

From the Cradle to the Labor Market?
The Effect of Birth Weight on Adult Outcomes^{*}

PRELIMINARY AND INCOMPLETE

by

Sandra E. Black
Department of Economics
UCLA, IZA and NBER
sblack@econ.ucla.edu

Paul J. Devereux
Department of Economics
University College Dublin and IZA
devereux@ucd.ie

Kjell G. Salvanes
Department of Economics
Norwegian School of Economics, Statistics Norway and IZA
Kjell.salvanes@nhh.no

October 2005

^{*} Black and Devereux gratefully acknowledge financial support from the National Science Foundation and the California Center for Population Research. Salvanes thanks The Norwegian Research council for financial support. We thank Anne Daltveit, Sue Dynarski, Per Magnus, and Ole Martin Sundet for very useful discussions and seminar participants at the Federal Reserve Bank of New York, Boston College, Brown University, and UCD for helpful comments.

Abstract

Lower birth weight babies have worse outcomes, both short-run in terms of one-year mortality rates and longer run in terms of educational attainment and earnings. However, recent research has called into question whether birth weight itself is important or whether it simply reflects other hard-to-measure characteristics. By applying within twin techniques using a unique dataset from Norway, we examine both short-run and long-run outcomes for the same cohorts. We find that birth weight does matter; very small short-run effects can be misleading because longer-run effects on outcomes such as height, IQ, earnings, and education are significant and similar in magnitude to OLS estimates. Our estimates suggest that eliminating birth weight differences between socioeconomic groups would have sizeable effects on the later outcomes of children from poorer families.

1. Introduction

Lower birth weight babies have worse outcomes, both short-run in terms of one-year mortality rates and longer run in terms of educational attainment and earnings. However, recent research has called into question whether birth weight itself is important or whether it simply reflects other hard-to-measure characteristics. If birth weight does not matter in the long-run, can government policies targeting the welfare of children through improved pre-natal care actually work?

Governments have assumed that birth weight is important and have implemented policies to improve the health of pregnant women in the hope that this will improve the outcomes of their babies; consider, for example, the Women, Infants, and Children Program (WIC) in the United States, a federally funded program that provides nutrition counseling and supplemental food for pregnant women, new mothers, infants and children under age five in order to prevent children's health problems and improve their long-term health, growth and development. A key presumption underlying this type of policy is that, by affecting children's birth weight through improved nutritional intake in utero, it will in turn affect the later health and ultimate success of the children. But is this presumption valid? Likewise, the Irish Government has a target of reducing the gap in low birth weight rates between the lowest and highest socio-economic group.¹ However, there is little evidence about whether being low birth weight (LBW) has a causal effect on later outcomes.

¹ The Irish National Anti-Poverty Strategy (NAPS) is a 10-year government plan to reduce poverty. The NAPS health targets aim to reduce health inequalities by meeting what are described as 3 key targets. One of these targets is 'To reduce the gap in low birth weight rates for children from the lowest and highest socio-economic group by 10% by 2007.'

Despite this, birth weight is very commonly used as the outcome variable of interest in studies of the effects of policy interventions such as welfare reform, health insurance, and food stamps on infant welfare (for example, Currie and Gruber, 1996). Likewise, birth weight is often used as an outcome variable in studies of the effects of inputs to the infant health production function and analyses of the impact of maternal behavior on infant health. (For example, Currie and Moretti, 2003 show that increased maternal education leads to a lesser incidence of LBW). Obviously, the degree to which LBW has true causal effects on later outcomes is critical to the interpretation of the results from this literature.

The principal difficulty in determining the effects of LBW on later outcomes arises because LBW is likely correlated with a range of socio-economic and genetic characteristics of families and their children. For example, LBW infants are more likely to be born to poor families; as a result, it is difficult to disentangle the effects of birth weight from that of family income. Our approach is to use twin comparisons to distinguish between LBW effects and the effects of other socio-economic and genetic factors.

We do this using a unique dataset on the population of births in Norway matched with later outcomes. We advance the recent literature by using twin fixed effects on a large sample of individuals and looking at both short- and long-run outcomes for the same cohort of individuals. The current literature has examined the effects of low birth weight using within-family and, most recently, within-twin estimates of the effect of birth weight on both early outcomes (Almond et al 2005) and later outcomes (Berhman and Rosenzweig 2004) separately. Almond et al suggest that cross-sectional estimates of the

effects of birth weight greatly overstate the true effects when they apply twin fixed effects and look at early outcomes such as one year mortality rates; in contrast, Behrman and Rosenzweig find exactly the opposite (the within-twin estimates are much larger than OLS estimates) when they look at longer run outcomes such as education, height, and wages. No research to date has studied both short-run and long-run outcomes for the same cohorts. This paper fills this void in the literature.

We find that birth weight does matter. Consistent with earlier findings, we find that twin fixed effects estimates of the effect of birth weight on short-run outcomes such as one-year infant mortality are much smaller than their cross-sectional equivalents. However, these short-run studies can be misleading, as we find that birth weight has a significant effect on longer-run outcomes such as height, IQ at age 18, earnings, and education, and the fixed effects estimates are similar in size to cross-sectional ones.

When studying long-run outcomes, an important selection issue arises because twin pairs that experience infant mortality are dropped from the analysis. Because, unlike previous studies, we have information on individuals from birth to the labor market, we can investigate the potential impacts of such bias. Our investigation concludes that selection bias most likely leads to an understatement of the effects of birth weight on adult outcomes.

The paper unfolds as follows. Section 2 reviews the relevant literature, Sections 3 and 4 discuss our methodology and data. Section 5 presents our results and robustness checks, and Section 6 focuses on the selection bias that arises when studying adult outcomes. Section 7 addresses issues of generalizability, Section 8 discusses the magnitudes of the estimates, and Section 9 concludes.

2. Relevant Literature

There is a long history of research across disciplines relating low birth weight to poor health, cognitive deficits, and behavioral problems among young children, as well as some evidence that this relationship persists for longer-term outcomes. For example, in the medical literature, Barker (1995) finds evidence that fetal under-nutrition is related to coronary heart disease later in life.² In the economics literature, Currie and Hyson (1998) find a relationship between birth weight and educational attainment, employment, wages, and health status at age 23 and age 33. More recently, Case et al. (2004) show that, controlling for family background measures, children with low birth weight and poorer childhood health indicators have significantly lower educational attainment, poorer health, and lower SES as adults. However, it is possible that there is no causal relationship underlying these correlations, as low birth weight may be correlated with many difficult-to-measure socio-economic background and genetic variables.

To ascertain the causal effect of birth weight on child outcomes, one requires variation in birth weight that is unrelated to genetic and socio-economic characteristics. Since researchers cannot randomly assign birth weights to children, the few papers that have not used cross-sectional approaches instead use sibling or twin fixed effects models.

The sibling fixed effects approach involves comparing the outcomes of siblings who differ in birth weight. This approach conditions out any characteristic that is family-

² Typically, medical studies have limited data on longer-run outcomes and small sample sizes. For example, Hack et al (1994) finds an effect of very low birth weights on school age outcomes using 68 treatment children using across family comparisons and Hack et al (2002) compare 242 very low birth weight young adults to 233 normal birth weight controls and find that the educational disadvantage associated with very low birth weight persists into early adulthood. Recent work in the Norwegian medical literature also finds a positive relationship between birth weight and adult outcomes. (Eide et al. (2005) and Grjibovski et al. (2005)).

specific and unchanging over time but is not robust to time-varying unobserved characteristics, such as maternal smoking behavior and the quality of pre-natal care received by the mother, which may vary across pregnancies. Also, siblings share only about 50% of their genetic material so there may be genetic differences across siblings that are correlated with birth weight.³

Most recently, the literature has moved to within twin variation to identify the effects of birth weight on children's outcomes. Both Conley, Strully, and Bennett (2003) and Almond, Chay, and Lee (2005) use U.S. data to identify the effects of birth weight on short run health outcomes, including mortality. Almond et al. conclude that the effects of low birth weight are substantially smaller than originally thought; Conley et al. have estimates of similar magnitudes. However, neither of these studies is able to look beyond short-run health outcomes.

Behrman and Rosenzweig (2004) use a subset of the Minnesota Twin Registry to do fixed effects using female monozygotic twins and examine the longer run effects of birth weight. They find evidence that the heavier twin goes on to be taller, have greater schooling attainment, and to have a higher wage. In contrast, they find no evidence of effects on adult body mass index. However, the sample sizes are small (804 cases) and because of the numerous surveys required, there is substantial attrition and item non-response that may not be random. Our study differs from Behrman and Rosenzweig in that we study both men and women, use larger nationally representative samples from

³ Conley and Bennett (2000) take a sibling fixed effects approach using data from the Panel Study of Income Dynamics (PSID). They find a negative association between LBW and timely high school graduation. Currie and Moretti (2005) use sibling comparisons to investigate the effects of birth weight on the birth weight of the child's future children. They also show that among mothers that were siblings, the women with the lower birth weight resided in a lower income zip code on average when she gave birth to her own child years later.

administrative data, and use outcome variables that are recorded in administrative registers rather than self-reported. Unlike them, we can include controls for birth order of twins and exclude twins that are born with congenital defects. Our sample also differs from theirs in that we study more recent cohorts (1967-81 compared to their 1936-55 cohorts). As such, the technology of birth and social conditions growing up should be more similar to those in the present day.

3. Conceptual Framework

Let

$$y_{ijk} = \alpha + \beta bw_{ijk} + x_{jk}'\gamma + f_{jk} + \varepsilon_{ijk} \quad (1)$$

where subscript i refers to the child, j refers to the mother, and k refers to birth. y_{ijk} is then the outcome of child j born to mother i in birth k , bw_{ijk} is birth weight, x_{jk} is a vector of mother- and birth-specific variables (for example, mother's education, the year of birth), f_{jk} refers to unobservables that are mother- and birth-specific (for example, the quality of pre-natal care, genetic factors), and ε_{ijk} is an idiosyncratic error term assumed independent of all other terms in the equation.⁴

The parameter of interest is β -- the effect of birth weight on the outcome variable holding constant all observed and unobserved mother- and birth-specific variables. Note that this is likely to be the policy relevant parameter as policies aimed at increasing birth weight cannot influence fixed mother-specific characteristics such as

⁴ In our estimation, we also include variables that may vary at the individual level within birth, such as the birth order of the twin and the sex of each child.

genetics and typically will not affect other mother- and birth-specific characteristics such as maternal education or the timing of the birth.

Cross-sectional estimation of equation (1) by OLS will generally lead to biased estimates of β because of the presence of elements of x_{jk} and f_{jk} that influence both birth weight and child outcomes (for example, genetics or maternal education).

Therefore, we take a twin fixed effect approach to estimation. That is, our sample is composed of twin pairs and we included dummy variables for each birth in the regression. Denoting the first-born twin as “1”, and the second-born as “2”, this can be written in differences as follows:

$$y_{i1k} - y_{i2k} = \beta(bw_{i1k} - bw_{i2k}) + (\varepsilon_{i1k} - \varepsilon_{i2k}) \quad (2)$$

Given the assumption that ε_{ijk} is independent of bw_{ijk} , the twin fixed effects estimator of β is obviously consistent. This assumption is more likely to hold in the case of monozygotic twins (who are genetically identical) than with fraternal twins (who on average share about 50% of genes). Our sample contains both monozygotic and fraternal twins and we cannot differentiate between the types of twins. While the medical literature suggests that adult health outcomes among fraternal twins are similar to those among identical twins (Christensen et al. 1995 and Duffy 1993), we are able to assess the seriousness of this assumption by also examining the subset of same-sex twin pairs. Given boy-girl twins must be dizygotic, we show later that fixed effects estimates for the same-sex pairs are similar to fixed effects estimates for the full sample of twins, suggesting that the results do not differ substantially by zygosity.⁵

⁵ Conley et al. (2005) take a more formal approach making use of what is known as the “Weinberg rule” that an equal number of fraternal twins should be in same sex pairs as in mixed sex pairs. Given one knows that all mixed-sex pairs are fraternal, this gives the total number of monozygotic twin pairs. Then if one

Why Does Birth Weight Differ Within Twin Pairs?

Low birth weight can arise either because of short gestational length (pre-term delivery) or because of low fetal growth rate, commonly known as intrauterine growth retardation (IUGR). When we look within twin pairs, gestation length is the same and differences in birth weight arise solely due to differences in fetal growth rates.⁶

Given that gestation is the same among twins, evidence suggests that much of the difference in birth weight is due to differences in nutritional intake. In the case where there are two placentas (called dichorionic, including all fraternal twins and about 30% of identical twins), nutritional differences can arise because one twin is better positioned in the womb. Among single-placenta (monochorionic) twins, nutritional differences are related to the location of the attachment of the two umbilical cords to the placenta and the placement of the fetus within the placenta. (Bryan 1992, Phillips 1993). Hence, since there are no genetic differences, birth weight differences within monozygotic twin pairs appear to come primarily from differences in nutritional intake.⁷

To the extent the differences in birth weight among twins results from these random environmental differences in the womb, twin differences is an excellent approach for studying the effects of birth weight on child outcomes. Of course, among fraternal twins, genetic differences may also play a role in determining birth weight.

4. Data

assumes that the effects of birth weight are constant across the birth weight distribution and that these effects are not a function of the gender of the other twin, one can back out the implied effect of birth weight amongst monozygotic twins (see Conley et. al. page 18). However, the more recent medical literature is quite skeptical of this approach. See Keith et al (1995) for more discussion.

⁶ While there are rare cases of twins who are not born at the same time, these twins are not included in our sample.

⁷ There is an extensive medical literature examining the determinants of birth weight differences (called discordance) among twins. See Blickstein and Kalish (2003) for a summary.

Our primary data source is the birth records for all Norwegian births over the period 1967 to 1997 obtained from the Medical Birth Registry of Norway. All births, including those born outside of a hospital, are included as long as the gestation period was at least 16 weeks.⁸ The birth records contain information on year and month of birth, birth weight, gestational length, age of mother, and a range of variables describing infant health at birth including APGAR score, malformations at birth, and infant mortality (defined as those who die within the first year).⁹ We are also able to identify twin births and the birth order of twins but cannot distinguish between fraternal and monozygotic twins. Because the dataset includes identifiers for mothers and fathers, we can also identify siblings. Furthermore, these identifiers allow us to identify multiple generations of families.

All persons in Norway have a unique personal identifier given at birth, so all administrative registers can be matched. To these birth records we matched several other administrative registers, including information on parents and family background for each person (such as parental education, age, and income) as well as adult outcomes for each person.

The main administrative register is from the Norwegian Registry Data, a linked administrative dataset that covers the entire population of Norwegians aged 16-74 in the 1986-2002 period, and is a collection of different administrative registers such as the education register, family register and the tax and earnings register. These data are maintained by Statistics Norway and provide information about educational attainment, labor market status, earnings, and a set of demographic variables (age, gender) as well as

⁸ The data also include stillbirths, which constitute approximately one percent of the sample.

⁹ Malformations are coded according to the international medical standard (ICD8).

information on families.¹⁰ When considering educational attainment, we use all individuals aged at least 21 in 2002 and use as our dependent variable a binary indicator for whether the person has at least 12 years of education.¹¹

Another source of data is the Norwegian military records from 1984 to 2005 which contains information on height, weight, and IQ. In Norway, military service is compulsory for every able young man. Before entering the service, their medical and psychological suitability is assessed; this occurs for the great majority between their 18th and 20th birthday.¹² For the cohorts of men born from 1967 up to 1987, we have information on height, weight, and Body Mass Index (BMI), all of which were measured as part of the medical examination.¹³

We also have a composite score from three speeded IQ tests -- arithmetic, word similarities, and figures (see Sundet et. al. 1988, 2004, Thrane, 1977, and Notes from the Psychological Services of the Norwegian Armed Forces, 1956 for details). The arithmetic test is quite similar to the Wechsler Adult Intelligence Scale (WAIS) (Sundet, 2005, Cronbach, 1964). The word test is similar to the vocabulary test in WAIS, and the figures test is similar to the Raven Progressive Matrix test (Cronbach, 1964). The composite IQ test score is an unweighted mean of the three subtests. The IQ score is

¹⁰ Our measure of educational attainment is reported by the educational establishment directly to Statistics Norway, thereby minimizing any measurement error due to misreporting. The educational register started in 1970; for parents who completed their education before then we use information from the 1970 Census. Thus the register data are used for all children and for all parents who had any education after 1970. Census data are self reported but the information is considered to be very accurate; there are no spikes or changes in the education data from the early to the later cohorts. See Møen, Salvanes and Sørensen (2003) for a description of these data.

¹¹ While we describe this as high school completion, in Norway many individuals with 12 years of education obtain vocational rather than academic qualifications.

¹² Of the men in the 1967-1987 cohorts, 1.2 % percent died before 1 year and 0.9 percent died between 1 year of age and registering with the military at about age 18. About 1 percent of the sample of eligible men had emigrated before age 18, and 1.4% of the men were exempted because they were permanently disabled. An additional 6.2 percent are missing for a variety of reasons including foreign citizenship and missing observations. See Eide (2005) for more details.

¹³ BMI is calculated as kilograms divided by meters squared.

reported in stanine (Standard Nine) units.¹⁴ We match these data with our other data files and use the height, BMI, and test score data as outcome variables for men.

Labor Market Outcomes

While education is one measure of human capital, it does not capture many less observable forms of human capital. Because these may be revealed in the later labor market outcomes of children, we report estimates of the effects of birth weight on the earnings of young adults.

We look at both the earnings of all labor market participants and the earnings of full-time employees. Earnings are measured as total pension-qualifying earnings reported in the tax registry. These are not topcoded and include all labor income of the individual. We restrict attention to individuals aged at least 21 who are not in full time education. In this group, about 86% of both men and women have positive earnings. Given this high level of participation, our first outcome is log(earnings) conditional on having non-zero earnings.

Since the results for this variable encompass effects on both wage rates and hours worked, we also study the earnings of individuals who have a strong attachment to the labor market and work full-time (defined as 30+ hours per week). To identify this group, we use the fact that our dataset identifies individuals who are employed and working full time at one particular point in the year (in the 2nd quarter in the years 86-95, and in the 4th

¹⁴ Stanine units are a method of standardizing raw scores into a nine point standard scale with a normal distribution, a mean of 5, and a standard deviation of 2. A scale is created with nine intervals, each interval representing half of a standard deviation. The 5th stanine straddles the midpoint of the distribution, covering the middle 20% of scores. Stanine 6, 7, and 8 cover the top end of the distribution and 4, 3, 2, and 1 fall below the mid-point with lower scores. For scores expressed in stanines, normalizing will put 4% of the sample in the first stanine, 7% in the second, and so on through 12%, 17%, 20%, 17%, 12%, 7%, and 4%. This method of standardization assumes that whatever ability the test measures is evenly distributed around a central peak

quarter thereafter).¹⁵ We label these individuals as full-time workers and estimate the earnings regressions separately for this group. About 63% of male participants and 46% of female participants in our sample are employed full time over this period.

Sample and Summary Statistics

When we analyze mortality, we use birth data from the entire 1967-1997 period. However, for the other outcomes, we cannot use data from the later part of this period because we need to observe individuals as adults. The ranges of cohorts that we can use differ by outcome and are reported in the tables and in the results section.

We drop twin pairs for which gestation length is unknown (about 4% of cases). We also dropped twin pairs where one or both of the twins were born with a congenital defect (approximately 2.1%). One argument for using the within-twin and within-same-sex-twin approaches is to make the two children as similar as possible with the exception of their birth weight. Congenital defects suggest an underlying difference between the twins. Results without dropping these individuals were slightly stronger for mortality but quite similar for the later outcomes.¹⁶

Tables 1 and 2 present summary statistics for our sample.¹⁷ Statistics are broken down into twin and non-twin samples in Columns 1 and 2 and the twin sample is reduced to same-sex twin pairs by sex in Columns 3 and 4.

¹⁵ An individual is labeled as employed if currently working with a firm, on temporary layoff, on up to two weeks of sickness absence, or on maternity leave.

¹⁶ When analyzing later outcomes, we also tried adding controls for the 5-minute APGAR score and an indicator for whether there were complications at birth. Adding these controls had a negligible effect on birth weight estimates, suggesting that the birth weight effects are not picking up unobserved health differences between twins.

¹⁷ Note that we are missing observations on APGAR scores for the earliest years in our sample. APGAR scores became available in 1977.

When estimating the effect of birth weight on high school graduation and earnings, we are limited to using the birth data from the earliest period in our sample (1967-1981) so individuals are aged at least 21 when the outcome is measured. Summary statistics from this period are presented in Table 2. However, we know that over the entire time period, fertility and infant mortality have been changing. In Figure 1, we see that the rate of twin births relative to all births was roughly constant up until the late 1980s; with the advent of fertility treatments, the incidence of twins rose. Figure 2 provides further evidence of this phenomenon; we see that, as of the late 1980s, the fraction of twin births that are same sex is declining. The increase we observe in the incidence of opposite sex twins, who cannot be identical, is consistent with fertility treatments having a larger effect on the incidence of fraternal twins.¹⁸ In addition, infant mortality has declined; see Figures 3 and 4. We discuss the implications of these changes for our results in Section 6, when we consider selection of individuals into our samples.

Differences between Heavier and Lighter Twins

There is substantial variation in birth weight within twin pairs; 21% of the variation in birth weight is within-twin. Figure 5 shows the distribution of the within-twin pair differences in birth weight.

A simple comparison of means between heavier and lighter same sex twins for the early sample period (1967-81) presents a preview of our results. (See Table 3.) The average difference in birth weight between heavier and lighter twins is about 320 grams, with the lighter twin being more likely to be LBW by a margin of .44 to .26. If birth

¹⁸ Later in the paper, we examine whether there are differential effects of birth weight in the earlier versus later periods of our data.

weight matters, we would expect there to be commensurate average differences between heavier and lighter twins in the outcome variables. Average mortality is slightly *lower* for the lighter twins, suggesting increased birth weight does not reduce mortality. However, there is evidence of differences in later outcomes that favor heavier twins. For example, the heavier twin has a higher probability of finishing at least 12 years of schooling and is, on average, almost a centimeter taller as a young adult.

More formally, we can look at the Wald estimates, calculated as the average difference in outcome between heavier and lighter twins divided by the average difference in log birth weight between heavier and lighter twins. While this approach is obviously less efficient than the twin fixed effects we will do later, it is a useful “first pass” as it is less parametric (as it just compares differences in means) and is less susceptible to measurement error.¹⁹ The Wald estimates are presented in Column 3 of Table 3.²⁰ They demonstrate a statistically insignificant effect of birth weight on mortality but positive and statistically significant effects on high school graduation, height, BMI, and ability. The earnings estimates are positive but statistically insignificant. These results are quite suggestive and preview our ultimate conclusions based on twin fixed effects analysis.

5. Results

Choice of Birth Weight Variable

¹⁹ Unlike fixed effects, the Wald estimator is consistent as the number of twin pairs goes to infinity even in the presence of measurement error in birth weight.

²⁰ The Wald estimator can be implemented with covariates by instrumenting log birth weight with an indicator variable for whether the person is the heavier twin. Including covariates has a negligible effect on the estimates.

In the literature, different variants of birth weight have been used as the primary variable of interest. These include birth weight, log(birth weight), fetal growth (defined as birth weight divided by weeks gestation), and an indicator for low birth weight (<2500 grams). Given that there is no obvious choice a priori, we have examined the explanatory power of these variables in the twin fixed effects regressions. The resulting R^2 statistics with the various dependent variables and each variant of birth weight run separately are presented in Appendix Table 1. They indicate that the three continuous measures perform quite similarly with the exception that log(birth weight) is a better fit for mortality. Thus, we use this variable in our analysis.²¹ It is interesting to note that the LBW indicator fits very poorly for most outcomes. This suggests that using cutoffs such as <2500 grams as the variable of interest may not be appropriate for this type of analysis.²²

OLS and Fixed Effects Estimates

We first examine the sample of all twins and compare the results when we use pooled OLS versus a twins fixed-effect estimation strategy. The control variables we use in the OLS estimation are year- and month-of-birth dummies, indicators for mother's education (one for each year), indicators for birth order (which is known to be correlated with birth weight and also a strong predictor of outcomes in Norway, see Black, Devereux, and Salvanes 2005a), indicators for mother's year of birth (one for each year to allow for the fact that age of mother at birth may have independent effects on child outcomes), and an indicator for the sex of the child. With twin fixed effects, all controls

²¹ Estimates are very similar when either of the other two continuous measures are used. To demonstrate this, we have included the results for all three variants for the basic specifications in Appendix Table 2.

²² We have tried including both log(birth weight) and an indicator for LBW (<2500 grams) in the same specifications. The continuous measure dominates for all outcomes and the effect of LBW is always statistically insignificant and often has the wrong sign.

are differenced out except the indicators for sex and birth order (either 1st born or 2nd born twin).

Table 4 presents these results. Each coefficient represents the results from a separate regression. We present the results in approximate chronological order so that outcomes measured earlier in the life-cycle come first. The outcomes we examine include one year mortality rates, height, BMI, and ability measures for men aged 18-20, an indicator for high school completion (at least 12 years of education) for all individuals aged at least 21, the log of earnings for all individuals who are aged at least 21 and working, and the log of earnings if the individual is working full time (and aged 21+).

Mortality

We begin by carrying out an analysis similar to Almond et al. using one year mortality (a dummy variable that equals one if a child dies within one year of birth and zero otherwise) as our outcome. The pooled OLS coefficient of -279 implies that a 10 percent increase in birth weight would reduce one-year mortality by approximately 27.9 deaths per 1000 births (the mortality rate for twins is about 50 deaths per 1000 births). The fixed effects coefficient of -41 is statistically significant but only one fifth the size of the OLS coefficient. When we use linear measures of birth weight, our estimates are almost exactly identical to the estimates of Almond et al. for the U.S., suggesting that the infant health production function may be similar in the U.S. and Norway.²³ (See Appendix Table 2.)

²³ Because infant mortality is a rare outcome, estimated derivatives may be sensitive to functional form. When we assume other functional forms and estimate logit or probit equations instead of linear probability models, we get very different marginal effects (smaller by a factor of 6) in the pooled estimation. Marginal effects from a fixed effects conditional logit model are also very different from the linear twin fixed effects

Height, BMI, and IQ at age 18-20 for Men

As discussed earlier, we also have information on height, weight, and IQ test scores for draft age men. Because individuals are at least 18 when they take the test and our latest test date is in 2005, all men come from the 1967-1987 cohorts. To take account of the fact that men take the test in different years and at different ages, we add dummies for the test year to the controls used earlier.

Table 4 shows the strong positive effects of birth weight on height, BMI, and IQ. Height is measured in centimeters so the OLS estimate suggests that a 10 percent increase in birth weight translates into about .75 extra centimeters of height at around age 18, and an increase in BMI of around .06.²⁴ Fixed effects estimates are quite similar, with a 10% increase in birth weight leading to a .57 centimeter increase in height and a .10 increase in BMI. Our IQ measure is on a scale from one to nine; the estimated fixed effects coefficient of 0.62 suggests that an increase in birth weight by 10 percent will increase the score by .06 (about 1/20th of a stanine).

Given that BMI is an ambiguous health measure, as health may be adversely affected if BMI is too high (so men are overweight) or BMI is too low (so men are underweight), we have used the Center for Disease Control (CDC) cutoffs for overweight (BMI greater than or equal to 25 – 11% of the twins sample) and underweight (BMI less than 18.5 – 8% of the twins sample) to analyze the effect of birth weight on the

estimates (not very surprising given the selection problem induced by the fact that the logit only includes cases in which one twin lives and one twin dies). Given the sensitivity to functional form, one is left questioning the credibility of twin fixed effects estimates in the case of these rare events. It is reassuring, however, that the fixed effects estimates are similar to the Wald estimates discussed earlier.

²⁴ There is an extensive literature suggesting that height is a useful indicator of health, both in developed as well as developing nations. See Strauss and Thomas (1998) for references.

probability of being in either of these two groups. The fixed effects estimates show that increased birth weight significantly increases the probability of being overweight but significantly decreases the probability of being underweight.²⁵

Educational Attainment

We find that the within twin estimates of the effect of birth weight on education are similar to the OLS estimates and statistically significant.²⁶ The magnitude implies that an increase in birth weight of 10 percent increases the probability of high school completion by a bit less than 1 percentage point. This suggests that, although OLS estimates greatly overestimate the effect of birth weight on mortality, the relationship between birth weight and later education remains strong.²⁷

Labor Market Outcomes

To maximize efficiency, we use all observations on individuals in the 1986-2002 panel, provided they are aged at least 21. We exclude observations from any year in which the relevant dependent variable is missing for either twin i.e. in the earnings regressions each twin must have positive earnings in a particular year for it to be included. Because we have people from many different cohorts, individuals are in the

²⁵ Compared to Behrman and Rosenzweig (2004), we find smaller effects of birth weight on height and larger effects of birth weight on BMI. The estimates are not directly comparable as theirs are for middle-aged women while ours are for young men.

²⁶ We also tried looking at completed years of education of individuals aged 25 or more. However, this required us to substantially reduce our sample size. As a result, although the results were consistent with the conclusions derived from the high school graduate results, standard errors were too large to make any independent inference.

²⁷ Unlike with mortality, logit and probit marginal effects for high school graduation are very close to those from the linear probability model. However, fixed effects logit marginal effects are larger than the fixed effects linear probability model estimates. The Wald estimates in Table 3 apply a more non-parametric approach and find estimates that are a little larger than the fixed effects estimates.

panel for different sets of years and at different ages. Therefore, as before, we control for cohort effects. Also, we augment the previous specification by adding indicator variables for the panel year. This takes account of cyclical effects on earnings etc.

The standard errors are adjusted to take account of the fact that there are multiple observations on individuals. To make this more tractable in the fixed effects case, we implement the fixed effects by first calculating twin differences for each year. This washes out the twin fixed effects and leaves one observation per twin pair per year.²⁸

The estimates imply that the OLS and fixed effects estimates are similar; both suggest that an extra 10 percent of birth weight raises earnings by about 1%. Given the return to education in Norway has been estimated to be about 4% for men (Black et al. 2005b), this suggests that 10% more birth weight is as valuable in the labor market as a quarter of year of education.

Same-Sex Twins

One concern with this estimation is that we may be comparing fraternal twins who are not genetically identical and may have different optimal birth weights; as a result, differences in birth weight would not reflect deviations from the optimum that result from nutritional differences. In addition, male and female babies might be differentially sensitive to birth weight differences. To investigate these issues, we first split our sample into same sex twin pairs. While we are unable to limit our sample to monozygotic twins only, by eliminating opposite twin pairs (which are clearly not monozygotic), the sample now contains a larger fraction of identical twin births. If the

²⁸ Some individuals are present in more periods than others and, hence, have greater weight in estimation. We have verified that if we weight each individual equally in estimation, we get similar but less precisely estimated coefficients. None of our conclusions change.

results are similar for same sex and all twins, this will suggest that it isn't underlying genetic differences driving our results. Table 4a reports OLS and fixed effects estimates for all same-sex twin pairs.²⁹ The estimates are very similar to those for the full twin sample, suggesting there are not large differences in estimates by zygosity.

We next separate the same sex twins sample by gender; the results by gender are presented in Table 5. As we believe the twin fixed effects estimates are more credible than the pooled OLS, from here on we only report the fixed effects results. For both men and women, there are significant effects of birth weight on one-year mortality. Note that the magnitude of these effects is similar for both men and women and is almost identical to the estimates from the whole twins sample, suggesting that the results for fraternal and monozygotic twins may not differ much. Consistent with Almond, et al, the effects of birth weight on infant mortality are quite small.

However, among men, birth weight has no statistically significant effect on educational attainment but a significant effect on earnings. For women, the opposite is true, as birth weight has a significant effect on educational attainment but not on earnings. Of course, the earnings of women in our sample are particularly prone to selection problems due to non-participation and this may be partly responsible for this result.³⁰

Heterogeneous Effects across the Birth Weight Distribution

²⁹ The outcomes from the military data are excluded from Table 4a because the estimates in Table 3 are for a sample of same-sex male twins.

³⁰ In contrast, Behrman and Rosenzweig (2004) find very large effects of fetal growth on female earnings (6 times as large as their OLS estimates).

While using the natural log of birth weight does allow for non-linear effects, it is possible to allow the effects of birth weight to be more flexible. Figures 6-10 do this graphically. For example, Figure 6 illustrates the differences between the OLS estimates for mortality and those with the twin fixed effects across the birth weight distribution by presenting the average mortality rate by birth weight, both with and without twin fixed effects.³¹ It is clear that not only are the twin fixed effects estimates much smaller than the OLS, but there is also evidence of significant nonlinearities in the relationship. In particular, increased body mass has a negative effect on mortality at low birth weights but little discernable effect at weights above 1500 grams.

Figure 7 demonstrates that, unlike with mortality, OLS and twin fixed effects estimates for height are very similar. Once again, there is some evidence of a non-linear relationship with the positive relationship between birth weight and height flattening out after about 1500 grams. The equivalent figures for BMI (Figure 8), ability (Figure 9), and high school graduation (Figure 10) shows once again that OLS and fixed effects estimates are very similar across the distribution. In all cases, there is a hint of more positive effects at very low weights. However, there is little evidence of strong nonlinearities. In all figures, the estimates are noisy at very low and very high weights, reflecting the paucity of data in these regions.

An alternative approach to allow for non-linearities is to estimate the effect of birth weight for these outcomes allowing for splines in birth weight with less than 1500, 1500-2500, and 2500 or more as the cutoffs. These results are presented in Table 6 for the full sample. (Appendix Tables 3a and 3b present the results when we break the sample by sex.) It is clear there are substantial non-linearities in mortality, with a large

³¹ Note that the averages only include controls for twin indicators and no other controls.

marginal benefit for additional grams among very low birth weight babies in terms of mortality. However, as was demonstrated in Figures 6-10, there is little evidence of significant non-linearities in later outcomes.

6. Selection into the Later Outcomes Sample

When looking at the effect of birth weight on later outcomes, we are inherently only including those individuals for whom we observe later outcomes. In particular, individuals who did not live will not be included in our sample. To the extent that birth weight may affect mortality, this may bias our results estimating the effect of birth weight on later outcomes. Given that there is evidence that selection into the sample may be changing over time (See Figures 3 and 4 for evidence of declining infant mortality rates over time), it is important to understand how it may be affecting our results.³² Note that, unlike previous twin studies of later outcomes, we can observe characteristics of twin pairs that are not used in the later outcomes sample due to infant mortality.

Though it is inherently impossible to know what the effects of birth weight would have been on the later outcomes of the individuals we do not observe, we do think about this selection from a number of perspectives. If there are heterogeneous effects of birth weight across twin pairs, and if there is a positive correlation between the effects of birth weight on early and later outcomes, we would expect that twin pairs that experience mortality are pairs for which birth weight would also be disproportionately important for

³² We have estimated the effect of birth weight on mortality separately for the sample used in our analysis of later outcomes. These results are presented in Appendix Table 4 and are similar to those from the full sample period.

later outcomes. This reasoning suggests that early mortality will tend to reduce the estimated effect of birth weight on later outcomes.

To examine this, we consider another intermediate outcome that is observed for the entire sample and see if the effect of birth weight on the sample that die is significantly different from the effect of birth weight on the sample that live. If the effect of birth weight on this intermediate outcome is the same for the sample who lived and those who did not, then one might believe that the effect of birth weight on the later outcomes would also be the same. Beginning in 1977, we observe the 5-minute APGAR score for all individuals. As a check, we separately estimate the relationship between birth weight and APGAR for the full sample and the sample of twin pairs where both twins live; when we do this using twin fixed effects, we find that log birth weight has a significantly larger positive effect on the APGAR score for the full sample of twin births. The difference is large -- .35 (.07) for the full sample, versus .19 (.06) for the sample without mortality. If this relationship is also true of other, later outcomes, then we may be underestimating the true effect of birth weight on later outcomes by a substantial amount.

Finally, a formal approach to the missing data problem is to model the probability that a twin pair will experience mortality within the first year and hence attrit from the later outcomes sample. We allow the probability of attrition to depend on the variables that are always observed which we denote as x_{it} (which includes all the usual control variables plus the birth weight of each twin and indicator variables for whether each twin is LBW), but do not allow dependence on the variables that are missing for some units (the later outcomes). The estimation of the model is carried out in two steps. In the first stage, we estimate a probit model that conditions the probability of attrition on x_{it} . The

predicted probabilities from the probit model are used to form weights and these weights are used to weight the observations in the twin fixed effects estimation in the second step. The weights are equal to the inverse of the probability of not attriting due to mortality in the first year.³³ When we do this reweighting, we again find that our estimates are likely underestimating the true effect of birth weight on later outcomes.³⁴

7. External Validity

While using within twin variation allows us to credibly identify the causal effect of birth weight on later outcomes, the question as to how generalizable these results are to the general population of births remains.

From Table 1, we can see that there are substantial differences between twin and singleton births. Not surprisingly, non-twins are on average heavier, with only 4 percent classified as low birth weight (less than 2500 grams), while 39 percent of twins are low birth weight. Gestation is also longer for singletons, with the average at 39.8 weeks versus 36.9 for twins. Both one minute and five minute APGAR scores are also higher, there are a lower fraction with complications, and the one-year mortality rate is only 7.7 per 1000 births as opposed to 33 for twins. Parental education is similar but the mothers of twins tend to be older.

One of the most notable differences is that twins come disproportionately from the lower part of the birth weight distribution; this can be seen in Figure 11, which shows

³³ This is referred to as Missing at Random (Little and Rubin, 1987). The argument is that there is nothing in the data that suggests that units that drop out are systematically different from units who do not drop out once we condition on all observed variables. This model has some intuitive appeal. Consider a unit that drops out in the first year with values of the observed variables equal to $X_{it} = x_{it}$. The Missing at Random assumption implies that for our best guess of the value of the missing later outcome variables, we should look at values of the later outcomes for units with the exact same values of x_{it} .

³⁴ The differences between the weighted and unweighted estimates are not large being of the order of 3-5%.

the distribution of birth weight for twins and non-twins. The question then becomes, are the outcomes of twins and singletons similar controlling for birth weight? To examine this, we have graphed the relationship between birth weight and mortality, education, height, BMI, and IQ for the sample of twins and non-twins. (See Figures 6-10.) It is interesting to note that the twins and non-twins actually have quite similar outcomes conditional on birth weight, suggesting that our results may be generalizable to the rest of the population.³⁵ This is consistent with findings in the medical literature that suggest that the primary cause of disparities in outcomes between twins and singletons is due to differences in size at birth. Allen (1995) notes that, in a sample of pre-term births, no differences were present between twins and singletons with respect to neurodevelopmental outcomes at 18 months from due date, after adjusting for confounding social, obstetric and neonatal factors.³⁶

8. Magnitudes

Many studies have used birth weight as an outcome variable in the belief that it has important implications for longer-run outcomes. Our estimates can be used to shed light on how sizeable the magnitudes in these types of studies are. For example, Currie and Moretti (2003) find that one year of extra maternal education reduces the probability of having a low birth weight child by 1 percentage point from a baseline of about 5 percent of children being born LBW. If we use this simple indicator and estimate the

³⁵ Of course, we cannot rule out the possibility that twins and singletons have very different causal relationships between birth weight and outcomes but that they are subject to different confounding factors that happen to cancel each other out so that the cross-sectional profiles are similar.

³⁶ Differences were only found when they examined pre-term infants with birth weights of <800 grams, suggesting greater vulnerability of twins born at the limit of viability. See also Hoffman and Bennett (1990).

effect of being low birth weight on our outcome measures, we can then calculate the effect of increasing mother's education on children's outcomes that occurs through increased birth weight. When we do this, the fixed effects estimates imply that 1 year of maternal education increases male earnings by approximately .067 percent and has an even smaller effect on annual earnings of full time workers. For men, fixed effect estimates imply one year of education increases height by .01 centimeters. Likewise, for women, one year of education increases the probability of graduating high school by .0004. Overall, these numbers suggest that the impact of one extra year of maternal education coming through birth weight on later outcomes is very small. However, magnitudes may be larger if we had estimates of the effects of maternal education on a more continuous measure of birth weight, as our evidence suggests there is little effect of the 2500 gram cut-off on longer run outcomes.

While we find that birth weight effects for later outcomes are statistically significant and similar in magnitude to cross-sectional relationships, it is difficult to determine whether these are big effects without reference to specific policy scenarios. One such scenario is the Irish Government's policy of reducing disparities in birth weight across socio-economic groups. In Norway (and in the United States), the average difference in birth weight between children of college graduate women and high school dropout women is about 200 grams. Our estimates suggest that if this disparity was eliminated by increasing the birth weight of the children of high school dropouts, the effects on male children of high school dropouts would be substantial. Earnings of male children would rise by about 2%, full time earnings would rise by about 1%, height would rise by about half a centimeter, and IQ would rise by $1/25^{\text{th}}$ of a standard

deviation. For female children, the proportion completing high school would increase by about .01. Thus, our estimates suggest that interventions to reduce socio-economic disparities in birth weight may have sizeable impacts on adult outcomes.

9. Conclusions

In this paper, we have examined the effect of birth weight on adult outcomes using within-twin variation in birth weight to control for other, often unobservable, parental and environment factors. Consistent with the recent literature, we find little if any relationship between birth weight and mortality, and OLS estimates greatly overestimate the true causal relationship. However, conclusions drawn from these results can be misleading, for we find a significant impact of birth weight on later outcomes of children, including height, BMI, and IQ, all at age 18, education, and earnings. In contrast with infant mortality, twin fixed effects estimates of the effects of birth weight for the later outcomes are similar in size to OLS estimates. Additionally, we find that the relationship is fairly linear, suggesting that earlier work using indicators for low birth weight (<2500 grams) may be misspecified.

Given that there are clearly long-run effects of birth weight differences, a natural question for future work is how investments of parents and society interact with birth weight to determine these effects. For example, parents with more resources may be better able to mitigate the negative effects of being born smaller.³⁷ We have found no statistically significant differences in the effects of birth weight by mother's education, by family income, or by birth order of the children. However, this in part reflects the

³⁷ There is some evidence of this in the cross-sectional and sibling fixed effects context, but not controlling for twin fixed effects. See Loughran et al (2004) for one example.

smaller sample sizes that arise when disaggregating the sample and so this area remains an important avenue for future research.

References

- Allen, M.C. (1995), "Factors Affecting Developmental Outcome" in *Multiple Pregnancy: Epidemiology, Gestation & Perinatal Outcome*. Keith, Papiernik, Keith, and Luke, eds. Parthenon Publishing: New York.
- Almond Douglas, Kenneth Y. Chay and David S. Lee (2005), "The Costs of Low Birth Weight", *Quarterly Journal of Economics*, 120(3).
- Barker, D. J. (1995), "Fetal Origin of Coronary Heart Disease", *British Medical Journal*, Vol. 317, 171-174.
- Behrman Jere R., and Mark R. Rosenzweig (2004), "Returns to Birth weight", *Review of Economics and Statistics*, May, 86(2), 586-601.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2005a), "The More the Merrier? The Effects of Family Size and Birth Order on Children's Education", *Quarterly Journal of Economics*, May.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2005b), "Why the Apple doesn't Fall Far: Understanding Intergenerational Transmission of Education", *American Economic Review*, March.
- Blickstein, Isaac and Robin B. Kalish, (2003). "Birthweight Discordance in Multiple Pregnancy." *Twin Research*, Volume 6, Number 6, pp. 526-531.
- Bryan, Elizabeth (1992), *Twins and Higher Order Births: A Guide to their Nature and Nurture*, London, UK: Edward Arnold, 1992).
- Case, Anne, Angela Fertig, and Christina Paxson, "The Lasting Impact of Childhood Health and Circumstance." Princeton Center for Health and Wellbeing Working Paper, April 2004.
- Christensen K, J.W. Vaupel, N. V. Holm, A.J. Yashin (1995). "Mortality Among Twins After Age 6: Fetal Origins Hypothesis Versus Twin Method." *British Medical Journal* Vol. 319, 432-36.
- Conley, Dalton, and Neil G. Bennett (2000), "Is Biology Destiny? Birth Weight and Life Chances", *American Sociological Review*, 65 (June), 458-467.
- Conley, Dalton, Kate Strully, and Neil G. Bennett (2003), "A Pound of Flesh or Just Proxy? Using Twin Differences to Estimate the Effect of Birth Weight on (Literal) Life Chances", mimeo.
- Cronbach, L. J. (1964), *Essentials of psychological testing*, 2.nd ed. London: Harper and Row.

- Currie, Janet and Jonathan Gruber (1996). "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women." *Journal of Political Economy*, Vol 104, No 6, pp. 1263-1296.
- Currie, Janet, and Rosemary Hyson (1998), "Is the Impact of Health Shocks Cushioned by Socio-Economic Status? The Case of Low Birth weight", mimeo.
- Currie, Janet, and Enrico Moretti (2003), "Mother's education and the Intergenerational Transmission of Human Capital: Evidence from College Openings," *Quarterly Journal of Economics*, CXVIII, 1495-1532.
- Currie, Janet, and Enrico Moretti (2005), "Biology as Destiny? Short and Long-Run Determinants of Inter-Generational Correlations in Birth Weight", mimeo.
- Duffy, D.L. (1993). "Twin Studies in Medical Research." *Lancet* Vol 341, 1418-19.
- 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 of Adult Height and Weight," *Epidemiology*, Vol. 16 (2), 175-181.
- Grjibovski, A. M., J. H. Harris and P. Magnus (2005). "Birthweight and adult health in a population-based sample of Norwegian twins," *Twin Results and Human Genetics*, Vol. 8(2), 148-155.
- Hack, Maureen, Daniel J. Flannery, Mark Schluchter, Lydia Cartar, Elaine Borawski, and Nancy Klein (2002). "Outcomes in Young Adulthood for Very-Low-Birth weight Infants." *New England Journal of Medicine*, Volume 246, Number 3, January 17, 2002, 149-157.
- Hack, Maureen, H. Gerry Taylor, Nancy Klein, Robert Eiben, Christopher Schatschneider, and Nori Mercuri-Minich (1994) "School-Age Outcomes in Children with Birth Weights under 750 grams." *New England Journal of Medicine*, Volume 331: 753-759. Number 12.
- Hoffman, E.L. and F.C. Bennett (1990). "Birth Weight Less than 800 Grams: Changing Outcomes and Influences of Gender and Gestation Number." *Pediatrics*, 86, 27.
- Keith, Louis G., Emile Papiernik, Donald M. Keith, and Barbara Luke, editors. (1995) *Multiple Pregnancy: Epidemiology, Gestation, & Perinatal Outcome*. Parthenon Publishing: New York.
- Little R.J.A. and D.B. Rubin (1987). *Statistical Analysis with Missing Data*. Wiley. Chichester.

- Loughran, David S., Ashlesha Datar, and M. Rebecca Kilburn (2004). "The Interactive Effect of Birth Weight and Parental Investment on Child Test Scores." RAND Labor and Population Working Paper # WR-168.
- Møen, Jarle, Kjell G. Salvanes and Erik Ø. Sørensen. (2003). "Documentation of the Linked Employer-Employee Data Base at the Norwegian School of Economics." Mimeo, The Norwegian School of Economics and Business Administration.
- Phillips, D.I.W. (1993). "Twin Studies in Medical Research : Can They Tell Us whether Diseases are Genetically Determined? *Lancet* Vol. 334, 1008-9.
- Strauss, John and Duncan Thomas. (1998) "Health, Nutrition, and Economic Development." *Journal of Economic Literature*, Vol. 36, No. 2 (June) 766-817.
- Sundet, Martin Jon, Dag G. Barlaug, and Tore M. Torjussen (2004), "The End of the Flynn Effect? A Study of Secular Trends in Mean Intelligence Test Scores of Norwegian Conscripts During Half a Century", *Intelligence*, 32, 349-362.
- Sundet, Jon Martin, Kristian Tambs, Jennifer R. Harris, Per Magnus, and Tore M. Torjussen (2005), "Resolving the genetic and environmental sources of the correlation between height and intelligence: A study of nearly 2600 Norwegian male twin pairs. *Twin Research and Human Genetics*, Vol 8(4), 1-5.
- Thrane, Vidkunn Coucheron (1977), *Evneprøving av utskrivingspliktige i Norge 1950-53. Arbeidsrapport nr. 26*, INAS.

Figure 1
Fraction of Twin Births Out of All Births

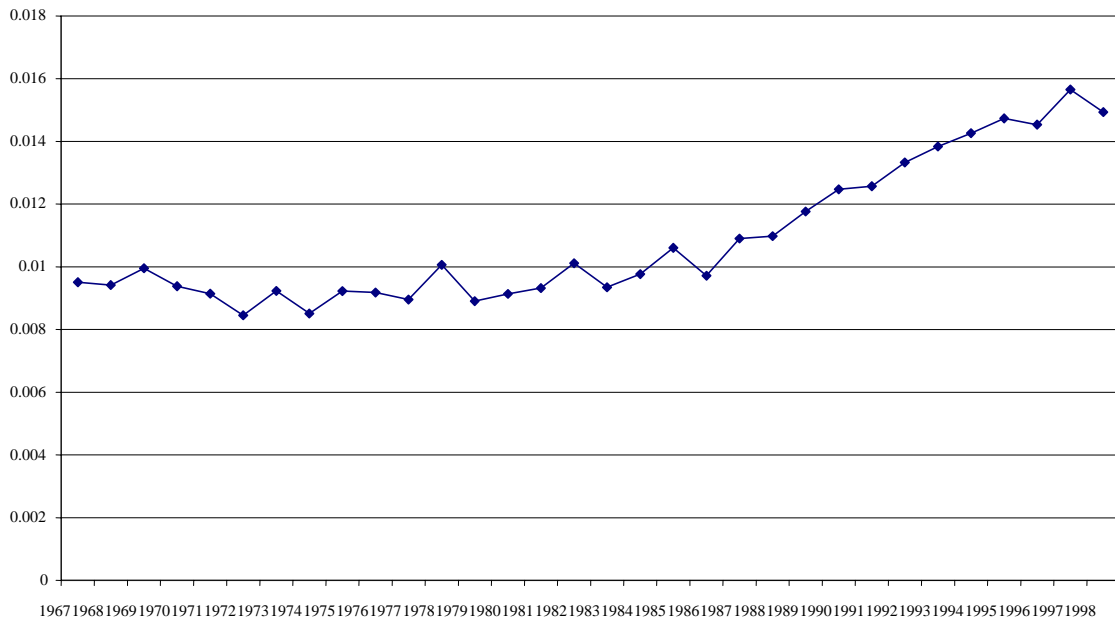


Figure 2
Fraction of Twin Births That Are Same Sex Twins

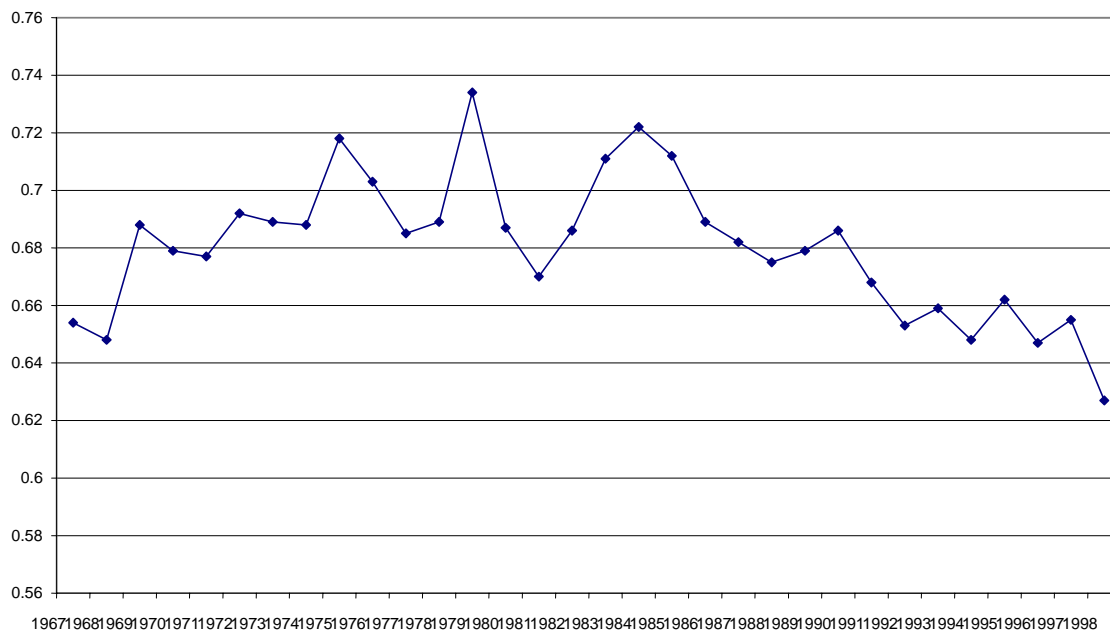


Figure 3
One-Year Mortality Rates
Per 1000 Births

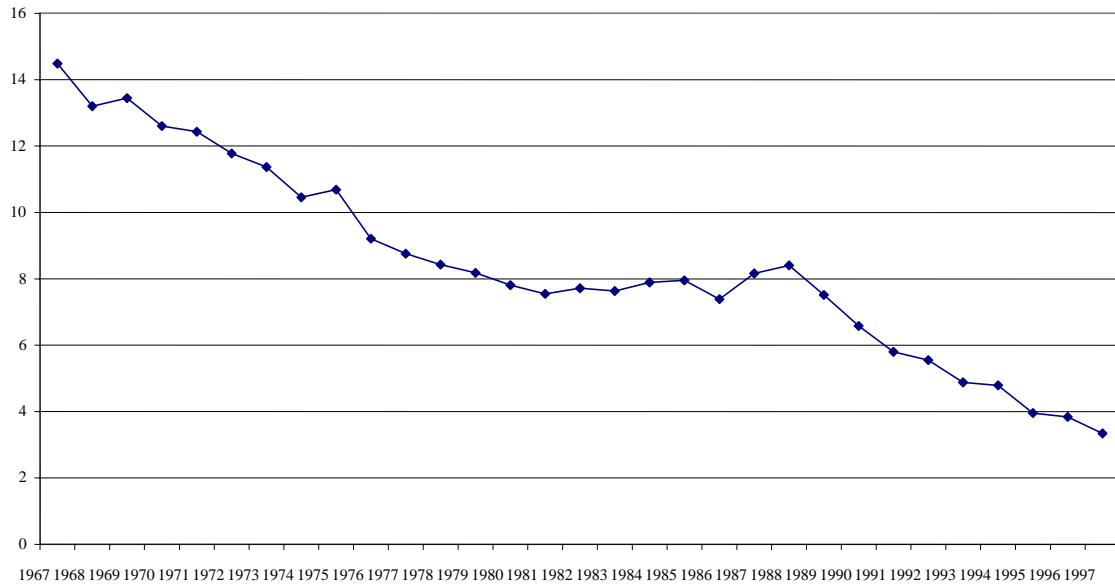


Figure 4
Mortality Rates by Twin Status
per 1000 births

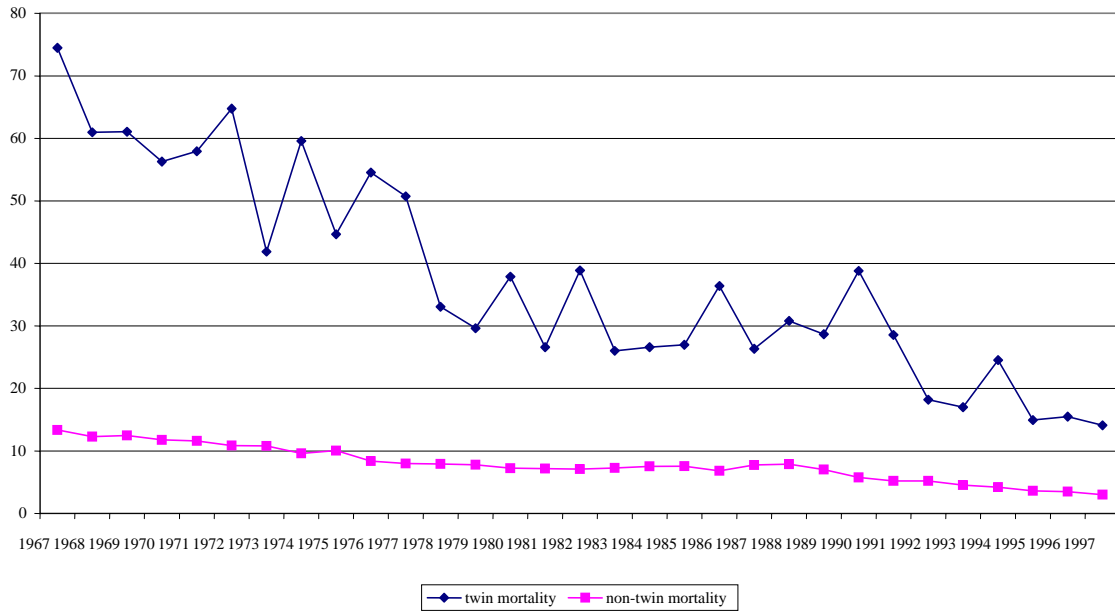


Figure 5
Distribution of Differences in Birth Weight of Twins

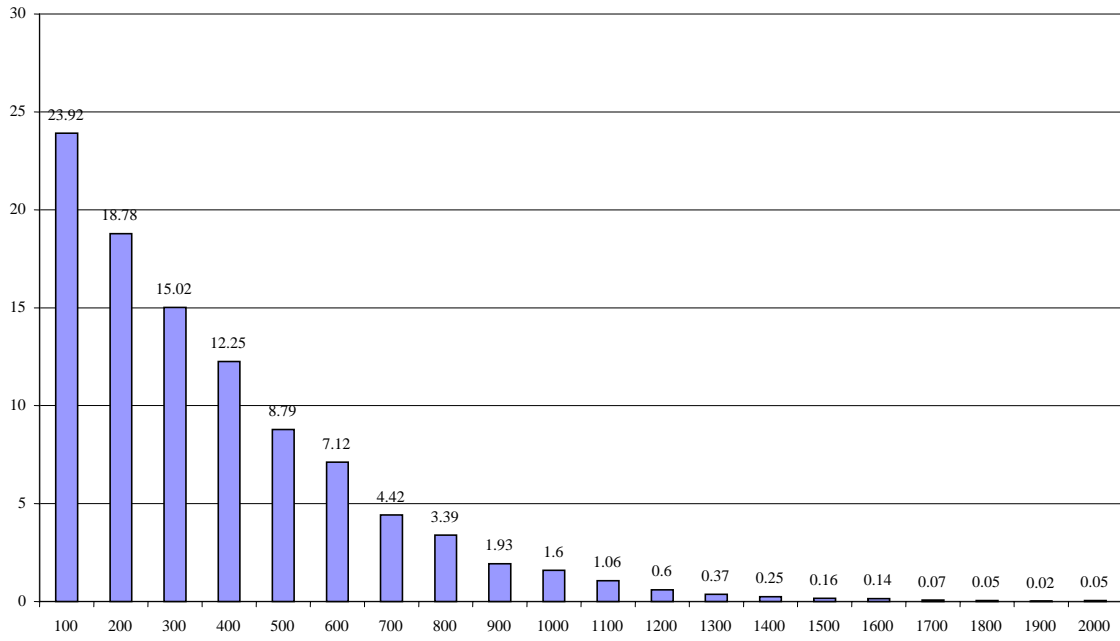


Figure 6
Mortality Rate by Birth Weight

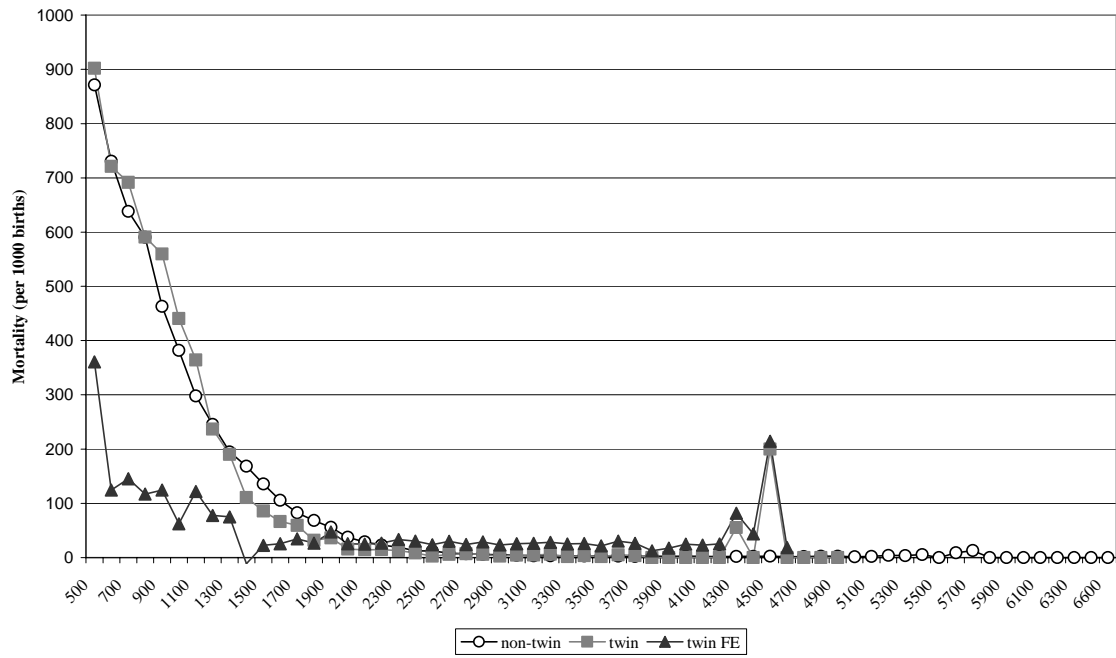


Figure 7
Height by Birth Weight

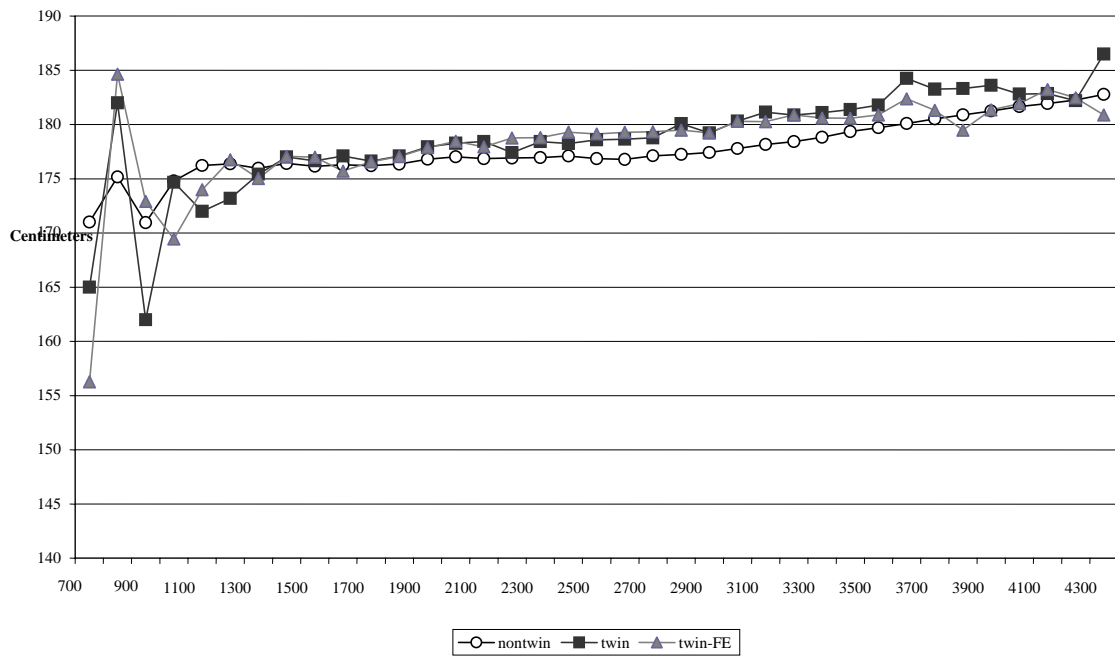


Figure 8
BMI by Birth Weight

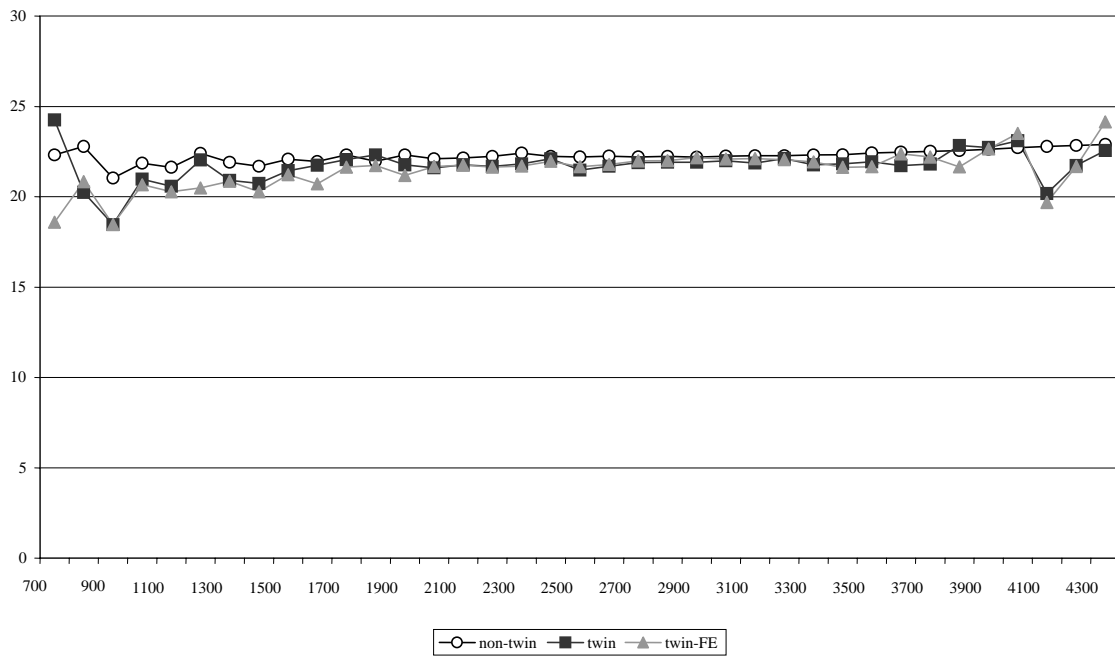


Figure 9
IQ by Birth Weight

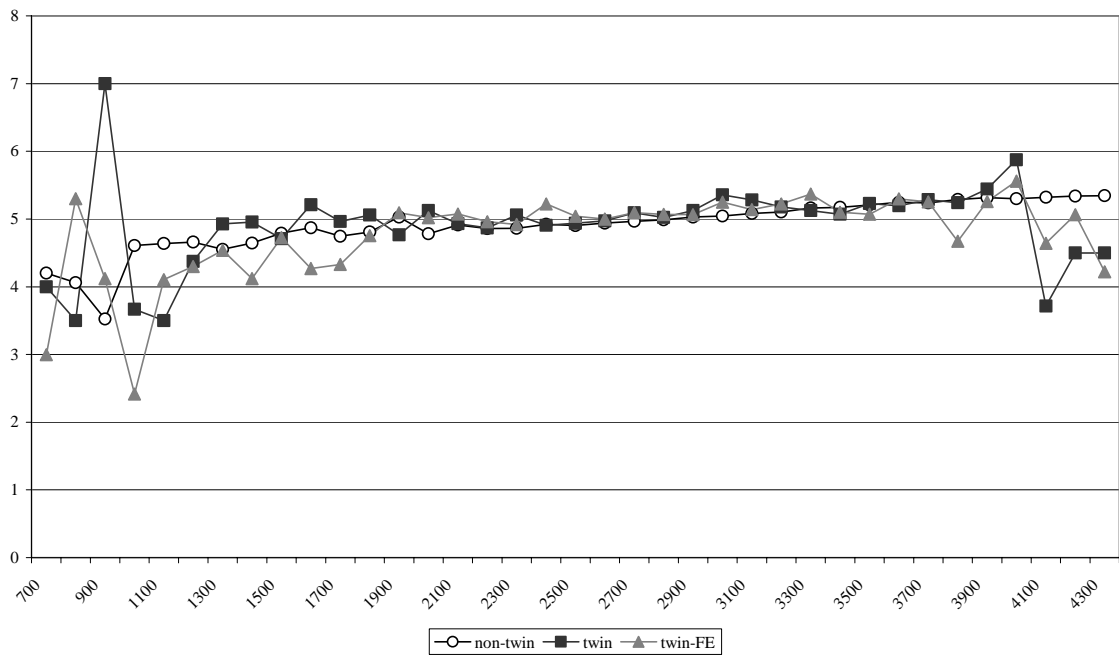


Figure 10
High School Graduation by Birth Weight

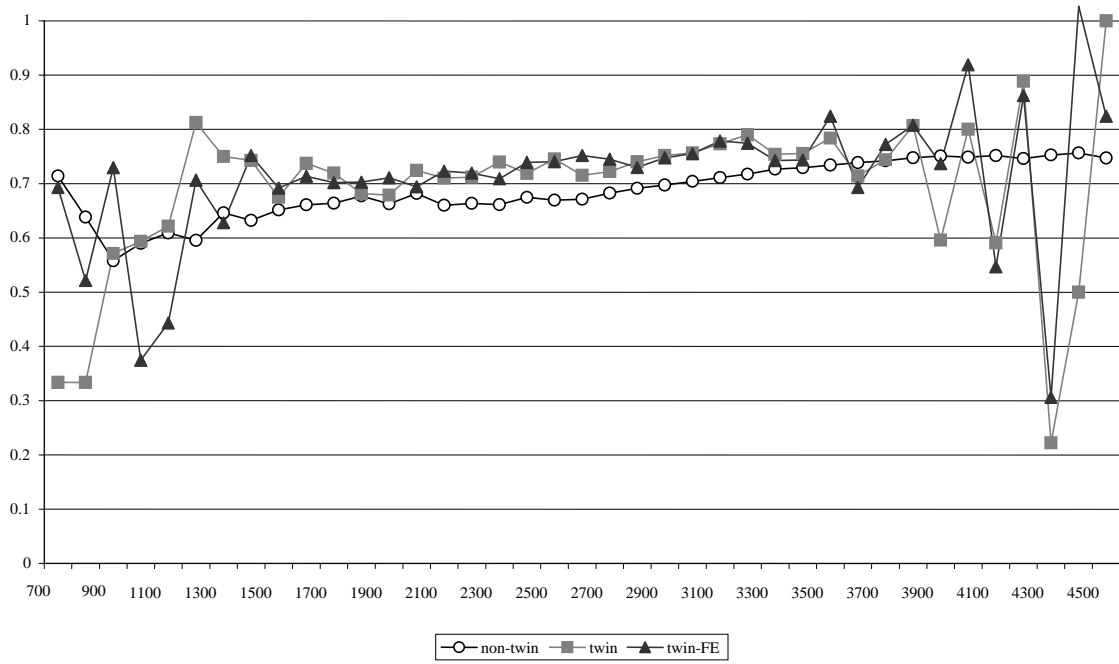


Figure 11
Distribution of Birth Weight

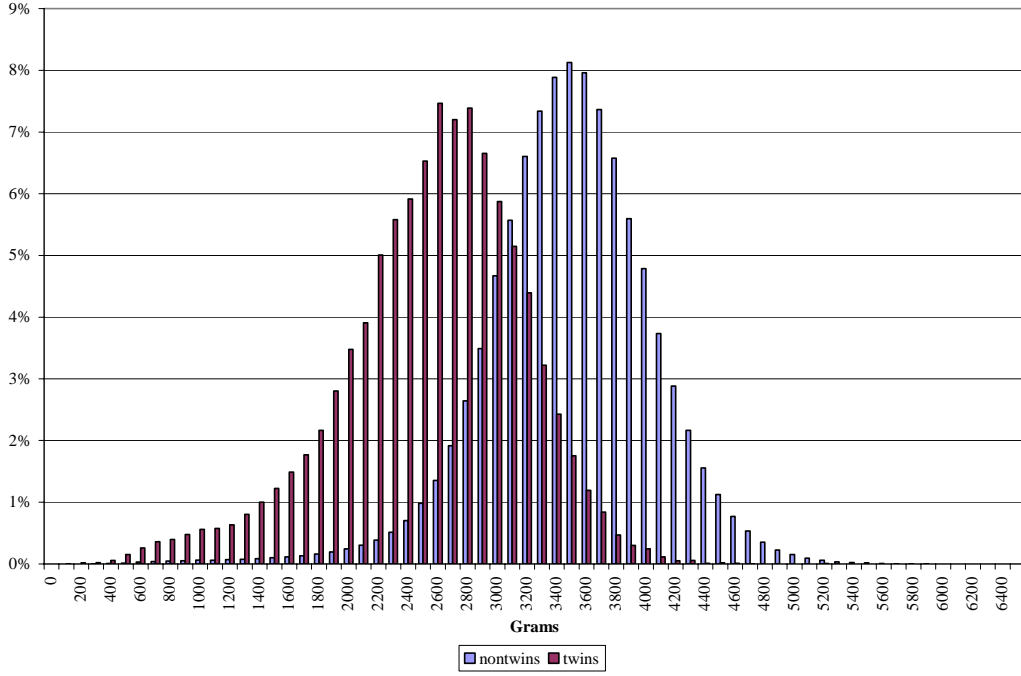


Table 1
Summary Statistics

	Non-twins Sample	Twins Sample	Same Sex Twins Male	Same Sex Twins Female
Child's Characteristics				
Infant Birth Weight				
Mean	3528 (558)	2598 (613)	2594 (639)	2540 (600)
Median	3540	2660	2660	2600
25 th percentile	3210	2250	2240	2200
10 th percentile	2880	1800	1750	1750
5 th percentile	2640	1470	1380	1430
1 st percentile	1860	820	760	800
Fraction low birth weight (<2500 Grams)	.03 (.17)	.33 (.47)	.33 (.47)	.36 (.48)
Gestation in weeks	39.83 (2.17)	36.90 (3.18)	36.62 (3.30)	37.02 (3.20)
Fetal Growth	88.46 (13.07)	69.83 (13.81)	70.14 (14.38)	68.05 (13.48)
1 minute APGAR score	8.68 (1.03)	8.21 (1.61)	8.14 (1.70)	8.22 (1.59)
5 minute APGAR score	9.29 (.75)	9.01 (1.10)	8.95 (1.19)	9.01 (1.10)
Fraction Female	.49 (.50)	.50 (.50)	0	1
Fraction with Complications	.31 (.46)	.49 (.50)	.49 (.50)	.49 (.50)
1 Year mortality rate (per 1000 births)	6.23 (78.69)	31.13 (173.67)	41.20 (198.75)	28.11 (165.30)
Mother's Characteristics				
Education	11.43 (2.56)	11.53 (2.62)	11.55 (2.60)	11.53 (2.63)
Fraction <12 Years of Education	.57 (.50)	.55 (.50)	.54 (.50)	.55 (.50)
Age	26.66 (5.23)	28.09 (5.11)	27.84 (5.11)	27.76 (5.18)
Fraction 30 or Older	.28 (.45)	.38 (.49)	.36 (.48)	.36 (.48)
N	1,595,233	33,346	11,530	11,276

APGAR scores are only available after 1977; as a result, we have APGAR scores for 959,518 nontwins, 21,708 twins, 7,540 same-sex male twins, and 7,243 same-sex female twins.

Table 2
Summary Statistics
Early Period (1968-1981)

	Non-twins Sample	Twins Sample	Same Sex Twins Male	Same Sex Twins Female
Child's Characteristics				
Infant Birth Weight				
Mean	3511 (549)	2607 (616)	2650 (643)	2531 (602)
Median	3520	2660	2670	2590
25 th percentile	3200	2250	2250	2180
10 th percentile	2870	1810	1750	1740
5 th percentile	2630	1480	1380	1440
1 st percentile	1900	860	810	810
Fraction low birth weight (<2500 Grams)	.03 (.17)	.33 (.47)	.32 (.47)	.38 (.48)
Gestation in weeks	39.89 (2.16)	37.30 (3.28)	37.07 (3.38)	37.37 (3.32)
Fetal Growth	87.92 (12.85)	69.33 (13.78)	69.75 (14.35)	67.18 (13.38)
Fraction Female	.49 (.50)	.50 (.50)	0	1
Fraction with Complications	.24 (.43)	.44 (.50)	.44 (.50)	.43 (.50)
1 Year mortality rate (per 1000 births)	8.10 (89.63)	46.03 (209.55)	59.72 (236.98)	41.94 (200.47)
Percentage Completing High School	.73 (.45)	.73 (.43)	.73 (.44)	.75 (.44)
Earnings Data				
Earnings	260,132 (367,588)	257,092 (138,950)	300,639 (149,416)	210,709 (108,849)
Earnings for Full Time Workers	311,616 (159,583)	307,463 (123,373)	337,016 (133,156)	262,239 (149,416)
Military Data				
Height (Male Sample)	179.96 (6.51)	-	179.33 (6.57)	-
BMI (Male Sample)	22.50 (3.38)	-	21.84 (2.90)	-
IQ (Male Sample)	5.20 (1.79)	-	5.06 (1.82)	-
Mother's Characteristics				
Education	10.76 (2.53)	10.69 (2.59)	10.77 (2.62)	10.69 (2.58)
Fraction <12 Years of Education	.69 (.46)	.70 (.46)	.69 (.46)	.69 (.46)
Age	25.80 (5.27)	27.07 (5.22)	26.88 (5.23)	26.78 (5.30)

Fraction 30 or Older	.22 (.42)	.30 (.46)	.29 (.45)	.28 (.45)
N	813,497	14,882	5,074	5,198

Table 3
Summary Statistics: Same Sex Twins
Early Sample

	Heavier	Lighter	Wald Estimates [standard errors]
Infant Birth Weight			
Mean	2732 (615)	2412 (585)	
Median	2800	2470	
25 th percentile	2390	2070	
10 th percentile	1940	1630	
5 th percentile	1580	1320	
1 st percentile	910	780	
Fraction low birth weight (<2500 Grams)	.26 (.44)	.44 (.50)	
Fetal Growth	72.74 (13.30)	64.25 (13.12)	
5 minute APGAR score	9.05 (1.13)	8.92 (1.29)	
Fraction with Complications	.43 (.49)	.45 (.50)	
Ln(Birth Weight)	.97 (.28)	.84 (.30)	
Outcomes:			
1 Year mortality rate (per 1000 births)	51.27 (220.57)	47.69 (213.14)	27.68 [21.10]
Height (Males only)	179.67 (6.57)	178.99 (6.56)	5.61** [.75]
BMI (Males only)	21.90 (2.89)	21.77 (2.86)	1.05** [.40]
Ability (Males only)	5.10 (1.81)	5.04 (1.82)	.49** [.24]
High School Graduation Rate	.75 (.43)	.73 (.44)	.16** [.05]
Ln(Earnings)	11.96 (.86)	11.95 (.86)	.08 [.11]
Ln(Earnings) for Full Time Workers	12.28 (.51)	12.27 (.52)	.11 [.08]
N	10,064		

APGAR scores are only available for 2,364 twins in this sample. Standard deviations in parentheses. Wald estimates are the difference in outcome divided by the difference in ln(birth weight).

Table 4
Regression Results
Twins Sample
Coefficient on Ln(Birth Weight)

Dependent Variable:	OLS	FE
Mortality N=33,346	-279.16** (9.11)	-41.15** (7.64)
Height (Males Only) N=5,388	7.51** (.55)	5.69** (.56)
BMI (Males Only) N=5,378	.59** (.23)	1.12** (.30)
Underweight N=5,378	-.08** (.02)	-.11** (.04)
Overweight N=5,378	.03 (.02)	.10** (.04)
IQ (Males Only) N=4,926	.49** (.14)	.62** (.18)
High School Completion N=13,472	.08** (.02)	.09** (.04)
Ln(Earnings) N=34,788 (5,858 Twin Pairs)	.10** (.03)	.09* (.05)
Ln(Earnings) FT N=16,214 (3,893 Twin Pairs)	.07** (.02)	.10** (.04)

The control variables we use in the OLS estimation are year- and month-of-birth dummies, indicators for mother's education (one for each year), indicators for birth order, indicators for mother's year of birth, and an indicator for the sex of the child. Twin fixed effects regressions include indicators for sex and birth order of the twin (either 1st born or 2nd born twin).

Table 4a
Regression Results
Same-Sex Twins Sample
Coefficient on ln(Birth Weight)

Dependent Variable:	OLS	FE
Mortality N=22,806	-293.88** (10.61)	-39.99** (9.46)
High School Completion N=9,248	.07** (.02)	.10** (.04)
Ln(Earnings) N=24,088 (4,060 Twin Pairs)	.08** (.03)	.09 (.06)
Ln(Earnings) FT N=11,712 (2,730 Twin Pairs)	.08** (.02)	.12** (.05)

The control variables we use in the OLS estimation are year- and month-of-birth dummies, indicators for mother's education (one for each year), indicators for birth order, indicators for mother's year of birth, and an indicator for the sex of the child. Twin fixed effects regressions include indicators for sex and birth order of the twin (either 1st born or 2nd born twin).

Table 5
Regression Results
Same Sex Twins Sample
By Sex
Coefficient on Ln(Birth Weight)

Dependent Variable:	Males	Females
Mortality	-34.40** (14.08)	-45.77** (12.59)
High School Completion	.06 (.06)	.13** (.05)
Ln(Earnings)	.24** (.08)	-.05 (.10)
Ln(Earnings) FT	.15** (.07)	.06 (.06)

The control variables we use in the OLS estimation are year- and month-of-birth dummies, indicators for mother's education (one for each year), indicators for birth order, indicators for mother's year of birth, and an indicator for the sex of the child. Twin fixed effects regressions include indicators for sex and birth order of the twin (either 1st born or 2nd born twin).

Table 6
Fixed Effects Regression Results
All Twins Sample

	1-Year Mortality	High School Completion	Height (Males Only)	BMI (Males Only)	IQ (Males Only)	Ln(Earn)	Ln(Earn) FT
Splines:							
BW<1500	-260.45** (22.94)	.26 (.17)	6.25** (2.39)	1.79 (1.28)	1.11 (.80)	-0.45 (.33)	0.41* (.22)
1500<BW<2500	4.17 (6.73)	.05 (.03)	2.67** (.50)	.69** (.27)	.47** (.16)	.08* (.04)	.05 (.03)
BW>2500	-4.62 (4.30)	.02 (.02)	1.86** (.29)	.22 (.16)	.07 (.09)	.03 (.03)	.01 (.02)
N	33,346	13,472	5,388	5,378	4,926	34,788 (5848 Twin Pairs)	16,214 (3893 Twin Pairs)

Regressions include indicators for sex and birth order of the twin (either 1st born or 2nd born twin).

**Appendix Table 1
Goodness of Fit**

(R-Square Statistics for Various Measures of Birth Weight with Twin Fixed Effects)

Dependent Variable:	Mortality	HS Graduation	Earnings	Earnings (F/T)	Height	BMI	IQ
Birth Weight	.0835	.0015	.0425	.0576	.0167	.0128	.0003
Ln(Birth Weight)	.1834	.0015	.0425	.0579	.0168	.0129	.0003
Fetal Growth	.0731	.0025	.0425	.0576	.0182	.0131	.0003
Indicator for <2500 grams	.0110	.0010	.0424	.0570	.0019	.0123	.0001
Splines	.3212	.0014	.0428	.0584	.0168	.0128	.0003

Numbers represents R-Squared Statistics from a regression with all twins including twin fixed effects, birth order, sex, and the indicated measure of birth weight.

Appendix Table 2
Regression Results
Twins Sample

Dependent Variable:	Ln(Birth Weight)		Birth Weight		Fetal Growth	
	OLS	FE	OLS	FE	OLS	FE
Mortality N=33,346	-279.16** (9.11)	-41.15** (7.64)	-103.39** (3.92)	-9.67** (3.16)	-4.29** (.17)	-.41** (.12)
Height (Males Only) N=5,388	7.51** (.55)	5.69** (.56)	3.08** (.21)	2.18** (.22)	.15** (.01)	.08** (.01)
BMI (Males Only) N=5,378	.59** (.23)	1.12** (.30)	.23** (.09)	.39** (.12)	.015** (.004)	.015** (.004)
IQ (Males Only) N=4,926	.49** (.14)	.62** (.18)	.18** (.06)	.21** (.07)	.01** (.002)	.008** (.003)
High School Completion N=13,472	.08** (.02)	.09** (.04)	.03** (.01)	.03** (.01)	.002** (.0003)	.0012** (.0005)
Ln(Earnings) N=34,788 (5,858 Twin Pairs)	.10** (.03)	.09* (.05)	.04** (.01)	.04* (.02)	.002** (.0004)	.0014* (.0008)
Ln(Earnings) FT N=16,214 (3,893 Twin Pairs)	.07** (.02)	.10** (.04)	.03** (.01)	.03** (.016)	.001** (.0003)	.0012** (.0006)

The control variables we use in the OLS estimation are year- and month-of-birth dummies, indicators for mother's education (one for each year), indicators for birth order, indicators for mother's year of birth, and an indicator for the sex of the child. Twin fixed effects regressions include indicators for sex and birth order of the twin (either 1st born or 2nd born twin).

Appendix Table 3a
Fixed Effects Regression Results
Men-Same Sex Twins Sample

	1-Year Mortality	High School Completion	Ln(Earn)	Ln(Earn) FT
Splines:				
BW<1500	-188.96** (39.90)	.47* (.27)	.17 (.24)	.21 (.23)
1500<BW<2500	3.75 (12.91)	-.01 (.05)	.11* (.07)	.11* (.05)
BW>2500	-3.34 (8.12)	.02 (.03)	.08* (.04)	.02 (.03)
N	11,530	4,486	12,057 (1990 Twin Pairs)	7,520 (1537 Twin Pairs)

Regressions include indicators for birth order of the twin (either 1st born or 2nd born twin).

Appendix Table 3b
Fixed Effects Regression Results
Female-Same Sex Twins Sample

	1-Year Mortality	High School Completion	Ln(Earn)	Ln(Earn) FT
Splines:				
BW<1500	-341.87** (37.41)	.09 (.22)	-1.02 (.72)	1.57* (.68)
1500<BW<2500	17.01 (10.46)	.08* (.04)	.05 (.07)	-.01 (.04)
BW>2500	-9.81 (7.71)	.03 (.03)	-.04 (.05)	.01 (.04)
N	11,276	4,762	12,031 (2070 Twin Pairs)	4192 (1193 Twin Pairs)

Regressions include indicators for birth order of the twin (either 1st born or 2nd born twin).

Appendix Table 4
Regression Results: One Year Mortality
Twins Sample
Early Period

(Mortality is measured as number of deaths per 1000 births)

	OLS	FE
Birth weight	-146.45** (6.35)	-1.99 (5.31)
Ln(Birth weight)	-390.13** (13.42)	-9.54 (12.98)
Fetal Growth	-5.96** (.28)	-.06 (.20)
N	14,882	

The control variables we use in the OLS estimation are year- and month-of-birth dummies, indicators for mother's education (one for each year), indicators for birth order, indicators for mother's year of birth, and an indicator for the sex of the child. Twin fixed effects regressions include indicators for sex and birth order of the twin (either 1st born or 2nd born twin).