Delivery

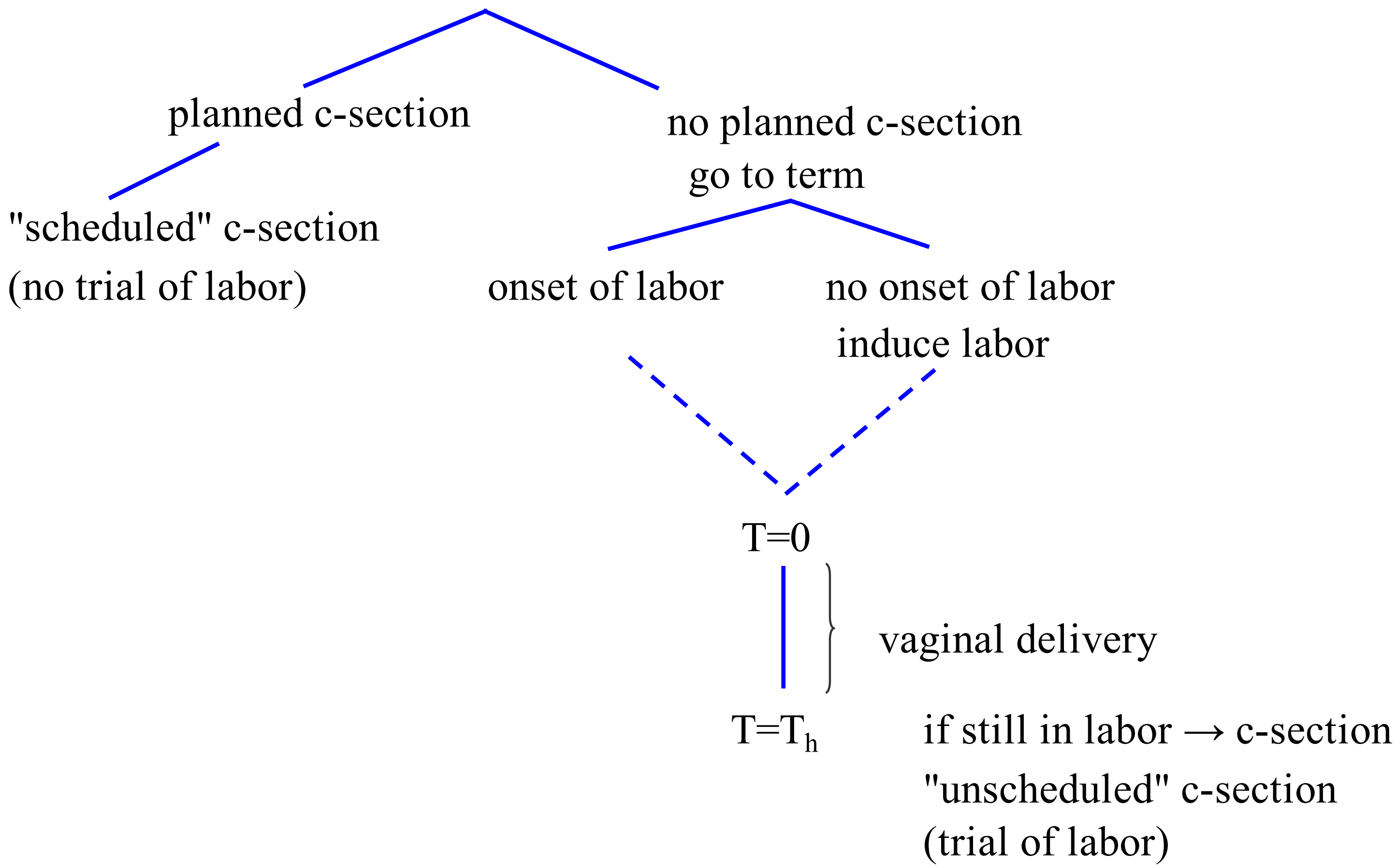
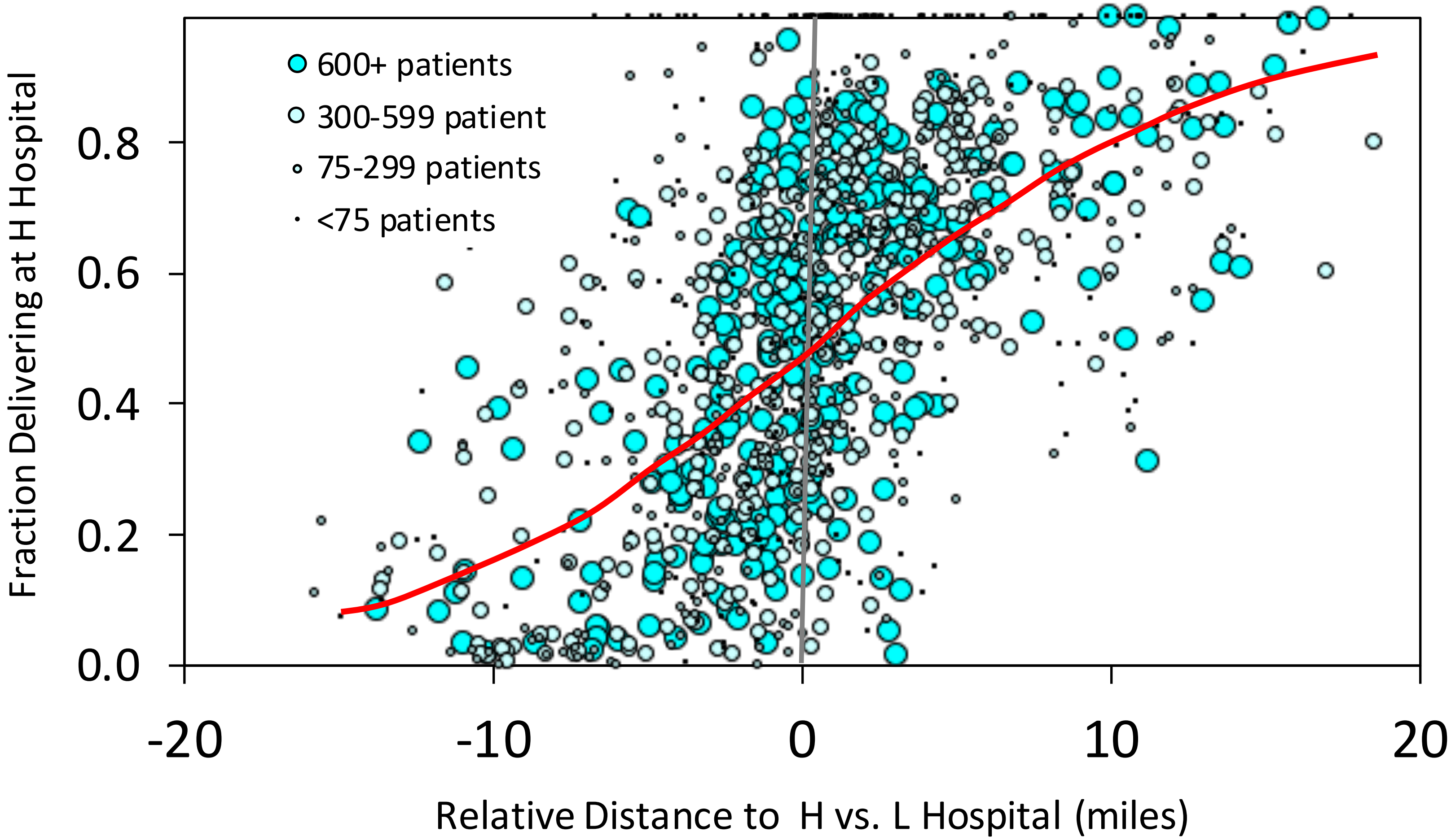


Figure II: Relative Distance and Probability of Delivery at High C-Section (H) Hospital



Note: each point represents a home zip code. Fitted logistic shown in red.

Figure III: Estimated Impact of Delivery at High C-Section Hospital on Probability of Inpatient Stay at Various Horizons

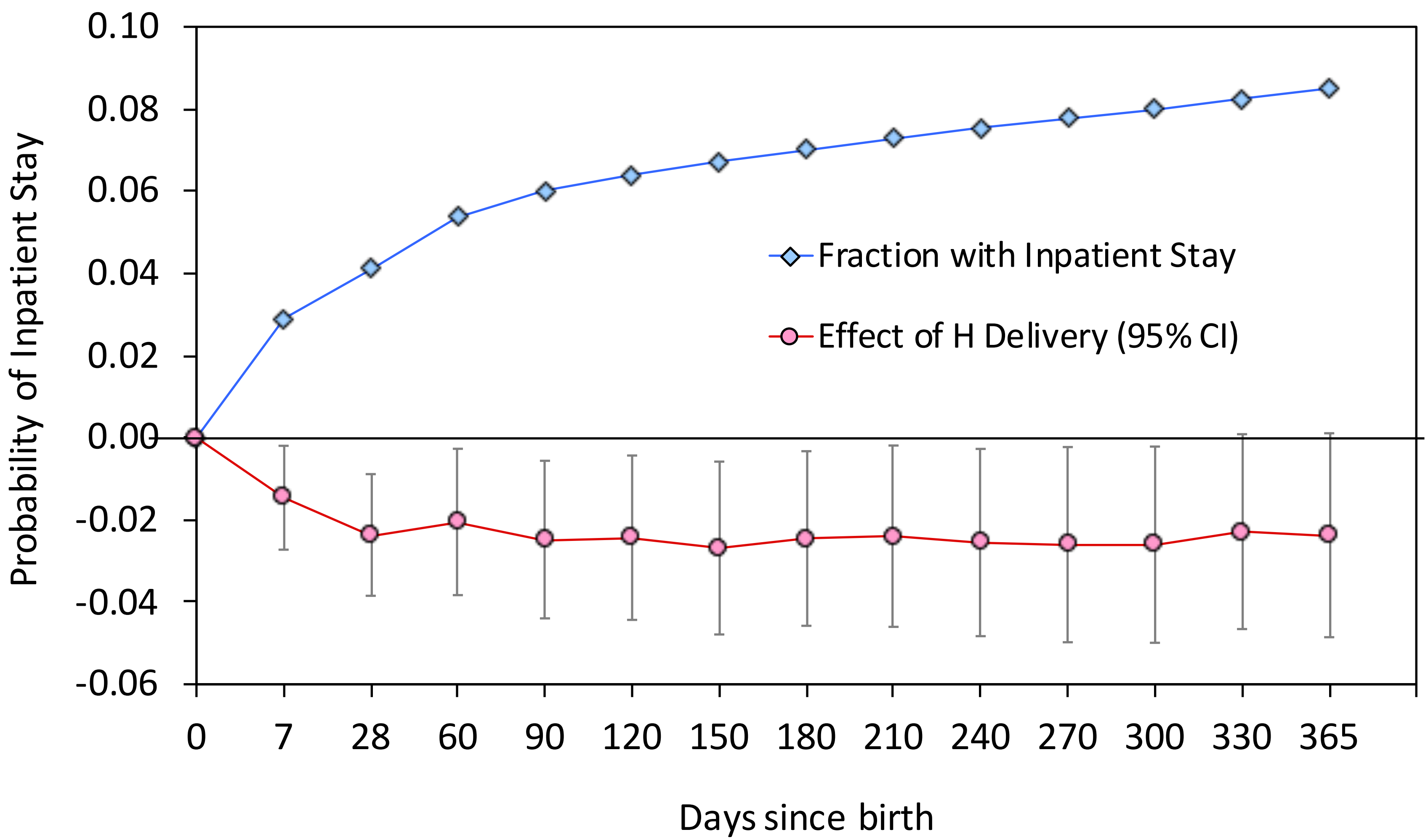
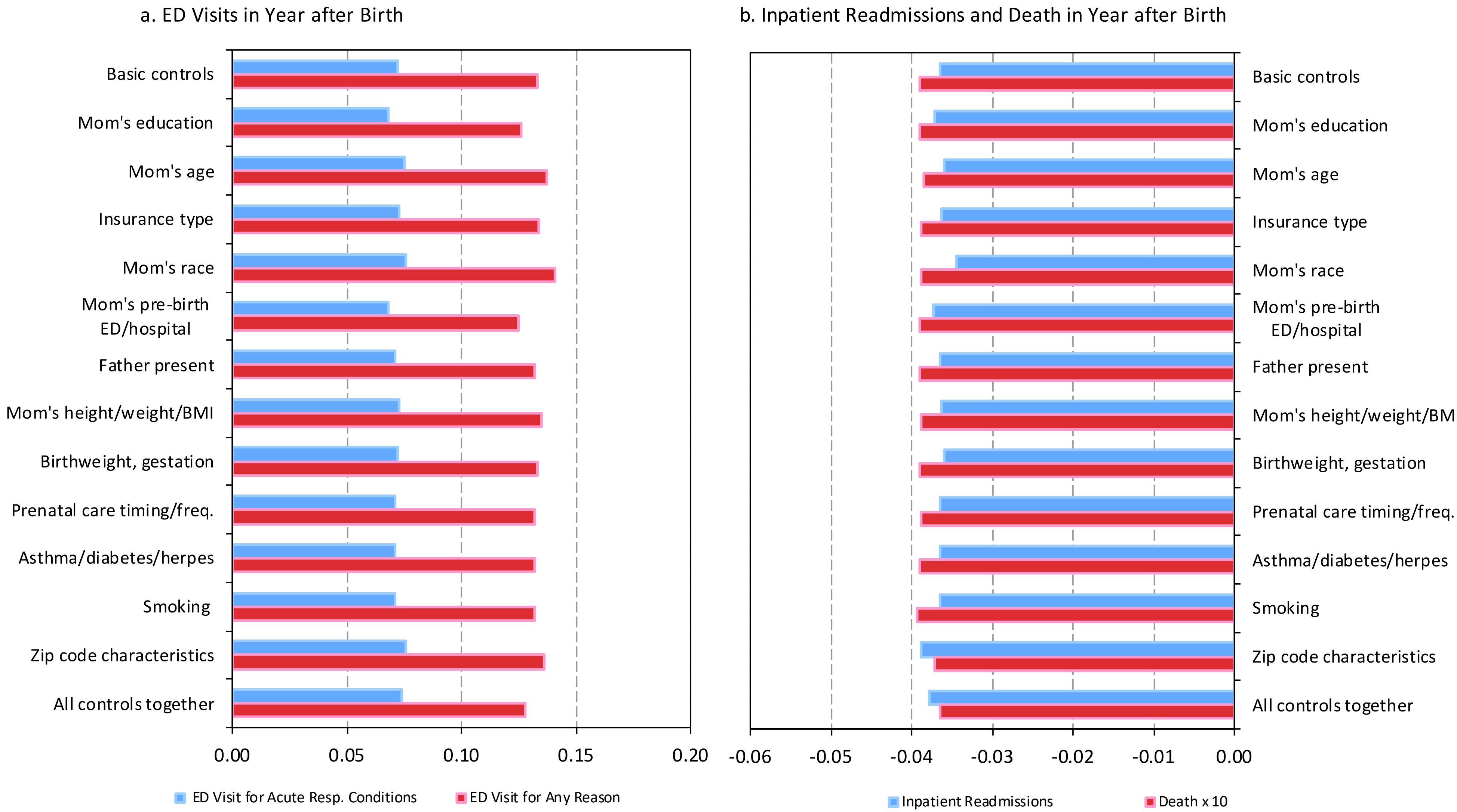


Figure IV: Sensitivity of Reduced-Form Effects of Relative Distance on Infant Health Outcomes



Notes: Figures report estimated reduced form effects on ED visits (panel a) or inpatient readmissions and death (panel b) from models that include basic set of controls described in Table 2 plus additional controls described on figure axes. "All controls together" estimate at bottom of figure include all control variables simultaneously.

Figure V: Sensitivity of Reduced-Form Effects of Relative Distance on Infant Outcomes -- Distributions of Possible Estimates

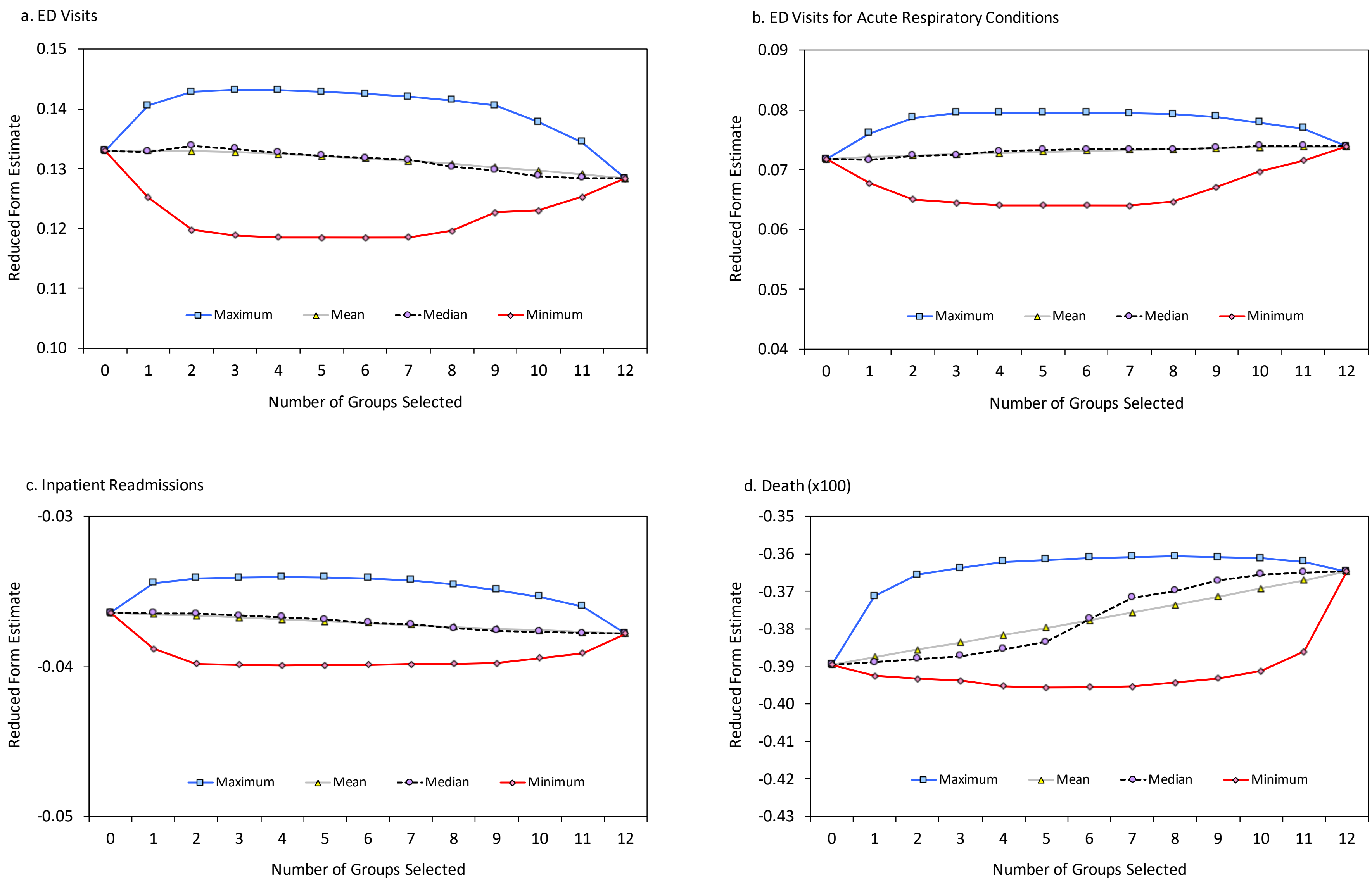
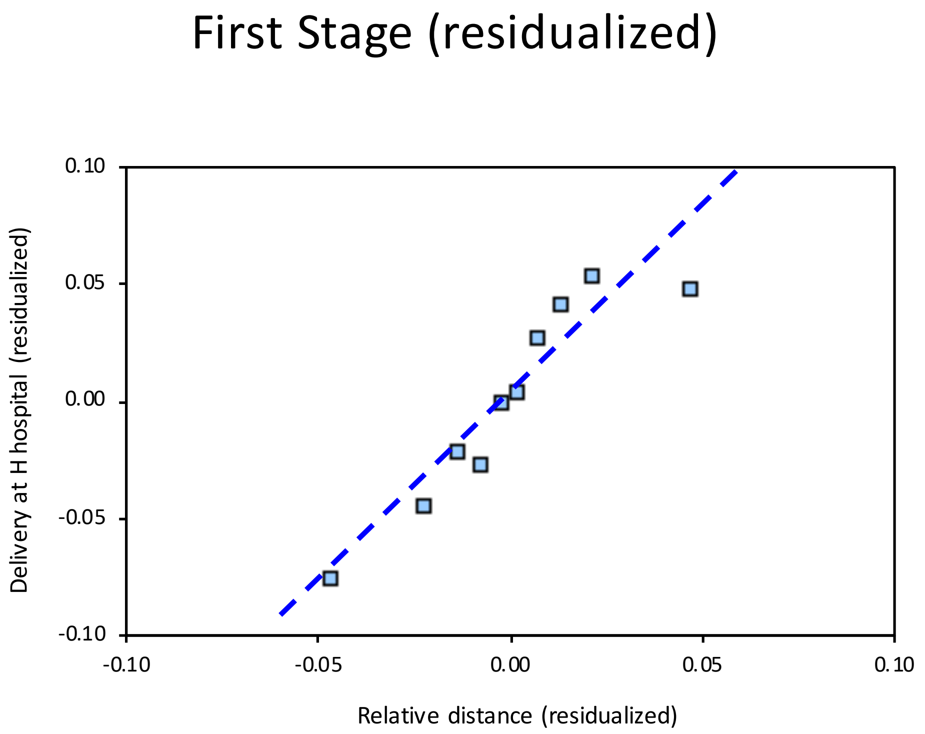


Figure VI: Plots of First Stage and Reduced Form for Infant Death



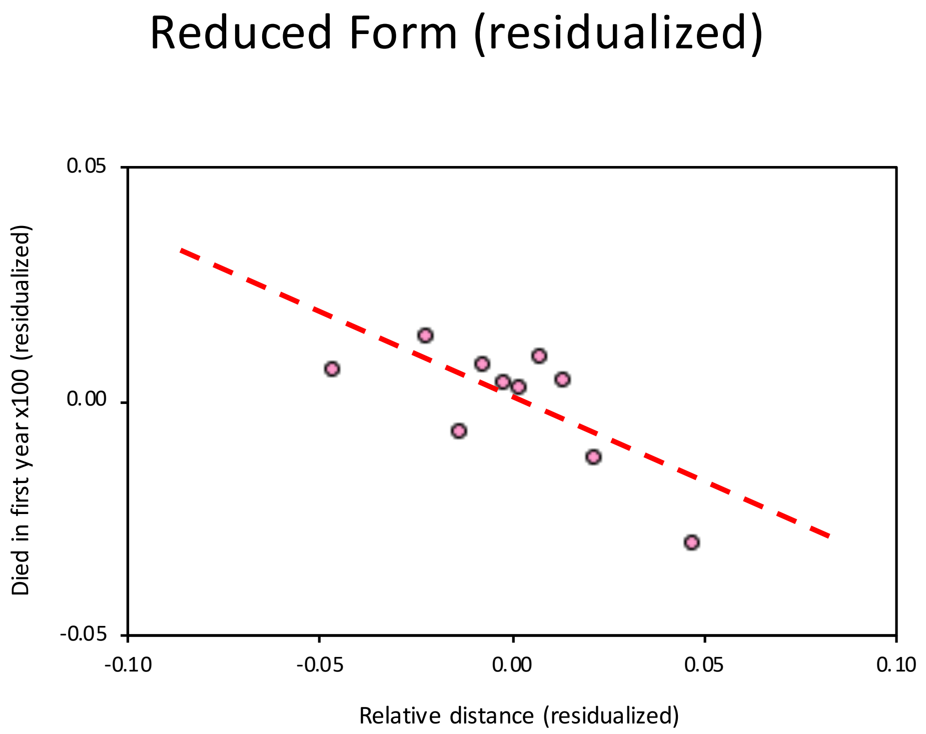


Table I: Characteristics of All Births, Low-Risk First Births and Analysis Sample

	All Births	Low Risk First Births	Analysis Sample
Mother's characteristics			
Age	28.3	25.4	25.6
High School or Less (%)	49.6	41.5	41.2
Mean weight (pounds)	149	137	137
Race/eth: Hispanic (%)	50.9	44.2	44.2
White (nonhispanic) (%)	27.7	32.2	31.7
Asian (nonhispanic) (%)	12.2	15.5	17.6
Black (nonhispanic) (%)	5.8	5.5	5.6
Insurance: Medi-Cal (%)	46.0	41.1	39.8
Private non-Kaiser (%)	34.3	38.0	39.9
Private Kaiser (%)	13.1	14.2	14.9
Birth risk factors and characteristics			
Mean parity	2.1	1.0	1.0
Previous c-section (%) (among parity>1)	24.7	0.0	0.0
Breech presentation (%)	3.4	0.0	0.0
Mean number prenatal care visits	12.0	12.1	12.2
Maternal ED visit year prior to birth (%)	20.6	19.9	19.5
Maternal IP visit year prior to birth (%)	5.5	4.0	3.9
Mean gestation (weeks)	39.2	39.9	39.9
Mean birthweight (grams)	3,309	3,347	3,348
Delivery outcomes			
C-section delivery (%)	32.7	25.5	25.6
C-section without indicated trial of labor (%)	23.7	9.4	9.2
Delivered at H hospital (%)	51.8	51.6	51.5
Postpartum outcomes			
Infant readmitted to ED (%)	33.9	34.0	33.8
Infant readmitted as inpatient (%)	9.3	8.2	8.2
Mother readmitted (any type) (%)	16.8	15.3	14.9
Births	2,699,302	631,506	491,604

Notes: All births include all live in-hospital births in California, 2007-2011. Low risk first births include singleton nonbreech full term (37+ weeks) first births to mothers age 18-35 with no indications of eclampsia, pre-eclampsia, or growth restrictions, mother's BMI < 33.83 (90th percentile), and ≤ 20 prenatal visits. Analysis sample includes mothers with valid home zip code, distance to nearest high or low c-section hospital ≤ 20 miles, and distance to actual hospital of delivery ≤ 20 miles.

Table II: Orthogonality of Relative Distance to Maternal Characteristics and Risk Factors

	Effect of moving 10 miles closer to H hospital on row variable (s.e. in parentheses)			Effect of delivering at H hospital on row variable (s.e. in parentheses)		
	(1)			(2)		
Maternal Characteristics						
Mother's Age	0.021	(0.068)		-0.098	(0.055)	*
Mother's Education	-0.040	(0.048)		-0.035	(0.040)	
White (non-Hispanic)	0.016	(0.013)		0.005	(0.004)	
Black	-0.003	(0.007)		-0.010	(0.002)	***
Asian	-0.008	(0.012)		0.009	(0.006)	
Hispanic	-0.006	(0.014)		-0.002	(0.007)	
Father Present	-0.001	(0.002)		-0.003	(0.001)	**
Gov't Insurance	-0.001	(0.004)		0.067	(0.011)	***
Private Insurance	0.002	(0.005)		-0.087	(0.011)	***
Mother's Height (inches)	0.023	(0.043)		0.000	(0.022)	
Mother's Weight (pounds)	0.039	(0.302)		-0.270	(0.170)	
BMI Pre-pregnancy	-0.011	(0.041)		-0.047	(0.027)	*
Mother's Use of Hospital in Year Before Birth						
Any ED Visit	0.006	(0.004)		0.007	(0.003)	**
Number ED Visits	0.010	(0.007)		0.010	(0.004)	***
Inpatient Stay	-0.001	(0.001)		0.003	(0.001)	***
Prenatal Care:						
Prenatal Visits (#)	0.046	(0.055)		0.251	(0.048)	***
Month Started Pre. Care	0.021	(0.021)		-0.060	(0.018)	***
Late Prenatal Care (>4th mo)	0.003	(0.002)		-0.002	(0.002)	
Other Risk Characteristics						
Diabetes	-0.001	(0.001)		-0.003	(0.001)	***
Herpes	-0.001	(0.001)	*	-0.001	(0.000)	**
Asthma	0.001	(0.001)	**	-0.004	(0.001)	***
Smoked When Pregnant	0.002	(0.001)		-0.001	(0.001)	
Cigs/Day Pre-pregnancy	0.024	(0.025)		0.006	(0.013)	
Infant Characteristics						
Gestation (days)	0.044	(0.076)		-0.711	(0.052)	***
Birth Weight (grams)	1.129	(4.023)		-13.981	(1.993)	***
Low Birth Weight (<2500 g)	0.000	(0.001)		0.002	(0.001)	***
Characteristics of Mother's Home Zip Code						
Mean Income (1000 US $)	1.349	(1.059)		-0.296	(0.259)	
Zip Mean Mother Educ.	0.010	(0.045)		0.004	(0.008)	
Zip Mean Dropout	0.005	(0.008)		0.000	(0.001)	
Zip Mean Black	-0.004	(0.006)		-0.003	(0.001)	**
Zip Mean Hispanic	-0.006	(0.014)		0.001	(0.002)	
Logit predictions based on above 31 covariates						
Predicted Pr(Infant ED visit)	0.001	(0.003)		0.006	(0.002)	***
Predicted Pr(Infant readmission)	0.000	(0.001)		0.002	(0.000)	***
Predicted Pr(Infant death) x 100	-0.001	(0.002)		0.004	(0.001)	***
F-tests based on above 31 covariates						
Joint F-statistic: F(31,1249)		1.041			17.714	
Joint F-test *p*-value		0.406			0.000	***

Notes: Table shows estimated coefficients and standard errors from regression of row variable on relative distance in 10s of miles to a high c-section (H) hospital (column 1) or delivery at an H hospital (column 2). All models include HSA and year effects, distance from home to nearest hospital, and fraction of mothers in zip code with government insurance. Logit predictions from logit model of respective outcome on all demographic and risk factors listed above. Bottom two rows present F-statistics and p-values from the joint F-test for all 31 row variables in reverse regression with relative distance or delivery hospital type as dependent variable. Standard errors clustered by zip code. * indicates p <0.10; ** indicates p <0.05; *** indicates p <0.01.

Table III: Estimated Effects of Relative Distance on Place, Mode, and Timing of Delivery

		Instrument=Relative Distance to H Hospital Coefficients × 100		Instrument= Indicator for Closer to H Hospital Coefficients × 10	
Outcome Variable	Mean	Baseline controls only	All controls	Baseline controls only	All controls
Deliver at High C-Section (H) Hospital	0.515	1.586	1.600	1.014	1.014
		(0.154)	(0.153)	(0.122)	(0.119)
C-section Delivery	0.256	0.186	0.183	0.109	0.118
		(0.030)	(0.030)	(0.023)	(0.022)
Scheduled C-section (no indicators of trial of labor)	0.092	0.056	0.053	0.026	0.023
		(0.021)	(0.019)	(0.014)	(0.013)
Unscheduled C-section (indications of trial of labor)	0.163	0.130	0.130	0.082	0.095
		(0.029)	(0.023)	(0.019)	(0.018)
Delivered 1+ Days After Arrival	0.479	-0.054	-0.052	-0.044	-0.049
		(0.036)	(0.037)	(0.025)	(0.025)
Delivered 2+ Days After Arrival	0.046	-0.041	-0.041	-0.022	-0.022
		(0.014)	(0.014)	(0.010)	(0.010)
C-section on Day of Arrival	0.125	0.126	0.120	0.084	0.088
		(0.022)	(0.025)	(0.016)	(0.016)
C-section 1+ Days After Arrival	0.132	0.061	0.067	0.024	0.029
		(0.021)	(0.020)	(0.016)	(0.015)
Vaginal Del. on Day of Arrival	0.396	-0.072	-0.068	-0.040	-0.039
		(0.032)	(0.031)	(0.024)	(0.023)
Vaginal Del. 1+ Days After Arrival	0.347	-0.114	-0.119	-0.068	-0.078
		(0.031)	(0.033)	(0.022)	(0.022)
Breakdown of C-Section Deliveries:					
C-Section at H Hospital	0.149	0.491	0.494	0.314	0.320
		(0.048)	(0.050)	(0.039)	(0.039)
C-Section at L Hospital	0.106	-0.305	-0.312	-0.205	-0.202
		(0.036)	(0.032)	(0.030)	(0.028)
Fractions of Complier Groups -- Moving 7 mi. closer to H hospital (col. 2-3) or closer to H hospital (col. 4-5)					
P(H Complier)		0.111	0.112	0.101	0.101
P(C&H Complier)		0.013	0.013	0.011	0.012
P(H Complier & C Always-Taker)		0.021	0.022	0.021	0.020
P(H Complier & V Always-Taker)		0.077	0.077	0.070	0.069
P(C Complier|H Complier)		0.117	0.114	0.107	0.116
P(C Always-Taker|H Complier)		0.192	0.195	0.203	0.200
P(V Always-Taker|H Complier)		0.690	0.691	0.690	0.684

Notes: Sample=491,604 low-risk first births. Standard errors in parentheses clustered at 5-digit ZIP code level. Baseline controls are dummies for Hospital Service Area and year of birth, and controls for distance to closest hospital, and fraction of new mothers in ZIP code covered by Medi-Cal or other public insurance. All controls include 59 additional controls: mother's age (17 dummies), mother's education (8 dummies), race (4 dummies), father present, insurance type (3 dummies), cubic in mother's height, cubic in mother's weight, pre-pregnancy BMI, mother's pre-birth hospital use (3 variables), prenatal care (3 variables), mother's diseases and smoking (5 variables), birthweigh and gestation (3 variables) and ZIP code characteristics (5 variables).

Table IV: Characteristics of Compliers

	Low-Risk First Births	Hospital (H) Compliers		Hospital and Prodedure (H&C) Compliers	
	(1)	(2)		(3)	
A. Socio-Economic Characteristics of Mother					
Race/Ethnicity					
White	31.7%	26.8	(3.7)	21.1	(8.7)
Black	5.6%	3.1	(1.1)	7.4	(4.0)
Asian	17.6%	11.1	(3.3)	12.0	(7.7)
Hispanic	44.2%	58.8	(4.2)	59.8	(9.4)
Education					
High School or Less	41.2%	51.7	(3.5)	68.5	(9.9)
Some College	20.1%	15.0	(1.4)	17.9	(5.8)
BA or Higher	38.7%	33.3	(3.9)	13.6	(10.1)
Home Zip Code Characteristics					
Income < Median	50.0%	49.7	(8.2)	88.5	(14.7)
Gov Insurance Rate > Mean	51.8%	51.9	(8.3)	88.5	(14.5)
B. Mother's Insurance Coverage at Delivery					
Government-provided	43.3%	60.9	(4.3)	82.0	(11.2)
Private	52.9%	35.6	(4.3)	18.0	(10.7)
Other	3.8%	3.5	(0.7)	0.0	(2.8)
C. Other Maternal/Infant Characteristics					
Mother height < 5 ft.	4.2%	5.4	(0.6)	8.0	(3.4)
Mother visit ED prepartum	19.5%	21.6	(1.6)	32.3	(6.4)
Number prepartum ED visits	25.8%	29.6	(2.4)	46.3	(9.2)
Male baby	51.0%	53.2	(1.0)	59.8	(7.3)
Birth weight < median	50.0%	51.1	(1.2)	50.4	(6.8)
Low birth weight (<2500 g)	2.3%	2.7	(0.3)	3.3	(2.1)
D. Predicted Probabilities of Infant ED Visits and Inpatient Admissions					
Prob(ED Visit) > median	50.0%	60.7	(4.9)	79.6	(11.5)
Prob(Inpat. Adm) > median	50.0%	55.2	(4.4)	71.6	(10.3)

Notes: Column 1 shows estimated means (in percents) for analysis sample of LRFBs. Column 2 shows means for births that are delivered at H hospitals as a result of being relatively closer to such hospitals; column 3 shows means for births that are delivered by c-section as a result of being closer to an H-hospital. Models used to estimate complier characteristics include all controls plus characteristic itself. Standard errors, clustered by maternal zip code, in parentheses.

Table V: Effects of Delivery at High C-Section Hospital on Infant and Maternal Outcomes

Outcome Variable	Mean	OLS Coefficients: Deliver at H-hospital	OLS Coefficients: C-Section	RF Coefficient (x 100)	IV Estimate (Scaled per delivery at H-hospital)
First-stage coefficients	--	1.600	0.183	--	--
		(0.153)	(0.030)		
Infant Outcomes:					
Apgar (5 minute)	8.915	0.024	-0.021	0.089	0.056
		(0.002)	(0.002)	(0.030)	(0.019)
Birth Injury (×100)	0.094	-0.005	-0.035	-0.146	-0.091
		(0.014)	(0.010)	(0.127)	(0.080)
NICU	0.034	-0.016	0.021	-0.028	-0.017
		(0.001)	(0.001)	(0.013)	(0.008)
NICU including transfers (with control for NICU at delivery hospital)	0.040	-0.012	0.025	-0.011	-0.007
		(0.001)	(0.001)	(0.013)	(0.008)
Ventilation	0.015	-0.002	0.010	0.044	0.027
		(0.001)	(0.001)	(0.018)	(0.011)
Length of stay (raw)	2.354	0.065	1.356	-0.120	-0.074
		(0.012)	(0.012)	(0.140)	(0.085)
Length of stay (top-coded at 6 days)	2.242	0.101	1.245	-0.058	-0.036
		(0.008)	(0.007)	(0.090)	(0.055)
Length of stay ≥ 5 days	0.037	-0.006	0.030	-0.018	-0.011
		(0.001)	(0.001)	(0.012)	(0.007)
Maternal Outcomes:					
Trauma to perineum and vulva during labor	0.461	-0.122	-0.612	-0.159	-0.099
		(0.004)	(0.005)	(0.051)	(0.028)
Perineal laceration (2nd degree or higher)	0.290	-0.074	-0.404	-0.145	-0.090
		(0.003)	(0.005)	(0.037)	(0.020)
Length of stay (days)	2.637	0.086	1.353	-0.025	-0.015
		(0.007)	(0.007)	(0.094)	(0.058)
Length of labor (days) (birth - admission)	0.530	-0.056	0.063	-0.106	-0.065
		(0.003)	(0.003)	(0.047)	(0.028)
Post-birth stay (days) (discharge-birth)	2.105	0.142	1.287	0.085	0.052
		(0.007)	(0.006)	(0.073)	(0.045)

Notes: Sample=491,604 first births, except models for 5 minute Apgar, which includes 487,643 observations, and models for length of stay, length of labor and length of post-birth stay, which have 482,187 observations. Length of labor is measured by number of days from mother's admission to birth, top-coded at maximum of 3 days. Mother's length of stay is top-coded at 5 days. Post birth stay is length of stay minus length of labor. Standard errors in parentheses clustered at 5-digit ZIP code level. All models (OLS and IV) include the full set of controls described in note to Table 2, except second set of models for NICU, which include additional control for whether hospital of delivery has its own NICU. First stage coefficients with this additional control are 1.518 (0.148) and 0.173 (0.029). Instrumental variable in all cases is relative distance to high c-section hospital.

Table VI: Effects of Delivery at High C-Section Hospital on Subsequent Hospital Visits

	Mean	OLS Coefficients: Deliver at H Hospital	OLS Coefficients: C-section	RF Coefficient (x 100)	IV Estimate (Scaled per delivery at H-hospital)
First-stage coefficients	--	1.600 (0.153)	0.183 (0.030)	--	--
Infant Outcomes:					
Any inpatient stay or ED/ASC visit	0.385	0.005 (0.004)	0.001 (0.002)	0.088 (0.052)	0.055 (0.032)
Any ED visit	0.338	0.008 (0.004)	0.006 (0.002)	0.128 (0.050)	0.080 (0.031)
ED visit for acute respiratory condition	0.126	0.010 (0.002)	0.005 (0.001)	0.074 (0.028)	0.046 (0.018)
Inpatient stay in neonatal period	0.041	0.001 (0.001)	-0.009 (0.001)	-0.038 (0.011)	-0.024 (0.007)
Inpatient stay in neonatal period, excluding immediate transfers	0.035	-0.003 (0.001)	-0.015 (0.001)	-0.050 (0.011)	-0.031 (0.007)
Inpatient stay in year after birth	0.085	0.002 (0.001)	-0.005 (0.001)	-0.038 (0.019)	-0.024 (0.012)
Maternal Outcomes:					
Any inpatient stay or ED/ASC visit	0.149	0.002 (0.001)	0.030 (0.001)	0.025 (0.025)	0.016 (0.015)
Any ED visit	0.129	0.001 (0.001)	0.026 (0.001)	0.025 (0.025)	0.015 (0.015)
Inpatient stay in year after birth	0.022	0.000 (0.000)	0.009 (0.001)	0.006 (0.007)	0.004 (0.005)

Notes: Sample is 491,604 low-risk first births. All models (OLS and IV) include the full set of controls described in note to Table 2. Instrumental variable in all cases is relative distance to high-c-section hospital.

Table VII: Effects of Delivery at High C-Section Hospital on ED Visits Controlling for Delivery at High-ED Use Hospital

	First Stage Models				Models for any ED visit				Models for Acute Respiratory ED visit			
	Delivery at High C-Section Hospital		Delivery at High ED Use Hospital		RF Coefficients (x100)		IV Estimates (1- and 2-Channel Models)		RF Coefficients (x100)		IV Estimates (1- and 2-Channel Models)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
First-stage & Reduced Form Coefficients												
Relative distance to high c-section hospital	1.600	1.591	--	0.034	0.128	0.133	--	--	0.074	0.077	--	--
	(0.153)	(0.157)		(0.107)	(0.050)	(0.049)			(0.028)	(0.028)		
Relative distance to high infant ED use hospital	--	-0.217	1.313	1.314	--	0.115	--	--	--	0.087	--	--
		(0.130)	(0.095)	(0.095)		(0.049)				(0.027)		
IV Coefficients												
Deliver at high c-section hospital	--	--	--	--	--	--	0.080	0.081	--	--	0.046	0.047
							(0.031)	(0.031)			(0.018)	(0.017)
Deliver at high infant ED use hospital	--	--	--	--	--	--	--	0.101	--	--	--	0.074
								(0.038)				(0.021)

Notes: Sample=491,604 first births. All models (OLS and IV) include the full set of controls described in note to Table 2. Instrumental variables are relative distance to high c-section hospital and relative distance to high infant ED use hospital.

Table VIII: Effects of Delivery at High C-Section Hospital on Inpatient Admissions Controlling for Delivery at High-Inpatient Use Hospital

	First Stage Models				Models for inpatient visit in neonatal period				Models for inpatient visit in first year			
	Delivery at High C-section Hospital		Delivery at High Infant-Inpatient Use Hospital		RF Coefficients (x100)		IV Estimates (1- and 2-Channel Models)		RF Coefficients (x100)		IV Estimates (1- and 2-Channel Models)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
First-stage & Reduced Form Coefficients												
Relative distance to high c-section hospital	1.600	1.613	--	-0.064	-0.038	-0.033	--	--	-0.038	-0.031	--	--
	(0.153)	(0.154)		(0.120)	(0.011)	(0.011)			(0.019)	(0.019)		
Relative distance to high infant-inpatient use hospital	--	0.138	2.112	2.104	--	0.048	--	--	--	0.076	--	--
		(0.163)	(0.138)	(0.139)		(0.012)				(0.020)		
IV Coefficients												
Deliver at high C-section hospital	--	--	--	--	--	--	-0.024	-0.020	--	--	-0.024	-0.017
							(0.007)	(0.007)			(0.012)	(0.012)
Deliver at high infant-inpatient use hospital	--	--	--	--	--	--	--	0.024	--	--	--	0.037
								(0.006)				(0.009)

Notes: Sample=491,604 first births. All models (OLS and IV) include the full set of controls described in note to Table 2. Instrumental variables are relative distance to high c-section hospital and relative distance to high infant-inpatient use hospital.

Table IX: Effects of Delivery at High C-Section Hospital on Adverse Outcomes

Outcome Variable	Mean	OLS Coefficients: Deliver at H Hospital	OLS Coefficients: C-Section	RF Coefficient (x 100)	IV Estimate (Scaled per delivery at H hospital)
First-stage Coefficients	--	1.613	0.178	--	--
		(0.154)	(0.029)		
6+ days in hospital or death in neonatal period (x100)	6.746	-0.464	4.387	-4.226	-2.530
		(0.114)	(0.109)	(1.434)	(0.877)
Death within 1st year (x100)	0.121	-0.006	0.084	-0.381	-0.239
		(0.013)	(0.014)	(0.160)	(0.098)
Late delivery and 6+ days in hospital/death as neonate (x100)	3.431	-0.655	2.310	-2.635	-1.596
		(0.074)	(0.083)	(0.967)	(0.573)
Late delivery and death within 1st year (x100)	0.058	-0.014	0.017	-0.261	-0.162
		(0.008)	(0.009)	(0.123)	(0.078)

Notes: Overall sample=491,604 first births, with probabilities of delivery at high-c-section hospital and c-section of 0.515 and 0.255, respectively. Sample of low-risk births has 327,736 observations with probabilities of delivery at high-c-section hospital and c-section of 0.515 and 0.282, respectively. Sample of high-risk births has 163,868 observations with probabilities of delivery at high-c-section hospital and c-section of 0.516 and 0.203, respectively. Late delivery and death is a combined outcome of delivery 1 or more days after mother's admission to hospital×death. All models include the full set of controls described in note to Table 2, and separately control for delivery at a high-inpatient hospital (for the OLS coefficients), or control for relative distance to a high-inpatient hospital (for the reduced form models), or instrument for delivery at a high-inpatient hospital using relative distance. Instrumental variable in all cases is relative distance to high-c-section hospital.

Table X: Generalized Control Function Models for Apgar Scores, Adverse Event in Neonatal Period, and Death

	Five-Minute Apgar Score			More than 6 days in Hospital or Death in Neonatal Period			Death in First Year (×100)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Delivered at H Hospital	0.054	0.053	0.053	-0.026	-0.025	-0.025	-0.239	-0.232	-0.233
	(0.022)	(0.023)	(0.020)	(0.011)	(0.011)	(0.010)	(0.117)	(0.117)	(0.114)
Delivered at H Hospital	--	-0.023	-0.024	--	0.022	0.023	--	0.251	0.246
× 1st stage residual		(0.010)	(0.010)		(0.005)	(0.006)		(0.064)	(0.065)
Delivered at H Hospital	--	--	0.005	--	--	0.000	--	--	0.014
× birthweight (standardized)			(0.002)			(0.001)			(0.015)
Delivered at H Hospital	--	--	-0.001	--	--	-0.001	--	--	-0.003
× gestation (standardized)			(0.002)			(0.001)			(0.012)
Delivered at H Hospital	--	--	0.001	--	--	-0.002	--	--	0.009
× HRR c-section rate (standardized)			(0.003)			(0.001)			(0.014)
1st stage residual	-0.029	-0.018	-0.017	0.235	0.109	0.110	0.235	0.109	0.112
	(0.022)	(0.024)	(0.021)	(0.119)	(0.125)	(0.119)	(0.119)	(0.125)	(0.120)

Notes: see note to Table 9. All models include the full set of controls described in note to Table 2 as well as control for delivery at high inpatient readmission hospital and residual from first stage model for delivery at high inpatient readmission hospital. Birthweight and gestation interaction terms are expressed in (demeaned) standard deviation units. Sample for Apgar scores is 487,643 births with non-missing 5-minute Apgar scores. Standard errors are bootstrapped (500 repetitions) and clustered at the mother's zip code.

Table XI: Approximate Social Costs and Benefits of Delivering at High C-Section Hospital

Parameter description	Parameter value	Source
Extra c-sections per H delivery	0.114	Table 3
Extra ED visits per H delivery	0.080	Table 6
Fraction of extra visits for respiratory disease	0.575	Table 6
Fewer inpatient stays per H delivery	0.024	Table 6
Fewer deaths per H delivery	0.00236	Table 9
More births to each LRFB mother	1.1	Johnson and Li (2007)
Path dependence of delivery mode: excess rate of c-section in 2nd birth if c-section in 1st	1	Unreported analysis
Average cost of delivery, by mode and complications		
Vaginal delivery without complication	$3,400	Podulka et al. (2011)
Vaginal delivery with complication	$4,400	Podulka et al. (2011)
Cesarean delivery without complication	$5,700	Podulka et al. (2011)
Cesarean delivery with complication	$7,600	Podulka et al. (2011)
Average cost of ED visit	$2,168	Caldwell et al. (2013)
Average cost of inpatient stay	$10,800	Wier et al. (2011)
Annual medical costs of childhood asthma	$1,039	CDC (2009)
Present discounted cost of childhood asthma (3% discount; no remission to age 80)	$31,379	
Value of a statisical life	$5,800,000	US DOT (2008)

Summary: costs and benefits of reallocating 1000 mothers to high CS hospitals

	Assumption on Infant Morality Effect	
	Use point estimate of reduced mortality	Use 0.5 × point estimate of reduced mortality
Benefits: reduced infant mortality, inpatient stays	$13,947,200	$7,103,200
Costs: direct medical costs, future ED visits, asthma burden, and higher-parity health effects	$6,072,000	$6,072,000
Benefits net of costs	$7,875,200	$1,031,200

Note: see text.

Appendix A: An Overview of the Literature on the Health Effects of Cesarean Delivery

(i) Infant Outcomes

Table A-I summarizes a selection of recent studies on the short and medium-run health effects of cesarean delivery for infants. We review studies on injury or death of the baby; lung function and respiratory problems; asthma; immune system; and breastfeeding. Not included in the table are several other active areas of research that study impacts of cesarean delivery on longer-term outcomes such as the probability of adult obesity (see the recent review by Darmasseelane et al., 2014).

Across the board a general finding is that babies delivered by c-section fare worse: higher neonatal and post-neonatal death; elevated risks of respiratory system problems including asthma; evidence of digestive system disorders, and lower rates of breastfeeding. An unusually detailed prospective study by Villar et al. (2007) of births in eight Latin American countries illustrates the general nature of these findings and the difficulty in interpreting the results as causal.[1] The authors show that neonatal death rates for cephalic fetuses delivered by c-section after trial of labor are substantially higher than rates for those delivered vaginally (0.65% versus 0.38%). Eliminating the roughly 30% of intrapartum c-sections performed after indications of fetal distress, the neonatal death rate of the remaining c-section group falls to 0.51% -- not statistically different from the rate for the vaginal births (but still higher), and indicative of a potentially large endogeneity bias in the overall comparison.

Our reading of the literature is that the most widely documented correlation is between c-section delivery and respiratory problems. Such a pattern has been documented in large-scale cohort studies in several Nordic countries (e.g., Hansen et al., 2008; Tollanes et al., 2008) and in

[1] This study is unusual in collecting detailed data on reasons for c-section, gathered immediately after the birth by trained survey staff.

meta analyses of the literature (e.g., Thavagnanam et al., 2008). As discussed in a recent review by Hyde et al. (2012), there is clinical evidence that babies born by c-section have worse lung function immediately after birth -- possibly attributable to a therapeutic effect of the labor process (including release of hormones and clearance of lung liquid). A number of researchers also hypothesize that there is a transfer of microbes from mother to infant during labor that aid in the development of the immune and digestive systems (e.g., Neu and Rushing, 2012).

(ii) Maternal Outcomes

Table A-II presents a parallel summary of the literature on the health effects of cesarean delivery on mothers. Here the literature is less numerous: our reading is that the major health risks include complications at birth and maternal death; reduction in future fertility; abnormal placentation in subsequent pregnancies; and risk of future stillbirths. Most studies find that mothers who deliver by c-section have higher risk of birth-related complications (such as need of a blood transfusion), higher risk of severe morbidity and mortality in the period after the birth, reduced future fertility, higher risk for placenta previa (placenta near or covering the cervix) and placenta accreta/increta/percreta (abnormal placental attachment). Evidence on future stillbirths is less clear.

As with the literature on infant health effects, most of these studies are based on observational designs, making it difficult or impossible to assert causality, though some of the potential effects are grounded in clinic evidence (see for example the review of studies on abnormal placentation by Clark and Silver, 2011). An interesting exception is the study by Halla et al. (2019) on future fertility, which uses day of the week of the birth as an instrument for c-section. We find that there appear to be more pre-scheduled c-sections on weekdays, leading to concerns over this instrument in our setting.

Table A-I: Summary of Literature on Infant Health Effects of C-Section Delivery

Health Issue	Study authors; design; main findings
1. Delivery injuries and death	a. Rouse and Owen (1999): prophylactic CS for large fetuses (>4000g) has small impact on permanent brachial plexus injury
	b. Alexander et al. (2006): 1.1% of CS babies have some birth injury - mostly cuts from the incision
	c. Villar et al. (2007): CS might decrease death for cephalic pregnancies, definitely for breech; increased NICU, but rupturing of membranes may be protective
	d. MacDorman et al (2008): CS has 1.7-2.4 higher risk of infant neonatal mortality for primary, low-risk births. Intention to treat analysis combines CS after TOL with vaginal births as intended vaginal
	e. Molina et al. (2015): cross-national analysis of CS and infant morality; neonatal mortality rates decline until CS rate of 20%, then stable across countries
2. Lung Function and Repiratory Problems	a. Hansen et al. (2008): Danish cohort study (cov-adj); scheduled CS increases risk of respiratory illness 200-400%
	b. Moore et al. (2012): Australian register study (cov-adj); elective CS increases risk of hospitalization for bronchiolitis by 10% in first year of life
	c. Hyde et al. (2012): review of clinical literature; CS without TOL associated with reduced lung function after birth
	d. Kristensen and Hendriksen (2016): Danish register study (cov-adj); elective CS associated with 20% higher risk of pneumonia and other mucosal system disorders
3. Asthma	e. Salam et al (2006): retrospective study of California youth; CS raises incidence of allergy by 26% (cov-adj)
	b. Roduit et al. (2008): Dutch cohort study (cov-adj). CS associated with 20% increase in risk of childhood asthma, higher effect for allergic parents
	c. Thavagnanam et al. (2008): meta analysis of 23 studies of CS and asthma; CS associated with 45% increase in risk at age 8
	d. Tollanes et al. (2008): Norwegian register study (cov-adj); CS raises risk of asthma by age 18 by 50%
	e. Jachetta (2014): IV study using MSA-level malpractice premiums instrument; CS associated with higher rate of hospitalization for asthma and lung disease
4. Immune System	a. Neu and Rushing (2011): review of clinical literature; CS without TOL affects micobial colonization/immune response
	b. Sevelsted et al. (2016): Danish register study (cov-adj); CS associated with higher risk of immune deficiency, inflammatory bowel disorders
	c. Stokholm et al. (2016): prospective study of Copenhagen births; CS associated with different gut microbes in first year
5. Breastfeeding	Prior et al (2012): meta-analysis of 48 studies; CS without TOL associated with lower rate of early initiation of breastfeeding; CS after TOL same as vaginal births

Notes: CS = c-section delivery; OR = odds ratio; TOL=trial of labor; cov-adj = covariate adjustment; IV=instrumental variables

Table A-II: Summary of Literature on Maternal Health Effects of C-Section Delivery

Health Outcome	Study authors; design; main findings
1. Complications at birth; mortality	a. Lydon-Rochell et al. (2000): cohort of primiparous women in Washington State; 80% higher rate of rehospitalization in 60 days following CS
	b. Deneux-Tharaux et al. (2006): 3.5 times more likely for mom to die in CS
	c. Villar et al (2007): WHO-supported study of Latin American births; incidence of mother injury/death increases in CS
	d. Kuklina et al (2009): rise in CS explains rise in maternal morbidity at birth
	e. Curtin et al. (2015): US births in 2013; (no cov-adj); higher rates of tranfusion, ICU admission
	f. Molina et al. (2015): cross-national analysis of CS and maternal morality; mortality rates decline until CS rate of 20%, then stable across countries
2. Fertility	a. Hall et al. (1989): U.K. cohort study (cov-adj); 23% lower fertility
	b. Kjerulff et al. (2013): U.S. cohort study (covariate adustment); 16% lower fertility
	c. Gurol-Urganci et al. (2013): meta analysis of 18 cohort studies; mean effect = 9% reduction in fertility following CS
	d. Halla et al. (2018): IV based on day of delivery; lower fertility
3. Abnormal Placentation (previa, accreta, etc.)	a. Hemminki et al. (2005): Finish register (cov-adj); 90% higher risk
	b. Getahun et al. (2006): U.S. linked cohorts (cov-adj); 30-100% higher risks
	c. Gurol-Urganci et al. (2011): U.K. cohort study and meta analysis of 37 studies; CS at first birth raises risk of placenta previa in second by 50-60%
	d. Clark and Silver (2011): review of previous studies; increased risks
4. Future Stillbirth	Bahtiyar et al. (2006): large U.S. cross-section study (cov-adj); no effect

Note: see Table A-I

Additional References for Appendix A

Bahtiyar, M. O., et al. (2006). Prior cesarean delivery is not associated with an increased risk of stillbirth in a subsequent pregnancy: Analysis of US perinatal mortality data, 1995-1997. *American Journal of Obstetrics and Gynecology*, 195(5), 1373-1378.

Curtin, S. C., Gregory, K. D., Korst, L. M., and Uddin, S. F. (2015). Maternal Morbidity for Vaginal and Cesarean Deliveries, According to Previous Cesarean History: New Data from the Birth Certificate, 2013. National Vital Statistics Reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System, 64(4), 1.

Darmasseelane, K., et al. (2014). Mode of Delivery and Offspring Body Mass Index, Overweight and Obesity in Adult Life: A Systematic Review and Meta-Analysis. *PLoS ONE*, 9(2).

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

Getahun, D., Oyelese, Y., Salihu, H. M., and Ananth, C. V. (2006). Previous cesarean delivery and risks of placenta previa and placental abruption. *Obstetrics and Gynecology*, 107(4), 771-778.

Gurol-Urganci, I., et al. (2013). Impact of Caesarean section on subsequent fertility: a systematic review and meta-analysis. *Human Reproduction*, 28(7), 1943-1952.

Hall, M. H., Campbell, D. M., Fraser, C., and Lemon, J. (1989). Mode of delivery and future fertility. *BJOG: An International Journal of Obstetrics and Gynaecology*, 96(11), 1297-1303.

Hansen, A. K., Wisborg, K., Uldbjerg, N., and Henriksen, T. B. (2008). Risk of respiratory morbidity in term infants delivered by elective caesarean section: cohort study. *British Medical Journal*, 336(7635), 85.

Hemminki, E., Shelley, J., and Gissler, M. (2005). Mode of delivery and problems in subsequent births: A register-based study from Finland. *American Journal of Obstetrics and Gynecology*, 193(1), 169-177.

Kjerulff, K. H., Zhu, J., Weisman, C. S., and Ananth, C. V. (2013). First birth Caesarean section and subsequent fertility: a population-based study in the USA, 2000–2008. *Human Reproduction*, 28(12), 3349-3357.

Kristensen, K., and Henriksen, L. (2016). Cesarean section and disease associated with immune function. *Journal of Allergy and Clinical Immunology*, 137(2), 587-590.

Kuklina, E. V., et al. (2009). Severe Obstetric Morbidity in the United States: 1998–2005. *Obstetrics and Gynecology*, 113(2), 293-299.

Lydon-Rochelle, M., Holt, V. L., Martin, D. P., and Easterling, T. R. (2000). Association between method of delivery and maternal rehospitalization. *JAMA*, 283(18), 2411-2416.

Molina, G., et al. (2015). Relationship Between Cesarean Delivery Rate and Maternal and Neonatal Mortality. *JAMA*, *314*(21), 2263-2270.

Moore, H. C., de Klerk, N., Holt, P., Richmond, P. C., and Lehmann, D. (2012). Hospitalisation for bronchiolitis in infants is more common after elective caesarean delivery. *Archives of Disease in Childhood*, 97(5), 410.

Neu, J., and Rushing, J. (2011). Cesarean versus vaginal delivery: Long-term infant outcomes and the hygiene hypothesis. *Clinics in Perinatology*, 38(2), 321-331.

Prior, E., et al. (2012). Breastfeeding after cesarean delivery: A systematic review and meta-analysis of world literature. *The American Journal of Clinical Nutrition*, 95(5), 1113-1135.

Roduit, C., et al. (2008). Asthma at 8 years of age in children born by cesarean section. *Thorax*, 64(2), 107-113.

Rouse, D. J., and Owen, J. (1999). Prophylactic cesarean delivery for fetal macrosomia diagnosed by means of ultrasonography—A Faustian bargain? *American Journal of Obstetrics and Gynecology*, 181(2), 332-338.

Salam, M. T., et al. (2006). Mode of delivery is associated with asthma and allergy occurrences in children. *Annals of Epidemiology*, 16(5), 341-346.

Sevelsted, A., Stokholm, J., and Bisgaard, H. (2016). Risk of Asthma from Cesarean Delivery Depends on Membrane Rupture. *The Journal of Pediatrics*, (171), 38-42.

Stokholm, J., et al. (2016). Cesarean section changes neonatal gut colonization. *Journal of Allergy and Clinical Immunology*, 138(3), 881-889.

Thavagnanam, S., Fleming, J., Bromley, A., Shields, M. D., and Cardwell, C. R. (2008). A meta-analysis of the association between Caesarean section and childhood asthma. *Clinical and Experimental Allergy*, 38(4), 629-633.

Tollånes, M. C., Moster, D., Daltveit, A. K., and Irgens, L. M. (2008). Cesarean section and risk of severe childhood asthma: A population-based cohort study. *The Journal of Pediatrics*, 1(153), 112-116.

Villar, J., et al. (2007). Maternal and neonatal individual risks and benefits associated with caesarean delivery: Multicentre prospective study. *British Medical Journal*, 335(7628), 1025-1029.

Appendix B: Interpretation of First Stage, Reduced Form and IV Estimates

Consider the case where individuals (indexed by i) belong to mututally exclusive subgroups. Let X_i represent a vector of indicators for membership in each of J subgroups, let y_i represent an outcome of interest, let D_i represent an endogenous treatment indicator, and let Z_i represent an instrumental variable.

Suppose we estimate a pooled first stage model for D_i that includes Z_i and the vector X_i:

$$D_i = \pi_0 + \pi_1 Z_i + \pi_X X_i + v_i.$$

By standard Frisch-Waugh arguments the OLS estimate of π_1 is:

$$\widehat{\pi}_1 = \frac{\sum_i (D_i - \overline{D}_{j(i)})(Z_i - \overline{Z}_{j(i)})}{\sum_i (Z_i - \overline{Z}_{j(i)})^2}$$

where $j(i)$ is $i's$ subgroup, and $\overline{D}_j$ and $\overline{Z}_j$ represent the means of D and Z within subgroup j. Let N represent the combined sample size and N_j the sample size for group j. Then

$$\begin{aligned}
\widehat{\pi}_1 &= \frac{\sum_j \sum_{i\in j} (D_i - \overline{D}_{j(i)})(Z_i - \overline{Z}_{j(i)})}{\sum_j \sum_{i\in j} (Z_i - \overline{Z}_{j(i)})^2} \\
&= \sum_j \left(\frac{N_j}{N}\right) \left(\frac{\frac{1}{N_j} \sum_{i\in j} (Z_i - \overline{Z}_j)^2}{\frac{1}{N} \sum_j \sum_{i\in j} (Z_i - \overline{Z}_{j(i)})^2} \right) \frac{\sum_{i\in j} (D_i - \overline{D}_{j(i)})(Z_i - \overline{Z}_{j(i)})}{\sum_{i\in j} (Z_i - \overline{Z}_{j(i)})^2} \\
&= \sum_j \left(\frac{N_j}{N}\right) \frac{V_{Zj}}{V_Z} \widehat{\pi}_{1j}
\end{aligned}$$

where V_{Zj} is the variance of Z within group j, V_Z is the overall variance of Z and $\widehat{\pi}_{1j}$ is the first stage regression coefficient for group j.

By the same argument if we estimate a pooled reduced form model for y_i that includes Z_i and the vector X_i:

$$y_i = \delta_0 + \delta_1 Z_i + \delta_X X_i + u_i.$$

the OLS estimate of δ_1 is

$$\widehat{\delta}_1 = \sum_j \left(\frac{N_j}{N}\right) \frac{V_{Zj}}{V_Z} \widehat{\delta}_{1j}$$

where $\widehat{\delta}_{1j}$ is the reduced form coefficient for group j. Finally, the pooled IV estimate of the effect of D on y using Z as an instrument and controlling for X is:

$$\begin{aligned}
\widehat{\beta}_1 &= \frac{\widehat{\delta}_1}{\widehat{\pi}_1} \\
&= \sum_j \left(\frac{N_j}{N}\right) \left(\frac{V_{Zj}}{V_Z}\right) \left(\frac{\widehat{\pi}_{1j}}{\widehat{\pi}_1}\right) \frac{\widehat{\delta}_{1j}}{\widehat{\pi}_{1j}} \\
&= \sum_j \left(\frac{N_j}{N}\right) \left(\frac{V_{Zj}}{V_Z}\right) \left(\frac{\widehat{\pi}_{1j}}{\widehat{\pi}_1}\right) \widehat{\beta}_{1j}
\end{aligned}$$

where $\widehat{\beta}_{1j} = \widehat{\delta}_{1j}/\widehat{\pi}_{1j}$ is the IV estimate within subgroup j.

Appendix C: Data

a. Overview of PDD/ED/AS/Linked Birth Cohort Data

California OSHPD has created a linked file that combines in-patient discharge records for delivering mothers and newborns with Vital Statistics (VS) data (i.e., information collected from birth certificates and death records) and information on in-patient, Emergency Department (ED), and Ambulatory Surgery Center (ASC) records for each mother in the period from one year before to one year after the birth, and for each infant in the period up to one year after the birth. We use a version of this file that has information on live hospital delivered births for the period from 2007 to 2011.

Appendix D of the data base gives the name, address, zip code, and Hospital Service Areas (HSA) for each hospital, ED, and ASC in the state. We also use external information from the Dartmouth Atlas website to assign HSA's and Health Referral Regions (HRR's). We add data from the US Census Bureau on average income in each zip code.

b. Construction of relative distance instruments

The procedure for constructing a mother's relative distance to high and low c-section hospitalscon-sists of 3 steps:

1. We estimate each hospital's risk-adjusted c-section rate among low-risk first births;
2. We classify hospitals as low (L) or high (H) c-section hospitals based on their risk-adjusted c-section rates from (1);
3. We calculate each mother's distances to the nearest L and H hospitals, from which we calculate our main relative distance measure.

In step 1 we fit a logistic regression model to our sample of low-risk first births that includes a baseline set of case risk factors $X_i$and indicators for the hospital $h(i)$ at which mother i delivered.

Specifically, using our LRFB sample, we estimate the model:

$$P(C_i = 1|X_i) = \Lambda(\alpha + \mathbf{X}_i'\beta + \gamma_{h(i)})$$

where Λ is the logistic CDF.

In step 2 we compare hospital h's estimated logit coefficient $\hat{\gamma}_h$ to the birth-weighted average hospital coefficient in each Hospital Referral Region (HRR) $\bar{\gamma}_{HRR} = \left[\sum_{j \in HRR} N_j\right]^{-1} \sum_{j \in HRR} N_j \hat{\gamma}_j$ (where N_h is the number of low risk first births delivered at hospital h in our analysis sample). We define a hospital to be a "high c-section hospital" (or H hospital) if $\hat{\gamma}_h \geq \bar{\gamma}_{HRR}$ and otherwise a "low c-section hospital."

In step 3 we use information on the centroid of each mother's home zip code and on the centroids of the zip codes for each hospital to define the distance from each mother to each hospital. We then define the distance to the nearest H hospital and the nearest L hospital.

Appendix Table I:
Characteristics of High & Low C-Section Hospitals

	Hospital Type:	
	High CS (H)	Low CS (L)
C-section rate (LRFB):		
All	0.289	0.220
Scheduled	0.104	0.081
Unscheduled	0.186	0.139
Ownership:		
For profit	0.180	0.086
Private non-profit	0.746	0.723
Government	0.068	0.140
Academic	0.006	0.051
Other Characteristics:		
Has NICU	0.741	0.858
NICU admit rate	0.027	0.042
Volume (births/yr.)	3,695	3,635
Weekend admit rate	0.240	0.262

Notes: see text and Appendix C for procedure to define H and L hospitals. Characteristics are based on low-risk first births (LRFBs).

Appendix Table II: Effect of Delivery at High C-Section Hospital on Joint Distribution of Timing/Mode of Delivery and Apgar Score

	All	Vaginal Delivery		Cesarean Delivery	
	Modes/Timing	Early	Late	Early	Late
A. Distribution of birth outcomes for compliers at low c-section hospitals					
Low Apgar (<9)	0.079	0.024	0.033	0.007	0.016
High Apgar (≥9)	0.921	0.387	0.361	0.080	0.095
B. Distribution of birth outcomes for compliers at high c-section hospitals					
Low Apgar (<9)	0.058	0.020	0.022	0.008	0.009
High Apgar (≥9)	0.945	0.365	0.284	0.156	0.140
C. Difference in outcomes caused by delivery at high c-section hospital					
Low Apgar (<9)	-0.021	-0.003	-0.012	0.001	-0.007
	(0.010)	(0.005)	(0.007)	(0.003)	(0.004)
High Apgar (≥9)	0.021	-0.022	-0.077	0.075	0.045
	(0.010)	(0.024)	(0.021)	(0.012)	(0.012)

Notes: Table entries are estimated shares of births with outcome described by row and column heading. Estimated standard errors in parentheses. "Early" delivery is delivery on day of arrival of mother at hospital. "Late" delivery is delivery 1 or more days after arrival. Estimates in panel A are IV estimates of the mean outcomes of compliers who deliver at low c-section (L) hospitals. Estimates in panel B are IV estimates of the mean outcomes of compliers who deliver at high c-section (H) hospitals. Estimates in panel C are IV estimates of the effect of delivering at H hospital on outcome. All models include control variables described in note to Table 5. Instrumental variable in all cases is relative distance to H hospital.

Appendix Table III: Characteristics of Infants that Die in First Year

	Low Risk First Births:	
	All	Deaths
Mother's characteristics		
Mean age	25.6	24.2
At most high school education (%)	41.2	54.4
Mean weight (pounds)	137	138
Mother obese (%)	8.1	9.6
Race/eth: Hispanic (%)	44.2	45.5
Asian (%)	17.6	16.3
Nonhispanic white (%)	31.7	30.0
Nonhispanic black (%)	5.6	7.0
Birth risk factors and characteristics		
Mean number prenatal care visits	12.2	11.6
18 or more prenatal visits (%)	4.3	3.9
Mother ED visit prior to birth (%)	19.5	23.7
Mean birthweight (grams)	3348	3091
Low birth weight (%)	2.3	14.4
Abn. fetal heart rate/rhythm (%)	20.2	33.9
Delivery outcomes		
C-section delivery (%)	25.6	34.6
Scheduled c-section (%)	9.2	22.5
Delivered at H hospital (%)	51.5	50.2
Apgar score <=7 (%)	1.6	30.4
Vacuum/forceps induction (%)	12.0	12.2
Postpartum outcomes		
Infant transferred to NICU unit (%)	3.4	30.5
Infant re-admitted to ED (%)	33.8	29.0
Infant re-admitted as in-patient (%)	8.2	31.2
Mother readmitted (any type) (%)	14.9	26.8
Sample size	491,604	596

Notes: See notes to Table I. Abnormal fetal heart rate/rhythm indicated by presence of secondary diagnosis code of 659.71.