DISCUSSION PAPER SERIES

IZA

IZA DP No. 6879

Saving Lives at Birth: The Impact of Home Births on Infant Outcomes

N. Meltem Daysal
Mircea Trandafir
Reyn van Ewijk

September 2012

Saving Lives at Birth: The Impact of Home Births on Infant Outcomes

N. Meltem Daysal
Tilburg University
and IZA

Mircea Trandafir
Université de Sherbrooke

Reyn van Ewijk
University Medical Centre Mainz
and University of Mainz

Discussion Paper No. 6879
September 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 6879
September 2012

ABSTRACT

Saving Lives at Birth: The Impact of Home Births on Infant Outcomes*

Many developed countries have recently experienced sharp increases in home birth rates. This paper investigates the impact of home births on the health of low-risk newborns using data from the Netherlands, the only developed country where home births are widespread. To account for endogeneity in location of birth, we exploit the exogenous variation in distance from a mother's residence to the closest hospital. We find that giving birth in a hospital leads to substantial reductions in newborn mortality. We provide suggestive evidence that proximity to medical technologies may be an important channel contributing to these health gains.

JEL Classification: I11, I12, I18, J13

Keywords: medical technology, birth, home birth, mortality

Corresponding author:

N. Meltem Daysal
Tilburg University
Warandelaan 2
P.O. Box 90153
5000 LE Tilburg
The Netherlands
E-mail: meltem.daysal@uvt.nl

* John Cawley, Erzo Luttmer, Ellen Meara, Daniel Parent, Jon Skinner, and seminar participants at Cornell, Dartmouth, HEC Montreal, Mainz, Sherbrooke, and the 2012 European Congress of Epidemiology provided helpful comments and discussions. We thank Perinatale Registratie Nederland (PRN) for making the data available and PRN staff, in particular Chantal Hukkelhoven, for assistance with the data. Tjeerd van Campen and Iris van Dam provided able research assistance. The authors bear sole responsibility for the content of this paper.

1 Introduction

Over the last few decades, most developed countries experienced an upsurge in medical expenditures. For example, health care expenditures in the United States increased from 5 percent of GDP in 1960 to over 16 percent in 2009 (OECD, 2012). Consequently, policy-makers around the globe are seeking ways to reduce medical expenditures without harming health outcomes. It is widely accepted that changes in medical technologies are the main driver of medical cost growth (Newhouse, 1992). Partly due to these changes in medical technology, spending for the very young increased substantially faster than spending for the average individual: during the period 1960–1990, per capita spending on infants under 1 year old increased by 9.8 percent per year whereas annual spending on individuals aged 1 to 64 increased by only 4.7 percent (Cutler and Meara, 1998). Therefore, a question central to current policy debates is the possibility of shifting low-risk births from more costly to less costly childbirth technologies such as midwifery care and home births. In this paper, we investigate the impact of home births on the health (7-day and 28-day mortality and 5-minute Apgar score) of low-risk newborns using a unique confidential dataset covering the universe of births in the Netherlands for the period 2000–2008.

The Netherlands is an ideal setting to study this question for several reasons. First, it is the only developed country where home births are widespread: between 2000 and 2008, approximately 25 percent of births took place at home, leading to sample sizes large enough to examine causal effects on rare health outcomes such as perinatal mortality. Second, the Dutch system is unique in its division between the primary care provided by midwives and the secondary care provided by obstetricians.[1] Low-risk women (women without known medical risk factors) begin their care at the primary level under the supervision of a midwife. They are referred to an obstetrician only if complications occur or new risk factors arise during the pregnancy, and these referrals are based on a fixed set of rules. Otherwise, the entire obstetric care of low-risk women

[1] We provide more detail on the institutional background in 2.1.

is provided by midwives and these women can choose between a home or a hospital birth. In both cases, their delivery is supervised by a midwife and no doctor is present. On the other hand, high-risk women (who are cared for by an obstetrician) are always required to give birth in a hospital. This institutional setup allows us to separate provider-effects (obstetrician versus midwife) from place-of-birth effects (home versus hospital).[2] Finally, the Netherlands is a country where childbirth technologies are a major policy issue because the Dutch perinatal mortality rate is one of the highest in Europe (Mohangoo et al., 2008) and the contribution of home births to this is hotly debated.

Empirical identification of the impact of home births is difficult due to the endogenous choice of location of birth: even among observably low-risk mothers, those who are at a higher risk of having an unhealthy infant for reasons unobservable to the midwife and to the researcher may choose to give birth in a hospital. In order to account for non-random selection into a home birth, we use an instrumental variables approach that exploits the exogenous variation in distance from a mother's residence to the closest hospital with an obstetric ward.[3]

Using the sample of low-risk women, all of whom are under the care of a midwife at the onset delivery, we find that giving birth in a hospital leads to economically large reductions in perinatal mortality and has no effect on Apgar scores. Consistent with negative selection into hospitals, these instrumental variable results are in sharp contrast to simple regression estimates which suggest negative effects of hospital births on Apgar scores and no effects on mortality.

The validity of our instrumental variable results rests on three assumptions:

[2]The use of physician extenders through midwifery care is another important policy question when shifting low-risk births from more costly to less costly technologies (Fuchs, 1998; Miller, 2006), but this is a different policy question that is examined in Daysal et al. (2012).

[3]The instrumental variables framework where the indicator for being treated with a given technology is instrumented with the differential distance between a patient's residence and the closest health care provider with that technology was first introduced by McClellan et al. (1994) in their study of returns to intensive heart attack treatments. We discuss this and other studies relying on a similar identification strategy in section 2.2.

instrument relevance, excludability and monotonicity. The relevance assumption is satisfied as the first stage results show that the likelihood of a hospital birth is significantly negatively correlated with the distance instrument. While the excludability and monotonicity assumptions cannot be tested formally, we bring suggestive evidence on their plausibility in several ways.

First, we show that the majority of the observable characteristics are balanced across the distribution of distance and, whenever there are differences, the statistics indicate that riskier infants reside closer to hospitals. Combined with the first stage results, this negative selection by distance implies that our estimates likely represent lower bounds of the true effect. Second, we show that the reduced form relationship between perinatal mortality and the instrument is strong in our analysis sample and tends to grow stronger as more controls are added. Finally, we conduct a falsification test that is made possible by the specifics of the Dutch institutional setup and that provides an arguably more direct check of the excludability assumption. Since high-risk women are required to give birth in a hospital, there is no relationship between the instrument and the likelihood of a hospital birth in this sample. We show that there is no reduced form relationship between perinatal mortality and the instrument in this sample, lending support to the claim that the instrument impacts the outcomes only through changes in the endogenous regressor.

In order to shed some light on the validity of the monotonicity assumption, we first show that women who deliver in a hospital (or at home) have similar observable characteristics across the distribution of distance, and hence we may be less concerned about them responding in different ways to the instrument. Second, we show that our results are robust when the sample is split by (average per capita) car ownership, a factor that potentially impacts directly the choice of type of birth location.

In addition to the checks on the validity of the identification assumptions, we also show that our results are robust to alternative sample selections, alternative definitions of the instrument, different classifications of referrals during delivery, the inclusion of different control variables as well as to using non-linear models.

The lack of an impact on the 5-minute Apgar score suggests that the general health of low-risk babies born in a hospital is similar to those born at home shortly after birth. Hence, any mortality reductions from a hospital birth are likely due to the medical care provided after delivery. A hospital birth may reduce infant mortality through various channels, such as the availability of better facilities and equipment, better hygiene or the proximity to other medical services. While data limitations constrain our ability to investigate many potentially important channels, we are able to examine a specific technology: admission to a neonatal intensive care unit (NICU) within the first seven days of life. Our results indicate that giving birth in a hospital leads to substantial increases in the probability of a NICU admission. This may be because transfers to a NICU are more feasible when the newborn is already in a hospital or because of differences in the length and intensity of post-natal monitoring between hospitals and home. Since NICUs have been shown to significantly improve the health and survival of new-borns, this finding suggests that access to medical technologies may be an important channel in explaining the lower mortality among hospital births.

Our instrumental variable strategy identifies the local average treatment effect for the subpopulation of low-risk women who give birth in a hospital because they reside close enough to it, but would give birth at home if they lived farther away. We show that compliers are more likely to be Dutch and younger than the median age of 29 and that their pregnancies are more likely to be within the normal range of gestational age, characteristics that are not associated with higher risk. At the same time, our results are entirely driven by women residing in below-median income postal codes. This finding is consistent with the previous literature documenting disparities in preventive behavior and quality of care by income and education (e.g., Smith, 1999; Cutler and Lleras-Muney, 2010). It also suggests that there can be substantial benefits from a hospital birth, at least for certain sub-populations, even in a health care system specifically geared toward risk selection and home births.

This paper makes several contributions to the growing literature in economics evaluating returns to medical technologies. First, as we summarize in section 2.2, the previous literature almost exclusively focuses on returns to medical technologies for high-risk individuals. In particular, the majority of this research examines heart attack technologies and the handful of papers on childbirth technologies generally investigate the returns to treating at-risk newborns. Our paper studies how medical technologies impact the health of low-risk newborns. Second, previous research on childbirth technologies examine exclusively the returns at the intensive margin (i.e., gains from incremental treatments). To our knowledge, our study is the first to analyze returns at the extensive margin, comparing two alternative technologies.

Although there are no economic studies on the safety of home births, there is a large medical literature on this topic. As we detail in section 2.2, these studies exclusively rely on simple regression models comparing outcomes among subsamples of low-risk women who (plan to) give birth at home or in the hospital, after controlling for observable differences in pregnant women. The major drawback of these studies is a potential selection bias due to the endogeneity in (planned) location of birth. In addition, the power of most of these studies is limited due to their small sample size.

Our results pertain directly to current policy debates on the health and safety of home births. After decades of continuous decline, many developed countries are now experiencing sharp increases in home birth rates.[4] For example, home births in the United States increased by almost 30 percent between 2004 and 2009 (MacDorman et al., 2012). Similarly, the fraction of home births in the United Kingdom almost tripled between 1990 and 2006 (Nove et al., 2008) and out-of-hospital births in Canada more than quadrupled between 1991 and 2009.[5] These trends are accompanied by policies and changes in the legal environment that bring the issue of home births to the forefront. For instance, several Canadian provinces legalized home births as

[4] Proponents of home births generally argue that labor is a natural process inherently safe in a majority of cases: medical interference in the delivery process might increase risk, e.g., because cesarean sections are more likely to be performed and this increases the risk of complications in subsequent pregnancies. Furthermore, they stress the importance of delivering in a comfortable, familiar environment (Borquez and Wiegers, 2006; Boucher et al., 2009; Janssen and Henderson, 2009; Henderson and Petrou, 2008; O'Connor, 1993).

[5] Authors' calculation using data from Statistics Canada, CANSIM Table 1024516.

early as 1994 (Janssen et al., 2002). In the United Kingdom, the Department of Health now asserts that home births are safe for women who have been properly assessed for risks and explicitly states that "[f]or the majority of women, pregnancy and childbirth are normal life events requiring minimal medical intervention. These women may choose to have midwifery-led care, including a home birth." (Department of Health, 2004, p. 6) In a joint statement, Royal College of Obstetricians and Gynaecologists and Royal College of Midwives (2007, p. 1) declare that "[t]here is ample evidence showing that laboring at home increases a woman's likelihood of a birth that is both satisfying and safe, with implications for her health and that of her baby." In the Unites States, The American College of Obstetricians and Gynecologists (2011, p. 1) notes that "[a]lthough the Committee on Obstetric Practice believes that hospitals and birthing centers are the safest setting for birth, it respects the right of a woman to make a medically informed decision about delivery" and a special Home Birth Consensus Summit was held in Virginia as recently as October 2011. Under these circumstances, the issue of home births is likely to be increasingly prominent in policy debates in the coming years.

The remainder of the paper is organized as follows. Section 2 presents a background on the Dutch obstetric care system and briefly summarizes the relevant literature. Section 3 outlines the empirical framework, while section 4 introduces the data and provides descriptive statistics. The results are presented in section 5 along with a discussion on heterogeneous effects, local average treatment effect, robustness checks and mechanisms. Finally, section 6 concludes.

2 Background

2.1 The Dutch obstetric system

The current Dutch maternity care system has its origins in the 1950s (Amelink-Verburg and Buitendijk, 2010). In an effort to cut healthcare expenditures, the Dutch National Health Insurance Board issued in 1958 a list of conditions that were deemed necessary for a hospital admission during childbirth. This list set the foundation for risk selection, the principle that uncomplicated births should stay in the primary care provided by a midwife or a general practitioner, and that hospital admissions into the secondary care provided by an obstetrician are necessary only in case of deviations from the normal course of a pregnancy. This list was updated in 1973 and became the official "List of Obstetric Indications" (LOI), which determines when referrals are made from primary to secondary care.

Subsequent updates to the LOI kept the same underlying idea: that pregnancy, delivery and puerperium are all natural processes. As a result, women are referred to an obstetrician only in specific cases. The LOI lists four main types of reasons for referral: non-gynecological pre-existing conditions, ranging from asthma, diabetes, hypertension and epilepsy to alcoholism and psychiatric disorders; gynecological pre-existing conditions (e.g., pelvic floor reconstructions); obstetric anamnesis, including items such as a C-section or complications in a previous delivery, previous preterm births or multiple miscarriages; conditions arising or first diagnosed during pregnancy, such as infections, hyperemesis gravidarum, plurality, gestational hypertension, blood loss and (threat of) preterm or postterm birth, defined as before 37 and after 42 completed gestation weeks, respectively (CVZ, 2003). It is sufficient to have only one of these reasons for referral (i.e., there is no continuous risk scale). Referrals for reasons other than those detailed in the LOI are not allowed and insurance plans do not cover doctor fees in these instances (CVZ, 2003). In addition, women are not allowed to contact directly an obstetrician. Between 2000 and 2008, about 47 percent of all pregnant women were deemed to have an increased risk and were referred to an obstetrician before the start of delivery. These high-risk women give birth in a hospital under the supervision of an obstetrician.

As long as there are no complications, women are not seen as patients and midwives supervise their entire pregnancy, perform all checks, and attend the birth (Bais and Pel, 2006). These low-risk women can choose the midwifery practice that cares for them as well as whether to deliver at home or in a hospital, but even hospital births are supervised by midwives in a polyclinic setting with no obstetrician present.[6] At the onset of labor (when contractions occur with a certain frequency or there is loss of amniotic fluid), the woman contacts her midwife, who then either comes to the woman's home for a home birth or notifies the hospital that they will be arriving for a hospital birth. Thus, women choosing a hospital birth will have to be transported to the hospital during the contraction phase and they have to arrange their own transportation.[7] If complications arise during delivery, the delivery takes too long, or the need for pain medication arises, the midwife refers the woman to an obstetrician. This can be a within-hospital transfer, if the woman was already there, or it could entail transport from home to the hospital in the case of a home birth. Around 31 percent of all women who started delivering with a midwife between 2000 and 2008 were referred to an obstetrician during delivery and about 12 percent of referrals were due to the need for pain relief medication.

Following a low-risk (uncomplicated) hospital birth, the woman is generally discharged a few hours after delivery, irrespective of the time of the day. Postnatal care for both home births and hospital births is ensured by a system in which trained health workers intensively take care of the woman and child during home visits totaling three to eight hours per day (depending on personal and health circumstances) over a period of eight to ten days. This care includes prevention, instruction, detection of any (health) problems, ensuring good hygiene, verification that the child is properly cared for, and often even household chores.

It should be mentioned that midwives have no financial incentive to in-

[6] There are very few exceptions when a low-risk woman is not allowed to choose her place of (midwife-supervised) delivery. For example, she is not allowed to deliver at home if she cannot deliver on the ground floor and her floor can only be reached by a steep or narrow staircase, since labor laws would not allow ambulance personnel to carry her down.

[7] Moreover, women cannot go to the hospital until their midwife agrees to it. According to the Royal Dutch Organisation of Midwives (KNOV), "hopital deliveries start at home as well. You will consult with the midwife about the moment you will go to hospital. Usually this is when contractions are well underway. The midwife will join you at the hospital." (www.knov.nl/voor-zwangeren/zwanger/de-bevalling/thuis-of-in-het-ziekenhuis/, authors' translation, accessed on August 31, 2012).

fluence a woman's choice of delivery location. Midwives supervise low-risk deliveries regardless of where they take place and their salary is fixed irrespective of the number of births supervised. Their midwifery practice receives a fixed amount per delivery, which as of 2008 was €333.50 per birth (NZA, 2008). On the other hand, there are differences between home and hospital births in terms of the out-of-pocket cost for the mother. The default types of delivery, at home for midwife-supervised low-risk births and in a hospital for obstetrician-supervised high-risk births, are fully covered by universal healthcare insurance. Hospitals charge an additional fee for low-risk deliveries in their polyclinics and for the use of their facilities. As of 2008, this fee was €468.50 and it is only partially covered by universal health insurance and by supplementary health insurance, if any (NZA, 2011a; Latta et al., 2011). In conclusion, the Dutch obstetric care system is designed around risk selection and encourages the use of home births for low-risk deliveries.

2.2 Previous Literature

This paper is at the intersection of three strands of research. The first strand includes the economic studies of returns to medical technologies, the majority of which examine treatments for heart attack patients (e.g., McClellan and Newhouse, 1997; Cutler et al., 1998; McClellan and Noguchi, 1998; Skinner et al., 2006). The handful of papers analyzing the returns to childbirth technologies focus almost exclusively on at-risk newborns, particularly those with low and very low birth weight (Cutler and Meara, 2000; Almond et al., 2010; Bharadwaj et al., forthcoming; Freedman, 2012). One notable exception is the study by Almond and Doyle (2011) on the health benefits of longer hospitalizations for newborns following uncomplicated births. The common theme of this entire literature is that it estimates the returns to *incremental* technologies: more intensive heart attack treatments, additional procedures for low birth weight infants, additional days of hospitalization, higher level of neonatal intensive care, etc. In contrast, our paper examines the return to *alternative* childbirth technologies: hospital instead of home birth.

The second related line of literature examines the benefits of a hospital as compared to a home birth. The research comes entirely from medical studies using observational data as it proved impossible to randomize birth location (Dowswell et al., 1996; Hendrix et al., 2009). These studies generally compare average outcomes between samples of (low-risk) women *planning* to give birth at different types of location after controlling for observable characteristics. The use of planned rather than actual place of delivery is justified by the assumption that there is less endogeneity in planned than in actual birth place, since the actual birth place may deviate from the planned one due to changes in individual health and risk factors. The results, interpreted as an intention-to-treat effect, are mixed, with some studies showing higher perinatal mortality risk among home births (e.g., Bastian et al., 1998; Pang et al., 2002; Kennare et al., 2010; Malloy, 2010; Birthplace in England Collaborative Group, 2011) and others finding no significant differences (e.g., Ackermann-Liebrich et al., 1996; Murphy and Fullerton, 1998; Janssen et al., 2002; Lindgren et al., 2008; de Jonge et al., 2009; van der Kooy et al., 2011).[8] However, as the medical literature acknowledges, planned birth place may be endogenous (Wiegers et al., 1998; Gyte et al., 2009). In addition, the small sample sizes in several of these studies pose statistical power problems. Unlike these studies, our paper analyzes the returns to *actual* (rather than planned) hospital birth using a large sample of low-risk births. We also explicitly correct for the endogeneity of birth location using distance to the nearest obstetric ward as an instrument.

Several other studies use distance as an instrument and they form the third line of research related to our paper. Generally, these studies examine the return to a more intensive procedure while accounting for the endogeneity of access to this procedure. The particular instrument used is the relative distance between the closest provider of this procedure and the closest provider of a less intensive treatment (see, for example, McClellan et al., 1994 and Cutler, 2007 in the case of heart attacks, or Freedman, 2012 in the case of at-risk newborns). The two technologies compared in this paper are home and

[8]The studies by de Jonge et al. (2009) and van der Kooy et al. (2011) use the same data as this paper and find no significant differences between home and hospital births.

hospital births. Therefore, our instrument, the distance between a woman's residence and the nearest hospital where she can give birth, also represents a relative distance. We discuss our instrument in more detail in section 4.

3 Empirical Strategy

We are interested in estimating the impact of type of delivery place (home versus hospital) on infant health outcomes. The structural equation of interest can be described as follows:

$$Y_{izt} = \beta_0 + \beta_1 Hospital_{izt} + \beta_2 X_{izt} + \epsilon_{izt} \tag{1}$$

where the unit of observation is infant i who is born in year t to a mother residing in postal code z. Y_{izt} is an outcome variable capturing infant health, $Hospital_{izt}$ is a dummy variable indicating that the birth occurred in a hospital, X_{izt} is a set of control variables representing observable characteristics of the mother and of the infant, and ϵ_{izt} is an idiosyncratic error term. We provide detailed information on each of these variables in section 4, after describing the data sources.

The coefficient of interest in the structural equation, β_1, measures the average difference in the health outcomes of infants born in a hospital as compared to those born at home, after controlling for observed characteristics of the mother and the infant. The primary challenge in interpreting the ordinary least squares (OLS) estimates of β_1 as causal stems from the endogenous choice of location of birth: mothers who are at a higher risk of having an unhealthy infant (for reasons that are unobservable to the researcher) may choose to give birth in a hospital, leading to a biased estimate of β_1.

To address this endogeneity problem, we employ an instrumental variables (IV) approach. In particular, we estimate the causal effect of hospital births via two stage least squares (2SLS), instrumenting for $Hospital_{izt}$ with the distance between a mother's residence and the nearest hospital with an obstetric ward.[9]

[9] In section 5, we show that our results from the structural and reduced form equations

Our instrumental variable strategy identifies the local average treatment effect (LATE) for mothers who give birth in a hospital only because they live "close enough" to it, but would give birth at home if they lived farther away. This population of "compliers" is likely not a random draw from the population and thus the effect of hospital births may not reflect the average treatment effect. However, since our paper is the first to convincingly identify the causal effect of place of birth, our results are relevant nevertheless. In addition, although we cannot identify individual compliers, in section 5.4 we compare their characteristics to those of the entire sample.

In order for the 2SLS method to yield consistent estimates of the parameter of interest, three conditions must be satisfied. First, the instrument should be a strong determinant of delivery location (the relevance condition). Intuitively, home and hospital births are alternative choices for the same final outcome — a healthy birth — and expectant mothers compare the costs and benefits of each of these options when choosing their delivery location. The distance to the nearest hospital with an obstetric ward impacts this cost-benefit calculation by changing the perceived costs of a hospital birth. This motivates the following first stage equation capturing the impact of the proposed instrument on the choice of delivery location:

$$Hospital_{izt} = \alpha_0 + \alpha_1 Distance_{izt} + \alpha_2 X_{izt} + u_{izt} \tag{2}$$

and the following reduced form equation relating the instrument to health outcomes:

$$Y_{izt} = \delta_0 + \delta_1 Distance_{izt} + \delta_2 X_{izt} + v_{izt} \tag{3}$$

where $Distance_{izt}$ is a measure of the distance between a mother's residence and the nearest hospital with an obstetric ward.[10] The relevance condition is easily tested using the results of the first stage equation. As a rule-of-thumb, if the first-stage F-statistic testing the significance of the instrument is greater than 10, then the instrument is considered strong.

are robust to using non-linear models.

[10] We discuss the construction of the instrument in detail in section 4.

Second, the instrument should be conditionally uncorrelated with the error term in the structural equation (the excludability condition): $E[Distance_{izt}\epsilon_{izt}|X_{izt}] = 0$. Intuitively, the excludability condition states that the instrument affects infant health outcomes only through its impact on the likelihood of a hospital birth. This is a non-trivial assumption and it would be violated if, for example, mothers whose infants have better expected health outcomes select their residential location based on the distance to the hospitals with an obstetric ward. Similarly, distance to the nearest obstetric ward may directly impact the health outcomes of infants born to mothers who experience complications and need to be transferred during delivery. While the excludability condition cannot be tested directly, we bring suggestive evidence on its plausibility in several ways. First, we check whether the observed characteristics of mothers and newborns are similar across the distribution of distance, as any relationship between observables and distance may provide an indication of the relationship between unobservables and distance. Second, we verify if distance has a direct effect on health outcomes by estimating the reduced form relationship between the instrument and health outcomes in samples where there is no relationship between the instrument and the likelihood of a hospital birth (i.e., there is no first stage).

The final assumption needed for the 2SLS to yield consistent estimates is monotonicity. This assumption states that while the instrument may not impact all individuals, those who are impacted by it are all impacted in the same way. In particular, it rules out a scenario where living closer to a hospital makes some mothers more likely to give birth in a hospital while making others less likely to do so. The monotonicity assumption cannot be tested formally but it is possible to shed some light on its plausibility. For example, we investigate whether the observable characteristics of mothers and newborns in a given delivery location (home or hospital) are similar across the distribution of distance. Differences in these characteristics would suggest variations in the choice of a delivery location as a function of distance, a violation of the monotonicity assumption. Similarly, we examine the robustness of our results in different subsamples defined by average car ownership, a characteristic that arguably has a direct impact on the cost-benefit calculation of a hospital birth. If the health effects of a hospital birth are similar among individuals more and less likely to own a car, we might be less concerned about a violation of the monotonicity condition. We discuss these and other checks in section 5.2.

4 Data

4.1 Data Sources

Our primary data comes from the Perinatal Registry of the Netherlands (Perinatale Registratie Nederland, PRN) and covers the period 2000–2008. PRN is an annual dataset that links three separate datasets of individual birth records collected separately by midwifes (LVR-1), obstetricians (LVR-2) and paediatricians (LNR). It covers approximately 99 percent of the primary care and 100 percent of the secondary care provided during pregnancy and delivery in the Netherlands (de Jonge et al., 2009).[11] The data includes detailed information on the birth process including delivery place (home or hospital), birth attendant (midwife or obstetrician) and method of delivery (natural birth, C-section, labor augmentation, induction, etc.) as well as on the presence of any complications during pregnancy or delivery. For each newborn, PRN also provides rich data on demographics such as gender, gestational age in days, birth weight, parity and plurality, on short term health outcomes including mortality and the Apgar score, as well as limited information on diagnosis and treatment such as NICU admission within the first 7 days of life. While the PRN data includes basic demographic characteristics of mothers (age, ethnicity, residential postal code), it does not provide information on education or income. For that reason, we complement this individual-level data with postal code-level data published by Statistics Netherlands (Kerncijfers postcodegebieden 2004), providing a snapshot of postal codes characteristics as of

[11] As discussed in section 2.1, the primary care in the Netherlands is provided by midwives and qualified general practitioners. PRN data does not include information on births supervised by general practitioners. These are a very small share of all primary care deliveries (Amelink-Verburg and Buitendijk, 2010).

January 1, 2004. Our main analysis uses the average household income in the postal code of residence of the mother and some of our robustness checks use additional variables from this data source. Finally, we use the 2005 Dutch National Atlas of Public Health to obtain the exact address and the availability of an obstetric ward for each hospital in the Netherlands.[12] This information is used in combination with geocode data on the latitude and longitude of the centroid of each postal code to construct the instrument.

Our outcome variables include three measures that capture the short term health of newborns: 7- and 28-day mortality and 5-minute Apgar score.[13] The observable characteristics included in the regressions can be classified in four groups. The first group (time effects) includes fixed effects for the year, month and day of the week of the birth. The second group (maternal characteristics) includes mother's age and ethnicity.[14] The third group (infant characteristics) includes birth weight, gestational age, and indicators for gender, plurality, type of birth attendant and birth position.[15] Finally, we include the average household income in the postal code of residence of the mother.[16]

Our instrument is based on the straight-line distance between mother's

[12] There were no closures or openings of hospitals with an obstetric ward during our study period.

[13] We do not have information on longer term mortality rates. The Apgar score summarizes the health of newborns based on five criteria: appearance (skin color), pulse (heart rate), grimace response ("reflex irritability"), activity (muscle tone), and respiration (breathing rate and effort). Newborns are usually evaluated at 1 and 5 minutes after birth. The score ranges from zero to 10 with higher scores indicating better health. Common reasons for a low Apgar score include a difficult birth (e.g, a fast delivery, a prolapsed cord, preterm birth, maternal hemorrhage), C-section, and amniotic fluid in the baby's airway.

[14] We include indicators for six maternal age categories (less than 20, 20–24, 25–29, 30–34, 35–39, 40 and above) and three maternal ethnicity categories: Dutch, Mediterranean and others (Moroccans and Turks, commonly identified as "Mediterraneans," represent the majority of the immigrant population in the Netherlands).

[15] Specifically, we include indicators for male, multiple birth, obstetrician supervision, breech birth, and a third degree polynomial in birth weight. Gestational age is included as a continuous variable but in some of the robustness checks we include additional indicators for preterm or late births.

[16] Some of the control variables (newborn gender and birth weight, mother's age, and average household income) are missing for a very small number of observations (less than 0.3 percent). We replace these missing values with the sample average of the corresponding variable and we include as additional controls indicators for missing values for each variable.

residence and the nearest hospital with an obstetric ward with both locations defined using the centroids of their respective postal codes.[17] To allow for potentially non-linear effects of distance, we construct our instrument as a set of six mutually exclusive dummy variables indicating distances less than 1 km, 1–2 km, 2–4 km, 4–7 km, 7–11 km, and more than 11 km. The lower cutoffs of these categories correspond approximately to the 10th, 25th, 50th, 75th and 90th percentiles of the distribution of the distance variable, respectively.[18]

The analysis sample is constructed as follows. The initial sample includes data on 1,630,062 newborns. First, we exclude observations for which the mother's residential postal code, the type of birth location and the type of birth attendant are missing. Second, we exclude stillbirths, planned C-sections and infants with congenital anomalies. The resulting sample of 1,478,187 births can be divided into two groups based on the perceived risk of the birth. We define high-risk mothers ($N = 689{,}844$) as those who start their perinatal care directly with an obstetrician or are referred to an obstetrician during pregnancy (before delivery) due to newly found risk factors. These women are required to give birth in a hospital under the supervision of an obstetrician and thus are excluded from our analysis sample.[19] In our main analysis, we only consider the low-risk mothers, who start their deliveries under the supervision of a midwife.

[17] Our data includes 6-digit postal codes for hospitals and 4-digit postal codes for mothers. Postal codes in the Netherlands have 6 digits and are much smaller than zip codes in the United States. The average 6-digit area has only 40 inhabitants and a land surface of 0.078 square kilometers (0.030 square miles); the average 4-digit area has 4075 inhabitants and a land surface of 8.5 square kilometers (3.28 square miles).

[18] In section 5.3, we show that our results are robust to alternative definitions of the instrument.

[19] One concern is the violation of the exclusion restriction due to a correlation between distance and the probability of being classified as a high-risk pregnancy. This could happen, for example, if midwives are more likely to refer women who reside farther away from an obstetric ward to obstetricians. Indeed, when we use an indicator for being classified as high-risk as the dependent variable in our reduced form equation, we do find a positive but economically small relationship between distance and high-risk classification (results available upon request). As a result, we would expect that in our sample of low-risk women those who live closer to the hospital are on average "unhealthier" than those who live farther away. Since our first stage results indicate that women are more likely to give birth in a hospital when they live closer to it, this selective referral strategy would bias our results in such a way that any health gains from a hospital birth likely represent lower bounds.

We further restrict our analysis to low-risk mothers at their first birth because it is likely that mothers who gave birth before have additional information on their own risk and preferences that is unobserved to the researcher but that is used in their choice of location of birth. This leaves us with a final sample of 356,412 observations.[20,21]

4.2 Descriptive Statistics

Table 1 provides descriptive statistics for the overall analysis sample, as well as by type of location of birth. Around 68 percent of all infants in the analysis sample are born in a hospital. Panel A lists the outcome variables and shows that there are substantial differences in mortality rates by location of birth. Babies who are born in a hospital are approximately four times more likely to die within a week and about 3.5 times more likely to die within 28 days than babies born at home. Similarly, babies born in a hospital have lower Apgar scores, on average, than those born at home. Panels B–D show that with the exception of birth weight and gestational age the observable characteristics of mothers and infants differ according to birth location in important ways. For example, over 90 percent of the infants born at home have a Dutch mother, in contrast to 79 percent of the babies born in a hospital. Children of Mediterranean mothers, on the other hand, tend to be born at the hospital rather than at home. Infants born in a hospital are also more likely to come from more densely populated postal codes with slightly lower average monthly income.

The differences in characteristics and health outcomes between hospital and home births have two likely causes. First, low-risk mothers who suspect themselves to be of increased risk for reasons unobserved to the midwife (and to the researcher) may self-select into a hospital birth. Second, all women

[20]In the rest of the paper, we refer to the sample consisting of the 1,478,187 observations described above as the "full sample" and the final sample consisting of the 356,412 observations as the "analysis sample". In section 5.3, we show that our results are robust to each of the sample selection criteria described above.

[21]There are slight differences between the estimating samples for mortality indicators and for the Apgar score because the Apgar score is missing for a small number of observations (less than 0.2 percent of the sample).

who need to be referred to an obstetrician during delivery (either because of complications, slow progression, or the need for pain relief medication) have to give birth in a hospital. As the Table shows, these referrals make up over 70 percent of hospital births.[22] Overall, the evidence presented in Table 1 is consistent with riskier births selecting into hospitals.

5 Results

5.1 OLS Estimates

Table 2 provides OLS estimates of the effect of hospital births on infant mortality. Panel A provides results from the full sample, while Panel B uses our analysis sample. Each column adds successively more control variables. Column 1 presents estimates from a regression model without any control variables. Column 2 adds time effects, column 3 maternal characteristics, and column 4 infant characteristics. Column 5 is our baseline specification which, in addition to time effects, maternal and infant characteristics, also controls for the average household income in the postal code of mother's residence. Here and in the rest of the paper, robust standard errors clustered at the postal code level are shown in parentheses.

Not controlling for any covariates, we find that giving birth in a hospital is associated on average with 4 to 5 more infant deaths per 1,000 births in the full sample. Among the sample of low-risk mothers who have their first child, the mortality rate among hospital births is higher by almost 2 deaths per 1,000 births. This roughly correspond to a 100 percent increase at the mean of the outcome variables. However, as more control variables are included successively in the model, the estimated effects diminish substantially. Looking across columns in Panel A, controlling for maternal characteristics reduces the OLS estimates by 70 percent and adding infant characteristics to the model decreases them again by 70–80 percent. Although controlling for mother's age and ethnicity does not impact the OLS estimates in the analysis sample

[22] We show in section 5.3 that our results are robust to different classifications of referrals.

in Panel B, adding infant characteristics completely wipes out the effects. Controlling for the average household income in the postal code of the mother's residence does not impact the results in either sample.[23]

Turning to the Apgar score results in the third row of Panels A and B, we find that in the absence of any control variables giving birth in a hospital is associated on average with a 0.23–0.27 lower Apgar score (column 1). Adding control variables reduces these raw correlations by 75–85 percent, indicating that giving birth in a hospital is on average associated with a 0.04–0.06 lower Apgar score (column 5).

In summary, an analysis relying on OLS methods would either conclude that giving birth in a hospital is detrimental to infant health or that there are no differential effects of home versus hospital births on newborn outcomes. However, the sensitivity of the results to the addition of control variables and, in particular, the monotonic decline in the estimated effects point again to the selection of mothers with riskier births into hospitals.

5.2 2SLS Estimates

5.2.1 Instrument Validity

In this section, we present the 2SLS estimates of the type of location of birth. As discussed in section 3, in order for the instrumental variable method to yield consistent estimates, the instrument must satisfy the relevance, the excludability and the monotonicity conditions. For that reason, we begin our discussion by providing evidence on the validity of these requirements.

Instrument Relevance The last panel of Table 1 (Panel E) provides descriptive statistics on the instrument. The average mother resides in a postal code that is 4.8 kilometers away from the nearest hospital with an obstetric ward. The distance from a woman's residence to the nearest hospital is correlated with the type of her delivery location: those who give birth in a hospital

[23] Table A1 in the appendix shows that these estimates are robust to using non-linear models.

reside in postal codes that are on average closer to hospitals (4.6 km) than those who give birth at home (5.3 km).

In Table 3 we show the first stage relationship between the endogenous regressor and the instrumental variable. Each column of the Table lists estimates from separate regressions of the main independent variable (an indicator for hospital birth) on the instrument.[24] The excluded category comprises postal codes at least 11 km away from an obstetric ward. Similar to Table 2, each column successively adds more controls. At the bottom of the Table, we also provide the F-statistic testing the joint significance of the distance indicators.

The results suggest that the distance between an expectant mother's home and the closest hospital with an obstetric ward is a strong predictor of whether she gives birth in a hospital or at home. The coefficient estimates are strongly statistically significant, regardless of the set of control variables included in the model, and they indicate that distance influences negatively the probability of a hospital birth. For example, our baseline specification (column 5) suggests that women living within 1 km of a hospital with an obstetric ward are 7.5 percentage points more likely to deliver in a hospital than those living at least 11 km away from a hospital. Although this effect diminishes as the distance between the mother's residence and the nearest hospital goes up, women located within 7–11 km of an obstetric ward are still 3 percentage points more likely to deliver in a hospital than those living farther away. The F-statistic testing the joint significance of the distance indicators is always well above the rule-of-thumb value of 10. Overall, the evidence provided in Table 3 confirms that the instrument satisfies the relevance condition.

Excludability Next, we turn to the issue of instrument excludability. While this assumption cannot be tested directly, we bring strong suggestive evidence on its plausibility in three ways. First, we examine whether the observable characteristics are balanced across the distribution of our instrument. We split the sample using the median distance between a mother's residence and

[24]The regressions in Table 3 use the estimating sample for mortality indicators. The results using the slightly smaller sample with available Apgar scores (available upon request) are virtually identical.

the nearest obstetric ward (4 km) and we provide sample means of selected covariates in Table 4.[25] We find that individuals located below and above the median distance are similar on many dimensions. For example, there are virtually no differences in maternal age, average monthly household income in the postal code, as well as the majority of infant characteristics (birth weight, gestational age, plurality, breech birth). However, mothers who live closer to a hospital are more likely to be immigrants residing in more densely populated postal codes and they are slightly more likely to be supervised by an obstetrician during birth (i.e., to be referred to an obstetrician during delivery).[26] The last panel of the Table summarizes these results by providing the average predicted outcomes based on a regression model that includes all the observable characteristics.[27] These statistics suggest that infants residing in areas close to a hospital are somewhat riskier in terms of observable characteristics. It is worth emphasizing that this *negative* selection combined with our first stage results imply that any bias in our 2SLS estimates would be in the direction of finding negative health effects from a hospital birth. As we show later in this section, our 2SLS estimates point to large health *gains* as a result of a hospital birth. Therefore, our findings likely represent lower bounds of the true effect.

Second, in columns 1–5 of Table 5 we show the reduced form relationship between the outcomes and the instrument for our analysis sample. The results indicate a strong relationship between the distance indicators and infant mortality. This relationship tends to grow stronger as more controls are added, pointing again to the negative selection of mothers into postal codes closer to hospitals. The baseline reduced form results (column 5) show an almost monotonic relationship between distance and infant mortality. For example, we find that the 7-day (28-day) infant mortality is lower on average by 0.701 (0.853) deaths per 1,000 births among those residing within 1 km of a hospital

[25] We present results by median distance for presentational clarity. Table A2 in the appendix provides a similar analysis for each distance category used in the construction of the instrument.

[26] Previous studies that use distance as an instrument when examining returns to heart attack technologies or NICUs also find a correlation between distance and ethnicity and average urbanization (McClellan et al., 1994; Cutler, 2007; Freedman, 2012).

[27] We use a probit model to predict infant mortality.

as compared to those who live at least 11 km away from a hospital. This is a large effect when compared to a sample mean of 1.779 (1.978) deaths per 1,000 births.[28] In contrast, the results in Panel C show a lack of relationship between the distance indicators and the Apgar score. The coefficients are generally statistically insignificant and always small in magnitude.

The instrument validity checks presented so far are similar to those conducted in previous studies that use a similar instrumental variables strategy. The specific institutional setup of the Netherlands allows us to perform a third validity check that arguably brings more direct evidence on the plausibility of the excludability condition. In particular, we compare the results of the reduced form relationship between the outcomes and the instrument from our analysis sample with those from the sample of high-risk women. As discussed in section 4, we define high-risk mothers as those who are under the care of an obstetrician at the start of delivery and have to give birth in a hospital. This means that there is no variation in type of delivery place among this sample (and so no relationship between the instrument and birth location). Hence, evidence of a relationship between the instrument and the outcomes in this sample would indicate a violation of the excludability assumption. We find no relationship between the instrument and outcomes, both for the sample of first-born children (column 6 of Table 5) and for the sample of all children born to high-risk mothers (column 7). The coefficient estimates of the distance indicators are always statistically insignificant and small in magnitude. An F-test of joint significance of the distance indicators in the sample of high-risk first births results in p-values of 0.555 for 7-day mortality, 0.521 for 28-day mortality and 0.769 for the Apgar score. The corresponding values for the sample of all high-risk births are 0.729, 0.725 and 0.575, respectively. In conclusion, although instrument excludability is an untestable assumption, the analyses presented in this section provide compelling evidence on its plausibility.

[28]Table A3 in the appendix checks the robustness of these reduced form estimates to non-linear specifications. As the Table shows, the average marginal effects from a probit model are almost identical to those produced by a linear probability model. For example, residing within 1 km of a hospital is associated with a 0.896 lower 7-day infant mortality per 1,000 births as compared to residing at least 11 km away from a hospital.

Monotonicity In our context, the monotonicity assumption states that all women who are affected by the instrument are less likely to choose a hospital birth as the distance to an obstetric ward increases. This is a non-trivial assumption because women choose their type of delivery location. Suppose women make their choice by comparing the comfort of a home birth to the risk of a negative outcome due to complications during delivery. As the distance to an obstetric ward goes up, the risk of a negative outcome increases due to longer travel times to a hospital. In this case, it is possible that women who live far away from a hospital prefer a hospital birth, violating the monotonicity assumption. Intuitively, we do not expect such a violation to be present in our sample for several reasons. First, the fact that women can only go to the hospital after contractions reach a certain frequency makes the trip increasingly uncomfortable for women living farther away from a hospital. Second, the fact that we observe a positive relationship between distance and the probability of being classified as high-risk suggests that midwives might refer more risk-averse women to an obstetrician in order to ensure a hospital birth.

While we cannot test directly the validity of the monotonicity condition, we provide suggestive evidence in Table 6 by comparing the means of selected covariates for women who deliver in a particular type of location when the sample is split using the median distance between their residence and the nearest obstetric ward (4 km).[29] To the extent that women who choose a hospital (or home) birth have similar observable characteristics when residing in different distance categories, we may be less concerned about them responding in different ways to the instrument. As the Table shows, within a particular type of delivery location, distance is not related to the majority of the average characteristics of women and their newborns: maternal age, newborn gender, birth weight and gestational age are virtually the same. Not surprisingly, we find differences in certain characteristics which closely mimic those found in the overall sample. For example, regardless of delivery location, women who

[29]We present results by median distance for presentational clarity. Table A4 in the appendix provides a similar analysis for each distance category used in the construction of the instrument.

live closer to a hospital are more likely to be immigrant and reside in more urban postal codes with a slightly lower average income. Finally, as in the overall sample, the average predicted outcomes listed in the last panel of the Table show that women living closer to a hospital are somewhat riskier in terms of observable characteristics for either type of delivery location. These findings may not be surprising given that 98 percent of the Dutch population lives within a 30-minute drive from an obstetrics ward (Nationale Atlas Volksgezondheid, 2011). In conclusion, we do not find any evidence of a violation of the monotonicity assumption.[30]

5.2.2 2SLS Results

Table 7 presents the instrumental variable estimates. Panel A reproduces the OLS results, while Panel B reports the 2SLS estimates. Each column provides results for a different outcome using a specification that includes our baseline set of controls. The last two rows report the Hausman statistic and the associated p-value testing the consistency of both the OLS and 2SLS estimates against the alternative of consistency of the 2SLS estimates only.

The instrumental variables results point to significant reductions in mortality from a hospital birth: giving birth in a hospital reduces infant mortality by 8 to 9 deaths per 1,000 births. Our results also indicate that hospital births do not have an impact on the Apgar score, as the point estimate is very small and statistically insignificant. These 2SLS results are in sharp contrast to the OLS estimates which indicate that giving birth in a hospital has no impact on mortality and significantly reduces the Apgar score. The Hauman tests further confirm these differences by strongly rejecting the equivalence of the OLS and 2SLS estimates.

The estimated mortality reductions from 2SLS are large when compared to sample means of 1.779 and 1.978 for 7-day and 28-day mortality, respectively.

[30]We provide additional evidence on the plausibility of the monotonicity assumption in section 5.3 by showing that our 2SLS results are robust when the sample is split by average per capita car ownership in the postal code, a factor likely to directly impact the choice of type of delivery location.

Two points are worth emphasizing. First, the 2SLS estimates have wide confidence intervals that include much smaller but still economically important effects. For example, the lower bounds of a 95-percent confidence interval indicate 2.1 and 2.6 fewer infant deaths per 1,000 births, respectively, for 7-day and 28-mortality.[31] Second, as noted in section 3, our instrumental variable strategy identifies a LATE and thus our results apply to a population of compliers: mothers who are induced to give birth in a hospital because they live "close enough" to it. We describe the characteristics of this compliant population in section 5.4 after demonstrating the robustness of our estimates to various checks.

5.3 Robustness Checks

Table 8 presents our robustness checks. In Panel A, we check the sensitivity of our 2SLS estimates to the selection of the analysis sample. The results in the first two rows show that our results are not driven by the exclusion of newborns with congenital anomalies or stillbirths. In both cases, a hospital birth still leads to significant mortality gains and has no impact on the Apgar score. In the third row, we add higher order low-risk births and still find substantial reductions in mortality. The estimated effects are somewhat smaller, consistent with the idea that women use information on their unobserved health risk from previous births to better select their delivery location.[32] In the last row of Panel A, we provide results from the full sample and confirm that giving birth in a hospital leads to lower mortality. While the coefficient estimates are slightly larger in absolute value than those in the low-risk sample (row 3), the standard errors also more than double and we cannot reject the equality of the two sets of estimates.

[31] The fact that, as mentioned before, the marginal effects are similar when we estimate the reduced form via linear and non-linear models (see Tables 5 and A3) suggests that the magnitude of our estimates is not due to nonlinearities.

[32] Mothers who have a risky first birth (and thus possibly worse unobserved risk) may become "always-takers" and always choose a hospital birth in subsequent pregnancies. As a result, the compliant population among higher-order births may have lower health gains from a hospital birth, leading to lower coefficient estimates among all low-risk births.

In our empirical strategy, we define birth location according to where the birth occurred rather than where the delivery started because this information is not available in our data. A delivery that starts at home ends in a hospital if the mother needs to be referred to an obstetrician during delivery. This can be due to complications during delivery, a delivery that is taking too long, or because of the need for pain relief medication. In order to shed some light on this issue, we use additional information on referrals available in our data. In particular, the PRN data has information on midwife-to-obstetrician referrals, which include home-to-hospital transfers as well as within-hospital referrals. In Panel B of Table 8, we investigate the sensitivity of our results to re-classifying all referrals as home births instead of hospital births (recall that in our baseline estimates, referrals are always hospital births). The estimated effects are very similar to our baseline results with slightly larger magnitudes, consistent with the referral of riskier births from midwives to obstetricians.[33]

In Panel C, we revisit the plausibility of the monotonicity condition. In order to address the concern that women may respond differently to the instrument as distance increases, we use information on car ownership, which arguably has a direct impact on the cost-benefit calculation of a hospital birth.[34] To the extent that the instrument influences the probability of a hospital birth in similar ways in postal codes with different levels of car ownership, we may be less concerned about a violation of the monotonicity assumption. We split the sample using the median number of cars per person in the postal code and check the robustness of our results in the two samples. The results from both samples suggest again mortality reductions from a hospital birth.[35] There is

[33] We also estimated regressions in which we replaced the actual birth place of referrals with the planned place of birth if this was provided (results available upon request). The planned place of birth is recorded by midwives at any time during pregnancy, and in a significant number of cases not at all or after delivery, making the variable potentially endogenous. The estimated effects of a hospital birth are somewhat smaller but still highly significant when we use this alternative classification, consistent with the idea that lower-risk women self-select into (planned) home births.

[34] We have data from Statistics Netherlands (CBS Statline, accessed on June 11, 2012) on the average car ownership per citizen in each postal code dating from January 1, 2004. The median number of cars per person is 0.435.

[35] In addition, the first stage in both samples confirms the negative relationship between

some loss of precision in the sample of postal codes with average car ownership greater than the median due to a smaller sample size but the magnitudes of the estimates are very similar across the two subsamples (and similar to the baseline estimates). The failure to reject the equality of the hospital effects lends further support to the plausibility of the monotonicity condition.

The final Panel of the Table checks the sensitivity of the results to the inclusion of additional control variables. Although our baseline model includes the average household income in the postal code, there might still be concerns about a potential bias due to omitted parental characteristics. In order to mitigate these concerns, we add controls for the average area density[36] and the share of 0-15 year-olds in the postal code. The results are close in magnitude to our baseline estimates but less precisely estimated. The loss in precision comes from the fact that both the postal code characteristics and the distance indicators do not vary over time and are thus highly correlated, leading to weaker first stage regressions. In conclusion, the various checks conducted in this section confirm the robustness of the baseline 2SLS results.[37]

5.4 Heterogeneous Effects and Complier Characteristics

To the extent that there is heterogeneity in the effect of a hospital birth, our 2SLS results represent a local average treatment effect that applies to the subpopulation of compliers: women who give birth in a hospital *because* of the particular distance between their residence and the closest hospital with an obstetric ward. Table 9 examines the heterogeneity in the 2SLS effects by selected observable characteristics of mothers and their newborns.

The estimated effects for Dutch mothers and for mothers at most 29 years old (the median age in the sample) are similar to the baseline results from the

distance and the likelihood of a hospital birth (results available upon request).

[36]This is the average number of addresses per square kilometer in a circle with a radius of 1 km around each address in the postal code.

[37]We also confirmed the robustness of our main results to the reclassification of the instrument as a continuous variable, to the inclusion of indicators for prematurity and late-term birth, to a potential finite-sample bias from weak instruments by using limited information maximum likelihood, as well as to clustering at larger geographic levels such as the municipality (results available upon request).

full analysis sample. While the effects on non-Dutch and older mothers also indicate substantial mortality reductions from a hospital birth, the results are estimated imprecisely. The findings in rows 5–8 show that the 2SLS estimates are similar when the sample is split by median gestational age (280 days) or median birth weight (3,410 grams). The largest treatment heterogeneity is found with respect to the average monthly income in the postal code of the mother's residence. The 2SLS estimates reported in the last Panel of the Table show that our baseline results are driven entirely by births to mothers residing in postal codes with less than the median of the average monthly household income in the postal code (€1,929). This confirms our prior that the baseline results represent a LATE.

In the remainder of this section we focus on the compliant subpopulation. While it is not possible to identify individual compliers, it is possible to calculate their share among the analysis sample as well as the distribution of their characteristics.[38] When the treatment variable and the instrument are both binary, the share of the compliant subpopulation is given by the first stage coefficient of the instrument. In addition, the relative likelihood of a complier having a binary characteristic is given by the ratio of the first stage coefficient of the instrument from the sample of individuals who have that characteristic to the first stage coefficient from the full analysis sample. When the instrument consists of a set of mutually exclusive indicators, as in our case, the estimated LATE is a weighted average of the LATEs using each indicator one at a time. In particular, there is a distinct compliant subpopulation corresponding to each distance indicator.[39] Therefore, the size and characteristics of compliers can be calculated separately for each indicator.

Table 10 provides the results. Each cell in columns 1-5 gives the likelihood that the compliers corresponding to the distance indicator in the column have the characteristic described in the row relative to the analysis sample. The last row shows the size of the compliant subpopulation associated with each

[38]For more details, see Imbens and Angrist (1994) and Angrist and Imbens (1995).

[39]Intuitively, for each distance category there is a distinct set of women who give birth in a hospital because they live within that particular distance interval from a hospital.

instrument. Compliers represent approximately 10.6 percent of all low-risk first births. Compliant mothers have observable characteristics generally not associated with higher risk: they are more likely to be Dutch and younger than the median age of 29, and their pregnancies are more likely to be within the normal range (i.e., gestational age between 37 and 42 weeks). It should be mentioned that although complier newborns tend to be lighter than the median newborn in the analysis sample (3,410 grams), the vast majority of babies in our sample are above the medical at-risk threshold of 2,500 grams.[40]

Our finding that the compliant subpopulation does not have higher observable risk but the results are entirely driven by births from lower-income postal codes is consistent with the previous literature documenting disparities in preventive behavior and quality of care by income and education (e.g., Smith, 1999; Cutler and Lleras-Muney, 2010). We do not have information that would allow us to distinguish whether it is unobserved preventive behavior or quality of care that drive our results. The tendency among complier newborns to be lighter than the median in the analysis sample fits with the poorer preventive behavior explanation.[41] Moreover, midwives serving lower-income postal codes frequently argue that expectant mothers residing in these areas have poorer health education and life styles. A recent survey by the Royal Dutch Organisation of Midwives reported that midwives needed on average 23 percent extra time when caring for lower-income women, leading to a 2009 policy change that increased the reimbursement for midwives by 23 percent in selected postal codes (NZA, 2011b).[42] This suggests that lower quality of care might also play a role in the higher mortality rate among home births.

[40]Fetal growth retardation is one of the reasons for referral to an obstetrician. As a result, only 2.7 percent of the newborns in the analysis sample weigh less than 2,500 grams, the standard definition of low birth weight.

[41]We should also note that maternal age is strongly positively correlated with education in the Netherlands (van Agtmaal-Wobma and van Huis, 2008). Therefore, the fact that compliers tend to be younger suggests that they might also be lower-educated.

[42]According to the report, the need for extra time was due to the difficulties in collecting relevant (medical) data, additional education on prevention, lifestyles and risk, more frequent home visits, consultations to exclude uncertainties, etc. (NZA, 2011b)

5.5 Mechanisms

Our 2SLS results indicate that the broadly measured general health condition of children born in different locations is similar shortly after birth, as captured by the 5-minute Apgar score. This indicates that the mortality reductions observed in the first 7 or 28 days of life following a hospital birth come from medical care provided after delivery. There are many channels through which a hospital birth may reduce infant mortality, such as the availability of better facilities and equipment, better hygiene (sterility) or the proximity to other medical services. Unfortunately, data are not available on most of these variables. In this section, we examine the impact of a hospital birth on a specific type of treatment for which reliable information is available: admission to a NICU within the first seven days of life.[43]

Using an indicator for NICU admission as dependent variable in our baseline specification, we find that giving birth in a hospital increases the probability of a NICU admission by 18.7 percentage points (s.e. 9.8). This is a large effect when compared to the average NICU admission rate of 10.8 percent in the analysis sample. There are several possible explanations for this. First, it is likely that transfers to a NICU are more feasible when the newborn is already in a hospital (there are only 10 NICUs in the Netherlands), either because the hospital has a NICU itself or because transportation by helicopter or ambulance to a hospital with a NICU can be arranged more easily. Second, there may be differences in the length and intensity of post-natal monitoring between hospitals and home. In particular, some newborns delivered in a hospital may be kept there for observation and eventually transferred to a NICU. Similar newborns delivered at home might not be transferred to a hospital for observation, and thus the necessity of a NICU transfer might not become apparent. In conclusion, the evidence we have suggests that access to medical technologies may be an important channel in explaining the lower mortality

[43]NICUs have been shown to significantly improve the health and survival of at-risk newborns through close monitoring of blood gases, heart rate and rhythm, breathing rate, and blood pressure; advanced diagnostic techniques such as CT scans and cardiac catheterization; and specialized treatments such as phototherapy and exchange blood transfusions, mechanical ventilation, and artificial surfactant (Cutler and Meara, 2000).

among hospital births.

6 Conclusions

In this paper, we examine the impact of home births on the health outcomes of low-risk newborns. We implement an instrumental variables strategy that exploits the exogenous variation in distance between a mother's residence and the nearest obstetric ward. Using data from the Netherlands for the period 2000–2008, we find that giving birth in a hospital leads to substantial reductions in infant mortality but has no effect on Apgar scores. We present evidence suggesting that our instrument satisfies the relevance, excludability and monotonicity assumptions and that the main estimates are robust to a host of specification checks.

Data limitations do not allow us to investigate many potentially important channels that may facilitate these health gains but our finding that a hospital birth is associated with substantial increases in the likelihood of a NICU admission indicates that proximity to medical technologies may be one of these channels. Our results represent a local average treatment effect that applies to the subsample of low-risk women who give birth in a hospital because they reside close enough to it, but would give birth at home if they lived farther away. We show that compliers generally have observable characteristics that are not associated with higher health risks — they are younger, more likely to be native, and more likely to give birth within the normal gestational age interval — but our results are entirely driven by those residing in below-median postal codes.

As high health care costs persist and out-of-hospital births keep rising sharply in many developed countries, understanding the impact of home births on newborn outcomes becomes even more important. Taken together, our results suggest that giving birth in a hospital leads to economically large mortality reductions even in a health care system that is specifically geared toward risk selection and home births.

References

Ackermann-Liebrich, Ursula, Thomas Voegeli, Katrin Gunter-Witt, Isabelle Kunz, Maja Zullig, Christian Schindler, Margrit Maurer, and Zurich Study Team (1996), "Home versus hospital deliveries: follow up study of matched pairs for procedures and outcome," *BMJ* 313(7068), 1313–1318.

Almond, Douglas, and Joseph Doyle (2011), "After midnight: A regression discontinuity design in length of postpartum hospital stays," *American Economic Journal: Economic Policy* 3(3), 1–34.

Almond, Douglas, Joseph Doyle, Amanda Kowalski, and Heidi Williams (2010), "Estimating marginal returns to medical care: Evidence from at-risk newborns," *Quarterly Journal of Economics* 125(2), 591–634.

Amelink-Verburg, Marianne, and Simone Buitendijk (2010), "Pregnancy and labour in the Dutch maternity care system: What is normal? The role division between midwives and obstetricians," *Journal of Midwifery & Women's Health* 55(3), 216–225.

Angrist, Joshua, and Guido Imbens (1995), "Two-stage least squares estimation of average causal effects in models with variable treatment intensity," *Journal of the American Statistical Association* 90(430), 431–442.

Bais, Joke, and Maria Pel (2006), "The basis of the Dutch obstetric system: risk selection," *European Clinics in Obstetrics and Gynaecology* 2(4), 209–212.

Bastian, Hilda, Marc Keirse, and Paul Lancaster (1998), "Perinatal death associated with planned home birth in Australia: population based study," *BMJ* 317(7155), 384–388.

Bharadwaj, Prashant, Katrine Løken, and Christopher Neilson (forthcoming), "Early life health interventions and academic achievement," *American Economic Review* .

Birthplace in England Collaborative Group (2011), "Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: the Birthplace in England national prospective cohort study," *BMJ* 343, d7400.

Borquez, HA, and TA Wiegers (2006), "A comparison of labour and birth experiences of women delivering in a birthing centre and at home in the Netherlands," *Midwifery* 22(4 Dec), 339–47.

Boucher, Debora, Catherine Bennett, and Barbara Mcfarlin (2009), "Staying home to give birth: Why women in the United States choose home birth," *American College of Nurse-Midwives* 54, 119–126.

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

Cutler, David, and Adriana Lleras-Muney (2010), "Understanding differences in health behaviors by education," *Journal of Health Economics* 29(1), 1–28.

Cutler, David, Mark McClellan, Joseph Newhouse, and Dahlia Remler (1998), "Are medical prices declining? Evidence from heart attack treatments," *The Quarterly Journal of Economics* 113(4), 991–1024.

Cutler, David, and Ellen Meara (1998), "The medical costs of the young and old: A forty-year perspective," in David Wise, ed., *Frontiers in the Economics of Aging*, 215–246, University of Chicago Press.

——— (2000), "The technology of birth: Is it worth it?" *NBER/Frontiers in Health Policy Research* 3(1), 33–67.

CVZ (2003), *Verloskundig Vademecum*, Diemen, The Netherlands.

Daysal, N. Meltem, Mircea Trandafir, and Reyn van Ewijk (2012), "Returns to childbirth technologies: Evidence from preterm births," Mimeo.

de Jonge, A, BY van der Goes, ACJ Ravelli, MP Amelink-Verburg, BW Mol, JG Nijhuis, J Bennebroek Gravenhorst, and SE Buitendijk (2009), "Perinatal mortality and morbidity in a nationwide cohort of 529,688 low-risk planned home and hospital births," *BJOG: An international journal of obstetrics and gynaecology* 116(9), 1177–84.

Department of Health (2004), *Maternity Standard, National Service Framework for Children, Young People and Maternity Services*.

Dowswell, T., JG Thornton, J. Hewison, RJ Lilford, J. Raisler, A. MacFarlane, G. Young, M. Newburn, R. Dodds, and RS Settatree (1996), "Should there be a trial of home versus hospital delivery in the United Kingdom?" *BMJ* 312(7033), 753.

Freedman, Seth (2012), "The effect of deregionalization on health outcomes: Evidence from neonatal intensive care," Mimeo.

Fuchs, Victor (1998), *Who Shall Live?: Health, Economics, and Social Choice*, World Scientific.

Gyte, G., M. Dodwell, M. Newburn, J. Sandall, A. Macfarlane, and S. Bewley (2009), "Estimating intrapartum-related perinatal mortality rates for booked home births: When the 'best' available data are not good enough," *BJOG: An International Journal of Obstetrics & Gynaecology* 116(7), 933–942.

Henderson, Jane, and Stavros Petrou (2008), "Economic implications of home births and birth centers: A structured review," *BIRTH* 35(2 June), 136–146.

Hendrix, M, M Van Horck, D Moreta, F Nieman, M Nieuwenhuijze, J Severens, and J Nijhuis (2009), "Why women do not accept randomisation for place of birth: feasibility of a RCT in the Netherlands," *BJOG: An International Journal Of Obstetrics And Gynaecology* 116(4), 537–542.

Imbens, Guido, and Joshua Angrist (1994), "Identification and estimation of local average treatment effects," *Econometrica* 62(2), 467–475.

Janssen, P.A., S.K. Lee, E.M. Ryan, D.J. Etches, D.F. Farquharson, D. Peacock, and M.C. Klein (2002), "Outcomes of planned home births versus planned hospital births after regulation of midwifery in British Columbia," *Canadian Medical Association Journal* 166(3), 315–323.

Janssen, Patricia, and Angela Henderson (2009), "The experience of planned home birth: Views of the first 500 women," *BIRTH* 36(4 December), 297–304.

Kennare, Robyn, Marc Keirse, Graeme Tucker, and Annabelle Chan (2010), "Planned home and hospital births in South Australia, 1991-2006: Differences in outcomes," *The Medical Journal of Australia* 192(2), 76–80.

Latta, J.M., J. Derksen, and F.M. van der Meer (2011), *Eigen bijdrage verloskundige zorg [Personal contribution obstetric care]*, Diemen, The Netherlands: CVZ, report for the Minister of Health, Welfare and Sport.

Lindgren, Helena, Ingela Rådestad, Kyllike Christensson, and Ingegerd Hildingsson (2008), "Outcome of planned home births compared to hospital births in Sweden between 1992 and 2004. a population-based register study," *Acta Obstetricia Et Gynecologica Scandinavica* 87(7), 751–759.

MacDorman, Marian, T.J. Mathews, and Eugene Declercq (2012), "Home births in the United States, 1990–2009," *NCHS data brief* 84, 1–9.

Malloy, M H (2010), "Infant outcomes of certified nurse midwife attended home births: United States 2000 to 2004." *Journal of perinatology: official journal of the California Perinatal Association* 30(9), 622–7.

McClellan, Mark, Barbara McNeil, and Joseph Newhouse (1994), "Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality?" *JAMA* 272, 859–866.

McClellan, Mark, and Joseph Newhouse (1997), "The marginal cost-effectiveness of medical technology: A panel instrumental-variables approach," *Journal of Econometrics* 77(1), 39–64.

McClellan, Mark, and Haruko Noguchi (1998), "Technological change in heart-disease treatment. Does high tech mean low value?" *American Economic Review* 88(2), 90–96.

Miller, Amalia (2006), "The impact of midwifery-promoting public policies on medical interventions and health outcomes," *B.E. Journal of Economic Analysis and Policy: Advances in Economic Analysis and Policy* 6(1), 1–34.

Mohangoo, AD, SE Buitendijk, CW Hukkelhoven, AC Ravelli, GC Rijninks-van Driel, P. Tamminga, JG Nijhuis, et al. (2008), "[Higher perinatal mortality in The Netherlands than in other European countries: the Peristat-II study]," *Nederlands tijdschrift voor geneeskunde* 152(50), 2718–2727.

Murphy, Patricia Aikins, and Judith Fullerton (1998), "Outcomes of intended home births in nurse-midwifery practice: a prospective descriptive study," *Obstetrics and gynecology* 92(3), 461–470.

Nationale Atlas Volksgezondheid (2011), "Reistijd naar ziekenhuis met afdeling verloskunde 2011," Available online at http://www.zorgatlas.nl/zorg/ziekenhuiszorg/algemene-en-academische-ziekenhuizen/aanbod/reistijd-naar-ziekenhuis-met-afdeling-verloskunde, accessed Sept. 4, 2012.

Newhouse, Joseph (1992), "Medical care costs: How much welfare loss?" *Journal of Economic Perspectives* 6(3), 3–21.

Nove, Andrea, Ann Berrington, and Zoë Matthews (2008), "Home births in the UK, 1955 to 2006," *Population trends* (133), 20–27.

NZA (2008), "Tariefbeschikking nummer 5700-1900-08-3," Available online at http://www.nza.nl/regelgeving/tarieven/, accessed Sept. 4, 2012.

——— (2011a), "DBC bedragen en overige bedragen medisch specialistische zorg door of vanwege de zorginstelling," Available online at http://dbc-tarieven.nza.nl, accessed Sept. 4, 2012.

——— (2011b), "Tariefbeschikking TB/CU-7007-01," Available online at http://www.nza.nl/regelgeving/tarieven/, accessed Sept. 4, 2012.

O'Connor, B.B. (1993), "The home birth movement in the United States," *Journal of Medicine and Philosophy* 18(2), 147–174.

OECD (2012), "Health data 2012. Health expenditure and financing," Data extracted on Sept. 4, 2012.

Pang, Jenny, James Heffelfinger, Greg Huang, Thomas Benedetti, and Noel Weiss (2002), "Outcomes of planned home births in Washington State: 1989-1996." *Obstetrics and gynecology* 100(2), 253–9.

Royal College of Obstetricians and Gynaecologists and Royal College of Midwives (2007), *Joint statement No. 2. Home Births*.

Skinner, Jonathan, Douglas Staiger, and Elliott Fisher (2006), "Is technological change in medicine always worth it? The case of acute myocardial infarction," *Health Affairs* 25(2), w34–w47.

Smith, James (1999), "Healthy bodies and thick wallets: The dual relation between health and economic status," *Journal of Economic Perspectives* 13(2), 145–166.

The American College of Obstetricians and Gynecologists (2011), *Committee Opinion Number 476. Planned Home Birth*.

van Agtmaal-Wobma, Elma, and Mila van Huis (2008), "De relatie tussen vruchtbaarheid en opleidingsniveau van de vrouw [The relationship between fertility and educational level of women]," *Bevolkingtrends* 56(2), 32–41.

van der Kooy, Jacoba, Jashvant Poeran, Johanna de Graaf, Erwin Birnie, Semiha Denktaş, Eric Steegers, and Gouke Bonsel (2011), "Planned home compared with planned hospital births in the Netherlands," *Obstetrics & Gynecology* 118, 1037–1046.

Wiegers, T.A., J. van der Zee, J.J. Kerssens, and M.J.N.C. Keirse (1998), "Home birth or short-stay hospital birth in a low risk population in the Netherlands," *Social Science & Medicine* 46(11), 1505–1511.

Table 1: Descriptive Statistics

	Analysis Sample		Hospital		Home	
	Mean	Sd	Mean	Sd	Mean	Sd
	(1)	(2)	(3)	(4)	(5)	(6)
A. Outcome Variables						
7-Day Mortality (per 1,000)	1.779	42.139	2.335	48.268	0.609	24.676
28-Day Mortality (per 1,000)	1.978	44.431	2.575	50.683	0.722	26.868
Apgar Score	9.660	0.818	9.585	0.898	9.818	0.586
B. Mother's Characteristics						
Age	28.380	4.652	28.332	4.828	28.482	4.255
Ethnicity: Dutch	0.827	0.378	0.786	0.410	0.913	0.281
Ethnicity: Mediterranean	0.064	0.246	0.082	0.274	0.029	0.166
C. Infant Characteristics						
Boy	0.509	0.500	0.519	0.500	0.488	0.500
Birth weight	3413	480	3416	503	3408	429
Gestational Age (days)	279	11	279	12	279	8
Obstetrician Supervision	0.482	0.500	0.712	0.453	0.000	0.000
Multiple Birth	0.001	0.023	0.001	0.028	0.000	0.000
Breech Birth	0.011	0.103	0.015	0.123	0.001	0.025
D. Average Postal Code Characteristics						
Monthly Household Income (euro)	1975	313	1970	322	1987	292
Density	1969	1889	2053	1907	1793	1840
Percent 0–15 years old	18.750	4.447	18.609	4.440	19.045	4.447
E. The Instrument						
Distance (km)	4.803	4.041	4.558	3.930	5.317	4.217
$<$ 1 km	0.092	0.289	0.099	0.298	0.079	0.269
1–2 km	0.225	0.418	0.240	0.427	0.194	0.395
2–4 km	0.243	0.429	0.250	0.433	0.226	0.419
4–7 km	0.198	0.399	0.191	0.393	0.214	0.410
7–11 km	0.139	0.346	0.129	0.335	0.159	0.365
$\geq$ 11 km	0.103	0.304	0.091	0.288	0.129	0.335
OBS	356,412		241,519		114,893	

Notes: The first two columns provide sample means and standard deviations for the full analysis sample. The remaining panels of columns provide descriptive statistics by type of location of birth.

Table 2: OLS Estimation of the Effects of Hospital Births on Infant Outcomes

	(1)	(2)	(3)	(4)	(5)
A. Full Sample					
Dependent Variable: 7-Day Mortality (per 1,000; N=1,478,187)					
Hospital	4.635***	4.645***	1.388***	0.362***	0.362***
	(0.086)	(0.086)	(0.089)	(0.092)	(0.091)
Dependent Variable: 28-Day Mortality (per 1,000; N=1,478,187)					
Hospital	5.029***	5.043***	1.484***	0.320***	0.320***
	(0.088)	(0.088)	(0.093)	(0.095)	(0.095)
Dependent Variable: Apgar Score (N=1,476,118)					
Hospital	−0.270***	−0.271***	−0.185***	−0.040***	−0.040***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
B. Analysis Sample					
Dependent Variable: 7-Day Mortality (per 1,000; N=356,412)					
Hospital	1.726***	1.721***	1.644***	−0.003	−0.001
	(0.129)	(0.131)	(0.133)	(0.155)	(0.155)
Dependent Variable: 28-Day Mortality (per 1,000; N=356,412)					
Hospital	1.853***	1.835***	1.754***	−0.074	−0.072
	(0.135)	(0.137)	(0.141)	(0.163)	(0.163)
Dependent Variable: Apgar Score (N=355,761)					
Hospital	−0.233***	−0.228***	−0.231***	−0.061***	−0.061***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Time Effects		X	X	X	X
Mother's Characteristics			X	X	X
Infant Characteristics				X	X
Average HH Income					X

Notes: Panel A provides the effect of hospital births on infant health outcomes using the full sample of births after excluding observations with missing information on the mother's residential postal code, the type of birth location and birth attendant as well as stillbirths, planned C-sections and infants with congenital anomalies. Panel B provides results from our analysis sample (first births to low-risk mothers). Time effects include indicators for year, month and day-of-week of birth. Mother's characteristics include indicators for maternal age and ethnicity groups. Infant characteristics include gestational age, a third degree polynomial in birth weight and indicators for male, multiple birth, obstetrician supervision and breech birth. Average household income refers to the average income in the postal code of mother's residence. For more information, see section 4. Robust standard errors clustered at the postal code level are shown in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Distance and Type of Location of Birth – First Stage Estimates

	(1)	(2)	(3)	(4)	(5)
Dependent Variable: Hospital					
< 1 km	0.127***	0.128***	0.087***	0.075***	0.075***
	(0.011)	(0.011)	(0.010)	(0.009)	(0.009)
1–2 km	0.124***	0.125***	0.082***	0.074***	0.073***
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
2–4 km	0.100***	0.100***	0.071***	0.060***	0.060***
	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)
4–7 km	0.053***	0.055***	0.047***	0.037***	0.037***
	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)
7–11 km	0.032***	0.036***	0.033***	0.030***	0.030***
	(0.010)	(0.009)	(0.009)	(0.008)	(0.008)
F-statistic	62.051	64.510	28.002	27.966	27.979
N	356,412	356,412	356,412	356,412	356,412
Time Effects		X	X	X	X
Mother's Characteristics			X	X	X
Infant Characteristics				X	X
Average HH Income					X

Notes: Each column lists estimates from separate regressions of the main independent variable (an indicator for a hospital birth) on the instrument. The excluded category is living in a postal code that is at least 11 km away from a hospital with an obstetric ward. Time effects include indicators for year, month and day-of-week of birth. Mother's characteristics include indicators for maternal age and ethnicity groups. Infant characteristics include gestational age, a third degree polynomial in birth weight and indicators for male, multiple birth, obstetrician supervision and breech birth. Average household income refers to the average income in the postal code of mother's residence. For more information, see section 4. The F-statistic refers to the test of joint significance of the distance indicators. Robust standard errors clustered at the postal code level are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Average Observable Characteristics by Median Distance

	Distance $\leq$ 4km (1)	Distance $>$ 4km (2)
A. Mother's Characteristics		
Age	28.269	28.522
Ethnicity: Dutch	0.758	0.914
Ethnicity: Mediterranean	0.095	0.026
B. Infant Characteristics		
Boy	0.509	0.509
Birth weight	3402	3427
Gestational Age (days)	279	279
Obstetrician Supervision	0.489	0.473
Multiple Birth	0.001	0.000
Breech Birth	0.010	0.012
C. Average Postal Code Characteristics		
Monthly Household Income (euro)	1964	1991
Density	2818	889
Percent 0–15 years old	17.576	20.243
D. Predicted Outcomes		
Predicted 7-Day Mortality	1.857	1.706
Predicted 28-Day Mortality	2.064	1.901
Predicted Apgar Score	9.656	9.664
N	199,510	156,902

Notes: For a description of the variables, see section 4.

Table 5: Distance and Infant Outcomes – Reduced Form Estimates

	Low-Risk First Births					High-Risk First Births	High-Risk All Births
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Dependent Variable: 7-Day Mortality (per 1,000)							
< 1 km	−0.286	−0.290	−0.492	−0.725**	−0.701**	0.058	0.102
	(0.342)	(0.342)	(0.344)	(0.324)	(0.324)	(0.615)	(0.414)
1–2 km	−0.501*	−0.505*	−0.718**	−0.724**	−0.702**	−0.089	−0.068
	(0.288)	(0.288)	(0.292)	(0.282)	(0.282)	(0.475)	(0.314)
2–4 km	−0.459	−0.466	−0.607**	−0.592**	−0.554**	0.054	−0.150
	(0.289)	(0.289)	(0.288)	(0.276)	(0.276)	(0.488)	(0.313)
4–7 km	−0.317	−0.323	−0.366	−0.383	−0.330	−0.430	−0.091
	(0.295)	(0.295)	(0.296)	(0.286)	(0.286)	(0.480)	(0.307)
7–11 km	−0.574*	−0.578*	−0.594**	−0.572*	−0.548*	0.402	0.279
	(0.301)	(0.301)	(0.302)	(0.294)	(0.294)	(0.525)	(0.320)
B. Dependent Variable: 28-Day Mortality (per 1,000)							
< 1 km	−0.403	−0.405	−0.620*	−0.873**	−0.853**	0.321	0.127
	(0.365)	(0.365)	(0.368)	(0.341)	(0.341)	(0.601)	(0.409)
1–2 km	−0.550*	−0.552*	−0.779**	−0.789***	−0.770***	0.013	−0.063
	(0.311)	(0.311)	(0.315)	(0.299)	(0.299)	(0.488)	(0.324)
2–4 km	−0.608*	−0.613**	−0.762**	−0.751**	−0.718**	0.133	−0.169
	(0.311)	(0.311)	(0.310)	(0.293)	(0.293)	(0.493)	(0.316)
4–7 km	−0.474	−0.478	−0.524*	−0.545*	−0.500*	−0.308	−0.069
	(0.315)	(0.316)	(0.316)	(0.301)	(0.301)	(0.489)	(0.311)
7–11 km	−0.650**	−0.652**	−0.669**	−0.649**	−0.629**	0.524	0.270
	(0.325)	(0.325)	(0.326)	(0.309)	(0.309)	(0.528)	(0.323)
C. Dependent Variable: Apgar Score							
< 1 km	0.006	0.006	0.006	0.013	0.020**	0.003	0.009
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.013)	(0.010)
1–2 km	−0.014*	−0.014*	−0.014*	−0.007	−0.003	−0.011	−0.008
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)
2–4 km	−0.003	−0.003	−0.003	0.001	0.006	−0.008	−0.004
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)	(0.008)
4–7 km	−0.002	−0.002	−0.002	−0.001	0.004	−0.007	−0.005
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)
7–11 km	0.014*	0.014*	0.014*	0.015*	0.016**	−0.009	−0.007
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)
Time Effects		X	X	X	X	X	X
Mother's Char.			X	X	X	X	X
Infant Char.				X	X	X	X
Avg. HH Income					X	X	X

Notes: See the text in section 5.2.1 for more explanations.

Table 6: Means of Observable Characteristics by Median Distance and Location of Birth

	Hospital		Home	
	Distance $\leq$ 4km (1)	Distance > 4km (2)	Distance $\leq$ 4km (3)	Distance > 4km (4)
A. Mother's Characteristics				
Age	28.129	28.622	28.616	28.350
Ethnicity: Dutch	0.713	0.889	0.870	0.957
Ethnicity: Mediterranean	0.114	0.035	0.046	0.012
B. Infant Characteristics				
Boy	0.518	0.520	0.487	0.488
Birth weight	3402	3435	3402	3414
Gestational Age (days)	279	279	279	279
Obstetrician Supervision	0.686	0.747	0	0
Multiple Birth	0.001	0.001	0	0
Breech Birth	0.013	0.018	0.001	0.001
C. Average Postal Code Characteristics				
Monthly Household Income (euro)	1954	1992	1986	1988
Density	2851	909	2736	855
Percent 0–15 years old	17.563	20.106	17.605	20.478
D. Predicted Outcomes				
Predicted 7-Day Mortality	2.335	2.308	0.684	0.660
Predicted 28-Day Mortality	2.585	2.561	0.779	0.755
Predicted Apgar Score	9.600	9.587	9.794	9.797
N	142,214	57,296	99,305	57,597

Notes: For a description of the variables, see section 4.

Table 7: 2SLS Estimation of the Effects of Hospital Births on Infant Outcomes

	7-Day Mortality (1)	28-Day Mortality (2)	Apgar Score (3)
A. OLS Estimates			
Hospital	-0.001	-0.072	-0.061^{***}
	(0.155)	(0.163)	(0.004)
B. 2SLS Estimates			
Hospital	-8.287^{***}	-9.219^{***}	-0.018
	(3.157)	(3.353)	(0.088)
Hausman	6.533	6.903	0.263
P-value	0.011	0.009	0.608
N	356,412	356,412	355,761

Notes: Each column presents results for a different health outcome using a specification that includes our baseline set of controls. Panel A provides OLS estimates, while Panel B provides 2SLS estimates. Robust standard errors clustered at the postal code level are shown in parentheses. The last two rows report the Hausman statistic and the associated p-value testing the consistency of both the OLS and 2SLS estimates against the alternative of the consistency of the 2SLS estimates only. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Robustness Checks

	7-Day Mortality	28-Day Mortality	Apgar Score
	(1)	(2)	(3)
A. Alternative Samples			
Adding Congenital Anomalies			
Hospital	-8.474^{***}	-9.489^{***}	-0.033
	(3.282)	(3.460)	(0.088)
N	360,817	360,817	360,151
Adding Stillbirths			
Hospital	-10.862^{***}	-11.742^{***}	0.010
	(3.791)	(3.905)	(0.090)
N	356,817	356,817	356,166
All Low-Risk Births			
Hospital	-3.650^{**}	-4.369^{***}	0.005
	(1.483)	(1.567)	(0.061)
N	788,294	788,294	787,010
All Births			
Hospital	-5.253^{*}	-5.971^{*}	0.012
	(3.161)	(3.245)	(0.116)
N	1,478,187	1,478,187	1,476,118
B. Referral Patients Classified As Home Births			
Hospital	-9.098^{***}	-9.886^{***}	-0.073
	(3.450)	(3.658)	(0.098)
N	356,412	356,412	355,761
C. Car Ownership			
Car ownership < median			
Hospital	-7.698^{**}	-9.050^{**}	0.113
	(3.543)	(3.827)	(0.099)
N	263,854	263,854	263,359
Car ownership > median			
Hospital	-6.407	-8.780	-0.296
	(6.944)	(7.145)	(0.214)
N	92,558	92,558	92,402
D. Additional Postal Code Characteristics			
Hospital	-6.364	-7.890^{*}	-0.046
	(3.917)	(4.138)	(0.115)
N	356,412	356,412	355,761

Notes: Each cell represents results for a different health outcome using a specification that includes our baseline set of controls. For a description of the robustness checks in each Panel, see section 5.3. Robust standard errors clustered at the postal code level are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Heterogeneous Results

	7-Day Mortality (1)	28-Day Mortality (2)	Apgar Score (3)
Mother's Ethnicity: Dutch			
Hospital	-8.219^{**}	-9.336^{**}	0.003
	(3.412)	(3.648)	(0.093)
N	294,671	294,671	294,099
Mother's Ethnicity: Non-Dutch			
Hospital	-12.866	-10.986	-0.185
	(12.069)	(12.081)	(0.233)
N	61,741	61,741	61,662
Mother's age $\leq$ median (29 years)			
Hospital	-8.210^{***}	-8.619^{***}	-0.001
	(2.837)	(3.003)	(0.080)
N	207,598	207,598	207,217
Mother's age $>$ median (29 years)			
Hospital	-9.618	-12.546	-0.107
	(10.534)	(10.993)	(0.214)
N	148,814	148,814	148,544
Gestational age $\leq$ median (280 days)			
Hospital	-7.267^{*}	-10.214^{**}	-0.069
	(4.365)	(4.730)	(0.092)
N	179,213	179,213	178,868
Gestational age $>$ median (280 days)			
Hospital	-8.497^{**}	-6.942	0.031
	(4.213)	(4.309)	(0.123)
N	177,199	177,199	176,893
Birth weight $\leq$ median (3,410 grams)			
Hospital	-6.444	-9.291^{*}	-0.012
	(4.689)	(5.063)	(0.097)
N	178,346	178,346	178,050
Birth weight $>$ median (3,410 grams)			
Hospital	-9.470^{**}	-8.197^{**}	-0.042
	(3.874)	(3.901)	(0.120)
N	178,066	178,066	177,711
Avg. HH Income $\leq$ median (€1,929)			
Hospital	-12.648^{***}	-15.634^{***}	0.102
	(4.743)	(5.069)	(0.123)
N	178,218	178,218	177,863
Avg. HH Income $>$ median (€1,929)			
	-1.027	0.122	-0.318^{*}
	(5.493)	(5.956)	(0.174)
N	178,194	178,194	177,898

Notes: Each cell represents results for a different health outcome using a specification that includes our baseline set of controls. Robust standard errors clustered at the postal code level are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Complier Characteristics by Distance

	< 1km (1)	1–2 km (2)	2–4 km (3)	4–7 km (4)	7–11 km (5)	N (6)
Mother's Ethnicity: Dutch	1.12	1.05	1.26	1.04	0.93	294,671
Mother's Age $\leq$ median (29 years)	1.25	1.34	1.36	1.13	1.40	207,598
37 wk $\leq$ Gestational Age $\leq$ 42 wk	1.02	1.04	1.05	1.03	1.05	337,830
Birth weight $\leq$ median (3410 grams)	1.05	1.08	1.26	1.15	1.29	178,346
Share of compliers	0.027	0.030	0.013	0.015	0.021	356,412

Notes: Each cell in columns 1–5 gives the relative likelihood that the compliers corresponding to the distance indicator in the column have the characteristic described in the row. The last row shows the fraction of compliers corresponding to the distance indicator in the column in the analysis sample.

Table A1: Hospital Birth and Infant Outcomes – Probit Models, Analysis Sample

	(1)	(2)	(3)	(4)	(5)
Dependent Variable: 7-Day Mortality					
Hospital	0.002284***	0.002266***	0.002195***	0.000300	0.000302
	(0.000234)	(0.000234)	(0.000235)	(0.000234)	(0.000233)
N	356,412	356,412	356,331	356,331	356,331
Dependent Variable: 28-Day Mortality					
Hospital	0.002409***	0.002371***	0.002298***	0.000240	0.000242
	(0.000237)	(0.000236)	(0.000239)	(0.000244)	(0.000244)
N	356,412	356,412	356,331	356,331	356,331
Time Effects		X	X	X	X
Mother's Char.			X	X	X
Infant Char.				X	X
Avg. HH Income					X

Notes: Each cell presents the average marginal effect of a hospital birth from a different regression and corresponds to the effects in Table 2 divided by 1,000. Time effects include indicators for year, month and day-of-week of birth. Mother's characteristics include indicators for maternal age and ethnicity groups. Infant characteristics include gestational age, a third degree polynomial in birth weight and indicators for male, multiple birth, obstetrician supervision and breech birth. Average household income refers to the average income in the postal code of mother's residence. For more information, see section 4. Robust standard errors clustered at the postal code level are shown in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Average Observable Characteristics by Distance

	< 1km (1)	1–2 km (2)	2–4 km (3)	4–7 km (4)	7–11 km (5)	≥ 11 km (6)
A. Mother's Characteristics						
Age	28.287	28.255	28.275	28.635	28.560	28.253
Ethnicity: Dutch	0.739	0.728	0.794	0.897	0.923	0.933
Ethnicity: Mediterranean	0.105	0.114	0.073	0.031	0.023	0.022
B. Infant Characteristics						
Boy	0.509	0.509	0.508	0.510	0.507	0.508
Birth weight	3400	3400	3404	3424	3425	3437
Gestational Age (days)	279	279	279	279	279	279
Obstetrician Supervision	0.493	0.487	0.490	0.482	0.467	0.463
Multiple Birth	0.001	0.001	0.000	0.000	0.000	0.000
Breech Birth	0.010	0.009	0.010	0.012	0.012	0.012
C. Average Postal Code Characteristics						
Monthly HH Income (euro)	1943	1935	1997	2057	1970	1892
Density	3338	3532	1957	1110	746	660
Percent 0–15 years old	16.116	16.578	19.056	20.453	20.072	20.067
D. Predicted Outcomes						
Predicted 7-Day Mortality	2.205	1.799	1.780	1.759	1.671	1.652
Predicted 28-Day Mortality	2.429	2.003	1.981	1.955	1.868	1.843
Predicted Apgar Score	9.652	9.655	9.658	9.663	9.665	9.664
N	32,887	80,193	86,430	70,619	49,446	36,837

Notes: For a description of the variables, see section 4.

Table A3: Distance and Infant Outcomes – Reduced Form Estimates, Probit Models

	7-Day Mortality (1)	28-Day Mortality (2)
< 1 km	-0.000896^{***}	-0.000988^{***}
	(0.000320)	(0.000335)
1–2 km	-0.000777^{***}	-0.000804^{***}
	(0.000288)	(0.000302)
2–4 km	-0.000703^{**}	-0.000799^{***}
	(0.000283)	(0.000296)
4–7 km	-0.000502^{*}	-0.000627^{**}
	(0.000295)	(0.000307)
7–11 km	-0.000690^{**}	-0.000787^{**}
	(0.000302)	(0.000311)
N	356,331	356,331

Notes: Each column represents a different regression. The dependent variable is the probability of death in the first 7 days (column 1) or 28 days (column 2) after birth. All specifications include our baseline set of controls. Each cell presents the average marginal effect of the corresponding variable and corresponds to the effects in Table 5 divided by 1,000. Robust standard errors clustered at the postal code level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Means of Observable Characteristics by Distance and Location of Birth

	Hospital						Home					
	< 1 km	1–2 km	2–4 km	4–7 km	7–11 km	≥ 11 km	< 1 km	1–2 km	2–4 km	4–7 km	7–11 km	≥ 11 km
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Mother's Characteristics												
Age	28.136	28.077	28.177	28.685	28.666	28.425	28.685	28.720	28.504	28.541	28.377	27.997
Dutch	0.694	0.681	0.752	0.870	0.900	0.913	0.856	0.851	0.891	0.949	0.962	0.963
Mediterranean	0.124	0.136	0.090	0.041	0.031	0.028	0.055	0.055	0.034	0.013	0.009	0.013
B. Infant Characteristics												
Boy	0.517	0.519	0.516	0.521	0.520	0.519	0.488	0.483	0.490	0.490	0.484	0.491
Birth weight	3399	3399	3405	3430	3435	3447	3402	3404	3400	3413	3406	3423
Gestational Age	279	279	279	279	279	279	279	279	279	279	279	279
OB/GYN Supervision	0.680	0.674	0.701	0.740	0.741	0.774	0.000	0.000	0.000	0.000	0.000	0.000
Multiple Birth	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
Breech Birth	0.013	0.012	0.014	0.018	0.018	0.019	0.001	0.000	0.001	0.001	0.001	0.001
C. Average Postal Code Characteristics												
Avg. HH Income	1934	1927	1989	2055	1971	1891	1967	1958	2017	2061	1968	1892
Density	3387	3539	1979	1132	765	650	3209	3513	1907	1067	714	675
Percent 0–15 y.o.	16.069	16.670	19.011	20.336	19.924	19.884	16.242	16.340	19.160	20.673	20.326	20.340
D. Predicted Outcomes												
7-Day Mortality	2.744	2.257	2.248	2.315	2.265	2.352	0.817	0.640	0.675	0.643	0.667	0.677
28-Day Mortality	3.014	2.505	2.493	2.563	2.524	2.607	0.921	0.728	0.772	0.743	0.758	0.773
Apgar Score	9.599	9.602	9.599	9.591	9.587	9.577	9.793	9.793	9.796	9.798	9.798	9.794
N	23,859	57,942	60,413	46,039	31,202	22,064	9,028	22,251	26,017	24,580	18,244	14,773

Notes: For a description of the variables, see section 4.