Understanding Returns to Birthweight

Shiko Maruyama *
ARC Centre of Excellence in Population Ageing Research (CEPAR)
The University of New South Wales
Sydney, NSW 2052 Australia
Phone: +61 2 9385 7685
Email: s.maruyama@unsw.edu.au

Eskil Heinesen
Rockwool Foundation Research Unit
Sølvgade 10, 2.tv. DK 1307
Copenhagen K Denmark
Phone: +45 3334 4806
Email: esh@rff.dk

*Corresponding author

THIS VERSION: September 17, 2013

Understanding Returns to Birthweight

Abstract

Worse future outcomes of low birthweight infants are well documented, but the causal relationship remains controversial, because birthweight itself might only be a composite variable of infant health, obstetric complications, and other unobserved factors. It might merely reflect the fact that children with worse future outcomes are more likely to be born with low birthweight. To quantify and understand the causal effects of birthweight on various outcomes, we apply an instrumental variable approach and analyze the population data of over 1.7 million newborns in Denmark since 1981. Our instrument is a diagnosis of placenta previa, an obstetric complication that often results in low birthweight. Placenta previa is highly unpredictable when its risk factors are controlled for, providing a desirable exogenous randomization. Unlike most other obstetric complications, it has limited long-term impacts on children except for its effects through low birthweight. Studying an extensive range of outcomes, we find the following: (1) birthweight has significant positive effect on health-related outcomes, especially reducing infant mortality and permanent disability, whereas a large part of correlation between birthweight and non-health outcomes is not causal; (2) the effect of birthweight diminishes as the child becomes older (catch-up effect); and (3) returns to birthweight in Denmark have significantly diminished over the last three decades.

Keywords: Birthweight, Infant Health, Placenta Previa, Instrumental variable

1. Introduction

A large body of medical and social science literature has documented worse future outcomes of low birthweight infants. Researchers have found this association for both short-term and long-term outcomes, including infant mortality and morbidity, long-term disabilities, chronic health problems, educational attainment, labor market outcomes, and even the birthweight of their children. These findings provide grounds for policy concerns. Low birthweight infants are expensive to treat in hospitals (Almond et al., 2005). They are more likely to experience health and social difficulties in leading an independent life (Conley, 2003) and have a higher chance of welfare take-up (Oreopoulos et al., 2008). Birthweight has implications for intergenerational mobility if it affects the health of children in the next generation (Royer, 2005; Black et al., 2007; Currie and Moretti, 2007).

In spite of the well-documented strong association between birthweight and later outcomes, its causal mechanism is still controversial. On the one hand, the medical literature has accumulated physiological understanding on how low birthweight and premature birth cause neonatal complications and associated disabilities, such as cerebral palsy and neurodevelopmental impairment. On the other hand, for long-term outcomes, innumerable potential transmission pathways exist, and hence disentangling causal mechanisms is a formidable task. Although not a few studies have compared the future outcomes of premature and normal infants, the interpretation of such results is not straightforward, because there are a number of variables that may affect both birthweight and future outcomes: e.g. physical conditions of mothers, genetics, mental health status, and various socioeconomic variables of the family. While low birthweight may have a negative effect on a future outcome, it is also plausible that those with worse future outcomes are more likely to be born with low birthweight.

The aim of this paper is to quantify and understand the *causal effects of birthweight* on various future outcomes. We analyze the population data of 1.7 million newborns in Denmark since 1981. We investigate an extensively wide range of outcomes including infant mortality, postnatal disabilities, hospital admission, educational attainment, future reliance on social welfare, teen pregnancy, the birthweight of their children, criminal tendency, and the results at military conscription examinations including IQ and body size measures.

We apply an instrumental variable approach to overcome bias due to unobserved factors that affect both birthweight and future outcomes. The use of mother fixed effects or grandmother fixed effects, another standard approach in the literature, is insufficient. Although mother fixed-effects models assist us in eliminating most mother-specific confounders (such as genetics, body size, and behavioral inclination), such estimates are highly likely to be confounded by birth-specific unobservables underlying low birthweight, such as congenital anomalies of the fetus, prenatal infection and complications, maternal mental stress, and readiness and affection toward the newborn. Our solution to overcome these likely sources of bias is the use of the instrumental variable (IV) approach. Our instrument is a diagnosis of placenta previa, an obstetric complication that often results in low birthweight. Placenta previa is highly unpredictable when its risk factors are controlled for, providing a desirable exogenous variation in birthweight. In addition, it has limited long-term impacts on children except for its effects through low birthweight, unlike most other obstetric complications, such as maternal cancer and malformation of the fetus.

In the field of social science, a number of researchers have attempted to quantify the causal effect of birthweight by applying within-twin techniques (Behrman and Rosenzweig 2004; Almond et al., 2005; Miller et al., 2005; Conley et al., 2006; Black et al., 2007; Oreopoulos et al., 2008; Royer, 2009), relying on the logic that the bias due to parent- and birth-specific unobservable factors, such as genetics and in-utero environment, can be eliminated by looking at within-twin variations. While their findings are fairly mixed, they tend to find that twin fixed-effects estimators significantly reduce the effect of birthweight on infant mortality, but long-term effects tend to remain large and significant, a pattern that puzzles researchers (Black et al., 2007).[1] In this paper we also show that the twin fixed-effects approach indeed generates substantially different misleading results, and discuss a possible reason.

Our findings are summarized in three main conclusions. First, birthweight has significant positive effect on health related outcomes, especially reducing infant mortality and permanent disability. Regarding other outcomes, such as social, behavioral, cognitive, and body size outcomes, however, while the OLS results show highly significant seemingly favorable role of birthweight, the causal estimates show substantially weaker evidence sometime with an opposite sign, indicating that the

[1] Black et al. (2007), applying the twin fixed-effects approach to rich Norwegian register data, find that while the use of twin fixed effects substantially reduces the birthweight effect on short-term outcomes, such as one-year mortality, the long-term effects are not affected by twin fixed effects and remain significant at the same level as the pooled OLS results.

correlation is mostly not causal. Those who experience desirable future outcomes are more likely to be born with a larger birthweight. Second, the role of birthweight diminishes as the child becomes older. Birthweight is not a critical initial condition of a person in the sense that once an infant survives without permanent disability, the influence of birthweight gradually attenuates and he or she can catch up. Birthweight itself does not have self-reinforcing multiplier effect for one's future life. Third, the results regarding infant mortality and disabilities indicate that returns to birthweight in Denmark have significantly diminished over the last three decades. This fact calls for caution from researchers who study birthweight: the implication of birthweight may significantly vary across countries and time periods.

2. Empirical Framework

2.1. Econometric Specification

We estimate the effect of birthweight on a future outcome by

$$(1) \qquad \mathrm{y}_{ijt} = \alpha + \beta \cdot \ln\left(\mathrm{birthweight}_{ijt}\right) + \mathrm{X}_{ijt}\gamma + \xi_j + \varepsilon_{ijt},$$

where y_{ijt} represents an outcome of child *i* born to mother *j* at time *t*. $\ln\left(\mathrm{birthweight}_{ijt}\right)$ is birthweight in log and X_{ijt} is a vector of child- and mother-specific variables that may vary over time. ξ_j refers to mother-specific unobservables (e.g., mother's height and genetic factors), and ε_{ijt} is an unobservable error term. The parameter of primary interest is β.

We study a large set of outcome variables, and depending on which outcome we study, y_{ijt} and $\ln\left(\mathrm{birthweight}_{ijt}\right)$ may have a highly non-linear relationship. For example, when we study a binary outcome such as infant mortality rate, our linear functional form specification does not coincide with the true data generating process. Nevertheless, we rely on this linear functional form for several reasons. First, the mechanics of the linear regression estimator is widely known and hence transparent. Second, it facilitates the comparison of birthweight effects on different types of outcome variables based on the same simple framework. Third, the use of instruments and fixed effects is straightforward compared to non-linear models. Fourth, and most importantly, it is standard and fairly reasonable to use the linear regression framework to estimate the *average causal effect* at the population level, which is the main interest of this paper (e.g. Cameron and Trivedi, 2005).

Cross-sectional estimation of equation (1) by OLS (pooled OLS) will generally lead to a biased estimate of β because of unobservable factors in $\xi_j + \varepsilon_{ijt}$ that influence both birthweight and outcomes. A mother fixed-effects model allows us to eliminate bias due to ξ_j, relying on within-siblings variation. In our setup, however, this is unlikely to be satisfactory because there may be time-varying birth-specific unobservable factors. While various prenatal conditions and complications create a predisposition to low birthweight, they may also affect future outcomes through channels other than low birthweight. For example, congenital anomalies of the fetus may lead to low birthweight and future health problems. Maternal smoking is also known to be a risk factor for low birthweight and may relate to the child's future outcomes through various channels. Other examples include a mother's readiness and affection toward a new baby, mental stress, and health knowledge. These unobservable birth-specific factors generate a correlation between birthweight and the error term and thus violate the condition necessary for consistent estimation of β.

Our solution is the use of placenta previa as an instrument. An instrument needs to be correlated with birthweight, BW_{ijt}, but uncorrelated with the error term. In our context, the latter condition consists of two requirements. First, the instrument offers exogenous variation, or randomness, for birthweight. Second, the instrument must not affect y_{ijt} except through the channel of birthweight. Below, we argue that our placenta previa instrument reasonably satisfies these conditions.

2.2. Placenta Previa

The placenta is an organ that develops during pregnancy to connect the developing fetus to the uterine wall of the mother. Placenta previa refers to a placenta that overlies or is proximate to the cervix, as illustrated in Figure 1. The placenta normally implants in the upper uterine segment. In placenta previa, the placenta either totally or partially lies within the lower uterine segment. A United States population-based study for the years 1979-1987 finds the overall annual incidence of placenta previa to be 0.48 percent (Iyasu et al., 1993). Oyelese and Smulian (2006) provide a review of placenta previa.

[Insert Figure 1]

Women with placenta previa often present with painless hemorrhage, and placenta previa can be confirmed with an ultrasound. There is consensus that a placenta previa that totally or partially overlies the cervix requires cesarean delivery under controlled, scheduled conditions at an institution with adequate blood banking facilities. For most stable patients, to increase the chance of neonatal survival, it is standard to schedule cesarean delivery after 36 weeks of gestation. Significant bleeding, however, may necessitate further preterm delivery. More than half of births with placenta previa are preterm with low birthweight (<2,500 grams). Sugimoto (2007) reports 59 percent for preterm births, 41 percent for less than 32 weeks, 15 percent for below 28 weeks, 51 percent for low birthweight (<2,500 grams), 12 percent for less than 1,500 grams, and 3 percent for stillbirth and neonatal mortality.

The use of placenta previa as an instrument is motivated by the following two key facts. First, the occurrence of placenta previa is largely random, and in spite of numerous previous attempts in the literature, the reason some placentas implant in the lower uterine segment remains unclear (Benirschke and Kaufmann, 2000). For many obstetric complications that significantly affect birthweight, such as premature rupture of membranes, infections, and placental abruption, the medical literature has identified risk factors, some of which are related difficult-to-measure maternal variables, such as mental stress, alcohol intake, and health literacy. The incidence of placenta previa, however, is highly unpredictable once certain risk factors are controlled for. Risk factors for placenta previa include prior cesarean delivery, pregnancy termination, uterine surgery, smoking, cocaine use, multiple birth, increasing parity, and maternal age (Oyelese and Smulian, 2006; Aliyu et al., 2011A). Except for certain clinical risk factors, such as past cesarean section, most risk factors have relatively small relative risk. Overall, placenta previa occurs to every mother fairly randomly regardless of their socioeconomic characteristics. Aliyu et al. (2011A) report the relative risk of smoking to be 1.5, which is fairly small for our purposes. Aliyu et al. (2011B) find no association between alcohol consumption and placenta previa. The clinical information available in the fertility register allows us to control for most important risk factors of placenta previa, and the rich register data further assists us to control for unknown risk factors by including numerous characteristics of mothers.

Second, compared with other neonatal and obstetric complications that induce low birthweight, it is fairly reasonable to assume that placenta previa has limited direct long-term impact on children except for its effect through the channel of preterm birth and low birthweight. This is because, while many complications originate from pathological factors either in the body of mothers or fetus (e.g., maternal cancer and malformation of the fetus), placenta previa occurs merely as an abnormal position of placenta, a temporary organ that is discarded after birth. It may lead to serious hemorrhage and complications, which are certainly of clinical importance, but the incidence rates of these complications can be regarded as small enough for our purposes. For example, in the United States, maternal mortality occurs only in 0.03 percent of the cases of placenta previa (Oyelese and Smulian, 2006). Regarding the effect of cesarean section on the child and mother, there exists a line of research with mixed results. While most of them are observational studies that do not address selection bias, there are a few recent causal studies, which document insignificant or sometimes favorable causal effect of cesarean section (Jensen and Wust, 2012; Hannah et al., 2000).[2]

2.3. Causal Framework and Determinants of Birthweight

A causal study requires the researcher to conceptualize the causal effect. One possible way to define birthweight effect is to follow the ceteris paribus notion as strictly as

[2] Currie and MacLeod (2008) also provide an indirect evidence for the negligible causal effect of caesarean delivery.

possible and conceptualize birthweight effect as the pure effect of body size at birth, where we consider newborns who have different body weight, e.g. 1500 versus 4000 grams but share exactly the same characteristics other than weight, such as the same lung capacity and the same functioning of other organs. This puristic approach, however, is not an option for us because our instrument of placenta previa influences by nature not only birthweight but also the overall maturity of the baby. For this reason, taking a more practical approach, our birthweight effect reflects not only weight but also the development of newborns naturally associated with fetus growth. Very low birthweight infants are typically associated with prematurity in many organs such as brain, lung, and eye. The effect of such prematurity is included in our birthweight effect.

Low birthweight arises either due to preterm delivery (short gestation length) or due to low fetal growth, which is known as intrauterine growth retardation (IUGR). If these two causes have totally different impacts on future outcomes and if placenta previa influences only one of these, our causal estimate yields a misleading result. To verify this point, the relationship between birthweight, gestation days, intrauterine growth, and placenta previa is shown in Table 1. Intrauterine growth here is defined as birthweight divided by gestation days. Based on all singleton births between 1981 and 2010, the table shows that both gestation days and intrauterine growth are highly correlated with birthweight, they are also positively correlated with each other, and placenta previa reduces both gestation length and intrauterine growth. For this reason, we focus this paper on the overall birthweight effect and abstract from these two mechanisms, although we acknowledge that the distinction between preterm delivery and IUGR is important for medical research because it has significant implications for medical intervention.

[Insert Table 1]

3. Data

3.1. Data and Population Selection

We use data from the Danish administrative registers. The birth registry contains the population data of newborns in Denmark with information on birth date, unique person identifiers of the newborns and biological parents, and a range of clinical variables about the mother and infant. The register also contains stillbirths and abortions, and by using the mother's identifier, the complete fertility history of the mother can be constructed. The validity and coverage of the birth registry are considered to be of high quality (Blenstrup and Knudsen, 2011).[3] Using the person identifiers of newborns and parents, we match birth records to other Danish registers that provide information about demographics, families, hospital admissions, death, labor market outcomes, education, crime, and military conscription examinations.

Because many variables in the Danish registers become available only after 1980, we study births from 1981 and onwards. For one-year mortality analysis, we use all births over the period 1981-2010. When studying outcomes that have limited availability and long-term outcomes that require a longer period to follow, we use birth cohorts from narrower time windows accordingly. We exclude the following from our analysis:

[3] The validity of the reference linkage between parents and children is also confirmed: for children born between 1973 and 1989, fewer than 1 percent do not have a reference to a parent.

stillbirths, children whose mother is not identified in the birth registry, children born overseas (because birth-related information is not available for this group), and adopted children and their siblings. Children whose father is not identified are kept in our population; we include an indicator for newborns missing father information. A very small number of observations with missing values, highly unrealistic values, and other data problems are discarded. For clear interpretation, we also exclude births from multiple pregnancy because the distribution of birthweight is very different in multiple pregnancy.[4] We keep singletons who are siblings of twins.

In the literature, variants of the birthweight measure have been used. These include birthweight, log(birthweight), fetal growth (defined as birthweight divided by weeks of gestation), and an indicator variable for low birthweight (typically less than 2,500 grams and 1,500 grams). Given that there is no obvious choice a priori, we use log(birthweight) following Black et al. (2007). They examine the explanatory power of these variables in their twin fixed-effects regressions, and find that log(birthweight) fits best for all of their outcome variables, although other measures also tend to provide results consistent with the results based on log(birthweight). The use of log(birthweight) also conforms with our intuition –diminishing returns to birthweight. Lastly, while most medical studies employ certain categories of low birthweight or pre-term births mainly for clinical concerns, we are more interested in quantifying the overall birthweight effect at the population level.

[4] Multiple pregnancy is based on the number of fetuses, not the number of live-born infants.

An indicator variable for placenta previa – our instrument – is constructed by combining two registers: the birth register and the hospital admission register. We combine them so that the placenta previa takes on the value of one if either of the two data sources indicates placenta previa. In most cases, both registers provide consistent information. In the hospital admission register, for each pregnant mother, we collect all the inpatient episodes that overlap the time period between nine months prior to the birth and three days after the birth date. Each admission record can have multiple diagnoses. Our placenta previa indicator takes the value of one if we observe the ICD code of placenta previa as one of the diagnoses in any of those admission episodes within this time window.

3.2. Control Variables

The regression analysis below includes a number of control variables. In the literature, it is standard to use control variables that are defined at birth, but we construct control variables at conception when it is appropriate. This is because birthweight is highly related to the timing of birth and hence control variables defined at birth may cause endogeneity bias. The estimate of the conception date is available for most observations. In early years, the date is constructed based on the gestation week.[5]

The birth registry offers the basic characteristics of the mother and child, as well as obstetric and clinical information. From this registry, we construct the following explanatory variables: the sex of the child, year- and month-of-birth dummies,

[5] For a very few cases whose information about gestation is missing, we impute the gestation length based on birthweight, sex, maternal age, and other variables. This imputation has almost no effect on the results.

indicators for mother's age at conception (one indicator for every two years), an indicator for the first child, birth order, indicators for past pregnancies (1, 2, 3, and 4+), the number of past cesarean sections, and indicators for the mother's past spontaneous and induced abortions (1 and 2+), past stillbirths, and smoking habits. From the hospital admission registry, we construct the number of days the mother spent in the hospital during the 180 days around the conception except for obstetrics-related admission. We construct this variable in this particular way to capture the mother's general health that is not the consequence of placenta previa.

The other demographic and socioeconomic variables are constructed from various administrative registers: indicators for the mother's highest education completed (less than 9 years, 9 years, upper secondary, low- and mid-tertiary, and high tertiary education), indicators for formal marital status and whether the biological father lives together on January 1st prior to birth, and an indicator for the immigrant status of either the mother or father. The mother's working status and gross income in the previous year are constructed and further interacted with conception month dummies to account for the effect of pregnancy on the previous year's labor supply. If the biological father lives with the mother, the father's age at conception, income and working status in the previous year, and highest education completed (the same grouping as the mother's education indicators above) are also included. We also include county dummy variables in all the regressions.[6]

[6] Under the legislation until 2006, Denmark consists of thirteen counties and three major municipalities, for which we construct fourteen dummy variables with Copenhagen and Frederiksberg municipalities being the reference group. Although the Municipal Reform of 2007 replaces these countries with five regions, we maintain the same county definitions for the entire period of analysis.

Table 2 reports the summary statistics of the characteristics of children and parents. Statistics are broken down into births without and with placenta previa in columns 1 and 2, respectively. In our population that covers from 1981 to 2010, 6,298 births are associated with placenta previa and its incidence rate is 0.35 percent, which is within the range found in the literature. For example, Oyelese and Smulian (2006) report that placenta previa complicates 0.3-0.5 percent of pregnancies. The group with placenta previa shows a 16.2 percent lighter mean birthweight.

[Insert Table 2]

Figure 2 shows the distributions of birthweight by placenta previa status. Comparing the two distributions highlights a higher frequency of low birthweight neonates when placenta previa complicates pregnancy. Variance also becomes larger, and thus the distribution still covers the possible range of birthweight. The fact that placenta previa provides exogenous variation in birthweight all over the possible range of birthweight is essential in identifying the average causal effect of birthweight at the population level.

[Insert Figure 2]

The comparison of the control variables between the two groups suggests that most of the variables identified in the literature as risk factors of placenta previa show expected differences. In the mean time, some variables that are not identified as risk factors in the literature show some differences. However, most of these unexpected differences either disappear or substantially reduce when we regress the placenta previa indicator on these control variables.

3.3. Outcome Variables

Table 3 summarizes the outcome variables with their definitions and descriptive statistics. The table also reports the population used in the analysis of each outcome. For example, we study most outcomes conditional on the child's survival up to a particular age to facilitate interpretation. Although this conditioning may result in selection bias, in general such bias does not significantly change our results because of the very low child mortality rate. The birth cohorts used for each outcome variable are determined by the availability of data and the number of years after birth necessary to observe the outcome variable. In the rest of this section, we discuss the details of each outcome variable.

[Insert Table 3]

3.3.1. Outcome Variables: Health

We study the following health-related outcome variables:

- *Infant mortality.* This is defined as mortality within the first 365 days of life, conditional on live birth.
- *Permanent disability*. The medical literature identifies low birthweight as a risk factor for permanent disability. We construct an indicator variable for permanent disability that consists of three conditions: cerebral palsy, loss of vision, and hearing impairment. These are identified as a diagnosis in the hospital admission

register. We identify cerebral palsy from diagnosis records regardless of the age of the child, while the other two conditions are identified based on the diagnosis records up to the second birthday, as vision and hearing impairment may occur later for reasons unrelated to birth.

- *Number of days in hospital before the second birthday.* This variable includes days spent in the hospital immediately after birth.
- *Hospitalization*. A series of indicator variables are constructed for different age brackets to investigate how the birthweight effect changes as a child grows older.

3.3.2. Outcome Variables: Education, Social Welfare, and Other Socioeconomics

We also study a number of medium- and long-term socioeconomic outcomes:

- *Grade 9 completion*. We use an indicator for whether a child has completed the ninth grade (i.e., compulsory education) by the year the child reaches age 16. In Denmark, the vast majority of children start the first grade in elementary school in the calendar year in which they turn seven years of age. Some parents, however, choose to start one year later and thus complete the ninth grade one year later. Given that drop-outs and grade retention are rather rare in Denmark, this variable primarily reflects children’s school entry decision at age 7.
- *Test scores*. In Denmark, all students take the compulsory exam at the end of the ninth grade, and the register data of exam marks exists in 2002 and onward. We use the standardized scores in four mandatory subjects: Danish, Mathematics, English, and Science (Physics and Chemistry). In addition to the average marks of the four subjects, we also use the mark in each individual subject. For Danish, the marks from the written test and oral test are available separately; thus, in

total we report the results of the five individual marks in addition to the overall average. Although these subjects are mandatory for all students, marks are occasionally missing, so the number of observations varies across subjects. Even if a student has a missing subject, we retain his/her other scores and compute the overall mean, as long as we observe the majority of the exam marks of the student.[7] In addition, the age of students at the end of the ninth grade varies for numerous reasons. We include only students whose exam scores are recorded when they are 15, 16, or 17 years old.[8]

- *Disability pension*. In Denmark, various social welfare supports are available for individuals of age 18 or older. Disability pension is paid for individuals who have serious permanent disability that prevents them from working. We use an indicator variable for the receipt of disability pension during the three calendar years of age 19 to 21. We also construct the number of weeks of disability pension received and the total amount of disability pension transferred during the same time window.
- *Other welfare assistance*. This variable indicates the receipt of welfare benefits other than disability pension that is related to labor market attachment. We use the same time window – the three years of age 19 to 21. These benefits are means-tested. Self-funded unemployment insurance and maternal benefits are

[7] To be more precise, the exam score variables are constructed based on seven exam marks available in the raw data: Danish oral, Danish written presentation, Danish written spelling, math written arithmetic, math written problem solving, English oral, and science. If a student lacks three or more marks out of these seven marks, we do not use the rest of the marks, considering that this student failed.
[8] In the raw data, 4.5 percent are excluded because of too few marks. 1.2 percent are excluded because they are 18 years or older. Fewer than 0.1 percent graduate before age 15.

not included. We also construct the number of weeks of welfare support and the total cash amount.

- *Work and student status.* We construct this variable for individuals aged 22 whose status is either working or student.
- *Gross income in the calendar year of age 22*. Although this variable is of considerable interest, its interpretation is not straightforward because at age 22, many individuals are still in tertiary education.
- *Teen pregnancy*. We construct a dummy variable that indicates pregnancy that started by the 20th birthday. This variable is constructed for males and females separately, and this does not depend on the birth outcome, i.e., induced and spontaneous abortions and stillbirths are included.
- *Marital status on January 1st in the year of age 22*. Similar to the gross income outcome, the interpretation of this outcome is not straightforward because in modern Denmark, marriage by age 22 is rather rare, and hence it does not necessarily indicate a successful marriage market outcome.
- *Birthweight of the first child by age 22*. This outcome variable is constructed only for those whose first child is born by the end of the year they reach age 22.
- *Criminal offense*. Based on the crime register, we construct three indicators for whether a person has a criminal charge and sentence by the 20th birthday: (1) any criminal sentence; (2) any criminal sentence that is probation or unconditional (this is to analyze more serious criminal offenses); and (3) any charge of a violent crime (regardless of its sentence status).[9] Because the age of

[9] In this violent crime variable, we include the cases in which the person is eventually acquitted.

criminal responsibility in Denmark is 15 years, criminal charges before age 15 are not included in our data.[10] Traffic offences are not included.

3.3.3. Outcome Variables: Military Conscription Variables

In Denmark, all men are required to attend an examination session for military conscription within a year after they turn 18. The sessions take place almost every day all year round. The purpose of the conscription examination is to assess the suitability of individuals for military duty. Although attendance at a conscription examination within a year after they turn 18 is mandatory, there are some exceptions. First, if one is physically handicapped, has serious psychological disorders, or has been in jail for more than 30 days, the army can exclude the person. Second, it is also possible to defer the session until the end of the year one turns 26 if the person is a student. Although these are a potential source of bias, the majority of males take the examination within two years after their 18th birthday,[11] and thus, we expect that the military variables still provide a better understanding of birthweight effect. The register also includes females who take the examination, but we do not include them due to the very small number of observations.

[10] Criminal charges before age 15 are extremely rare. The age of criminal responsibility is changed to 14 in 2010 and reverted to 15 in 2011. We focus on criminal offense after age 15 for consistency across years, though this change does not affect the results.

[11] In our data, 76 percent of males of age 18 or older have records of session attendance. Among these session attendants, national statistics documents that approximately 48 percent are 18 years old, 27 percent 19 years old, and 25 percent 20 years old or older. In Denmark, an individual can choose to serve duty as a conscientious objector, but the conscription session attendance is mandatory. For those who do not show up for a required session, there are penalties including fines and arrest. In a small number of cases where multiple session records are observed for one person, we employ the data of the session with the earliest date.

The military conscription register is available only from 2006 to 2011. From this data, we use whether a person is qualified for military duty (including qualification with some restriction), IQ test, height, weight, BMI, and color vision deficiency. IQ is measured by a test called Børge Priens test. The scores go from 0 to 78, with its distribution reasonably close to a normal distribution. We use its standardized score. We also study color vision deficiency as a placebo test to examine whether our regression model behaves as expected and whether the above-mentioned selection process of the conscription examination creates a significant bias. Because the occurrence of color vision deficiency is predominantly hereditary, it provides an opportunity to verify potential bias by testing whether we observe a significant effect of birthweight on color vision deficiency.

4. Results

Before turning to the causal effect estimates, it is useful to discuss the first-stage regressions. Table 4 reports the estimated coefficients of selected control variables based on the entire cohorts (those born 1981-2010) for the three regression models used below: [1] OLS; [2] OLS with grandmother fixed effects; and [3] OLS with mother fixed effects. Standard errors that are robust to arbitrary heteroskedasticity and correlation within grandmother cluster are shown in parentheses. As shown in the first row, the presence of placenta previa reduces birthweight by approximately 20 percent in all regression models even after controlling for a wide range of potential risk factors for placenta previa. The coefficients are statistically significant in all the three specifications at the 0.1 percent level, once again indicating the identification power of placenta previa as an instrument. This strong correlation implies that a tiny violation of the validity condition of the instrument is unlikely to cause a significantly biased result. The F statistics for the relevance of this IV is over 1500, and thus a weak instrument and associated finite sample bias are not a concern.

[Insert Table 4]

The large number of observations also provides precise estimates for the other controls, and the factors identified in the medical literature show the expected signs. Birthweight tends to be low for female newborns of mothers with past cesarean sections, poorer health, and smoking habits. Birthweight increases with birth order, in particular from the first to the second child. Importantly, socioeconomic factors have significant predictive power for birthweight. The presence of the father increases birthweight. Being married also increases birthweight though to a lesser extent. Birthweight increases with parental education, where maternal education plays a larger role than paternal education. Birthweight increases with the work status and income of the father. The same relationships hold for the mother's work status and income, although they are not reported here because the maternal work and income terms are interacted with conception month dummies and thus requiring more space in the table. Similarly, though this is not reported either due to the large number of dummy variables, birthweight decreases with parental age. Lastly, the infants of immigrants have lower birthweight.

Clearly these socioeconomic factors influence the future outcomes of the child through various transmission mechanisms. Although we can control for the above-mentioned variables, the results shown in this table suggest the existence of other innumerable unobservable confounders that affect both birthweight and future outcomes of the child, and thus an appropriate econometric method is crucial to disentangling causal effects.

4.1. Returns to Birthweight: Infant Mortality

Tables 5 to 10 report the estimates of the birthweight effects by outcome variable, starting with infant mortality in Table 5. Each row in each table reports the coefficient of ln(birthweight) for each outcome variable, and each column represents a different regression model. We report the results from six regression models: (1) OLS; (2) instrumental variable regression; (3) OLS with grandmother fixed effects; (4) instrumental variable regression with grandmother fixed effects; (5) OLS with mother fixed effects; and (6) instrumental variable regression with mother fixed effects. Though not reported in the tables, the control variables discussed earlier are included in all regression models. In parentheses below each coefficient estimate, we report standard errors that are robust to arbitrary heteroskedasticity and clustering by the grandmother identifier, which allows statistical dependence among siblings and cousins.[12]

[Insert Table 5]

The first row, [1], in Table 5 shows the effect of ln(birthweight) on one-year mortality. The OLS coefficient of -0.0761 implies that a 10 percent increase in birthweight would reduce one-year mortality approximately by 7.61 deaths per 1,000 births. When mother fixed effects are applied, this number becomes 13.3 deaths, and the coefficient in the

[12] Compared to clustering by the mother identifier, the use of grandmother identifier provides more conservative standard errors.

grandmother fixed-effects model is between these two estimates. The use of the placenta previa instrument considerably increases standard errors, yet produces estimates largely consistent with the OLS estimates. The positive effect of birthweight on one-year survival remains highly significant, indicating that a 10 percent increase in birthweight would reduce one-year mortality by 7.60 to 8.24 deaths per 1,000 births. This finding is rather different from the findings in twin fixed-effect studies: including twin fixed effects considerably reduces the birthweight effect on one-year mortality to less than half of OLS estimates.

The OLS and fixed-effects estimates allow us a further comparison with the literature. Our OLS and fixed-effects estimates are smaller than those that Black et al. (2007) find using Norwegian register data. While the main focus of their study is on twins, they also report the coefficients based on singleton siblings. Although Denmark and Norway share many similarities, their estimates are approximately 50 percent larger than ours: 12.3 and 18.7 deaths per 1,000 births for the pooled OLS and the mother fixed-effects model, respectively. This difference may arise because they use older cohorts – those born between 1967 and 1997. To clarify this point, we conduct the same estimation by separating the entire cohorts into two periods, one from 1981 to 1993 and the other from 1994 to 2010, and the results are reported in Rows [1A] and [1B] in Table 5. All six estimates show substantial reduction in the magnitudes over time. The main contributing factor for this reduction is the trend in the infant mortality rate, which has decreased from 0.66 percent in Period 1 to 0.34 percent in Period 2, reflecting the advancement in neonatal care. Given this time trend and given that the period used in

Black et al. (2007) is even older, our OLS and fixed-effects estimates are in line with their estimates, validating our data.

There is also a notable contrast between the time trends in the OLS and IV estimates. The reduction over time in the OLS estimates is not as large as the reduction in the infant mortality rate, while we see a much larger decrease in the IV estimates. A possible explanation for this finding is that low birthweight itself becomes less important than it used to be because of improved neonatal care, but unobserved confounders that link birthweight and infant mortality are still at work. This notable reduction in returns to birthweight calls for caution from researchers who study birthweight: the implication of birthweight may heavily depend on contexts such as cohorts and countries.

Rows [2A] and [2B] provide a robustness test that concerns variables that are supposed to be highly relevant to birthweight and infant mortality – maternal body size measures. Because the information of maternal height and weight are available only from 2004, we do not include these variables in the rest of our regression analyses. To verify the size of potential bias due to the omission of body size measures, we conduct the same infant mortality regressions with and without maternal height, BMI, and BMI squared. Though not reported here, the first-stage regression confirms strong positive effects of height and BMI on birthweight, implying that omitting these variables may cause non-negligible bias in causal effect estimates for the outcomes on which maternal body size has direct influence. Rows [2A] and [2B] show a mild difference in the OLS and IV estimates, indicating the overall reliability of our IV estimator. However, once grandmother- or mother-fixed-effects are applied, the exclusion of body size measures effectively makes no difference.

4.2. Returns to Birthweight: Health Outcomes

Table 6 reports the estimated birthweight effects on other health outcomes. Rows [1] – [1B] concern the long-term disability that originates from the perinatal period, and Rows [2] – [2B] the number of days in hospital before the second birthday. The negative and highly significant IV estimates highlight the importance of extra birthweight in avoiding permanent disabilities and long hospital stays.

[Insert Table 6]

For these two health outcomes, we again study the cohort effect by comparing the two periods. Unlike infant mortality, there is no reduction in the prevalence rate of permanent disability over time. This result is most likely due to the two opposing forces of improved neonatal care: while improved neonatal care enables newborns to avoid disabilities, it also saves the life of neonates with a greater risk who would have died in earlier periods. Nevertheless, we observe significant reduction in the magnitudes of the birthweight effects, in particular in IV estimates (Rows [1A] and [1B]). This provides a story consistent with the previous table: the current role of birthweight is not as important as it used to be prior to 1993. We do not observe this time trend for the length of the hospital stay (Rows [2A] and [2B]), which does not necessarily contradict with the previous findings because hospital days should be considered as not only health outcomes but also health input. Improved medical treatments may require a longer stay at the hospital to reduce infant mortality and morbidity.

It is also of interest to investigate how the birthweight effect varies over time, not only across cohorts but also within a cohort. In particular, whether the influence of this initial condition at birth grows over time or diminishes as the child becomes older is a potentially important empirical question. Rows [3A] to [3D] report the birthweight effect on hospital admissions by age of children. To delineate the age effect, we limit our population to those born between 1981 and 1991. The comparison of Rows [3A] to [3D] illustrates that, while the hospital admission rate does not vary much by age, both association and causation diminish as a child becomes older. Thus, although birthweight is crucial to infant mortality and infant health, surviving children can catch up to some degree.[13]

As we have seen so far, the IV estimator is accompanied with large standard errors, primarily because of the nature of our instrument – the rare occurrence of placenta previa. This insignificance makes the interpretation of results difficult. Throughout the paper, when we face insignificant IV estimates, we base our argument on the following logic. First, when an IV estimate is not significant at the conventional levels but nearly significant, we compare the estimate with the estimates of other regression models. If results do not vary much across the regression models, this fact makes the IV estimate somewhat more credible, even if the IV regression alone does not offer a significant estimate. Similarly, if an IV estimate shows a magnitude much larger than the other

[13] This effect can partly be attributed to attrition due to death, but we believe this selection plays only a minor role at most, given the low mortality rate after age 2.

estimates that are significant in the other regression models, we regard it as an indication of a birthweight effect. Second, we compare supposedly similar outcome variables, such as test scores of various subjects. If IV regressions for similar outcomes yield similar estimates, the results are considered to be more plausible. Another practical question when coefficient estimates vary across models is which model is more reliable. We favor fixed-effects estimates when the outcome variable is supposed to heavily depend on mother-specific factors, such as maternal body size and genetics, although the IV estimate without fixed effects should provide a consistent estimate as long as placenta previa occurs randomly. For outcomes for which genetics and maternal constitution are not supposed to play a major role, the IV estimate without fixed effects is preferred because of its smaller standard errors. Many observations without a sibling within the group of cohorts used are discarded in the fixed-effects models, which reduces the precision of the estimator. Furthermore, as shown below when we discuss the twin fixed-effects estimator, the use of mother fixed effects may yield a misleading estimate that is different from the population average birthweight effect, placing larger weights on low birthweight siblings.

4.3. Returns to Birthweight: Child Socioeconomic Outcomes

The outcome variables reported in Table 7 concern the educational attainment of children. The first outcome shown in Row [1] is whether a child completes the ninth grade by the year the child reaches age 16. As discussed in the data section, this outcome variable effectively captures the school entry decision at age 7, rather than later educational attainment. Similar to the results of infant health, all the models indicate the favorable effect of birthweight. The size of this effect can be interpreted more easily by comparing it with the effect of birth month, a strong predictor of ninth grade completion; naturally, children born in January are more likely to complete the ninth grade by the year they reach age 16 compared to children born in December of the same year. The estimated coefficients in Row [1] indicate that January-born children with 10 percent less birthweight have the same propensity to complete the ninth grade by the year of age 16 as those born between September and October with average birthweight. In the previous table, we report evidence of the catch-up effect, but the result here suggests that a significant impact of birthweight still remains at around age 7.

[Insert Table 7]

The outcome variables reported in the rest of Table 7 are the standardized scores of the national exam held at the end of the ninth grade. The OLS and IV estimates show a clear contrast. While the OLS estimates consistently show a highly significant positive impact of birthweight, the IV estimates show *negative* signs with a few exceptions. Those negative coefficients are imprecisely estimated because of large standard errors, but they have magnitudes comparable to the OLS results, and three of them are large enough to be statistically significant (the grandmother fixed-effects estimates in Rows [2], [4], and [5]). By and large, birthweight and exam scores are highly correlated most likely because of birth-specific confounders, such as congenital disorders and parental divorce, that are associated both with birthweight and test scores. Birthweight itself has no positive effect on exam scores, and it might even have a negative effect on test scores.

Table 8 reports the birthweight effects on disability pension and other welfare benefits. Although our results so far indicate the smaller effects of birthweight on longer-term outcomes, Rows [1A] to [1D] reveal a large birthweight effect on the disability pension take-up at around age 20. The IV estimates with mother fixed effects state that a 10 percent increase in birthweight reduces the amount of disability pension transfers during the three years by 3,132 Danish Kroner (≈ 420€) on average. Some serious disabilities originate from the perinatal period, and as we have seen, birthweight has a significant effect on cerebral palsy and neurodevelopmental impairment. The catch-up effect discussed above does not benefit those affected by these long-term disabilities. These children are at the lowest end of the distribution of the birthweight effect, heavily depending on social welfare assistance usually throughout the entire life.

[Insert Table 8]

Largely consistent results are found in the second half of Table 8. The causal effect estimates in Row [2A] are barely significant, most likely because this outcome variable is a noisy measure of welfare dependence as it includes various welfare benefits and the majority of the recipients receive a small amount of transfers in a short period. Rows [2B] to [2D], however, show that once we take into consideration how heavily an individual depends on social welfare, we observe significant influence of birthweight on welfare dependence. Note that, however, the results in this table are based on those born between 1981 and 1991. Because the impact of birthweight on infant health and permanent disabilities is smaller in recent years, we may well have a much smaller effect of birthweight on social assistance by now.

Table 9 reports the birthweight effects on the other socioeconomic outcomes. As shown in Row [1], birthweight increases the propensity that an individual works or studies as a student at age 22. Row [2] reports that birthweight and gross income at age 22 are positively correlated, but there is no strong evidence of a positive causal effect: the birthweight effect is estimated with a large standard error with different signs. The interpretation of these results is difficult because at age 22, many individuals are still students and income at this age is a weak predictor of lifetime economic success.

[Insert Table 9]

Teen parenthood is a major cause of poverty in early life. In Row [3A], the pooled OLS result shows a negative and highly significant association between birthweight and teen motherhood. This association, however, disappears when we apply fixed effects, and on the contrary, the estimates of IV and fixed-effects models have a positive sign even though they are insignificant. Row [3B] shows that some estimates for males are statistically significant at the 5 or 10 percent level, suggesting a positive birthweight effect on teen pregnancy. We observe larger effects for males than for females despite the fact that teen pregnancy is a much rarer event for males than females, indicating that the observed birthweight effect of males is relatively robust.

Similar results are found in the next four rows. Row [4] concerns marital status at age 22. The OLS results show that birthweight is significantly correlated with marital status, but once we apply the instrument, the estimates show a different sign or small magnitude with a large standard error, and thus the association is unlikely to reflect the causal effect of birthweight. In Rows [5], [5A], and [5B], we observe a highly significant positive correlation between the birthweight and their children's birthweight, but when the instrument is applied, much of the correlation disappears or turns to a negative effect. These results show a picture quite different from the conclusions of Black et al. (2007), Currie and Moretti (2007), and Royer (2005). These studies utilize within-sibling and within-twin variation and find a positive birthweight effect on the birthweight of the next generation. Our results indicate that much of this positive correlation is unlikely to be causal. Our results, however, are not definitive because of two limitations. First, the outcome we use here is conditional on that the child becomes a parent by age 22, which is considerably earlier than the average age, and hence this selection may cause bias. Second, because of the small number of parents by age 22, the number of observations is too small for the IV estimation with mother fixed effects. Validating these results is left for future work.

Rows [6A] – [6C] in Table 9 provide answers to whether birthweight reduces criminal offense. While the pooled OLS estimates in Rows [6A] and [6B] show a highly significant deterrence effect of birthweight on criminal offense by the age of 20, the fixed-effects and IV results indicate a *positive* birthweight effect on crime. The fixed-effects estimates are always positive and highly significant. The IV estimates are not statistically significant except for one, but the standard errors are relatively small and the magnitudes are consistently even larger than the fixed-effects estimates. Thus, low birthweight is likely to decrease criminal inclination.

4.4. Returns to Birthweight: Military Conscription Variables

Table 10 reports the birthweight effects on the outcome variables constructed from the military examination register. The first row states a significant causal effect of birthweight on the attendance at a conscription examination. Similarly, the second row shows an even larger birthweight effect on the qualification for military service. A 10 percent increase in birthweight raises the probability that the person is qualified for military by 2.2 to 7.6 percentage points. This effect can be explained by the birthweight effect on permanent disabilities to some degree, but not fully, given the prevalence rate of permanent disabilities. Those large magnitudes in Table 10 may suggest a possible birthweight effect on long-term general health and physical strength.

[Insert Table 10]

The rest of Table 10 requires caution in interpretation. As discussed earlier, the empirical setup for these variables are less clean than the outcome variables used so far because a quarter of the population do not attend a conscription exam session for various reasons and some of the attendants attend the session years after they turn 18. Nevertheless, it would be worthwhile to investigate these unique outcome variables. To provide some assurance regarding the credibility of the results below, we conduct a "placebo" test using color vision deficiency. Color vision deficiency is supposed to be completely hereditary, except for extremely rare trauma cases, and birthweight must have no predictive power for this variable. At the same time, color vision deficiency influences educational attainment, and tertiary education attendance is one of the main causes of missing conscription exam records. When we regress the average exam score on color vision deficiency, we find a negative coefficient that is significant at the 0.1 percent level. For this reason, the correlation between educational attainment and conscription session attendance may create bias in our estimate of the birthweight effect on color vision deficiency. The results in Row [3] document that this is not the case, showing no estimates with statistical significance. Row [3] also shows a typical behavior of our IV estimates: they tend to have larger standard errors and larger magnitudes (possibly in the opposite directions) compared to the OLS estimates, even when there should not be any effect.

Now we consider whether birthweight increases IQ, and our results show no strong support for this hypothesis. Row [4] shows a salient difference between association and causation: the results of the pooled OLS and fixed-effects OLS all show positive coefficients that are statistically significant at the 0.1 percent level, whereas all IV estimates are substantially smaller and not even close to statistical significance. This result provides evidence that those with a higher IQ are more likely to be born with higher birthweight.

The last three rows report the results of body size measures. Height, weight, and BMI are all highly significantly correlated with birthweight, but most of the correlation disappears when we use the instrument, and the estimates of the birthweight effect changes to negative when we apply mother fixed effects. The use of fixed effects makes a notable difference here because mother-specific factors, such as maternal constitution and genetics, significantly influence outcome variables, and thus we favor the mother fixed-effects estimator here. The results reveal that, against our common sense, birthweight has no major influence on body size at age 18. This finding suggests that the significant birthweight effects on the conscription session attendance and military qualification shown in Rows [1] and [2] are not driven by body size but something else, such as general health and physical strength.

5. Interpretation of IV Estimator under Heterogeneous Birthweight Effects

The effect of an increase in $\ln(\text{birthweight}_{ijt})$ is highly unlikely to be constant across births. This heterogeneity may capture a non-linear relationship between the outcome variable and $\ln(\text{birthweight}_{ijt})$ or may reflect the fact that birthweight has differential effects on the outcome of children with different characteristics (either X_{ijt}, ξ_j, or ε_{ijt}).

The aim of this study is to estimate the population average effect of $\ln(\text{birthweight}_{ijt})$. When $\ln(\text{birthweight}_{ijt})$ is statistically independent of unobservables, OLS regression of linear functional form (1) consistently estimates the population average effect of $\ln(\text{birthweight}_{ijt})$. However, when $\ln(\text{birthweight}_{ijt})$ is endogenous and an instrument is used, the IV estimator is in general no longer a consistent estimator of the population average treatment effect, with its bias depending on the specific characteristics of the instrument.

Although we cannot formally eliminate this potential bias in our IV estimator, we argue that our estimates are reasonably close to the population average birthweight effects because of the following desirable characteristics of our instrument. First, as discussed earlier, we can reasonably claim that placenta previa occurs to everybody. This is essential in estimating population average effects. If placenta previa occurred only to a certain subgroup, the results might not be applicable to individuals outside this subgroup.

Second, although the impact of placenta previa on birthweight varies considerably, we argue that this heterogeneity is largely orthogonal to observable and unobservable variables, because the seriousness of placenta previa depends on the position of placenta, which is highly unpredictable. The impact of placenta previa on birthweight is unlikely to be correlated with maternal characteristics also because there is not much a mother can do to reduce the impact of placenta previa, especially in a country like Denmark with a universal health care system with free access. Third, placenta previa never increases birthweight, and therefore the monotonicity assumption is satisfied. For these reasons, we expect that our IV estimator produces estimates that are close to population average birthweight effects.

6. Comparison of Twins and Singletons

As discussed earlier, most recent causal studies of birthweight effect rely on the twin fixed-effects estimator. An important remaining task for us is to quantify and understand the difference between the twin fixed-effects approach and our IV approach, if any. To quantify the difference between the two approaches, we apply the same empirical framework to twin observations extracted from the same data source. For a clear interpretation, we only include twins who are live-born from pregnancy with two fetuses.

Table 11 shows the comparison for selected outcome variables. In each row, the first two columns report the estimated birthweight effects based on the twin sample; the first column reports OLS estimates and the second column fixed-effects estimates. For comparison, the remaining four columns show singleton estimates that have already been presented so far. In each row, the short description of each outcome variable is followed by the mean values of the outcome variable and the numbers of observations for both twin and singleton samples. In addition to five selected outcome vibrations, the results regarding infant mortality from Black, Devereux, and Salvanes (2007) are presented in the second row, Row [1-BDS], for comparison.

[Insert Table 11]

Overall, Table 11 shows that the twin fixed-effects estimator may result in considerably different conclusions from singleton estimators. Rows [1] – [3] report the estimated birthweight effects on infant health outcomes and Row [4] reports the effect on disability pension receipt, which is fairly related to birth outcome. Typical patterns found in these rows can be summarized as follows. First, when we compare the twin and singleton OLS estimators, the former shows considerably larger effects of birthweight. However, this result is simply because the health outcomes of twin infants are significantly worse than those of singleton infants. Second, applying fixed effects creates a significant difference between twins and singletons: the twin fixed-effects estimator leads to substantially smaller birthweight effects compared to those of OLS on the twin sample, providing us an impression that birthweight is not as important as it appears from OLS estimates, whereas applying mother fixed effects to singleton infants yields birthweight effects substantially larger than those of OLS, providing us a totally converse impression. Row [1-BDS] reports the results from Black, Devereux, and Salvanes (2007) to verify the robustness of these findings. Because they use older cohorts than ours and one year mortality is higher in their sample, their estimates are all larger than ours. Nevertheless, Row [1-BDS] shows a consistent pattern – fixed effects generates opposite effects between twin and singleton samples.

Similarly to singleton fixed-effects estimators, the singleton IV estimator leads to larger birthweight effects, though it may or may not be larger than the singleton fixed-effects estimator, depending on each outcome. Therefore, for these outcomes related to infant health, our IV estimates suggest the considerably larger role of birthweight compared to the twin fixed-effects estimator.

The last row in Table 11 reports the comparison for national exam scores in the 9th grade. This is one of the outcome variables for which we find a substantial difference between IV and non-IV estimators. Nevertheless, we observe the consistent pattern – the use of fixed effects generates opposite effects between twin and singleton samples, and thus conclusions from twin fixed-effects results are quite different from conclusions implied by fixed-effects or IV estimates based on singleton sample.

Our further investigation indicates that the reason for this remarkable difference between the twin and singleton fixed-effects estimators is a combined effect of (1) the non-random assignment of birthweight between sibling pairs and (2) non-linearity in the birthweight effect. In the language of the treatment effect literature, the difference arises because the two estimators concern treatment effects at different margins when the birthweight effect is heterogeneous. Regarding (1), we find that a difference in birthweight between a pair of siblings has quite different meanings for twins and singleton siblings. Table 12 reports a simple regression model in which we regress the mean birthweight of two siblings on the difference in birthweight between the two siblings. The regression is run separately for singletons and twins. The results highlight a simple fact: singleton siblings born with similar birthweights are heavier, whereas twin siblings born with similar birthweights are lighter, probably because the difference in weight between twins in utero tends to increase as twin fetuses grow. This fact itself does not necessarily imply any bias, especially when the marginal effect of birthweight (in log) is constant. However, if the relationship between birthweight and an outcome variable involves sizable non-linearity, or if the birthweight effect involves heterogeneity, the twin fixed-effects estimator and singleton fixed-effects estimator may capture different local birthweight effects both of which are different from the population average birthweight effect.

[Insert Table 12]

The previous studies based on the twin fixed-effects estimator show mixed results and their results are different from our results most likely because of this combined effect of non-linearity and non-random assignment of birthweight between siblings. The twin fixed-effects approach also has other caveats. First, because twins are special in many aspects, for example their health substantially poorer than singleton infants, the generalizability of the findings is questionable. Second, some part of the low birthweight effect may arise through its effect on parental resources and family environments, which are shared by twins (and by siblings), so studying twins and siblings may not be able to capture the entire effect.

7. Conclusion

A new instrument, placenta previa, and the population data of 1.7 million newborns since 1981 allow us to quantify the causal effects of birthweight. Studying an extensive range of outcomes, we find the following. First, birthweight has significant positive effect on health related outcomes, especially reducing infant mortality and permanent disability. Regarding other non-health outcomes, however, while the OLS results show highly significant seemingly favorable role of birthweight, the causal estimates show substantially weaker evidence, indicating that the correlation is mostly non-causal. Second, the effect of birthweight diminishes as the child becomes older. Third, returns to birthweight in Denmark have significantly diminished over the last three decades.

Our results confront the conventional statement that birthweight is a good measure of infant health. Not only is it a noisy proxy of maternal and infant health, it also represents the socioeconomic status of the family and the development of medical technology in the society. Although there is no doubt for the importance of birthweight as one of major birth outcome measures, our findings highlight the difficulty in interpreting birthweight, especially when researchers attempt to make a causal argument.

References

Almond, Douglas, Kenneth Y. Chay, and David S. Lee (2005) “The Costs of Low Birth Weight,” *Quarterly Journal of Economics*, 120(3), 1031-1083.

Aliyu, M.H., Lynch, O., Nana, P.N., Alio, A.P., Wilson, R.E., Marty, P.J., Zoorob, R., and Salihu, H.M. (2011A) “Alcohol Consumption During Pregnancy and Risk of Placental Abruption and Placenta Previa,” *Maternal and Child Health Journal*, 15, 670-676.

Aliyu, M.H., Lynch, O., Wilson, R.E., Alio, A.P., Kristensen, S., Marty, P.J., Whiteman, V.E., and Salihu, H.M. (2011B) “Association Between Tobacco Use in Pregnancy and Placenta-Associated Syndromes: A Population-Based Study,” *Archives of Gynecology and Obstetrics*, 283, 729-734.

Behrman, Jere R. and Mark R. Rosenzweig (2004) “Returns to Birthweight,” *Review of Economics and Statistics*, 86(2), 586-601.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2007) “From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes,” *Quarterly Journal of Economics*, 122(1), 409-439.

Blenstrup, Lene Tølbøll, and Lisbeth B. Knudsen (2011) “Danish Registers on Aspects of Reproduction,” *Scandinavian Journal of Public Health*, 39(Suppl. 7), 79-82.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press.

Case, Anne, Angela Fertig, and Christina Paxson (2005) “The Lasting Impact of Childhood Health and Circumstance,” *Journal of Health Economics*, 24, 365-389.

Conley, Dalton, and Neil G. Bennett (2000) “Is Biology Destiny? Birth Weight and Life Chances,” *American Sociological Review*, 65, 458-467.

Conley, Dalton, Kate W. Strully, and Neil G. Bennett (2003) “A Pound of Flesh or Just Proxy? Using Twin Differences to Estimate the Effects of Birth Weight on Life Chances,” *NBER Working Paper*, No. 9901.

Conley, Dalton, Kate W. Strully, and Neil G. Bennett (2006) “Twin Differences in Birth Weight: The Effects of Genotype and Prenatal Environment on Neonatal and Post-Neonatal Mortality,” *Economics and Human Biology*, 4, 151-183.

Currie, J., and MacLeod, W.B. (2008) “First Do No Harm? Tort Reform and Birth Outcomes,” *Quarterly Journal of Economics*, 123(2), 795-830.

Currie, Janet, Mark Stabile, Phongsack Manivong, and Leslie L. Roos (2010) “Child Health and Young Adult Outcomes,” *Journal of Human Resrouces*, 45(3), 517-548.

Currie, J., and E. Moretti (2007) “Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Transmission of Birth Weight," *Journal of Labor Economics*, 25(2), 231-264.

Eide, Martha G., Nina Øyen, Rolv Skjærven, Stein Tore Nilsen, Tor Bjerkedal, and Grethe S. Tell (2005) “Size at Birth and Gestational Age as Predictors and Adult Height and Weight,” *Epidemiology*, 16(2), 175-181.

Grjibovski, Andrej M., Jennifer R. Harris, and Per Magnus (2005) “Birthweight and Adult Health in a Population-Based Sample of Norwegian Twins,” *Twin Research and Human Genetics*, 8(2), 148-155.

Hannah, M.E., Hannah, W.J., Hewson, S.A., Hodnett, E.D., Saigal, S., and Willan, A.R. (2000) “Planned Saesarean Section Versus Planned Vaginal Birth for Breech Presentation at Term: A Randomised Multicentre Trial,” *The Lancet*, 356(9239), 1375-1383.

Jensen, V.M., and Wust, M. (2012) "Can Caesarean Section Improve Child and Maternal Health? The Case of Breech Babies," working paper available at: http://sole-jole.org/13404.pdf.

Joy, Saju, Lorene Temming, and Ryan A. Stone (2012) "Placenta Previa," M*e*dscape reference, available at: http://emedicine.medscape.com/article/262063-overview.

Miller, Paul, Charles Mulvey, and Nick Martin (2005) "Birth Weight and Schooling and Earnings: Estimates from a Sample of Twins," *Economics Letters*, 86, 387-392.

Mortensen, Laust H., Finn Diderichsen, George Davey Smith, and Anne Marie Nybo Anderson (2009) "The Social Gradient in Birthweight at Term: Quantification of the Mediating Role of Maternal Smoking and Body Mass Index," *Human Reproduction*, 24(10), 2629-2635.

Oreopoulos, Philip, Mark Stabile, Randy Walld, and Leslie L. Roos (2008) "Short-, Medium-, and Long-Term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins," *Journal of Human Resources*, 43(1), 88-138.

Oyelese, Yinka, and John C. Smulian (2006) "Placenta Previa, Placenta Accreta, and Vasa Previa," *Obstetrics and Gynecology*, 107(4), 927-941.

Royer, Heather (2009) "Separated at Girth: US Twin Estimates of the Effects of Birth Weight," *American Economic Journal: Applied Economics*, 1(1), 49-85.

Table 1: Correlation between weight, gestation, intrauterine growth, and placenta previa

	[1] Birthweight	[2] Gestation days	[3] Intrauterine growth (=[1]/[2])	[4] Placenta previa
Birthweight	1.000			
Gestation days	0.562	1.000		
Intrauterine growth	0.974	0.373	1.0000	
Placenta previa	-0.059	-0.090	-0.044	1.0000

Based on 1,783,467 singleton births between 1981 and 2010. All the correlation coefficients are statistically different from zero at the 0.1 percent significance level.

Table 2: Characteristics of Children and Parents

	Births without placenta previa		Births with placenta previa	
All singleton births (1981-2010)				
N and proportions	1,777,169	(99.65%)	6,298	(0.35%)
Birthweight (in grams)	3,494.2	(566.2)	2,929.6	(759.4)
Female	0.49	(0.50)	0.45	(0.50)
Birth order	1.79	(0.91)	1.94	(1.00)
Mother’s age at conception	27.99	(4.91)	29.97	(5.08)
Parity (number of past pregnancies)	1.14	(0.95)	1.36	(1.13)
Number of past cesarean sections	0.072	(0.278)	0.14	(0.40)
Number of past abortions recorded (both spontaneous and induced)	0.333	(0.724)	0.47	(0.87)
Indicator for stillbirth in the past	0.008	(0.091)	0.014	(0.117)
Number of days in hospital between 90 days before and after conception, exclusive of obstetrics-related admission	0.12	(1.80)	0.13	(1.06)
Indicator for mother’s smoking (available only for 1991 and onwards)	0.212	(0.408)	0.212	(0.408)
Cohabitation with biological father on Jan 1 before birth	0.85	(0.36)	0.86	(0.35)
Formally married on Jan 1 before birth	0.55	(0.50)	0.60	(0.49)
Mother’s education				
Less than 9 years	0.04	(0.19)	0.04	(0.20)
9 years	0.19	(0.40)	0.18	(0.39)
Upper secondary education	0.44	(0.50)	0.42	(0.49)
Tertiary education (short and medium)	0.23	(0.42)	0.25	(0.43)
Tertiary education (long)	0.10	(0.30)	0.11	(0.31)
Mother’s working status in previous year	0.82	(0.39)	0.82	(0.38)
Mother’s gross income in previous year	161,466	(109,672)	168,345	(109,502)
Father’s age at conception	31.3	(5.75)	33.2	(5.94)
Father’s education				
Less than 9 years	0.04	(0.19)	0.04	(0.20)
9 years	0.16	(0.36)	0.14	(0.35)
Upper secondary education	0.46	(0.50)	0.44	(0.50)
Tertiary education (short and medium)	0.17	(0.38)	0.19	(0.39)
Tertiary education (long)	0.11	(0.31)	0.12	(0.33)
Father’s working status in previous year	0.86	(0.35)	0.87	(0.33)
Father’s gross income in previous year	262,037	(214,663)	282,034	(176,285)
Immigrant indicator (either parent or both)	0.13	(0.34)	0.13	(0.34)

Note: Based on singleton births only. Standard deviations are in parentheses. Parental income is measured in Danish Krone (100 Kr $\approx$ 13.4€). The variables regarding smoking, mother’s education, and father’s characteristics are not observed for every child. In the regression analysis, these variables are set to zero when missing, and indicators for a missing value are used. The summary statistics in this table is based on observations without missing values.

Table 3: Definitions and Summary Statistics of Outcome Variables

Outcome variables	Birth cohort	Mean	Std Dev	Population conditional on:
Child outcomes: health				
Infant mortality (within 365 days from birth)	1981-1993	0.0066	0.0810	Live birth
Infant mortality (within 365 days from birth)	1994-2010	0.0034	0.0586	Live birth
Permanent disability (CP, vision, hearing)	1981-1993	0.0024	0.0489	Survival up to 2nd birthday
Permanent disability (CP, vision, hearing)	1994-2010	0.0024	0.0487	Survival up to 2nd birthday
Number of hospital days: up to 2nd birthday	1981-1993	2.713	10.795	Survival up to 2nd birthday
Number of hospital days: up to 2nd birthday	1994-2009	2.470	12.037	Survival up to 2nd birthday
Hospital admission: 2nd to 5th birthday	1981-1991	0.174	0.379	Survival up to 5th birthday
Hospital admission: 5th to 10th birthday	1981-1991	0.180	0.384	Survival up to 10th birthday
Hospital admission: 10th to 15th birthday	1981-1991	0.145	0.352	Survival up to 15th birthday
Hospital admission: 15th to 20th birthday	1981-1991	0.167	0.373	Survival up to 20th birthday
Child outcomes: education				
Completed Grade 9 by the year of age 16	1981-1995	0.852	0.355	Observed Jan 1 before age 17
Standardized score of national exam at Grade 9 (≈age 16): Mean of 4 mandatory subjects	1986-1994	0.017	0.766	Observed Jan 1 before age 17 and exam scores of at least 3 subjects at age 15, 16, or 17
Standardized exam score: Danish (oral)	1986-1994	0.033	0.990	Same as above
Standardized exam score: Danish (written)	1986-1994	0.052	0.891	Same as above
Standardized exam score: English	1986-1994	0.006	0.991	Same as above
Standardized exam score: Mathematics	1986-1994	0.052	0.951	Same as above
Standardized exam score: Science	1986-1994	0.020	0.994	Same as above
Child outcomes: social welfare				
Receipt of disability pension during the three calendar years of age 19 to 21	1981-1991	0.011	0.105	Observed Jan 1 before age 21
Number of weeks, disability pension: 19-21	1981-1991	1.307	13.51	Observed Jan 1 before age 21
Total amount, disability pension: 19-21	1981-1990	4,238	45,073	Observed Jan 1 before age 21
Receipt of other welfare assistance: 19-21	1981-1991	0.221	0.415	Observed Jan 1 before age 21
Number of weeks, other welfare: 19-21	1981-1991	10.47	29.30	Observed Jan 1 before age 21
Total amount, other welfare: 19-21	1981-1990	20,419	72,059	Observed Jan 1 before age 21
Child outcomes: other socioeconomic status				
Worked or student in the year of age 22	1981-1988	0.878	0.328	Observed Jan 1 before age 22
Gross income in the year of age 22	1981-1988	152,696	83,521	Observed Jan 1 before age 22
Pregnancy before 20th birthday (including induced and spontaneous abortions): females	1981-1990	0.051	0.220	Survival up to 20th birthday
Pregnancy before 20th birthday (including induced and spontaneous abortions if recorded in MFR): males	1981-1990	0.012	0.110	Survival up to 20th birthday
Married on Jan 1 in the year of age 22	1981-1990	0.013	0.113	Observed Jan 1 before age 22
Birthweight of the first child by age 22	1981-1990	3391.1	573.8	1st child by the year of age 22
Any criminal sentence by 20th birthday	1981-1991	0.092	0.290	Survival up to 20th birthday
Probation or unconditional sentence by 20th birthday	1981-1991	0.033	0.179	Survival up to 20th birthday
Any violent crime charge by 20th birthday	1981-1991	0.028	0.166	Survival up to 20th birthday
Military conscription variables (males only. All variables below are conditional on survival up to 18th birthday)				
Conscription session attendance	1988-1993	0.762	0.426	
Qualified or qualified with restriction	1988-1993	0.564	0.496	
Color vision deficiency	1988-1993	0.061	0.240	Session attendance
Standardized IQ score (Børge Priens test)	1988-1993	-0.003	1.004	Session attendance
Height (cm)	1988-1993	180.46	6.68	Session attendance
Weight (kg)	1988-1993	77.67	14.95	Session attendance
BMI (= Weight / $(\text{Height}/100)^2$)	1988-1993	23.82	4.25	Session attendance

Note: All statistics are based on singleton births only. Permanent disability consists of cerebral palsy, loss of vision, and hearing impairment. Hospital admission after age 10 excludes pregnancy and birth related admissions. Parental income is measured in Danish Krone (100 Kr ≈ 13.4€).

Table 4: Determinants of Birthweight – First Stage Regressions with Selected Variables

Dependent variable: ln(birthweight)	[1] OLS	[2] Grandmother FE	[3] Mother FE
Placenta Previa	-0.196****	-0.195****	-0.190****
	(0.00394)	(0.00446)	(0.00478)
Female	-0.0345****	-0.0365****	-0.0373****
	(0.000271)	(0.000295)	(0.000301)
First child	-0.0401****	-0.0367****	-0.0354****
	(0.000628)	(0.000694)	(0.000751)
Birth order	0.0104****	0.0159****	0.0177****
	(0.000509)	(0.000608)	(0.000758)
Past cesarean section	-0.0228****	-0.00998****	0.00182**
	(0.000627)	(0.000746)	(0.000874)
Number of days in hospital	-0.000868****	-0.000471****	-0.000217**
around conception	(0.0000978)	(0.000110)	(0.000110)
Mother smoking	-0.0583****	-0.0371****	-0.0189****
	(0.000488)	(0.000632)	(0.000690)
Cohabitation status	0.00677****	0.00686****	0.00504****
	(0.000511)	(0.000581)	(0.000606)
Formally married	0.00151****	0.00194****	0.00112**
	(0.000342)	(0.000440)	(0.000477)
Parental education (reference: 9 years)			
Mother: Less than 9 years	-0.00767****	-0.00298*	
	(0.000970)	(0.00169)	
Mother: Upper secondary	0.0172****	0.0110****	
	(0.000483)	(0.000779)	
Mother: Tertiary (short and mid)	0.0262****	0.0134****	
	(0.000593)	(0.000985)	
Mother: Tertiary (long)	0.0264****	0.0128****	
	(0.000901)	(0.00142)	
Father: Less than 9 years	-0.00486****	-0.000174	0.000604
	(0.000832)	(0.00126)	(0.00193)
Father: Upper secondary	0.00684****	0.00589****	0.00574****
	(0.000471)	(0.000700)	(0.000955)
Father: Tertiary (short and mid)	0.0131****	0.00861****	0.00722****
	(0.000593)	(0.000926)	(0.00132)
Father: Tertiary (long)	0.0145****	0.00747****	0.00487***
	(0.000709)	(0.00115)	(0.00163)
Father's work status	0.00757****	0.00652****	0.00548****
	(0.000586)	(0.000679)	(0.000723)
Father's annual gross income	0.00449****	0.00317**	0.00204
(in 1,000 Danish Krone ≈ 134€)	(0.00129)	(0.00124)	(0.00137)
Either parent immigrant	-0.0204****	-0.0127****	-0.0145****
	(0.000604)	(0.00138)	(0.00208)
F statistics for the relevance of IV	2460.3	1908.3	1575.2
R^2	0.0750	0.0713	0.0812
N	1,783,092	1,574,355	1,431,770

In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. The three models reported are: [1] OLS, [2] OLS with grandmother fixed effects, and [3] OLS with mother fixed effects. The mean of ln(birthweight) is 8.143. The other variables included in the regressions are year- and month-of-birth dummies, county dummies, indicators for mother's age at conception (one dummy for every two years), indicators for parity (1, 2, 3, and 4+), indicators for mother's past abortions (1 and 2+) and past stillbirth, mother's income and working status in the previous year, interacted with conception month dummies, and father's age. Indicators for missing county and missing parental education are also used. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Table 5: Estimated Coefficients of ln(birthweight) – One Year Infant Mortality

	OLS	IV	OLS Grandmother FE	IV Grandmother FE	OLS Mother FE	IV Mother FE
[1] Infant mortality (from birth to 365 days) (1981-2010) $\bar{Y}$=0.0047, N=1,783,340						
	-0.0761****	-0.0825****	-0.106****	-0.0776****	-0.133****	-0.0757****
	(0.00119)	(0.00886)	(0.00173)	(0.0118)	(0.00217)	(0.0144)
[1A] Infant mortality, period 1 (1981-1993) $\bar{Y}$=0.0066, N=723,817						
	-0.0883****	-0.133****	-0.140****	-0.151****	-0.175****	-0.164****
	(0.00187)	(0.0163)	(0.00323)	(0.0269)	(0.00398)	(0.0314)
[1B] Infant mortality, period 2 (1994-2010) $\bar{Y}$=0.0034, N=1,059,523						
	-0.0676****	-0.0413****	-0.100****	-0.0322**	-0.124****	-0.0287*
	(0.00155)	(0.00907)	(0.00243)	(0.0128)	(0.00298)	(0.0163)
[2A] Infant mortality: with maternal body size variables (2004-2010) $\bar{Y}$=0.0025, N=398,452						
	-0.0648****	-0.0412***	-0.104****	-0.0377	-0.133****	-0.0322
	(0.00268)	(0.0138)	(0.00502)	(0.0282)	(0.00630)	(0.0358)
[2B] Infant mortality: without maternal body size variables (2004-2010) $\bar{Y}$=0.0025, N=398,452						
	-0.0622****	-0.0400***	-0.103****	-0.0373	-0.133****	-0.0322
	(0.00258)	(0.0134)	(0.00498)	(0.0279)	(0.00630)	(0.0357)

The description of each outcome variable is followed by the birth years of the cohorts used, the mean value of the outcome variable, and the number of observations. The effective number of observations in the fixed-effects regressions is smaller than the number shown. In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. The control variables we use in the OLS and grandmother fixed-effects regressions are year- and month-of-birth dummies, county dummies, an indicator for the sex of the child, indicators for mother's age at conception (one dummy for every two years), an indicator for the first child, birth order, indicators for parity (1, 2, 3, and 4+), the number of past cesarean sections, indicators for mother's past abortions (1 and 2+), past stillbirth, and smoking habits, the number of days in hospital during 180 days around conception, indicators for mother's education (less than 9 years, upper secondary, low- and mid-tertiary, and high tertiary), indicators for formal marital status and whether biological parents live together on January 1st prior to the birth, the mother's income and working status in the previous year, and an indicator for immigrant status of either the mother or father. If the biological father lives together on January 1st prior to the birth, also included are the father's age, income and working status in the previous year, and indicators for father's education (the same grouping as mother's education above). The income and working status of the mother are also interacted with the month of conception. Indicators for missing values in the education of the mother and father are also used. Mother fixed-effects regressions include all of the above minus mother's education.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Table 6: Estimated Coefficients of ln(birthweight) –Health Outcomes

	OLS	IV	OLS Grandmother FE	IV Grandmother FE	OLS Mother FE	IV Mother FE
[1] Permanent disability diagnosis by 2nd birthday (1981-2009) $\bar{Y}$=0.0024, N=1,706,598						
	-0.0164****	-0.0274****	-0.0199****	-0.0209***	-0.0218****	-0.0229***
	(0.000537)	(0.00614)	(0.000782)	(0.00733)	(0.000965)	(0.00826)
[1A] Permanent disability diagnosis by 2nd birthday, period 1 (1981-1993) $\bar{Y}$=0.0024, N=716,658						
	-0.0189****	-0.0329****	-0.0233****	-0.0343***	-0.0240****	-0.0308**
	(0.000879)	(0.00996)	(0.00150)	(0.0133)	(0.00173)	(0.0157)
[1B] Permanent disability diagnosis by 2nd birthday, period 2 (1994-2009) $\bar{Y}$=0.0024, N=989,940						
	-0.0146****	-0.0229***	-0.0181****	-0.0148	-0.0194****	-0.0204*
	(0.000671)	(0.00771)	(0.00105)	(0.00975)	(0.00129)	(0.0104)
[2] Number of days in hospital before 2nd birthday (1981-2009) $\bar{Y}$=2.572, N=1,706,598						
	-21.08****	-39.43****	-24.64****	-38.07****	-26.07****	-35.92****
	(0.137)	(1.176)	(0.184)	(1.405)	(0.220)	(1.605)
[2A] Number of days in hospital before 2nd birthday, period 1 (1981-1993) $\bar{Y}$=2.713, N=716,658						
	-20.00****	-40.46****	-23.77****	-38.67****	-25.37****	-36.71****
	(0.200)	(2.014)	(0.307)	(3.216)	(0.367)	(3.593)
[2B] Number of days in hospital before 2nd birthday, period 2 (1994-2009) $\bar{Y}$=2.470, N=989,940						
	-21.85****	-38.62****	-25.40****	-37.10****	-26.36****	-35.92****
	(0.185)	(1.388)	(0.269)	(1.695)	(0.318)	(1.727)
[3A] Hospitalization: 2nd to 5th birthday (1981-1991) $\bar{Y}$=0.1744, N=586,397						
	-0.111****	-0.178****	-0.128****	-0.252***	-0.121****	-0.199**
	(0.00317)	(0.0495)	(0.00608)	(0.0860)	(0.00755)	(0.0975)
[3B] Hospitalization: 5th to 10th birthday (1981-1991) $\bar{Y}$=0.1803, N=584,147						
	-0.0837****	-0.121**	-0.0953****	-0.151*	-0.0875****	-0.0848
	(0.00316)	(0.0496)	(0.00607)	(0.0859)	(0.00748)	(0.0975)
[3C] Hospitalization: 10th to 15th birthday (1981-1991) $\bar{Y}$=0.1447, N=582,800						
	-0.0454****	-0.109**	-0.0385****	-0.0962	-0.0349****	-0.105
	(0.00286)	(0.0461)	(0.00559)	(0.0818)	(0.00696)	(0.0922)
[3D] Hospitalization: 15th to 20th birthday (1981-1991) $\bar{Y}$=0.1674, N=580,379						
	-0.0338****	-0.0802*	-0.0276****	-0.102	-0.0183**	-0.0299
	(0.00297)	(0.0475)	(0.00584)	(0.0858)	(0.00726)	(0.0976)

The short description of each outcome variable is followed by the birth years of the cohorts used, the mean value of the outcome variable, and the number of observations. In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. For the list of the control variables included in each regression, see the note to Table 5.

* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Table 7: Estimated Coefficients of ln(birthweight) – Education Outcomes

	OLS	IV	OLS Grandmother FE	IV Grandmother FE	OLS Mother FE	IV Mother FE
[1] Completed Grade 9 by year of age 16 (1981-1995) $\bar{Y}$=0.852, N=820,136						
	0.147****	0.139****	0.137****	0.155***	0.136****	0.119*
	(0.00241)	(0.0375)	(0.00410)	(0.0584)	(0.00502)	(0.0661)
[2] National exam score, Grade 9 (≈ age 16): overall mean (1986-1994) $\bar{Y}$=0.017, N=459,105						
	0.196****	-0.0121	0.135****	-0.289*	0.124****	-0.0964
	(0.00641)	(0.106)	(0.0112)	(0.163)	(0.0133)	(0.175)
[3] Standardized exam score: Danish (oral) (1986-1994) $\bar{Y}$=0.033, N=458,007						
	0.120****	-0.101	0.0734****	-0.229	0.0738****	0.0737
	(0.00849)	(0.142)	(0.0161)	(0.232)	(0.0203)	(0.262)
[4] Standardized exam score: Danish (written) (1986-1994) $\bar{Y}$=0.052, N=459,060						
	0.176****	-0.0612	0.118****	-0.393*	0.113****	-0.138
	(0.00752)	(0.126)	(0.0133)	(0.202)	(0.0160)	(0.217)
[5] Standardized exam score: English (1986-1994) $\bar{Y}$=0.006, N=449,444						
	0.117****	-0.109	0.0496***	-0.449*	0.0487**	-0.242
	(0.00877)	(0.143)	(0.0160)	(0.235)	(0.0193)	(0.255)
[6] Standardized exam score: Mathematics (1986-1994) $\bar{Y}$=0.052, N=457,662						
	0.337****	0.153	0.269****	-0.269	0.233****	-0.135
	(0.00817)	(0.135)	(0.0145)	(0.211)	(0.0174)	(0.228)
[7] Standardized exam score: Science (1986-1994) $\bar{Y}$=0.020, N=436,290						
	0.159****	-0.0689	0.102****	-0.0321	0.103****	0.119
	(0.00883)	(0.156)	(0.0169)	(0.253)	(0.0212)	(0.279)

The short description of each outcome variable is followed by the birth years of the cohorts used, the mean value of the outcome variable, and the number of observations. In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. For the list of the control variables included in each regression, see the note to Table 5.

* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Table 8: Estimated Coefficients of ln(birthweight) – Social Welfare Assistance

	OLS	IV	OLS Grandmother FE	IV Grandmother FE	OLS Mother FE	IV Mother FE
[1A] Receipt of disability pension during the three years of age 19-21 (1981-1991) $\bar{Y}$=0.011, *N*=578,987						
	-0.0353****	-0.0533***	-0.0438****	-0.0811***	-0.0496****	-0.0817**
	(0.00135)	(0.0177)	(0.00244)	(0.0283)	(0.00296)	(0.0323)
[1B] Receipt of disability pension: DKK100,000+, during 3 years of age 19-21 (1981-1990) $\bar{Y}$=0.009, *N*=519,772						
	-0.0330****	-0.0526***	-0.0400****	-0.0596**	-0.0460****	-0.0714**
	(0.00135)	(0.0175)	(0.00252)	(0.0282)	(0.00305)	(0.0319)
[1C] Number of weeks of disability pension receipt during 3 years of age 19-21 (1981-1991) $\bar{Y}$=1.307, *N*=578,987						
	-4.610****	-5.880***	-5.955****	-8.308**	-6.725****	-10.59**
	(0.181)	(2.253)	(0.333)	(3.574)	(0.405)	(4.341)
[1D] Total amount of disability pension transfers during 3 years of age 19-21 (1981-1990) $\bar{Y}$=4,238, *N*=519,772						
	-15701.2****	-23717.5***	-19770.9****	-26764.3**	-22383.4****	-31460.7**
	(653.0)	(8316.3)	(1228.8)	(12401.1)	(1464.6)	(14803.8)
[2A] Receipt of other welfare during the three years of age 19-21 (1981-1991) $\bar{Y}$=0.221, *N*=578,987						
	-0.0960****	-0.0929*	-0.0591****	-0.0387	-0.0433****	-0.0652
	(0.00328)	(0.0506)	(0.00599)	(0.0842)	(0.00719)	(0.0931)
[2B] Receipt of other welfare: DKK100,000+, during 3 years of age 19-21 (1981-1990) $\bar{Y}$=0.063, *N*=519,772						
	-0.0567****	-0.102***	-0.0433****	-0.196****	-0.0323****	-0.173***
	(0.00225)	(0.0338)	(0.00443)	(0.0576)	(0.00529)	(0.0624)
[2C] Number of weeks of other welfare receipt during 3 years of age 19-21 (1981-1991) $\bar{Y}$=10.47, *N*=578,987						
	-8.208****	-9.769**	-5.812****	-14.59**	-4.578****	-11.66*
	(0.262)	(3.870)	(0.478)	(6.333)	(0.571)	(7.059)
[2D] Total amount of other welfare transfers during 3 years of age 19-21 (1981-1990) $\bar{Y}$=20,419, *N*=519,772						
	-15285.0****	-17234.2*	-9823.8****	-39902.6**	-6223.3****	-31075.3*
	(636.0)	(8966.7)	(1285.4)	(16299.7)	(1578.7)	(18379.8)

The short description of each outcome variable is followed by the birth years of the cohorts used, the mean value of the outcome variable, and the number of observations. In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. For the list of the control variables included in each regression, see the note to Table 5. DKK100,000=USD17,298 as of May 28, 2013.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Table 9: Estimated Coefficients of ln(birthweight) – Socioeconomic Outcomes

	OLS	IV	OLS Grandmother FE	IV Grandmother FE	OLS Mother FE	IV Mother FE
[1] Worked or was a student in the year of age 22 (1981-1988) $\bar{Y}$=0.8776, N=401,944						
	0.0856****	0.111**	0.0730****	0.159	0.0771****	0.149
	(0.00334)	(0.0513)	(0.00721)	(0.0981)	(0.00882)	(0.103)
[2] Gross income in the year of age 22 (1981-1988) $\bar{Y}$=152,696, N=401,944						
	6001.7****	10099.8	7623.0****	-11773.7	6106.6***	-18634.9
	(742.1)	(11601.9)	(1580.9)	(21039.2)	(1942.1)	(22762.6)
[3A] Pregnancy as a mother by the birthday of age 20 (1981-1990) $\bar{Y}$=0.0512, N=253,631						
	-0.0152****	0.0132	0.00218	-0.00792	0.0142	0.00849
	(0.00259)	(0.0399)	(0.00749)	(0.0946)	(0.0101)	(0.120)
[3B] Pregnancy as a father by the birthday of age 20 (1981-1990) $\bar{Y}$=0.0123, N=267,010						
	-0.00191	0.0300**	0.00287	0.102**	0.00944*	0.0791*
	(0.00121)	(0.0146)	(0.00376)	(0.0446)	(0.00519)	(0.0476)
[4] Married on Jan 1 in the year of age 22 (1981-1990) $\bar{Y}$=0.0130, N=518,450						
	0.00269***	-0.00259	0.00339*	-0.00399	0.00436*	0.000187
	(0.000917)	(0.0142)	(0.00193)	(0.0260)	(0.00252)	(0.0280)
[5] Birthweight of the first child by age 22 (1981-1990) $\bar{Y}$=3391.1, N=25,902						
	585.7****	-225.5	382.6****	30.43	253.7**	237.8
	(22.97)	(295.3)	(91.83)	(502.2)	(108.5)	(525.4)
[5A] Birthweight of the first child by age 22: female (1981-1990) $\bar{Y}$=3394.3, N=18,421						
	672.5****	-314.0	463.3****	-481.0	285.7*	N too small
	(27.63)	(394.9)	(126.5)	(2425.1)	(156.6)	
[5B] Birthweight of the first child by age 22: male (1981-1990) $\bar{Y}$=3383.1, N=7,481						
	378.5****	-152.1	258.8	-313.5	-332.7	N too small
	(42.40)	(438.7)	(320.9)	(248.5)	(489.6)	
[6A] Any criminal sentence by the 20th birthday (1981-1991) $\bar{Y}$=0.092, N=580,492						
	-0.00860****	0.0481	0.0199****	0.0751	0.0300****	0.0707
	(0.00219)	(0.0347)	(0.00433)	(0.0591)	(0.00530)	(0.0662)
[6B] Probation or unconditional sentence by the 20th birthday (1981-1991) $\bar{Y}$=0.033, N=580,492						
	-0.00564****	0.0245	0.00929****	0.0472	0.0159****	0.0284
	(0.00136)	(0.0211)	(0.00282)	(0.0369)	(0.00344)	(0.0430)
[6C] Any charge of violent crime by the 20th birthday (1981-1991) $\bar{Y}$=0.028, N=580,492						
	-0.000120	0.0345*	0.00998****	0.0375	0.0129****	0.00454
	(0.00124)	(0.0184)	(0.00260)	(0.0344)	(0.00319)	(0.0386)

The short description of each outcome variable is followed by the birth years of the cohorts used, the mean value of the outcome variable, and the number of observations. In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. For the list of the control variables included in each regression, see the note to Table 5. DKK1,000=USD173.0 as of May 28, 2013.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Table 10: Estimated Coefficients of ln(birthweight) – Military Conscription Variables

OLS	IV	OLS Grandmother FE	IV Grandmother FE	OLS Mother FE	IV Mother FE
[1] Conscription session attendance (1988-1993) $\bar{Y}$=0.762, N=183,877					
0.0965****	0.185**	0.111****	0.546**	0.103****	0.521**
(0.00597)	(0.0890)	(0.0163)	(0.242)	(0.0231)	(0.255)
[2] Qualified for military service (inclusive of restricted qualification) (1988-1993) $\bar{Y}$=0.564, N=183,877					
0.119****	0.221**	0.124****	0.764***	0.132****	0.688**
(0.00657)	(0.0980)	(0.0185)	(0.295)	(0.0261)	(0.348)
[3] Color vision deficiency at conscription examination (1988-1993) $\bar{Y}$=0.0612, N=139,414					
0.000890	-0.0521	0.0127	0.0894	0.00507	0.0671
(0.00393)	(0.0653)	(0.0109)	(0.138)	(0.0155)	(0.121)
[4] IQ test standardized score at conscription examination (1988-1993) $\bar{Y}$= –0.003, N=138,924					
0.447****	0.0994	0.322****	0.0456	0.361****	0.117
(0.0156)	(0.227)	(0.0425)	(0.625)	(0.0574)	(0.613)
[5] Height at conscription examination (1988-1993) $\bar{Y}$=180.46, N=140,090					
10.39****	2.604*	7.976****	0.462	7.125****	-0.920
(0.117)	(1.558)	(0.301)	(4.047)	(0.367)	(4.202)
[6] Weight at conscription examination (1988-1993) $\bar{Y}$=77.67, N=139,973					
15.38****	5.989*	12.47****	4.348	9.631****	-4.911
(0.255)	(3.595)	(0.714)	(9.732)	(0.876)	(10.47)
[7] BMI at conscription examination (1988-1993) $\bar{Y}$=23.82, N=139,972					
1.994****	1.137	1.763****	1.102	1.113****	-1.344
(0.0707)	(1.027)	(0.205)	(2.809)	(0.254)	(2.968)

The short description of each outcome variable is followed by the birth years of the cohorts used, the mean value of the outcome variable, and the number of observations. In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. For the list of the control variables included in each regression, see the note to Table 5.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Table 11: Coefficients of ln(birthweight) – Difference Between Twin and Singleton Fixed-Effects Estimators

	Twin OLS	Twin Fixed Effects	Singleton OLS	Singleton Mother FE	Singleton IV	Singleton Mother FE IV
[1] Infant mortality (from birth to 365 days) (1981-2010) $\bar{Y}$=0.0181 / 0.0047, N=60,047 / 1,783,340						
	-0.183****	-0.0406****	-0.0761****	-0.133****	-0.0825****	-0.0757****
	(0.00659)	(0.00708)	(0.00119)	(0.00217)	(0.00886)	(0.0144)
[1-BDS] 1 year mortality, results from Black, Devereux, and Salvanes (2007), Tables I and III. (1967-1997) $\bar{Y}$=0.0311 / 0.0062, N=33,366 / 1,253,546						
	-0.2796****	-0.0411****	-0.1235****	-0.1867****		
	(0.00912)	(0.00764)	(0.00171)	(0.00069)		
[2] Permanent disability diagnosis by 2nd birthday (1981-2009) $\bar{Y}$=0.0061 / 0.0024, N=56,135 / 1,706,598						
	-0.0332****	-0.00669	-0.0164****	-0.0218****	-0.0274****	-0.0229***
	(0.00270)	(0.00464)	(0.000537)	(0.000965)	(0.00614)	(0.00826)
[3] Number of days in hospital before 2nd birthday (1981-2009) $\bar{Y}$=11.517 / 2.572, N=56,135 / 1,706,598						
	-57.71****	-6.259****	-21.08****	-26.07****	-39.43****	-35.92****
	(0.585)	(0.672)	(0.137)	(0.220)	(1.176)	(1.605)
[4] Receipt of disability pension during the 3 years of 19-21 (1981-1991) $\bar{Y}$=0.015 / 0.011, N= 13,132 / 578,987						
	-0.0433****	-0.0274*	-0.0353****	-0.0496****	-0.0533***	-0.0817**
	(0.00767)	(0.0140)	(0.00135)	(0.00296)	(0.0177)	(0.0323)
[5] National exam score, Grade 9 (≈ age 16): overall mean (1986-1994) $\bar{Y}$=0.012 / 0.017, N=11,941 / 459,105						
	0.112****	0.206****	0.196****	0.124****	-0.0121	-0.0964
	(0.0312)	(0.0559)	(0.00641)	(0.0133)	(0.106)	(0.175)

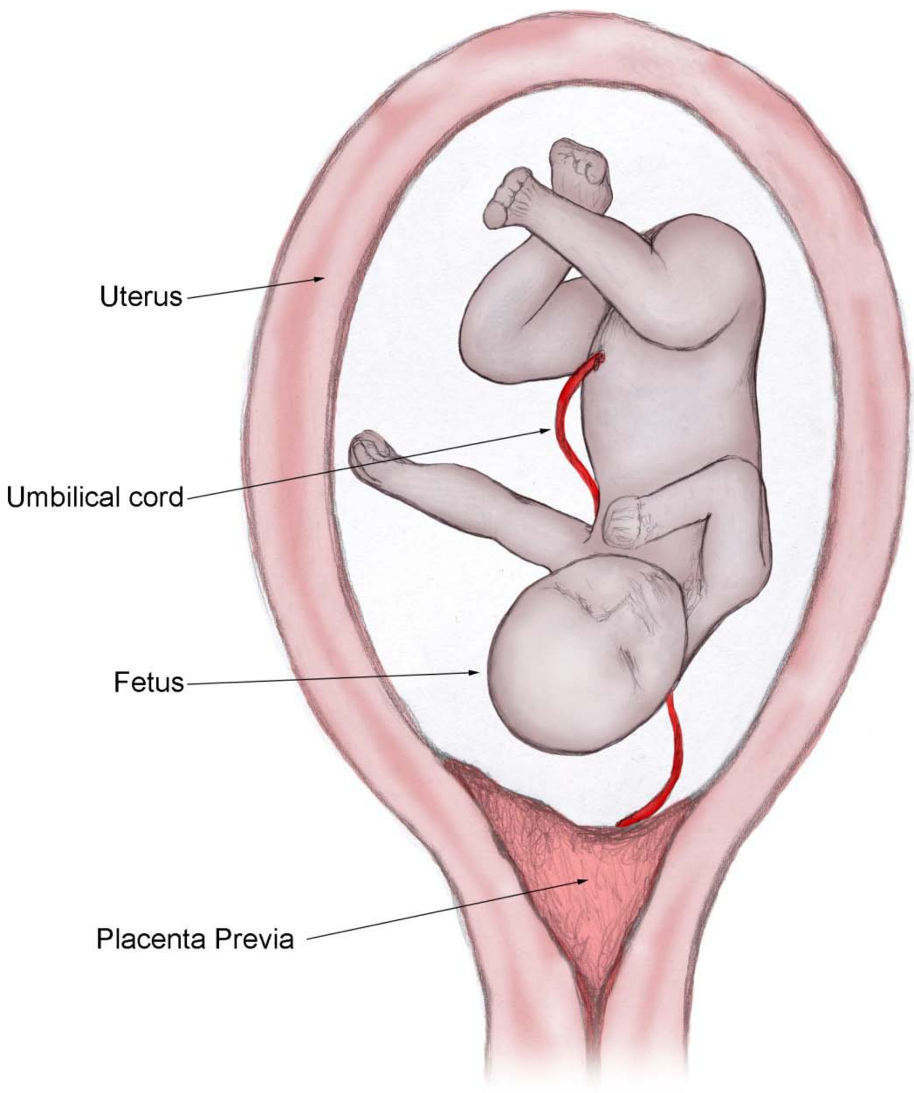
In each row, the short description of each outcome variable is followed by the birth years of the cohorts used, and then by the mean values of the outcome variable and the numbers of observations for twins and singletons, respectively. The effective number of observations in the fixed-effects regressions is smaller than reported in this table. In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. For the list of the control variables included in each regression, see the note to Table 5.
* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Table 12: Relationship between birthweight and birthweight difference among twins and singleton siblings

Dependent variable: Mean birthweight of two siblings	[1] Singleton siblings	[2] Twins
Difference in birthweight between two siblings (absolute value)	-0.185****	0.094****
	(0.0021)	(0.0136)
Constant	3608.9****	2498.3****
	(1.032)	(6.052)
R^2	0.0220	0.0019
N	729,643	26,298

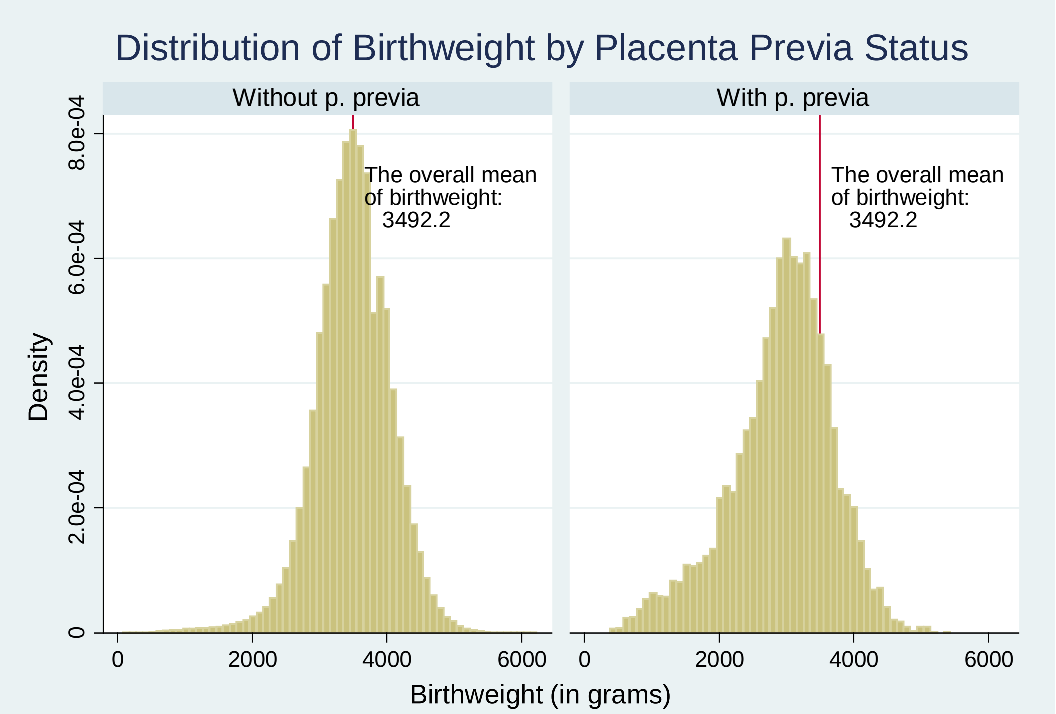
The unit of observation is a sibling pair. In case of more than two singleton siblings, average and difference are taken over two consecutive births. In parentheses are standard errors robust to heteroskedasticity and grandmother clusters. * p<0.1, ** p<0.05, *** p<0.01, **** p<0.001

Figure 1: Complete Placenta Previa



Source: Joy et al., 2010.

Figure 2: The Distribution of Birthweight



Based on singleton cohorts 1981-2010. Placenta previa accounts for 0.35% of the sample.