CENTER FOR LABOR ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY
WORKING PAPER NO. 53

Does Low Birth Weight Matter?

Evidence from the U.S. Population of Twin Births*

Douglas Almond
NBER

Kenneth Y. Chay
UC Berkeley and NBER

David S. Lee
UC Berkeley and NBER

September 2002

ABSTRACT

A widely held position is that spending resources to reduce the incidence of low birth weight (LBW) births represents one of the best investments in improving infant health and welfare. The emphasis of public policies (e.g. Medicaid, WIC, maternal smoking cessation campaigns) on LBW rests crucially on the claim that any policy that works to reduce the incidence of LBW births will improve infant health outcomes. This paper empirically evaluates this claim – that birth weight can reliably substitute for direct health outcomes of interest (e.g. mortality) when evaluating and predicting the effects of health policies and interventions.

If this claim is true, then the component of birth weight that is exclusively driven by environmental factors (and not by genetic factors) should be highly correlated with other known measures of infant health, including mortality. We examine this type of birth weight variation by focusing on within twin-pair correlations between birth weight and various health outcomes, using data on the population of twin births in the United States, which we construct using a new matching algorithm on available linked birth-infant death micro-data. While there are substantial differences in birth weights between twins, we find that the heavier twin is no more likely to survive past one-year of life than the lighter twin. This finding is consistent across demographic groups, causes of death, and years of birth and is insensitive to various econometric specifications. We find that another continuous proxy of health at birth, the 5-minute APGAR score, is more highly associated with infant mortality both between and within twin pairs, but is not correlated with within twin-pair differences in birth weight.

Furthermore, we examine a policy application – maternal smoking cessation – that provides a concrete example of the implications of our findings. While smoking during pregnancy has a substantial association with birth weight and LBW incidence, it has little relation with either infant mortality or the 5-minute APGAR score. The evidence suggests that birth weight may provide a misleading proxy for measuring the effectiveness of health interventions.

* We thank David Card, Janet Currie, Lorenz Goethe, Michael Grossman, Jon Gruber, Ted Joyce, Bob Kaestner, Karen Norberg, Doug Staiger, Paul Torelli, and participants of the Berkeley Labor Lunch and NBER Summer Institute in Children's Studies for helpful comments.

Introduction

It is a widely held position that, whether at the individual or governmental level, spending resources on reducing the incidence of low birth weight (LBW) births (defined to be less than 2,500 grams) represents one of the best investments in improving infant health and welfare. This position is based on the observation that LBW infants have been shown to have higher mortality rates, poorer health as children or adults, reduced cognitive ability, and eventually lower education levels and labor market earnings relative to their normal weight counterparts. In fact, reducing the incidence of LBW births has been a goal of several public policies targeting infant health, including Medicaid, the Women, Infants, and Children (WIC) program, and maternal smoking cessation campaigns. The emphasis of public policy on LBW rests crucially on the claim that any policy that works to reduce the incidence of LBW births will necessarily improve infant health outcomes, and in particular, reduce infant mortality.

This paper empirically evaluates this claim – that birth weight is a valid policy "marker" for infant health. By "valid policy marker", we mean a variable that can reliably substitute for the actual outcome of interest when evaluating and thus predicting the health effects of policies and interventions. For example, within a human capital framework, education could provide a valid policy marker for improved earnings capacities, since the presumed structural link between educational attainment and earnings implies that any intervention that works to increase educational attainment (and other measures of human capital) would generally result in higher individual earnings. On the other hand, a concrete example of a misleading policy marker from recent medical research is the level of cholesterol as a marker for cardiovascular heart disease risk among postmenopausal women. A recent, highly publicized randomized trial found higher heart disease rates among the women treated with estrogen plus progestin, even though previous studies had established that the therapy reduced cholesterol levels, a generally accepted indicator for increased risk of coronary heart disease. [1]

[1] An intermediate outcome clinical trial established that estrogen/progestin replacement therapy reduces cholesterol levels, a generally accepted indicator for increased risk of coronary heart disease, among postmenopausal women. As a result, it was widely believed that one of the health benefits of taking estrogen after menopause would be a

As we show below within a simple econometric framework, if birth weight constitutes a valid policy marker, then it must be true that the component of birth weight that is exclusively driven by environmental factors has a strong correlation with other known measures of infant health, including mortality. A strong correlation would be consistent with the notion that any policy that succeeds in reducing the incidence of LBW births would improve health and reduce infant mortality. On the other hand, a weak correlation between LBW and mortality would call into question the conjecture that birth weight is a sufficient index of overall infant health. The main confounding problem is that genetic factors undoubtedly contribute to the determination of birth weight and overall health in unknown ways, and hence may entirely explain the strong cross-sectional relation between birth weight and mortality. Thus, in order to account for genetic variation across families, and isolate birth weight variation that is primarily or entirely driven by environmental factors, we analyze the population of twin births in the U.S. and examine whether within twin-pair differences in health are correlated with within twin-pair differences in birth weight.

Using a new algorithm to match twins with their siblings in National Center for Health Statistics-linked birth-infant death microdata, we analyze the population of twin births in the United States for the years 1985-1986, 1989-1991, and 1995 – a sample of over 350,000 twins. While the average twin weighs only 2,400 grams, there are substantial differences in birth weights between twins of the same mother. For example, on average the heavier twin weighs 300 grams more than the lighter twin at birth. This differential is 50 percent larger than the average birth weight difference between the infants of mothers who did and did not smoke during pregnancy. The large variation in birth weights within twin pairs leads to precise fixed effects estimates of the effect of birth weight on infant mortality. It also allows us to examine the significant nonlinearities in the infant mortality-birth weight relation.

There is a strong and robust cross-sectional correlation between infant mortality and birth weight between twin pairs, particularly at birth weights below 1,000 grams. In addition, there are substantial

---

reduced incidence of heart attack and stroke. However, a recent, randomized controlled trial found higher heart

differences in several measures of health within twin pairs. Surprisingly, however, we find that within twin-pairs the heavier twin is no more likely to survive past one-year of life than the lighter twin, even among twins born at very low weights. This finding is consistent across demographic groups, causes of death, and years of birth and is insensitive to econometric specification. Two interpretations are possible: 1) the within twin-pair comparison yields a "true" structural estimate of the "effect" of birth weight on infant mortality, and that estimate is small and in many cases zero, or 2) the evidence is simply inconsistent with the notion of a "one-to-one" relation between birth weight and important measures of health: some interventions might simultaneously reduce LBW and mortality, while others might reduce LBW without impacting (or perhaps having a perverse impact on) mortality. Either interpretation calls into question the validity of using birth weight alone as a substitute for actual health outcomes of interest in evaluating and predicting the effects of policy interventions.

We find that another continuous proxy of health at birth, the 5-minute APGAR score, is more highly associated with infant mortality both between and within twin pairs, but is not correlated with sibling differences in birth weight. Twin differences in birth weight are also uncorrelated with the use of assisted ventilation after birth, the most common post-birth intervention for at-risk infants. On the other hand, twin differences in APGAR scores are highly associated with assisted ventilation incidence. In our analysis, birth weight stands out as the one infant outcome that appears uncorrelated with other known measures of health, after accounting for common genetic factors within twin-pairs.

Finally, we examine the implications of our findings for an important policy application: the validity of using birth weight to evaluate infant health benefits of maternal smoking cessation. Maternal cigarette smoking during pregnancy has been identified as the largest modifiable risk factor for LBW incidence in developed countries (Kramer 1987). Further, a recent editorial in the Journal of the American Medical Association advocates the prevention of all maternal smoking as an "optimal public health outcome", since "maternal smoking is a significant risk factor for LBW infants, which, in turn,

---

disease rates among the women treated with estrogen plus progestin (Women's Health Initiative Investigators, 2002).

influences infant mortality" (Vogler and Kozlowski, 2002). In our analysis of both singletons and twins, we find that while smoking during pregnancy has a substantial association with birth weight and LBW incidence, it has little relation with either infant mortality or the 5-minute APGAR score. These findings provide one concrete and important example of the primary implication of our analysis of the infant twins: even if a specific health policy is known to reduce the incidence of low birth weight, this contains little or no information, by itself, about the overall health consequences (and implicit costs) of the intervention.

The paper is organized as follows. We first provide a brief background on the biomedical, public health, and economics literatures that emphasize the importance of LBW. We then present our conceptual and econometric framework as a simple two-equation system. The construction of the data is discussed, followed by the results of our analysis, and then the policy application. The last section concludes.

Background on Birth Weight

The two factors that are generally recognized to govern infant birth weight are the duration of gestation and the intrauterine growth rate. Since gestation length has an independent effect on infant health, this study focuses on the effects of reduced birth weight for a fixed gestation length, also known as intrauterine growth retardation (IGR). Research on the etiology of birth weight suggests that both environmental and genetic factors play a role in the rate of fetal growth in utero. Among modifiable environmental factors, a widely-cited study found that maternal cigarette smoking during pregnancy is, by far, the most important factor with a direct impact on IGR in developed countries (Kramer, 1987). In developing countries, low caloric intake of the mother has been identified as one of several important factors leading to low birth weight.

In this section, we review the large biomedical, public health, and economics literatures that find that low birth weight (LBW) infants (those weighing less than 2,500 grams) have higher mortality rates, poorer health as a child and adult, and reduced cognitive ability, education levels and labor market

earnings relative to their normal weight counterparts. We then discuss how, as a result, reducing the incidence of LBW has been the objective of several policy interventions targeting infant health. Finally, we briefly describe several examples that seem to contradict the validity of using birth weight as a metric for evaluating the effectiveness of an intervention.

Birth Weight as Cause

Today, it is widely accepted that low birth weight is the single most important determinant of neonatal mortality (death within 28 days of birth), as well as being a significant predictor of post-neonatal mortality and infant and later childhood morbidity.[2] The supporting evidence comes from the strong cross-sectional correlation between birth weight and infant mortality. In 1989, for example, while only 7 percent of all infants were born at low birth weight, these infants accounted for over 60 percent of all infant deaths. Further, many have stated that the leading cause of the substantial black-white gap in infant mortality in the United States is the higher incidence of LBW among black infants.[3]

Several studies have also found a significant correlation between low birth weight and childhood health and developmental outcomes. For example, LBW has been found to be associated with high blood pressure, cerebral palsy, deafness, blindness, respiratory problems and lung disease among children.[4] It has also been shown to be correlated with reduced IQ, test scores, and cognitive development.[5] Research in the economics literature has also found that LBW is correlated with lower educational attainment, poorer self-reported health status, and reduced employment and earnings among adults.[6] For example,

[2] For example, see McCormick (1985), Institute of Medicine (1985), Rees, et al. (1996), and Alexander.
[3] See Schoendorf, et al. (1992), Iyasu, et al. (1992), JAMA (1999), NIH (1999), NICHD (2000), U.S. Department of Health and Human Services (2000), and Alexander. Miller (2002) questions the role of LBW differentials in explaining the black-white infant mortality gap.
[4] See Paneth (1995), Nelson and Grether (1997), McCormick, et al. (1992), and Brooks, et al. (2001). Ijzerman, et al. (2000) finds evidence that the birth weight-blood pressure relation may be an artifact of genetic omitted variables. Based on a sample of 200 twin pairs, Zhang, et al. (2001) find that differences between twins in birth weight are not associated with twin differences in blood pressure at age 7.
[5] See Kaelber and Pugh (1969), Lucas, Morley, and Cole (1998), Matte, et al. (2001), and Richards, et al. (2001). Li and Poirier (2002), using a Bayesian simultaneous equations framework, find no evidence of a structural effect of birth weight on early childhood outcomes.
[6] See, for example, Behrman, Rosenzweig, and Taubman (1994), Currie and Hyson (1999), and Behrman and Rosenzweig (2001). Schultz (2002) examines the association between height and wages.

one study finds that a 4-ounce increase in birth weight results in almost a half-year increase in educational attainment (Behrman, Rosenzweig, and Taubman, 1994).

Finally, over the past decade a burgeoning literature has found a significant association between infant birth weight and adult morbidity and mortality. The "fetal origins hypothesis" postulates that nutritional conditions during pregnancy, for which birth weight serves as a proxy, "program" the fetus for the development of chronic diseases in adulthood, such as diabetes, hypertension, and cardiovascular disease.[7] For example, several studies have documented a strong correlation between infant birth weight and adult death due to ischaemic heart disease (Barker, et al., 1989, Vagero and Leon 1994). Interestingly, this is the one area of research where the causality of the documented birth weight associations has been seriously questioned.[8]

Birth Weight as the Target of Policy

In the clinical and epidemiologic literatures, it has been stipulated that birth weight has a causal effect on infant mortality, child health and cognitive development, and adult morbidity and mortality. In the economics literature, birth weight is commonly viewed as being the most important indicator of infant health and a proxy for the initial health human capital of the infant.[9] Thus, it is not surprising that birth weight has often been the target for public health intervention and used as the outcome of interest when evaluating the effectiveness of the intervention.

In the United States, a motivation for the Medicaid expansion to pregnant women during the 1980s was the reduction of the incidence of low birth weight through expanded access to prenatal care (Currie and Gruber, 1996).[10] Increasing infant birth weights has also been the objective of the Women,

[7] See Barker (1992), Barker, et al. (1989, 1993), Vagero and Leon (1994), Leon, et al. (2000), and Innes, et al. (2002). Rasmussen (2001) provides a review of the literature.

[8] Comparing twins to singletons, Williams and Poulton (1999) and Christensen, et al. (1995) find no evidence of an effect of birth weight on either blood pressure or adult mortality, respectively. Phillips, et al. (2001) summarizes the studies of the effects of twins' birth size on long-run morbidity and mortality. Susser and Levin (1999) and Kramer (2000) provide commentary questioning the causal effect of birth weight. Rasmussen (2001) discusses several quasi-experiments, including Stein, et al. (1975), that provide mixed empirical evidence on the fetal origins hypothesis.

[9] For example, see Grossman and Joyce (1990) and Currie and Cole (1993).

[10] Behrman and Rosenzweig (2001) conclude that the beneficial returns to interventions that increase fetal growth supports programs that result in increased birth weights, such as Medicaid.

Infants, and Children (WIC) program, which provides caloric supplementation to women during pregnancy. In fact, most of the evaluations of the effectiveness of the WIC program have focused on birth weight as the sole outcome of interest, with little analysis of the impact of WIC on infant mortality.[11] The U.S. National Institutes of Health have proposed that infant birth weight be targeted for intervention due to the view that differences in LBW are the leading cause for the black-white gap in infant mortality.[12] Further, increasing birth weight in developing countries has been the objective of several international development projects.[13]

Several studies have calculated the economic cost of low birth weight. Lewit, et al. (1995) estimate that in 1988 the health care, education, and child care costs associated with the 3.5-4 million children aged 0 to 15 born at low birth weight was $5.5-6 billion more than if the children had been born at normal weight. They calculate that LBW accounts for 10 percent of all health care costs for children. Cutler and Meara (1999) conclude that the rate of return for the treatment of LBW infants is over 500 percent. Consequently, some have suggested that reducing the incidence of LBW would be a more cost-effective policy than investing in more neonatal intensive care, for example.[14]

Finally, research has singled out maternal smoking during pregnancy as the leading modifiable cause of LBW in the U.S., contributing 17-26 percent to the incidence of all LBW births during the 1980s (Kramer 1987). Based on the strength of the association between LBW and infant mortality, a recent editorial in JAMA states, "Maternal smoking is an ideal target for intervention and the optimal public health outcome would be prevention of all maternal smoking" (Vogler and Kozlowski, 2002).

[11] For example, see Kotelchuck, et al. (1984) and Kowaleski-Jones and Duncan (2002).
[12] For example, the chief goal of a program announcement for NIH-sponsored research (PA-99-045) is the "development of innovative strategies to prevent LBW in minority populations." In addition, Duane Alexander, director of the NICHD, has stated "[T]he increased black mortality from low birth weight is not due to weaker infants or poorer care. It is simply because there are so many more low birth weight and particularly very low birth weight African-American births. So the key is to try to prevent these low birth weight births."
[13] For example, the World Bank's $100 million "Second Tamil Nadu Integrated Nutrition Project" in India had "cut[ting] in half the incidence of low birth weight newborns" as a goal.
[14] Some studies have noted that most of the costs of LBW are the result of preterm delivery and not due to intrauterine growth retardation at a fixed gestation length. Thus, they conclude that policy should focus on strategies that reduce preterm birth and not IGR.

Further, Lightwood, et al. (1999) calculate that the direct medical cost for each pregnant smoker is $511 ($1995) per live birth, helping to justify the cost of California's Proposition 99 antismoking program.[15]

Puzzles and Anomalies Concerning Birth Weight

Although the conventional wisdom appears to be that birth weight is the key to and key indicator of infant health, several recent studies provide evidence that is inconsistent with this view. For example, Chay and Greenstone (2001a, 2001b) document a strong association between sharp reductions in particulates air pollution across counties (due to Clean Air Act regulations and the 1981-1982 recession) and declines in infant mortality, with particularly strong effects on neonatal mortality rates. However, they find only a small effect of the air quality improvements on birth weight, with no impact on the incidence of LBW.[16] This is particularly surprising given the view that LBW is the leading cause of neonatal mortality in the U.S.

Almond, Chay, and Greenstone (2001) document that the dramatic reduction in black infant mortality rates (IMR) from 1965-1971 in the U.S. accounts for the greatest convergence in black-white IMRs in the entire post-World War II era. However, they find a comparatively small change in the birth weight distribution of black infants relative to whites during this same period, bringing into doubt the view that LBW differences have historically accounted for much of the black-white IMR gap.[17] Finally, Torelli (2000) provides strong evidence that while maternal smoking during pregnancy is associated with large reductions in birth weight and an increase in LBW incidence, it has little correlation with infant mortality once parents' socioeconomic characteristics and other risk factors are accounted for.

These examples suggest: 1) there may be many interventions that have a significant effect on infant health that do not appear in birth weight; and 2) there may be interventions that have a large

[15] Based on their estimates of the effect of birth weight on earnings, Behrman and Rosenzweig (2001) estimate that a mother who smokes one pack per day during pregnancy reduces the lifetime earnings of her child by over 10 percent. They also calculate that twins will have 12 percent lower lifetime earnings than singletons due to their (28 ounces) lower weight at birth.

[16] Interestingly, they provide evidence that the pollution declines impact other indices of infant health as well, such as the 5-minute APGAR score.

[17] They find that in places such as Mississippi increased black access to hospital care played a significant role, particularly in the large reduction in post-neonatal mortality among black infants.

impact on birth weight that do not translate into a real effect on infant health. Thus, it is not clear whether reducing the incidence of LBW is an appropriate policy target or whether birth weight provides a valid metric for evaluating the effectiveness of public health interventions. In other words, birth weight may provide a misleading marker for policy.

Conceptual and Empirical Framework

Here, we develop a framework that provides a definition of what constitutes a valid policy marker – an intermediate variable that correctly predicts a change in the targeted outcome of interest when the marker is manipulated by an intervention. We show that the existing basis for focusing on birth weight as a target of health policy implicitly relies upon two sets of assumptions. First, it is assumed that *all* environmental factors affecting birth weight must have a corresponding and commensurate impact on overall health. Second, it is assumed that genetic determinants of birth weight and health do not contribute to the cross-sectional association between infant mortality and birth weight. Unless these two assumptions hold, the strength of the cross-sectional association between a "marker" and the outcome of interest provides no information on the marker's validity. Finally, we illustrate how the examination of the birth weight-mortality relationship among *twins* of the same mother – by accounting for the confounding genetic factors – can help evaluate the usefulness of birth weight as an underlying index of health for policy purposes.

Birth Weight as Target of Policy

Numerous studies on infant health rely upon birth weight as an index of overall health partly because of the limitations of other, more direct indicators of health. For example, evaluating policy on the basis of its effects on infant mortality can pose two limitations. First, since infant mortality is a rare event in the U.S., only extremely large-scale studies would be able to detect a meaningful effect of a health intervention. Second, the dichotomous nature of infant death makes it uninformative for studies that analyze infants' health outcomes conditional on their survival. As a result, birth weight has been

adopted as an alternative index of infant health because of the robust cross-sectional association between low birth weight and infant mortality.

In numerous studies on infant health, it has been implicitly conjectured that this strong cross-sectional association constitutes evidence that policies that minimize low birth weight will necessarily maximize infant health. As shown below, however, this conjecture relies heavily upon several strong assumptions about the relation between birth weight and health.

Assumption 1: Index Property

First, it is implicitly assumed that: 1) *any* environmental factor that contributes to low birth weight will also reduce overall "health", 2) *any* environmental factor that is harmful to infant health will cause low birth weight, and 3) the larger the impact of some policy on low birth weight, the larger will be the impact on overall health.

If any of these conjectures do not hold, low birth weight ceases to be a useful target for health policy. If the first conjecture does not hold, an expensive intervention that succeeds in raising birth weights may have no influence on mortality. If the second conjecture does not hold, inexpensive interventions that reduce infant mortality could be overlooked if they did not have commensurate impacts on birth weight. And if the third conjecture does not hold, policies designed to reduce low birth weight are unlikely to be optimal for improving health, as too many (too few) resources would be dedicated to programs that have little (large) impacts on infant health.

If there are environmental factors that affect birth weight, but not health (or vice versa), then birth weight ceases to have value as an index of health that should be targeted. This is because one could not determine which environmental factors influence only birth weight, and which influence only health. The *only* way to determine this would be to assess the impact of each intervention on *both* birth weight *and* health. But if this were possible, there would be no need for analyzing birth weight, since impacts on health could be evaluated directly.

The implicit "index property" can be formally expressed by the following two-equation system:

(1) $h = g'\alpha_0 + \beta(e'\theta)$

$$bw = g'\alpha_1 + e'\theta$$

where *h* and *bw* are overall infant "health" and birth weight, respectively. *g* is a vector of all genetic factors, and *e* is a vector of all environmental factors. Without the index restriction, the coefficient vectors associated with *g* and *e* in one equation need not have any relation to the coefficients in the other equation. However, if birth weight is a valid policy marker, then the coefficient vector on *e* in the health equation is directly proportional (by a factor of β, a scalar) to the coefficient vector on *e* in the birth weight equation. This formally captures the notion that any environmental factor that improves birth weight (health) will necessarily improve health (birth weight), and that larger impacts on birth weight imply larger impacts on health.

The above system of equations equivalently formalizes the view that birth weight has a "causal" influence on overall health (as has been suggested in existing studies). Substituting the second equation into the first yields,

(2) $$h = g'\alpha_0 - \beta g'\alpha_1 + \beta \cdot bw$$

where β is the structural or causal impact of birth weight on health. If β was large, birth weight would be a useful index of overall health, and in the opposite extreme, if β was zero, birth weight would be an inappropriate target of health policy.[18]

Assumption 2: No Genetic Confounding

Second, in order for the cross-sectional birth weight-mortality relation to validate birth weight as a policy marker, it must also be assumed that the cross-sectional birth weight-mortality relation is not generated by: 1) a correlation between genetic determinants of health and genetic determinants of birth weight, or 2) a correlation between fetal environmental factors and genetic factors. To see this, note that the regression of *h* on *bw* yields the expression:

(3) $$\frac{\text{cov}(h, bw)}{\text{var}(bw)} = \frac{\text{cov}(g'\alpha_0, g'\alpha_1) + \text{cov}(g'\alpha_0, e'\theta) + \beta\,\text{cov}(g'\alpha_1, e'\theta) + \beta\,\text{var}(e'\theta)}{\text{var}(bw)}$$

[18] This two-equation system highlights the notion that only the structural parameter β is useful for policy purposes, whereas the reduced-form correlation between health and birth weight is not.

Without additional assumptions, this regression coefficient suffers from several sources of bias, all of which make it difficult to evaluate the usefulness of birth weight as a proxy for health for policy purposes.

First, as reflected in the first term in the numerator, there could be genetic factors that simultaneously determine birth weight and overall health. For example, there may be a genetic component to overall lung development of the fetus, and under-development of lungs could simultaneously contribute both to low birth weight and health problems for the infant. In addition, as represented by the second and third terms in the numerator, there could be an incidental correlation between those same genetic factors and environmental determinants of health and birth weight. As one hypothetical example, suppose lung capacity is partially genetically determined, so that on average, fetuses with a genetic predisposition to have smaller lungs tend to be born to mothers with smaller lung capacities. This could induce a correlation in the population between fetal lung capacity and the amount of oxygen supplied to the fetus.

If either of the above two types of confounding are present, the observed cross-sectional relationship between birth weight and a measure of health (such as mortality) could be consistent with a positive, zero, or even negative β. Thus, the conjecture that the cross-sectional birth weight-mortality relation is evidence that birth weight is a useful "sufficient index" for health policy contexts relies upon ruling out these kinds of genetic confounding.

If the "no confounding" assumption were true, then the regression coefficient would equal:

$$\frac{\text{cov}(h, bw)}{\text{var}(bw)} = \beta \frac{\text{var}(e'\theta)}{\text{var}(bw)} \tag{4}$$

which would mean that the sign of the regression coefficient would be informative about the sign of β, even if the estimate itself would be attenuated. In this particular case, the downward bias would occur because birth weight would be considered to be $e'\theta$ measured with classical error $g'\alpha_1$.

Evaluating Birth Weight as an Index of Health

To what extent are the above sets of assumptions empirically relevant? The most straightforward and credible way to assess these assumptions is to conduct a multitude of randomized evaluations of the effect of various policy instruments on *both* birth weight *and* the health measure of interest (e.g., infant mortality). After accumulating a great deal of these experimental estimates, it would then be possible to directly examine whether policies that are more effective in reducing low birth weight are in fact also more effective in reducing infant mortality. If such a relation existed across these experiments, the birth weight-as-policy-index hypothesis would be based on sound empirical evidence.

In this paper, we argue that in the absence of such experimental trials, a "second-best" way to test the index hypothesis is to control for the genetic factors directly, by using data on twin-pairs.[19] The basic idea is that identical twins share the same genetic factors, so that twin differences in birth weight *must* be attributable to environmental factors.[20] And if the index property holds and birth weight is a useful policy marker, then those twin birth weight differences must be strongly correlated with twin mortality differences.

More formally, adding subscripts to denote the first- and second-born twin of the same mother, some manipulation of the above relation yields:

(5) $$h_1 - h_2 = \beta(e_1'\theta - e_2'\theta)$$

$$bw_1 - bw_2 = e_1'\theta - e_2'\theta$$

where the genetic components have been eliminated by computing twin-differences. As a result, the regression of twin-differences in health (mortality) on twin-differences in birth weight yields:

(6) $$\frac{\text{cov}(h_1 - h_2, bw_1 - bw_2)}{\text{var}(bw_1 - bw_2)} = \beta$$

[19] In the epidemiological literature, twins studies are viewed as one of the best designs for eliminating confounding due to socioeconomic conditions during infancy and genetic characteristics for monozygotic twins. See Kramer (2000) and Rasmussen (2001) for commentary advocating the use of twins as a natural experiment for evaluating the birth weight hypothesis.

[20] This study cannot distinguish monozygotic from dizygotic twin pairs. Below, we present evidence that, if anything, our twin-pair fixed effects estimates of the effect of birth weight are biased upward by genetic omitted variables.

Therefore, if the two sets of assumptions described above are true, then the "twins estimate" should be strong, and potentially *larger* than the cross-sectional, "between-family" relation. On the other hand, if the twins estimate is significantly smaller, it casts serious doubt on the assumptions of the "index property" and/or of "no genetic confounding" that constitute the basis for using birth weight as an index for health policy purposes.[21]

It is important to note that finding a small or zero within twin-pair correlation between birth weight and mortality does not imply that there exists no modifiable environmental factor that simultaneously reduces LBW and mortality. For example, suppose the single-index property does not hold, and for ease of exposition, assume there are only two environmental factors that determine birth weight and overall "health" so that we have

(7) $$h = g'\alpha_0 + \beta_1 e^{*} + e^{**}$$

$$bw = g'\alpha_1 + \theta_1 e^{*} + e^{**}$$

where e^{*} and e^{**} are the two environmental factors, e^{**} is assumed to have positive effects on both health and birth weight. The covariance between twin differences in health and twin differences in birth weight would then equal

$$\beta_1\theta_1 \text{ var}(\Delta e^{*}) + (\beta_1 + \theta_1)\text{cov}(\Delta e^{*}, \Delta e^{**}) + \text{var}(\Delta e^{**})$$

where Δ denotes twin differences. Clearly, a zero within twin-pair correlation can easily be consistent with one of the factors having positive effects on both birth weight and health, if for example, β_1 and θ_1 have opposite sign.

Data and Summary Statistics on Twins

This study uses comparisons of twin differences in birth weight and infant mortality among the population of twin births in the U.S. to test the validity of using birth weight as a policy marker. Although twins represent a small subpopulation of all infants born in the U.S., there are several reasons

[21] An example of a potentially good marker for policy purposes is educational attainment. Several studies have found that intervention-induced changes in education generally result in higher individual earnings. Ashenfelter and Krueger (1994) find that between-twin differences in education are highly correlated with between-twin differences in earnings, although Bound and Solon (1998) question the validity of using twins to evaluate the effect of education.

why they are of great interest in the health and biomedical literatures. First, while constituting 3 percent of all births in the U.S. in 1997, twins accounted for 21 percent of all LBW births, 14 percent of preterm births, and 13 percent of all infant deaths (Kogan, et al., 2000).[22] Further, Kogan, et al. (2000) find that multiple births have accounted for an increasing share of all LBW infants over the past 20 years. Since LBW is strongly associated with increased risk of infant mortality and subsequent developmental difficulties, several authors have suggested that the rising incidence of twin births is an important public health problem.[23] Finally, several studies suggest that interventions that decrease infant mortality among twins should be applicable to other high-risk groups (e.g., Fowler, et al., 1991).

Linked Birth-Infant Death Data

The datasets used in the empirical analysis are the annual linked birth and infant death microdata produced by the National Center for Health Statistics (NCHS). The microdata files provide detailed information on the universe of births occurring each year in the United States as reported on birth certificates. This natality data are then linked to death certificate information for infants who die before age 1. For the 1989 birth cohort, approximately 97.4 percent of the 38,605 infant death records were matched to one of the 4,045,881 natality records.

The natality portion of the linked microdata provides socioeconomic and demographic information for each mother giving birth in the United States. This information includes maternal age, race, educational attainment, marital status, child-bearing history, prenatal care, cigarette smoking during pregnancy, and geographic residence. Information collected on the father includes age, race, education, and Hispanic origin. The natality portion of the linked data also includes detailed information on the newborn infant, such as the baby's sex, race, gestational age, birth weight, and plurality (i.e., whether the infant was part of a multiple birth).

[22] In our sample frame, twins account for over 2 percent of all births, 15 percent of LBW births, and 10 percent of all infant deaths in the U.S.

[23] For example, Behrman and Rosenzweig (2001) calculate that twinning results in a 12 percent reduction in lifetime earnings when compared to singleton births due to their (28-ounce) lower birth weights. They conclude that the growing incidence of twin births resulting from the increased use of fertility procedures among older women imposes significant costs on children's future development.

For infants who die in the first year of life, NCHS matches detailed information drawn from death certificates to the corresponding birth record. This information includes the infant's age at death in days, where the death occurred, and the precise cause of death. Annual linked birth/infant death files are available from 1983 through the present, with the exception of births occurring in 1992, 1993, and 1994, when only "unlinked" natality and mortality detail files are available.

In our twins analysis, only information for matched twin pairs is used. Multiple birth records are identified with the plurality variable (DPLURAL). In 1989, 2.2 percent of all births (90,022) were twin births (DPLURAL=2). While the corresponding twin pairs for each mother are not explicitly identified in the linked data, beginning with the 1989 linked file, records for twin pairs are generally located next to each other in the "denominator-plus" file produced by NCHS. The adjacency of twin records can be inferred from the detailed information on parental and pregnancy characteristics, which generally are repeated the same number of times as the plurality variable would indicate. The precise procedure used to confirm twin matches is described in the Data Appendix.[24] Further, the birth order of the matched twins can be determined by the information provided on the number of children born to the mother at the time of birth. For the 1989 analysis, 58,132 matched twin pairs born to non-Hispanic, black and white native-born mothers are used in the analysis.

While birth weight is the most commonly used proxy for infant health, other health measures recorded on the birth certificate also reflect an infant's health. In particular, we focus on the five-minute APGAR score as a preferred alternative measure to birth weight. NCHS describes this measure as a "predictor of the infant's chances of surviving the first year of life" and a "summary measure of the infant's condition" (NCHS *Vital Statistics* Technical Appendix, 1990). APGAR score ranges from 0 to 10 and is calculated from five separate tests of newborn health made both one and five minutes after birth. The five factors are each scored a 0, 1 or 2, and then summed to calculate the APGRAR score. The five health factors are heart rate, respiratory effort, muscle tone, reflex irritability, and color.

[24] It should be noted that any "mismatching" of twin pairs will cause an upward bias in the twins fixed effects estimates of the effect of birth weight. Pollack, et al. (2000) use a different algorithm to match twins with their siblings in linked birth-infant death data. However, the matched twin pairs are only used to adjust their estimated standard errors for twin-pair correlation in the residuals.

While births occurring in 1989 are the primary focus of the analysis, twin pairs were also matched in 1985, 1986, 1990, 1991, and 1995. Before 1989, twin records are not located next to each other in NCHS's "denominator-plus" file. Therefore, the plurality indicator is used along with a string of eleven variables reflecting parental and pregnancy characteristics. Checks of this alternative matching algorithm using variables not included in the matching string indicate that the matching is nearly as good as the record location-based approach. Moreover, the 1989 results do not change when the covariate matching approach is used instead. For births occurring in 1995, father's education is no longer collected by NCHS. Therefore, adjacent twin-birth records for 1995 are not checked for consistency in father's education. However, maternal education, maternal age, race, and other factors are still checked for consistency within adjacent twin pairs, as described in the Data Appendix.

Summary Statistics on Twins

Table 1 presents summary information for the 2.7 million singletons and 58,000 twins born to non-Hispanic, black and white native-born mothers in the 1989 linked birth-infant death data. This sample is used in much of the analysis below. The first set of rows show that mothers of twins are more likely to be 30 or older and a college graduate and less likely to be a teenager and a high school dropout than mothers of singletons. While the smoking rates of the two sets of mothers are similar (about 20 percent), mothers of twins have higher incidences of risk factors such as pregnancy associated hypertension and anemia. In addition, twins are much more likely to experience a breech birth and require assisted ventilation after birth.

The next set of rows in Table 1 show that twins, on average, weigh over two pounds less than singletons at birth (2,417 versus 3,369 grams)and that the average twin has a birth weight below the LBW threshold of 2,500 grams. Further, while the incidences of LBW and very low birth weight (VLBW – less than 1,500 grams) among singletons are 6 percent and 1 percent respectively, the LBW and VLBW rates among twins are 50 percent and 10 percent. Twins also have substantially lower gestation lengths and one- and five-minute APGAR scores.

The final sets of rows show that the infant mortality rate within one-year of births of twins is over 4.5 times greater than that of singletons (38.71 versus 8.46 deaths per 1,000 live births). The twin-singleton gaps in death rates are greatest within 24-hours and 28 days of birth (neonatal mortality). While 57 percent and 76 percent of all infant and neonatal mortality among singletons occur among LBW infants, these figures are 93 percent and 97 percent for twins. Thus, it is not surprising that many have concluded that LBW is the predominant cause of infant mortality in the U.S. and also the leading explanation for the greater mortality rate among twins relative to singletons.[25] Twins appear to be a high-risk category of infants and should provide a meaningful subpopulation for testing whether birth weight has a structural effect on infant health.

Table 2 presents sample means on differences in the characteristics of siblings within twin pairs, sorted by two characteristics. The first three columns compare the heavier and lighter twin, while the final three columns compare the first- and second-born twins.[26] The first set of columns show that there are substantial differences in birth weights between siblings within a twin pair. On average, the heavier twin weighs 300 grams (10.6 ounces) more than the lighter twin at birth. This difference is 50 percent larger than the average birth weight difference between the infants of mothers who did and did not smoke during pregnancy that has been documented in the literature. Further, there is substantial variation in twin differences in birth weight across twin pairs. For example, 10 percent of the overall variation in birth weights is within-family, while over 20 percent of all twin pairs have greater than a 20 percent difference in birth weights.[27]

The underlying causes of differences in birth weights between twins are not well understood. The most commonly cited factors include discrepancies in the microintrauterine environment due to

[25] Studies that cite higher LBW incidence as a significant reason for higher mortality among twins include McCarthy, et al. (1981), Fraser, et al. (1994), and Gall, et al. (1996).

[26] Of the 29,066 twin pairs, 1,375 have identical birth weights according to their birth certificates. The birth-order of siblings could not be determined for 2,823 twin pairs.

[27] Misclassification error due to the assignment of twins to an incorrect birth or death certificate appears to be an unlikely source of variation in twin pair birth weight differences. Specific safeguards are in place to prevent this. For example, often two separate teams are present at each twin delivery to examine each newborn. Babies are labeled with tags almost immediately after birth and remain tagged until they leave the hospital. While it is conceivable that mixing could occur at home, we test for this possibility below by examining boy-girl twin pairs. We thank Dr. Christopher Almond for providing us with this information.

different placentation and blood perfusion (Zhang, et al., 2001), different nutritional sources at different parts of the uterus, and different genetic growth potentials among dizygotic twins.[28] Thus, among monozygotic twins differences in birth weight are presumably entirely attributable to the intrauterine environment. A separate literature attributes some of the difference in birth weight among discordant twin pairs (defined to be pairs with over a 25 percent birth weight differential) to twin-to-twin transfusion within the womb.[29]

Continuing with the first three columns in Table 2, the heavier twin is more likely to be male. Interestingly, the heavier twin is no more likely to require assisted ventilation after birth than the lighter twin and has only slightly higher APGAR scores. Thus, based on these measures, it is unclear whether the heavier twin is healthier. On the other hand, the lighter twin has a slightly higher infant mortality rate than the heavier twin. We find below that this gap is much smaller than would be predicted by the cross-sectional birth weight-infant death association and is driven almost entirely by differences between twins in the incidence of genetic disorders.

In the final three columns of Table 2, it appears that while there is only a small difference in birth weights between the first- and second-born twin, other measures of infant health suggest that the first-born twin is significantly healthier at birth. The first-born is less likely to be breech, less likely to require assisted ventilation, and has significantly higher APGAR scores. Further, given the small birth weight difference, the first-second born gap in infant mortality is surprisingly large.[30] There appear to be real differences in the health of first- versus second-born siblings that would be masked if one focused on their virtually non-existent birth weight differential.

Results

[28] Some of this follows from a conversation with Dr. Louis Keith in the Obstetrics and Gynecology Department of Northwestern University. Among many other books, Dr. Keith has edited a text on multiple pregnancies.
[29] The causes and effects of birth weight discordance are unresolved (Dube, et al., 2002, Cunningham, et al., 2001). Studies have found increased risk of neonatal death among discordant pairs (Fraser, 1994, Hollier, et al., 1999).
[30] Fowler, et al. (1991) document that the second-born twin is at increased risk of death due to asphyxia and newborn respiratory disease.

Summary of the Results

In this section, we report the results from our analysis. First, we present the basic regression analysis, which shows that within twin-pair, the correlation between infant mortality and birth weight is precisely estimated to be small, and in many instances, roughly zero. This finding is strongly inconsistent with the "index property" described above, and hence also inconsistent with the notion that health interventions that work to reduce the incidence of low birth weight necessarily lead to reducing infant mortality.

Second, we perform the analysis with a wide variety of specifications and samples. We allow for non-linearities in the mortality-birth weight relationship in a more flexible way, and in both linear probability and conditional logit specifications. We perform a test of the exchangeability of the twins implied by our fixed-effects specification. We examine the results by different demographic groups defined by education and race of the mother. We consider whether misclassification of twins could affect the results by examining boy-girl twin pairs. The analysis is broken down by different causes of death, and examined for stability across time. In addition, we substitute our primary measure of health (infant mortality) with other common measures of infant health (APGAR scores, and whether the newborn utilized a ventilator). In every instance, we obtain the same result: a strong cross-sectional relation between measures of health and birth weight that falls dramatically, and often to zero when accounting for fixed genetic factors among twin-pairs.

Third, we show that among the four candidates for measures of infant health – mortality, APGAR score, use of ventilator, and birth weight – birth weight is the only one of the four that is uncorrelated with the remaining measures, after accounting for the family fixed effects. By contrast, the relationships between mortality and APGAR, and between ventilator-use and APGAR score, survive even after accounting for common genetic factors within twin-pairs. This implies that health measures such as the APGAR score – in contrast to birth weight – may have a closer link to overall "health", and thus could be useful in evaluating the effectiveness of various health interventions.

Basic Regression Analysis

We first document the well-established cross-sectional relation between infant mortality and birth weight. Using our main sample of twins, Column (1) of Table 3 reports the estimated slope coefficient from an OLS regression of whether the infant died within the first year on birth weight (in grams). The coefficient is expressed in terms of the probability of dying per 1000 grams, or equivalently, the number of deaths per 1000 births per 1 gram. As is well-known, the cross-sectional association is strongly negative, and in our large sample, it is highly statistically significant. The coefficient remains at around -0.12, when mother's characteristics (race, education, age, martial status) are included in the regression in Column (2). The coefficient also does not change significantly, when pre-natal care status is included, and mother's characteristics are included in a more flexible way (interactions of separate categories of the variables).

Since there is a mechanical relation between birth weight and gestation length, it is natural to examine whether birth weight is associated with mortality after conditioning on gestation length. Columns (4) and (5) include gestation length (in weeks) both linearly and in week categories. Clearly gestation length is a strong predictor of mortality propensities, and substantially diminishes the coefficient on birth weight, which falls to about -0.056. The coefficient falls slightly more when individual gestation length week-dummies are included. Overall, Table 3 suggests that the simple cross-sectional relation between birth weight and infant mortality is partially driven by determinants of birth weight that have their own association with infant mortality. To the extent that a policy focuses on affecting birth weight at fixed gestation length (i.e. influencing the degree of IGR), Table 3 suggests that the impact of that policy on the IMR would be significantly smaller than that suggested by the simple regression in Column (1). Table 3 also raises the question of whether the coefficient on birth weight would fall even more, and possibly to zero, after conditioning on other characteristics of the mother and infants – in particular, immutable determinants such as genetic determinants of birth weight and overall "health".

Table 4 directly addresses that question, and presents the primary findings of our paper. Panel A focuses on the full sample of twins, and it shows that after accounting for common genetic factors between twins, the coefficient on birth weight falls dramatically, in some cases more than a factor of 10. The family fixed effects estimate falls to -0.02 for one-year mortality, and falls from -0.075 to -0.005 for mortality within 24 hours of birth. In Panel B, we exclude twin-pairs where at least one twin was coded as having a congenital anomaly or as having died due to a congenital anomaly, which is standard practice in the analysis of these data. This exclusion eliminates less than 2000 observations, but the impact on the birth weight coefficient is substantial, particularly for the fixed effects estimates. When these observations are dropped, the coefficient in the benchmark OLS regression falls about 15 percent, but falls by a factor of 3 for the fixed effects specification for the one-year mortality regression. The same is true for the other mortality measures, and the fixed effects estimate for 1-day mortality is no longer statistically significant, even though the standard errors are on the order of 0.002 in probability per 1000 grams.

As pointed out earlier, if the "single-index" property holds and the cross-sectional relationship is not subject to genetic confounding, the fixed effects estimate should actually be larger than the OLS estimate. Since the fixed effects specification accounts for common genetic factors within twin-pairs, the fact that the coefficient falls significantly, and in some cases to zero implies two possible interpretations: 1) the cross-sectional relation between infant mortality and birth weight is driven primarily by omitted genetic factors, and the true "structural" effect of birth weight is close to zero, or 2) the "single-index" property does not hold. The first interpretation would imply that any intervention that reduces the incidence of low birth weight is likely to have little or no impact on infant mortality. On the other hand, the latter interpretation would imply that – despite the undeniably strong cross-sectional relation between birth weight and mortality – birth weight is not a reliable "substitute" for infant's propensity to die in the evaluation of health interventions; there may be many interventions will affect birth weight but not infant mortality, or vice versa, and the only way to evaluate the effect of some policy on mortality is to directly

examine mortality. Although we do not have any direct evidence that can distinguish between these two interpretations, we favor the latter (and more cautious) interpretation.

Graphical Analysis

Since there is an important non-linearity in the birth weight-mortality relationship, we present our basic twins analysis graphically. Figure 1 illustrates our main results. The figure first plots the one-year infant mortality rates by birth weight for our sample of 52,486 twins (or 26,243 twin-pairs) from our 1989 data.[31] There is a striking negative relationship between infant mortality and birth weight; more striking is the apparent non-linearity in the relation. While 90 percent of infants born weighing around 500 grams die within one year, the rate falls rapidly to about 50 percent for newborns weighing 700 grams. By about 1000 grams, the one-year mortality rate falls to about 20 percent. After that, point the relationship is much flatter, with over 90 percent of infants weighing 1500 grams (which is conventionally considered "very low birth weight") at birth surviving at least one year after birth.

The non-linearity in the figure illustrates that the relation between "low birth weight" (conventionally defined as less than 2500 grams) and infant mortality is primarily driven by the 5-10 percent of twins that are born less than 1000 to 1500 grams. Overall, in this sample, the mortality rate for twins weighing less than 2500 grams is about 6.5 percentage points higher than the rest of the newborns, but when babies less than 700 grams are excluded from this computation, the difference falls by more than half.[32] When newborns less than 1000 grams are excluded, the difference is reduced to about 1.6 percentage points. This non-linearity is important to note for policy purposes because it suggests that even if the observed cross-sectional relation between mortality and birth weight were "causal", only policies that affected birth weights at extremely low birth weights would be effective in significantly impacting mortality.

[31] The sample excludes 5646 twins for whom it could not be determined which twin was born first. Each dot represents 0.5 percent of the sample. The y-coordinate is the infant mortality rate within the 0.5 percent cell and the x-coordinate is the average birth weight for the cell.

A flexible way of computing a within twin-pair estimate that accounts for the marked non-linearity is to estimate the following specification

(8) $$M_{ij} = \alpha_i + FB_{ij}\delta + x'_{ij}\pi + \varepsilon_{ij}$$

Where M_{ij} is an indicator variable for death within one-year for family *i* and twin *j=1,2*. a_i is the family fixed effect, *FBij* is an indicator that the infant was the first-born twin, and x_{ij} is a vector of dummy variables indicating to which birth percentile the infant belongs. The coefficients on the regressors are defined as the linear projection coefficients, so that ε_i is by construction orthogonal to all of the regressors (including the family dummy variables). The corresponding estimated coefficients π estimated at the average fixed effect is plotted by the solid squares in Figure 1A.

The figure shows that, after allowing for a family-specific fixed-effect, the relation between one-year mortality and birth weight does not increase, as predicted by the index property and no-genetic-confounding assumptions. Instead, the relationship is much *weaker*. Moving across from 300 to 1000 grams, rather than falling by 0.70, it falls by about 0.20. An important feature of the graph is that the mortality-birth weight gradient falls most in the range of birth weights where the cross-sectional gradient is steepest.

The fall in the gradient is even more evident when births involving congenital anomalies are excluded from the sample (as is standard practice in the analysis of these data). Focusing on the range of birth weights between 700 and 1100 grams, the steep profile evident in the cross-section becomes virtually flat when examining the within twin-pair estimates.

A concern that arises is that the analysis of the twins may not be readily extrapolated to a broader population of singleton births. In particular, if the birth weight-mortality relationship in the cross-section for twins appeared significantly different than that for singleton births, a question would arise as to which relationship should be used as a "benchmark" for the fixed effects estimates: the singletons cross-

[32] Calculated from a least squares regression of mortality on a dummy variable indicating first-born, and a dummy variable indicating birth weight less than 2500 grams, and successively dropping newborns weighing less than 200,

sectional relation, or the twins cross-sectional relation? Figure 1A shows that there is little evidence that is a problem in our context. The open circles represent the cross-sectional relation between mortality and birth weight, and the profile exhibits a very similar pattern to that of the twins, even though twins are born at significantly lower birth weights.

Figure 1B repeats the analysis for the 1-day mortality measure, and it yields an even more striking picture. First, there is a strong negative relation with birth weight at extremely low birth weights, and a marked non-linearity, with the relation significantly flattening out after 1000 grams. Second, the corresponding fixed-effects estimates reveal a virtually flat relation even for very low birth weights, whether or not births involving congenital anomalies are excluded. Again, the profile flattens out in the range of birth weights where the cross-sectional gradient is steepest. And finally, the cross-sectional relations for singleton and twins births are virtually coincident with one another.

Fixed Effects Specification Test and Alternative Non-linear Specifications

The fixed effects formulation permits a simple specification test (Chamberlain 1984), so we perform these tests here. Table 5A provides a basic quantification of the relationships depicted in Figure 1A. The first column reports the cross-sectional OLS estimate and is similar to the estimate reported in Table 4.[33] The second column reports that the inclusion of the family fixed effects reduces the coefficient on birth by more than a factor of 4.

Re-writing Equation 7 for the first-born and second-born twins separately, we obtain

$$(9) \qquad M_{i1} = \delta_1 + x'_{i1}(\pi + \lambda_1) + x'_{i2}\lambda_2 + \varepsilon^*_{i1}$$
$$M_{i2} = \delta_2 + x'_{i1}\lambda_1 + x'_{i2}(\pi + \lambda_2) + \varepsilon^*_{i2}$$

where λ_1 and λ_2 are the linear projection coefficients obtained from linearly projecting a_i onto the first-born and second-born birth weights, x_{i1} and x_{i2}, respectively. ε^*_{i1} and ε^*_{i2} are, by construction, orthogonal to all regressors in the equations. From the reduced-form regressions of an infant's mortality status on its

300 grams, and so forth.

[33] The samples in Tables 5A-5D exclude twins with missing APGAR scores. We do this to keep the sample constant, when we examine the relation between APGAR and birth weight.

own and twin's birthweights, two estimates of π can be obtained. For example, one estimate is obtained from subtracting from the coefficient on "own birth weight" in the first-born regression the coefficient on "twin's birth weight" in the other twin's regression. The other estimate is computed by doing the same thing in the second-born regression. If the two estimates are substantially different, it indicates that the fixed-effects specification may be inappropriate (Chamberlain, 1984).

Column (3a) reports the coefficients from regressing the first-born's mortality status on both twins' birth weights. Column (3b) does the same for the second-born twin. As evident from the table, the two estimates – (-.078+0.059=) -0.019 and (-0.073+0.048=) -0.025 are reasonably similar in magnitude, even though there is enough data to reject their equality at conventional levels of statistical significance. In addition, complete symmetry – where λ_1 and λ_2 are equal – can be tested, which is reported in Table 5A. The estimates are similar (-0.048 and -0.059).

Column (4) reports the coefficients of a logit specification. In the cross-section, the log-odds of death decrease by 0.31 per 100 grams of birth weight. Starting at a base of a 0.90 mortality rate, a 700 gram increase in birth weight would imply a decrease in the mortality rate to about 0.50. Comparing this drop to that implied by Figure 1A shows that the linear-in-birth weight logit specification does not fully capture the degree of non-linearity in the cross-sectional association.

Column (5) reports the coefficient from a "fixed-effects" logit specification – a conditional logit. The estimate in terms of log-odds is considerably smaller. Again, starting at a base 0.90 rate, the estimate implies that a 700 gram increase in birth weight would imply a decrease in the rate to about 0.80 – or about 20 percent of the size of the drop implied by the cross-sectional logit.

Table 5B reports the results from estimating a more flexible functional form to account for the non-linearity in birth weight. In particular, instead of including a linear term for x_{ij} in Equation (8), we include a piece-wise linear spline in birth weight, where the knot points are birth weights that correspond to the 5^{th}, 10^{th}, 15^{th}, etc. percentiles of the twins birth weight distribution.

Column (1) in Table 5B reports the slopes of the first three segments of the spline in a linear probability specification; the knot points defining the first three segments are 1085, 1503, and 1758 grams. The first segment captures most of the cross-sectional birth weight association, with a decrease in mortality of 0.12 in probability per 100 gram increase in birth weight. That slope changes falls substantially for the second and third segments.

The fixed-effects specification in Column (2) yields a slope of 0.032 per 100 grams in the first segment. Column (3) shows that the fixed-effects formulation passes the specification test suggested by Chamberlain (1984). The coefficients on separate regressions for the first- and second-born twin are quite similar. In fact, equality of the two sets of estimates that can be generated from a fixed effects specification cannot be rejected at conventional levels of statistical significance.[34] The further restriction of symmetry between the two twins also is also not rejected by the data.

Finally, Columns (4) and (5) report the results from the cross-sectional logit and fixed-effects logit, but utilizing the piecewise linear spline in birth weight.[35] The logit coefficient in the first segment implies that starting at a mortality base rate of 0.90, a 700 gram increase in birth weight reduces the mortality rate to about 0.07. By contrast the conditional logit coefficient, reported in Column (5), implies that the same 700 gram increase would reduce the mortality rate to about 0.75.

Starting at a base mortality rate of 0.40, the conditional logit estimate for the first segment implies that a 700 gram increase in birth weight would decrease the rate to about 0.17. It should be noted that this decrease is virtually identical to that implied by fixed-effects linear probability model (evaluated at the average fixed effect), as depicted in Figure 1A. Thus, the estimates summarized in Figure 1A are not an artifact of imposing linear specifications to a binary dependent variable.

[34] The coefficients on the 4th through 20th segments are not reported here, but the test of equal coefficients and the symmetry test utilize all 20 segments.

[35] For the logit-spline specifications, we utilize only a three-segment spline where we use the 5th and 10th percentile birth weights as knot points. The conditional logit necessarily excludes all twin-pairs where either both died or both survived. This meant that there were no infants in many of the segments of the 20-segment spline.

For completeness, Tables 5C and 5D report analogous results to Tables 5A and 5B, except we examine 24-hour mortality rates. The results are qualitatively similar: the fixed effects estimates are substantially smaller, and particularly so in the first segment of the spline, the over-identifying restrictions implied by the fixed effects specification are not rejected by the data, and the importance of the family effects is evident in the large difference between the logit and conditional logit estimates.

The most important difference found in the analysis of the one-day mortality rates is that the twins fixed effects estimates imply *no statistically detectable* association between birth weight and mortality. This is result is found in the linear, linear spline specification, and in the logit specifications, and it agrees with the graphical summary in Figure 1B.

Stability of Estimates: Sub-populations, Twin-types, Cause of Death, Over Time

Table 6 reports our basic estimates for different racial and education groups. It shows that even though the cross-sectional relationship between infant mortality and birth weight is stronger for blacks than for whites, it becomes weaker in the fixed effects specification. This implies that the reliance on birth weight as a "substitute" proxy for health may be even more suspect for the population of black mothers in the U.S. Furthermore, if one were to infer something meaningful from the fact that there is a race interaction in the birth weight gradient in the cross-section (e.g. that policies that improve birth weight will improve infant health more for the black population), one would also have to confront the fact that the sign of the interaction reverses when accounting for confounding common genetic factors within twin-pair. The rest of the table shows how the same basic pattern – fixed effects estimates being significantly smaller in magnitude than the OLS estimates – holds in every sub-population that we examine.

Table 7 presents evidence that mis-identification between twins is unlikely to be the source of the small fixed effects estimates. In principle, it is possible that twins are "switched" after birth weight measurement. If all twins are randomly switched with their sibling, then one would expect to observe a zero fixed effects estimate. However, if this were true, one would expect the fixed effects estimate for

same-gender twin-pairs to be significantly smaller (due to twin "mix-up"s) than the fixed effects estimate for boy-girl twin-pairs (where a mix-up is impossible). In fact, the fixed effects estimates for same-gender, both-boy, both-girl, and boy-girl twin pairs are virtually identical, for both the one-year and one-day mortality regressions.

Table 8 presents both the cross-sectional and fixed effects estimates by cause of death. For each category, the dependent variable equals 1 if the infant was recorded to have died of that cause, and equals 0 otherwise (died of another cause or survived). The first two causes of death – low-birth weight/short gestation and respiratory distress – are known to have the strongest association with birth weight. In the table, the slope coefficients are -0.75 and -4.53. In the fixed effects specification, the coefficients drop by more than a factor of 10, and are no longer statistically significant. The only cause of death for which the fixed effects estimate is larger in magnitude, and remains statistically significant, is death due to congenital anomalies. This is natural to expect in light of Figure 1A, which shows that the fixed effects gradient itself is non-trivially driven by the inclusion of twin-pairs where one of the twins obviously do not share similarly "healthy" genes with his/her twin sibling. As shown in Figure 1A and Table 4, the inclusion of these infants caused a significant upward bias in absolute magnitude. This raises the question of whether our fixed effects estimates are still upward biased, because of dizygotic twins, whose non-common genetic factors may be correlated with the environmental factors to which they are subjected. However, as mentioned above, we hesitate to interpret our fixed effects analysis as having identified the "true" structural relation between birth weight and infant mortality, but rather interpret the evidence as a rejection of the single-index hypothesis that is necessary to justify birth weight as an appropriate "substitute" for more direct measures of health when evaluating the effects of policy interventions.

Table 9 shows that our fixed effects estimates are stable over different time periods. We examined births in 1985-86, 1990-1991, and 1995, and the analysis shows that the coefficient on birth weight in the fixed effects specification is quite stable, for different mortality measures. Our estimates

are not an artifact of choosing one particular year of data. By contrast, the slope coefficient from the cross-sectional regressions are changing significantly over time, falling in absolute magnitude from -0.132 to -0.090 for one-year mortality over the ten-year period we examined. These patterns are fully depicted in Figure 2A and 2B.

Alternative Measures of Health

The choice of birth weight as a target of health policy is primarily motivated by the strong cross-sectional birth weight-mortality association. However, it is instructive to explore whether birth weight, after accounting for genetic determinants of birth weight and health, is related to *any* independent measure of health. A step in this direction is to conduct an analysis, replacing the mortality indicator with another measure of the health of the infant: the 5-minute APGAR score of the infant.

Figure 3 and Tables 10A and 10B report a parallel analysis to that in Figures 1A and Tables 5A and 5B, with the only difference that the five-minute APGAR score for the newborn is the dependent variable.[36] Figure 3 shows that there is a fairly strong cross-sectional association between the APGAR score and birth weight, throughout most of the birth weight range below 2500 grams. There is a marked non-linearity in the relation among birth weights less than 700 grams, which comprise less than 3 percent of twins births.

Again, if the "index property" and "no genetic confounding" assumptions hold, the within-family relation should be *stronger*. But Figure 3 shows that after accounting for genetic determinants of birth weight and APGAR scores, the relationship is significantly weaker. In fact, after about 700 grams, the relation is virtually flat.

Tables 10A and 10B quantify these findings. In the linear specification, the cross-sectional estimate implies a 0.10 increase in the APGAR score for every 100 gram increase in birth weight. The twins estimate, by contrast, implies a 0.01 increase in the score for every 100 gram increase in birth weight. The fixed-effects linear-spline estimates imply that there is no statistically detectable relation

[36] Since the APGAR score is not binary, the logit specifications are omitted.

between birth weight and APGAR score beyond the first segment – that is, among 95 percent of twins births. Similar results are found using the 1-minute APGAR score, as documented in the Appendix.

Table 11 conducts a similar analysis, but using an indicator for whether the infant required the use of assisted ventilation after birth as an alternative measure for health. As shown in the first 4 columns, while there is a strong negative relation between birth weight and the use of assisted ventilation (at all, or only more than 30 minutes). But after accounting for common genetic factors via the fixed effects specification, the coefficient drops to a precisely estimated effect that is close to zero, and statistically insignificant.

Overall, Tables 10A, 10B, and 11 show same basic pattern when we use these alternative measures of overall "health": the fragility of the cross-sectional relations to the inclusion of family fixed effects, and a negligible within twin-pair correlation between birth weight and measures of health.

Other Possible Targets of Policy: APGAR score?

An important concern about the twins analysis is that twins may be "too similar". That is, in principle, one of the reasons why there is weak between-twin relation between mortality and birth weight, and between APGAR score and birth weight, is that there is little between-twin variation in overall "health". That alone could explain the weak predictive power of birth weight.

However, this explanation has little empirical support for two reasons. First, empirically, there is substantial between-twin variation in health, as measured by infant mortality or APGAR score. For example, family effects explain about 57 percent of the cross-sectional variation in the one-year mortality indicator and about 65 percent of the cross-sectional variation in the five-minute APGAR score, leaving a fair amount of between-twin variation in those two measures of health.

Second, there exists a substantial between-twin relation between APGAR score and mortality rates. Figure 4A plots mortality rates by five-minute APGAR score (where scores of 0 and 1 are combined) for our population of twins births. In the same figure is the corresponding fixed-effects relationship, which is significantly reduced, but not zero, which would be the case if between-twin

variation in health were negligible.[37] The same is true for one-day mortality rates, where the gradient for the fixed-effect specification is almost as steep as the cross-sectional slope, as shown in Figure 4B. Thus, the APGAR-score-mortality relation is not as fragile as the birth weight-mortality relationship.

Table 11 provides further evidence that there is meaningful variation in health, and that APGAR score may be a useful measure of health. Fixed effects regressions of the use of assisted ventilation on 5-minute APGAR score reveal somewhat smaller coefficients than in the cross-section, but at the same time, remain statistically significant – in contrast the fixed effects estimates for birth weight (the left part of the table).

In summary, the within twin-pair analysis reveals: a correlation between APGAR score and mortality, a correlation between APGAR score and use of assisted ventilation, but little or no correlation between birth weight and mortality, no correlation between birth weight and APGAR score, and no correlation between birth weight and the use of assisted ventilation. These findings suggest: 1) there is meaningful variation in overall "health" within twin-pairs, and 2) APGAR score may be superior to birth weight as a(n) (imperfect) "substitute" for actual health outcomes of interest (e.g. propensity to die) in health intervention evaluation.

Policy Application: Maternal Smoking Cessation

In this section, we consider the implications of our empirical analysis of twins for an important policy context by directly examining the empirical relationship between maternal smoking behavior and infant health outcomes. It is conventionally accepted that maternal smoking causes an increase in the incidence of low birth weights. It has thus been argued that since birth weight is highly correlated with overall health, the most effective way to improve infant health and reduce infant mortality in the United States is to promote maternal smoking cessation policies. If the cross-sectional relation between birth weight and infant mortality represents a primarily structural link, then we should expect to see strong

[37] The fixed-effect relation are the coefficients from a regression of the mortality indicator on an indicator of first-born and dummy variables for each APGAR score (combining the 0 and 1 scores to be one group). The figure plots the coefficients evaluated at the average fixed effect.

associations between maternal smoking and infant mortality. On the other hand, our analysis of twins, suggests that there probably not a simple "one-to-one" link (the single-index property) between birth weight and mortality, and that if such a structural link exists at all, it may even be zero.

Below, we turn to a between-family analysis of infants, and show that simple multivariate regression and matching estimates confirm the well-established negative empirical relation between maternal smoking and birth weight. However, an exactly parallel analysis also reveals that there is virtually no empirical relationship between maternal smoking and any of the following infant outcomes: gestation duration, APGAR scores, death due to low birth weight, death due to respiratory distress, or overall mortality within 1-day, 28-days, or 1-year. These findings provide one concrete and important example of the primary implication of our analysis of the infant twins: even if a specific health policy is known to reduce the incidence of low birth weight, this contains little or no information, by itself, about the overall health consequences (and implicit costs) of the intervention.

Basic Regression Analysis

Table 13, Panel A reports estimates from a simple regression analysis of the association between smoking and various infant outcomes for the population of mothers/infants in Pennsylvania, 1989-1991. The first column implies that infants to smoking mothers are born weighing about 280 grams less than infants of non-smoking mothers. That difference falls to about 200 grams when mothers' characteristics are included.[38] This is clearly a non-experimental, purely observational study, so we recognize that maternal smoking and other health outcomes of the mother and infant are likely to be simultaneously determined, and that the coefficient on smoking may be biased. Nonetheless, for the purposes of this policy application, it is instructive to stipulate that this represents a causal effect since it is generally accepted in the biomedical and health policy literatures that maternal smoking has a negative causal impact on birth weight.

[38] For details on the covariates included, see Torelli (2000).

As Panel A shows, a parallel analysis – including the same set of covariates – of regressions of gestation length and 5-minute APGAR score on maternal smoking, reveals no effect of maternal smoking. Furthermore, as Panel B demonstrates, a parallel analysis reveals that the regression-adjusted "effect" of maternal smoking on infant mortality is slightly negative. This is the case for mortality within one-year, within 24 hours, or within 28 days. There is also a slightly negative "impact" of maternal smoking on death due to low birth weight or death due to respiratory distress.

Table 13 presents a similar analysis, but with the twins population in 1989 in most of the U.S.[39] It reveals a qualitatively similar pattern: a large impact of smoking on birth weight, on the order of 160 grams in the regression-adjusted specification, but either zero or slightly negative impacts on infant mortality. It should be noted, however, that the standard errors are much larger in magnitude than the standard errors for the sample utilized in Table 12. As does Table 12, Table 13 highlights that the prediction that smoking would have a significant positive impact on mortality has very little empirical support.

Propensity Score Analysis

Figure 5 graphically depicts a propensity score analysis of the empirical association between maternal smoking and birth weight for the Pennsylvania population of births. The crosses in Panel A represent the average birth weight of infants to non-smoking mothers by values of the estimated propensity score, where the propensity to smoke is modeled as a logit that includes a multitude of mothers' characteristics as covariates.[40] The predicted probabilities are then grouped into 200 equally sized bins from the overall distribution of the estimated propensity scores. The circles represent the average birth weight of infants to smoking mothers. The graph reveals a visible gap that appears constant across the whole range of estimated propensity scores. The difference in average birth weights is plotted in Panel B.

[39] The records for New York and California do not include information on smoking behavior in this year.
[40] For the details of the specification, see Torelli (2000).

Figure 6A and 6B present a parallel analysis using the 5-minute APGAR score as the outcome of interest. It shows a very different result. There appears to be no "effect" of maternal smoking on the health of the infant, as measured by APGAR score. Finally, Figures 7A, B, C, and D reveal that after conditioning on a rich set of covariates, there is little difference in infant mortality rates between infants of smoking and non-smoking mothers.

Overall, the evidence presented in Tables 12 and 13 and Figures 5-7 show that among the many measurable infant outcomes available in the birth certificate data, it appears that the only outcome with which maternal smoking is correlated is birth weight. There is little evidence that maternal smoking, after for adjusting for observable characteristics of the mother, is associated with gestation length, APGAR score, overall infant mortality, or with death due to low birth weight or respiratory distress.

This particular policy application provides a concrete example of the main implication of our earlier analysis of twin births: even though a health policy is known to manipulate birth weight, it may have no effect (or a perverse effect) on infant mortality. This conclusion does not follow from simple logic, but from an analysis of the twin-pair data. If there is a "one-to-one" or index relationship between birth weight and infant mortality, the relationship should survive a within-family analysis, where the variation in birth weight is primarily driven by environmental factors. The fact that the data do not support such a within-family correlation implies that either 1) the "one-to-one" slope on birth weight is roughly zero, or 2) there exists no "one-to-one" relationship, so that the knowledge that a health policy has an impact on the incidence of low birth weight provides no information on the impact of the policy on health, as measured by mortality propensities.

Conclusion

From our analysis of the conceptual issues and the empirical results, we draw the following conclusions. First, the claim that an outcome is a valid policy marker for a different, actual outcome of interest is a strong claim. It implicitly imposes a tight structural link between the two outcomes – a link that essentially assumes an "index property". If the index property does not hold, the variable (e.g. birth

weight) ceases to be a valid policy marker, as it implies that there may be interventions that manipulate the marker without any corresponding impact on the outcome of interest (or vice versa). In that case, the only way to assess whether an intervention has an impact on mortality, other measures of health, or costs, is by directly examining mortality, other measures of health, or costs. If there are important omitted factors, any cross-sectional association between the marker and the outcome of interest is not sufficient to establish that structural link.

Second, the evidence we present suggests that the cross-sectional association between LBW and mortality is seriously biased by omitted genetic factors, and an analysis that accounts for those genetic factors reveals little correlation between birth weight and several measures of health. The notion of birth weight as a valid policy marker for infant mortality and other indicators of health appears to have weak empirical grounding. It is important to note that we are not concluding that there exists no intervention that will simultaneously reduce the incidence of LBW and reduce infant mortality. Rather, a more plausible (and cautious) interpretation of our results is that birth weight is simply not a reliable substitute for direct health measures of interest, and the reliance upon birth weight to evaluate and predict the effects of health interventions has very weak empirical justification.

While it may be true that birth weight is a convenient and relatively costless infant outcome to record, the benefits of minimizing data collection efforts must be weighed against the costs of utilizing a poor policy marker, which has the potential of leading to misleading inferences about the true magnitude of impacts of health interventions and public policies. APGAR scores could be collected at relatively low cost, and our analysis suggests that APGAR scores may actually be a superior policy marker. This points to the possibility that there may be relatively low-cost measures that do act as a valid policy marker for overall infant health. It seems that a fruitful avenue for future research is to explore that possibility.

Finally, while we do explore alternative measures of health, our analysis primarily focuses on whether birth weight is a valid policy marker for infant mortality. Our results indicate that birth weight may be a poor policy marker for mortality, but they imply nothing about the appropriateness of birth

weight as policy marker for other less health-related outcomes. For example, birth weight may have a structural link with – and hence be a valid policy marker for – cognitive development, eventual educational attainment, and eventual labor market outcomes. At a minimum, our study suggests that the standard of evidence for establishing that birth weight is a policy marker for those outcomes is the same as that for establishing a structural link between birth weight and those outcomes.

References

Almond, Douglas, Kenneth Y. Chay, and Michael Greenstone, "Civil Rights, the War on Poverty, and Black-White Convergence in Infant Mortality in Mississippi," mimeograph, University of California, Berkeley, 2001.

Alexander, Duane "Child Health: Birth Outcomes and Infant Mortality," Speeches and Reports, National Institute for Child and Human Development.

Ananth, C.V., G.S. Berkowitz, D.A. Savitz, and R.H. Lapinski, "Placental Abruption and Adverse Perinatal Outcomes," Journal of the American Medical Association, 282 (1999), 1646-1651.

Ashenfelter, Orley and Alan Krueger, "Estimates of the Economic Return to Schooling from a New Sample of Twins," American Economic Review, 84 (1994), 1157-1173.

Baker, Laurence C., and Ciaran S. Phibbs, "Managed Care, Technology Adoption, and Health Care: The Adoption of Neonatal Intensive Care," mimeo, Stanford University, 2001.

Barker, D.J.P., "The Fetal and Infant Origins of Adult Disease," British Medical Journal, 301 (1992), 1111.

Barker, D.J.P., P.D. Winter, C. Osmond, et al., "Weight in Infancy and Death from Ischaemic Heart Disease," Lancet, 2 (1989), 577-580.

Barker, D.J.P., P.D. Gluckman, K.M. Godfrey, et al., "Fetal Nutrition and Cardiovascular Disease in Adult Life," Lancet, 341 (1993), 938-941.

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman, "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment," Journal of Political Economy, CII (1994), 1131-1174.

Behrman, Jere R., and Mark R. Rosenzweig, "The Returns to Increasing Body Weight," PIER Working Paper No. 01-052, 2001.

Bound, John, and Gary Solon, "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling," NBER Working Paper 6721, 1998.

Brooks, Anne-Marie, et al., "Impact of Low Birth Weight on Early Childhood Asthma in the United States," Archives of Pediatric and Adolescent Medicine, 155 (2001), 401-406.

Bugalho, A., et al., "Outcomes of Twin Pregnancies at the Hospital Central of Maputo: Retrospective Study of 315 Consecutive Twin Deliveries, January 1 – September 30, 1987," International Journal of Gynecology and Obstetrics, XXIX (1989), 297-300.

Case, Anne, Darren Lubotsky, and Christina Paxson, "Economic Status and Health in Childhood: The Origins of the Gradient," NBER Working Paper No. 8344, 2001.

Chay, Kenneth Y., and Michael Greenstone, "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," mimeo, University of California, Berkeley, 2001a.

_____, "Air Quality, Infant Mortality, and the Clean Air Act of 1970," mimeo, University of California, Berkeley, 2001b.

Christensen, K., J.W. Vaupel, N.V. Holm, et al., "Mortality among Twins after Age 6: Fetal Origins Hypothesis versus Twin Method," British Medical Journal, 310 (1995), 432-436.

Cunningham, F. Gary, et al., Williams Obstetrics (New York: McGraw-Hill Medical Publishing Division, 2001).

Currie, Janet, and Nancy Cole, "Welfare and Child Health: The Link Between AFDC Participation and Birth Weight," American Economic Review, LXXXIII (1993), 971-985.

Currie, Janet, and Jonathan Gruber, "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women," Journal of Political Economy, CIV (1996), 1263-1296.

Currie, Janet and Rosemary Hyson, "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight," AEA Papers and Proceedings, LXXXIX (1999), 245-250.

Cutler, David M., and Ellen Meara, "The Technology of Birth: Is It Worth It?" NBER Working Paper 7390, 1999.

Cutler, David M., and Ellen Meara, "Changes in the Age Distribution of Mortality over the 20th Century," NBER Working Paper 8556, 2001.

Dube, Johanne, et al., "Does Chorionicity or Zygosity Predict Adverse Perinatal Outcomes in Twins?," American Journal of Obstetrics and Gynecology, CLXXXVI (2002), 579-583.

Fowler, Mary Glenn, et al., "Double Jeopardy: Twin Infant Mortality in the United States, 1983 and 1984," American Journal of Obstetrics and Gynecology, CLXV (1991), 15-22.

Fraser, D., et al., "Birth Weight Discordance, Intrauterine Growth Retardation and Perinatal Outcomes in Twins," Journal of Reproductive Medicine, XXXIX (1994), 504-508.

Gall, Stanley A., et al., Multiple Pregnancy and Delivery (St. Louis: Mosby, 1996).

Goodman, David C., E.S. Fisher, G.A. Little, T.A. Stukel, C. Chang, and K.S. Schoendorf, "The Relations between the Availability of Neonatal Intensive Care and Neonatal Mortality," New England Journal of Medicine, 346 (2002), 1538-1544.

Grossman, Michael, and Theodore J. Joyce, "Unobservables, Pregnancy Resolutions, and Birth Weight Production Functions in New York City," Journal of Political Economy, XCVIII (1990), 983-1007.

Hollier, Lisa M., et al., "Outcome of Twin Pregnancies According to Intrapair Birth Weight Differences," Obstetrics and Gynecology, XCIV (1999), 1006-1010.

Ijzerman, Richard G., et al., "Evidence for Genetic Factors Explaining the Birth Weight-Blood Pressure Relation: Analysis in Twins," Hypertension, XXXVI (2000), 1008-1012.

Institute of Medicine, Preventing Low Birthweight (Washington, D.C.: National Academy Press, 1985).

Innes, Kim E., et al., "Association of a Woman's Own Birth Weight with Subsequent Risk for Gestational Diabetes," Journal of the American Medical Association, 287 (2002), 2534-2541.

Iyasu, S., J.E. Becerra, D.L. Rowley, C.J.R. Hogue, "Impact of Very Low Birth Weight on the Black White Infant Mortality Gap," American Journal of Preventive Medicine, 8 (1992), 271-277.

Journal of the American Medical Association, "Healthier Mothers and Babies – 1900-1999," 282 (1999), 1807-1810.

Joyce, Theodore J., Michael Grossman, and Fred Goldman, "An Assessment of the Benefits of Air Pollution Control: The Case of Infant Health," Journal of Urban Economics, XXV (1989), 32-51.

Kaelber, Charles T., and Thomas F. Pugh, "Influence of Intrauterine Relations on the Intelligence of Twins," New England Journal of Medicine, CCLXXX (1969), 1030-1034.

Kogan, M.D., G.R. Alexander, M. Kotelchuck, et al., "Trends in Twin Birth Outcomes and Prenatal Care Utilization in the United States, 1981-1997," Journal of the American Medical Association, 284 (2000), 335-341.

Kotelchuck, M., J.B. Schwartz, M.T. Anderka and K.S. Finison, "WIC Participation and Pregnancy Outcomes: Massachusetts Statewide Evaluation Project," American Journal of Public Health, 74 (1984), 1086-1092.

Kowaleski-Jones, Lori and Greg J. Duncan, "Effects of Participation in the WIC Program on Birthweight: Evidence from the National Longitudinal Survey of Youth," American Journal of Public Health, 92 (2002), 799-804.

Kramer, Michael S., "Intrauterine Growth and Gestational Duration Determinants," Pediatrics, 80 (1987), 502-511.

Kramer, Michael S., "Invited Commentary: Association between Restricted Fetal Growth and Adult Chronic Disease: Is It Causal? Is It Important?" American Journal of Epidemiology, 152 (2000), 605-608.

Leon, D.A., M. Johansson, and F. Rasmussen, "Gestational Age and Growth Rate of Fetal Mass Are Inversely Associated with Systolic Blood Pressure in Young Adults," American Journal of Epidemiology, 152 (2000), 597-604.

Lewit, Eugene M., "The Direct Cost of Low Birth Weight," The Future of Children, V (1995), 35-56.

Li, Kai and Dale J. Poirier, "The Roles of Birth Inputs and Outputs in Predicting Health, Behavior, and Test Scores in Early Childhood," mimeo, UC-Irvine, 2002.

Lightwood, James M., Ciaran S. Phibbs, and Stanton A. Glantz, "Short-term Health and Economic Benefits of Smoking Cessation: Low Birth Weight," Pediatrics, 104 (1999), 1312-1320.

Lucas, A., R. Morley, and T.J. Cole, "Randomised Trial of Early Diet in Preterm Babies and Later Intelligence Quotient," British Medical Journal, 317 (1998), 1481-1487.

Matte, T.D., M. Bresnahan, M.D. Begg, and E. Susser, "Influence of Variation in Birth Weight within Normal Range and within Sibships on IQ at Age 7 Years: Cohort Study," British Medical Journal, 323 (2001), 310-314

McCarthy, BJ, et al., "The Epidemiology of Neonatal Death in Twins," American Journal of Obstetrics and Gynecology, CXLI (1981), 252-256.

McCormick, MC, "The Contribution of Low Birth Weight to Infant Mortality and Childhood Morbidity," New England Journal of Medicine, 312 (1985), 82-90.

McCormick, MC, J Brooks-Gunn, et al., "The Health and Developmental Status of Very Low-Birth weight Children at School Age," Journal of the American Medical Association, 267 (1992), 2204-2208.

Miller, Douglas, "What Underlies the Black-White Infant Mortality Gap? The Importance of Birthweight, Behavior, Environment, and Health Care," mimeo, University of California, Berkeley, 2002.

National Center for Health Statistics, State Definitions and Reporting Requirements for Live Births, Fetal Deaths, and Induced Terminations of Pregnancy (Hyattsville, MD, 1981).

National Institute for Child and Human Development, Health Disparities: Bridging the Gap (Bethesda, MD, 2000).

National Institutes of Health, "NIH Guide: Low Birth Weight in Minority Populations," PA-99-045, 1999.

Nelson, K.B. and J.K. Grether, "Cerebral Palsy in Low-Birthweight Infants: Etiology and Strategies for Prevention," Mental Retardation and Developmental Disabilities Research Reviews, 3 (1997), 112-117.

Paneth, Nigel S., "The Problem of Low Birth Weight," The Future of Children, V (1995), 19-34.

Permutt, T. and J.R. Hebel, "Simultaneous-Equation Estimation in a Clinical Trial of the Effect of Smoking on Birth Weight," Biometrics, 45 (1989), 619-622.

Phillips, David I. W., et al., "Fetal Growth and the Fetal Origins Hypothesis in Twins – Problems and Perspectives," Twin Research, IV (2001), 327-331.

Pollack, Harold, Paula M. Lantz, and John G. Frohna, "Maternal Smoking and Adverse Birth Outcomes Among Singletons and Twins," American Journal of Public Health, 90 (2000), 395-400.

Rasmussen, Kathleen M., "The 'Fetal Origins' Hypothesis: Challenges and Opportunities for Maternal and Child Nutrition," Annual Review of Nutrition, 21 (2001), 73-95.

Rees J.M., S.A. Lederman, and J.L. Kiely, "Birth Weight Associated with Lowest Neonatal Mortality: Infants of Adolescent and Adult Mothers," Pediatrics, XCVIII (1996), 1161-1166.

Richards, Marcus, et al., "Birth Weight and Cognitive Function in the British 1946 Birth Cohort: Longitudinal Population Based Study," British Medical Journal, CCCXXII (2001), 199-203.

Rosenzweig, Mark R., and T. Paul Schultz, "Estimating a Household Production Function: Heterogeneity and the Demand for Health Inputs, and Their Effects on Birthweight," Journal of Political Economy, XCI (1983), 723-746.

Rosenzweig, Mark. R., and Kenneth I. Wolpin, "Inequality at Birth: The Scope for Policy Intervention," Journal of Econometrics, L (1991), 205-228.

_____, "Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions," American Economic Review, 76 (1986), 470-482.

Schoendorf, Kenneth C., C. J. Hogue, J.C. Klein, and D. Rowley, "Mortality among Infants of Black as Compared with White College-Educated Parents," New England Journal of Medicine, 326 (1992), 1522-1526.

Schultz, T. Paul, "Wage Gains Associated with Height as a Form of Health Human Capital," Center Discussion Paper No. 841, Yale University, 2002.

Sexton, Mary, and J. Richard Hebel, "A Clinical Trial of Change in Maternal Smoking and Its Effect on Birth Weight," Journal of the American Medical Association, CCLI (1984), 911-915.

Stein, Z., M. Susser, G. Saenger, and F. Marolla, Famine and Human Development: The Dutch Hunger Winter of 1944-45 (New York, NY: Oxford University Press, 1975).

Susser, Mervyn and Bruce Levin, "Ordeals for the Fetal Programming Hypothesis," British Medical Journal, 318 (1999), 885-886.

Thompson, Lindsay A., David C. Goodman, and George A. Little, "Is More Neonatal Intensive Care Always Better? Insights from a Cross-National Comparison of Reproductive Care, Pediatrics, 109 (2002), 1036-1043.

Torelli, Paul A., "Maternal Smoking and Birth Outcomes: A Nonparametric Approach," mimeo, University of California, Berkeley, 2000.

U.S. Congress, Office of Technology Assessment, Healthy Children: Investing in the Future, OTA-H-345 (Washington D.C.: U.S. Government Printing Office, 1988).

U.S. Department of Health and Human Services, Racial and Ethnic Disparities in Infant Mortality (Washington D.C.: Office Of Minority Health, 2000).

Vagero, D., and D. Leon, "Ischemic Heart Disease and Low Birth Weight: A Test of the Fetal-Origins Hypothesis from the Swedish Twin Registry," Lancet, 343 (1994), 260-263.

Vogler, George P., and Lynn T. Kozlowski, "Differential Influence of Maternal Smoking on Infant Birth Weight," Journal of the American Medical Association, 287 (2002), 241-242.

Williams, Sheila, and Richie Poulton, "Twins and Maternal Smoking: Ordeals for the Fetal Origins Hypothesis? A Cohort Study," British Medical Journal, 318 (1999), 897-900.

Winn, Hung N., and W. R. Gerber, "Multiple Gestation," in Hung N. Winn and John C. Hobbins, eds., Clinical Maternal-Fetal Medicine (London: The Parthenon Publishing Group, 2000).

Women's Health Initiative Investigators, "Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women," Journal of the American Medical Association, 288 (2002), 321-333.

Zhang, Jun, Ruth A. Brenner, and Mark A. Klebanoff, "Differences in Birth Weight and Blood Pressure at Age 7 Years among Twins," American Journal of Epidemiology, 153 (2001), 779-782.

Data Appendix

Matching twins in linked birth/infant death data

Two approaches were used to match twins in the annual linked birth/infant death data. The primary approach uses the fact that beginning in 1989, multiple-birth records are generally located next to each other in the "denominator-plus" microdata file provided by NCHS. Additionaly, with the rich set of covariates included on the micordata file, twin pairs can be matched when two unique sets of parental and pregnancy covariates are identified. This second approach is necessary for births occurring before 1989, when multiple births were not located next to each other in the "denomnator-plus" file.

The record location-based matching approach begins by selecting all the the "higher-order" pregnancies identified with the DPLURAL variable. For 1989, the DPLURAL variable indicates that 90,222 of the 4,045,881 births (2.2 percent) were twin births. The repetition of paternal and pregnancy characteristics indicate that adjacent twin records are part of the same twin set. Adjacent twin records are considered matched twins if the following covariates are each identical: mother's age, state of birth, state of residence, race, and education, marital status of the parents, father's age, father's education[41], and the month in which prenatal care began. Of the 81,757 adjacent twin records that have both information on maternal education and birth weight, 74,472 (91%) have identical information for these nine covariates. Additional information on these files not used in matching suggests that incorrect twin matches are uncommon: of the 74,742 matched twin pairs for 1989, all but 37 of the pairs were born in the same month. While a longer string of covariates could be used to increase the likelihood that adjacent twin records are indeed part of the same twin set, incomplete or miscoding of the second twin record can cause the incorrect rejection of true twin matches.

This risk of Type 1 error needs to be balanced against the risk generated by relying more on the adjacency of twin records and less on the correspondence of parental and pregnancy covariates. Relying more on adjacency increases the risk of Type 2 error for several reasons. First, there appear to be singleton birth records that were coded as twins. (This could possibly result from the fact that the plurality field, DPLURAL, should be coded as "2" for twins, whereas "2" is often used as a code for "no"). It is also possible that one twin in the pair was not born alive, and therefore will appear in the fetal death data rather than in the "denominator-plus" natality data. Finally, it could be the case that while both twin birth records are present in the natality data, they are not adjacent in the "denominator-plus" file. All of these possibilities increase the risk of incorrectly matching adjacent twin records when a more sparse set of covariate checks is used.

Before 1989, multiple birth records are not located together in the "denominator-plus" files. Therefore, an alternative matching algorithm is needed. A long string of parental and pregnancy covariates can instead be used to search across the entire set of twin records (i.e. where DPLURAL = 2) for a unique match. A string of eleven covariates was created that attempts to balance the risk of incorrectly matching twin pairs from different mothers (by including fewer covariates) against the risk of not matching correct twin pairs where there is incomplete or incorrect information for one of the twin pairs (by including more covariates in the matching). The eleven covariates used to match twins are: maternal education, maternal age, maternal race, marital status, mother's state of birth, state of birth occurrence, county of birth occurrence, father's age, father's education, month of pregnancy prenatal care began, and the number of prenatal care visits. Twins are considered matched when there are exactly two records with DPLURAL = 2 and with equal values for this set of eleven covariates. For 1989, 74,886 twin pairs are matched,

[41] Note for births occurring in 1995, father's education is no longer collected and therefore this covariate is dropped from the matching algorithm.

which is nearly identical to number of pairs matched with the location-based method (74,472). Moreover, the econometric results using these two different matching techniques are nearly identical.

The matching procedures described above were necessary because NCHS does not include information on the set to which a multiple birth belongs. However, NCHS has recently created the Matched Twins File for twin births in 1995. This file is scheduled to be made publicly available in the Fall of 2002. Because NCHS has access to additional identifying information which is suppressed on the public-use linked files used here, their matching algorithm can presumably be more accurate. However, it should be noted that incorrect matching of twin births would generally generate a bias against the results found in this paper.

Finally, the birth order of the matched twins is not explicitly noted in the linked birth/infant death data. As pronounced health differences exist between first- and second-born twins, birth order should be allowed for in the econometric estimation. On the "denominator-plus" file, the DTOTORD variable gives the number of children born to the mother at the time of birth as well as the number of terminations of pregnancy. In most cases, DTOTORD differs by 1 for matched twin pairs. In these cases, the matched twin with the lower value for DTOTORD is considered the first-born twin.

Figure 1A: Infant Mortality Rate (1-year) and Birthweight

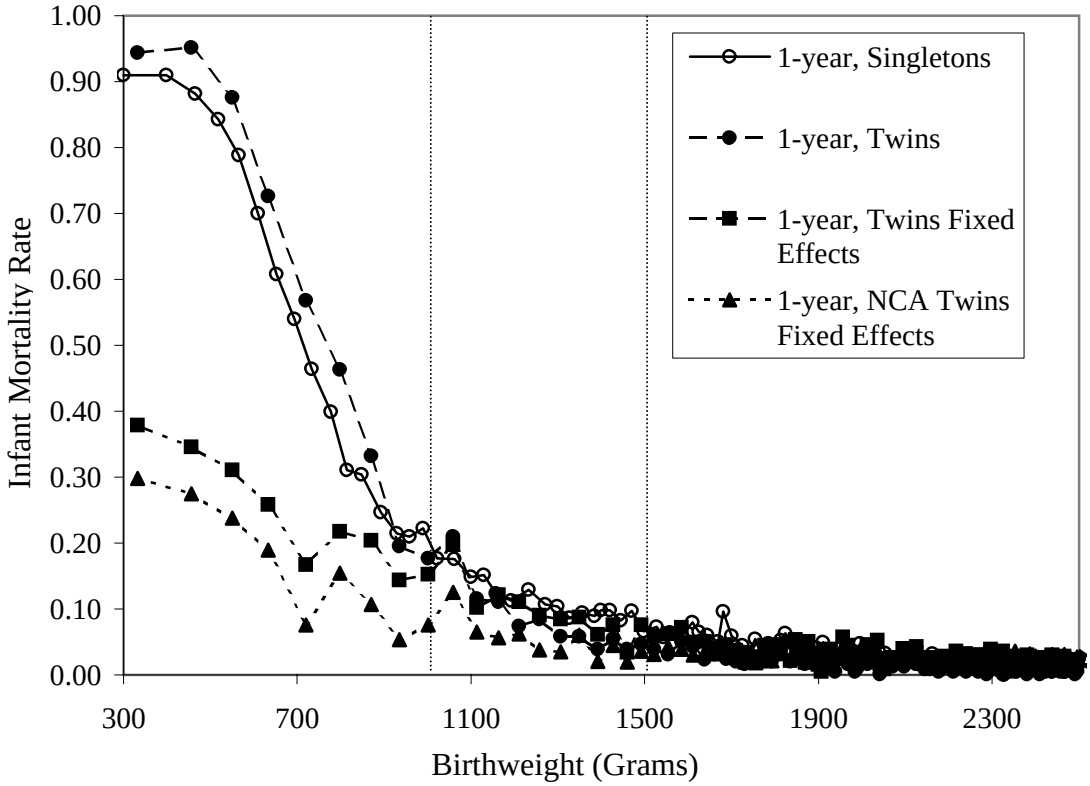


Figure 1B: Infant Mortality Rate (1-day) and Birthweight

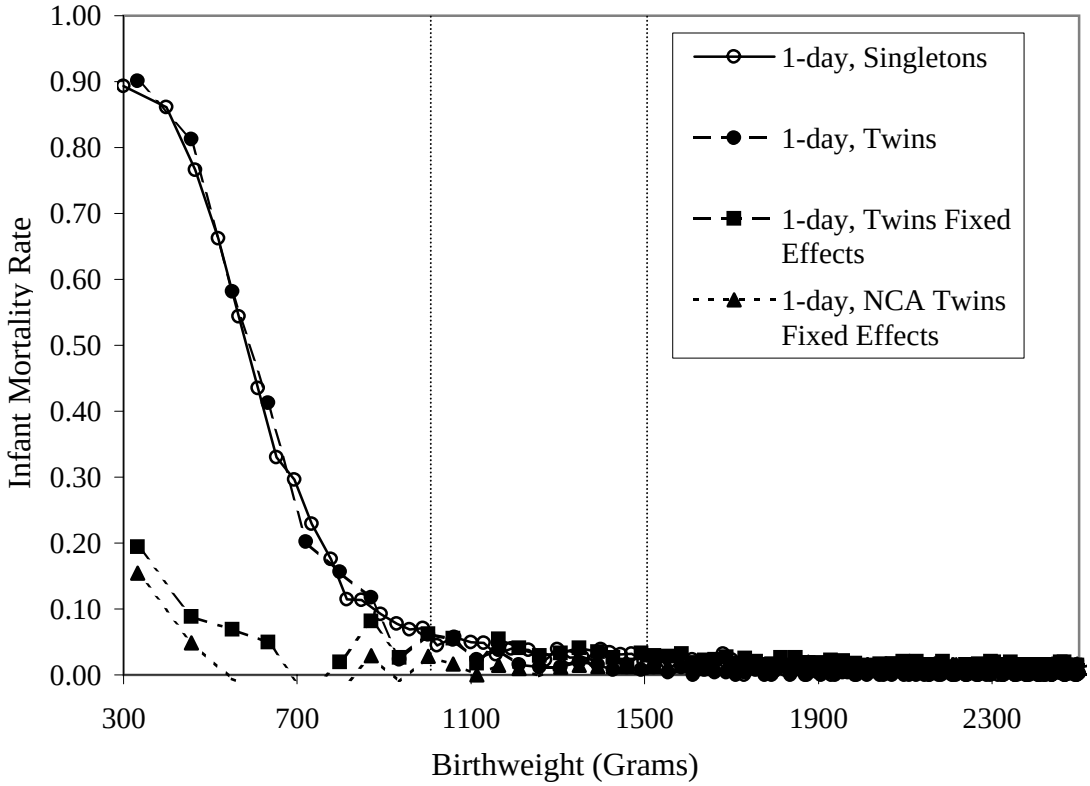


Figure 2A: Infant Mortality and Birthweight, Twins Pooled

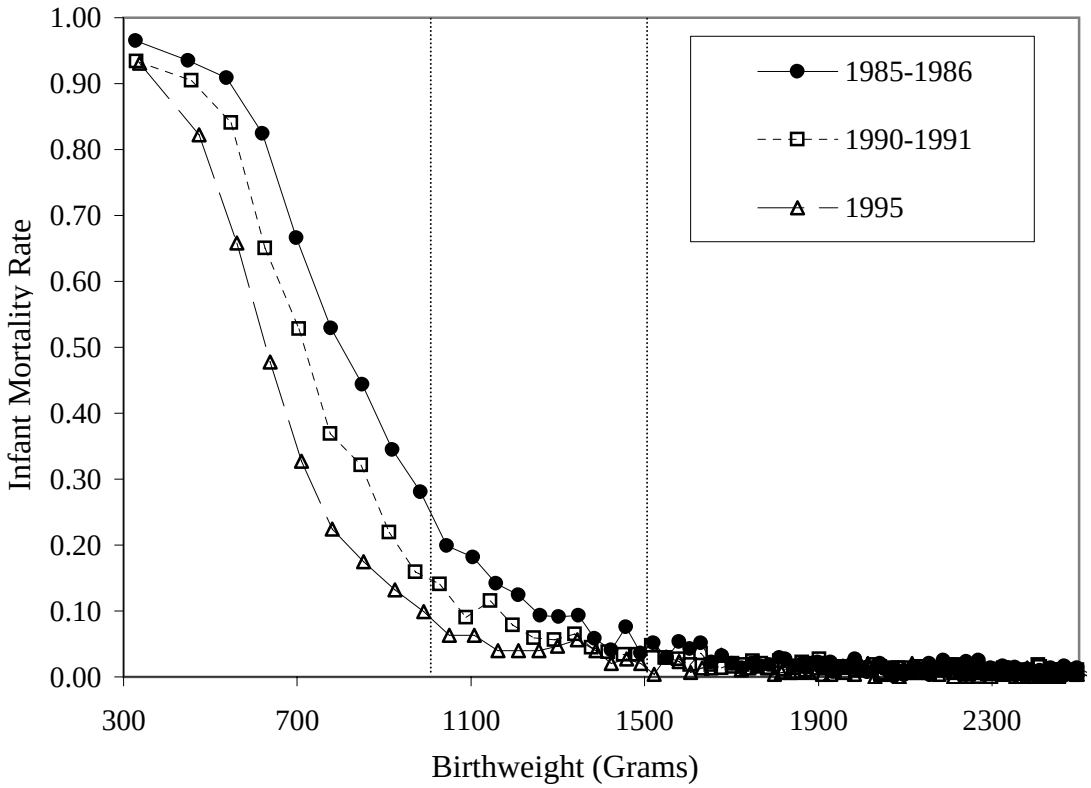


Figure 2B: Infant Mortality and Birthweight, Twins Fixed Effects

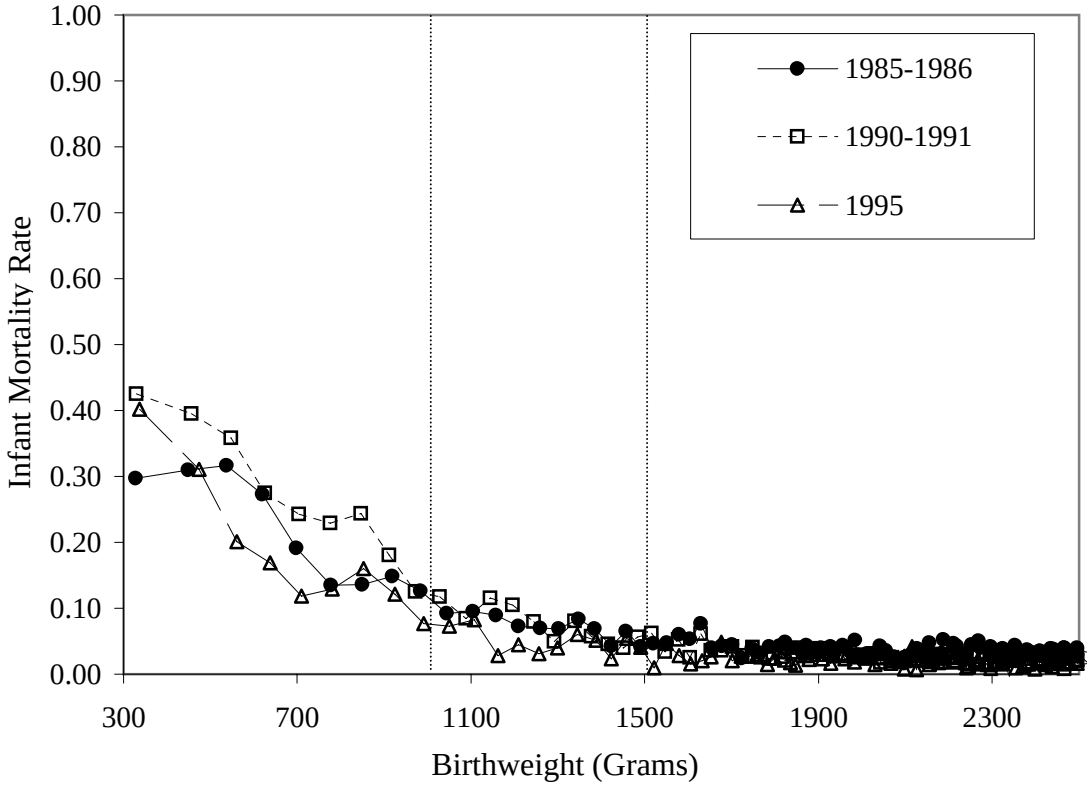


Figure 3: 5-minute APGAR Scores and Birthweight

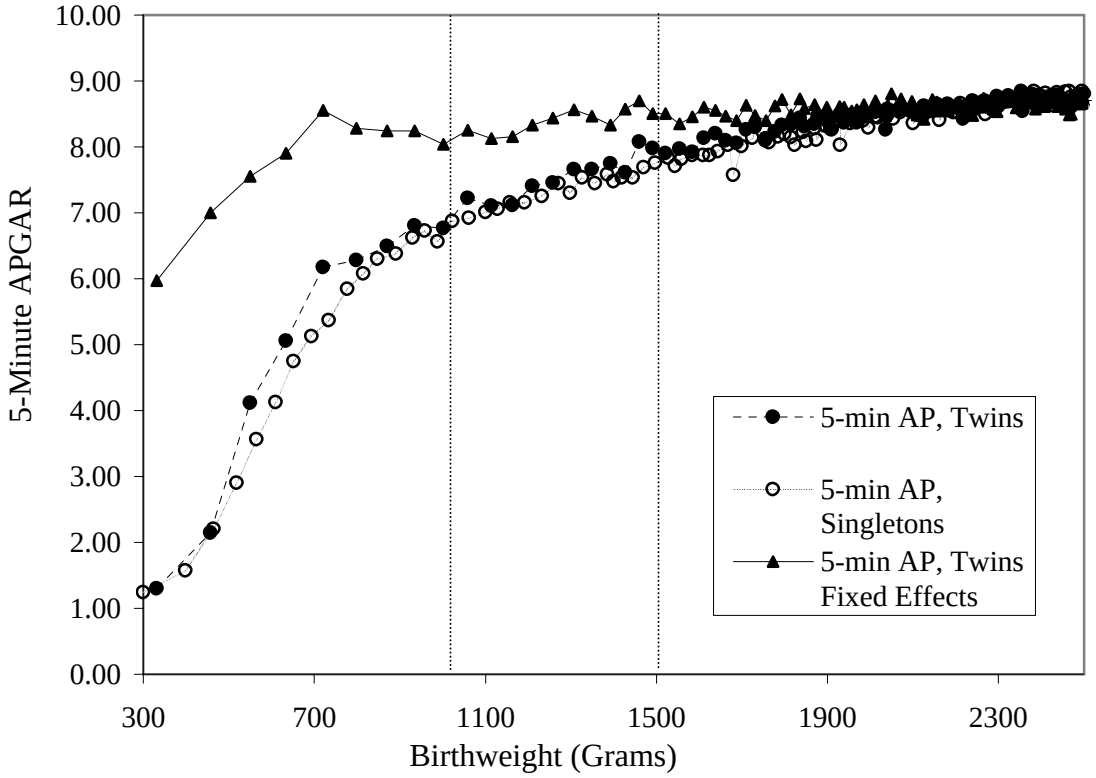


Figure 4A: Infant Mortality Rate (1-year) and APGAR

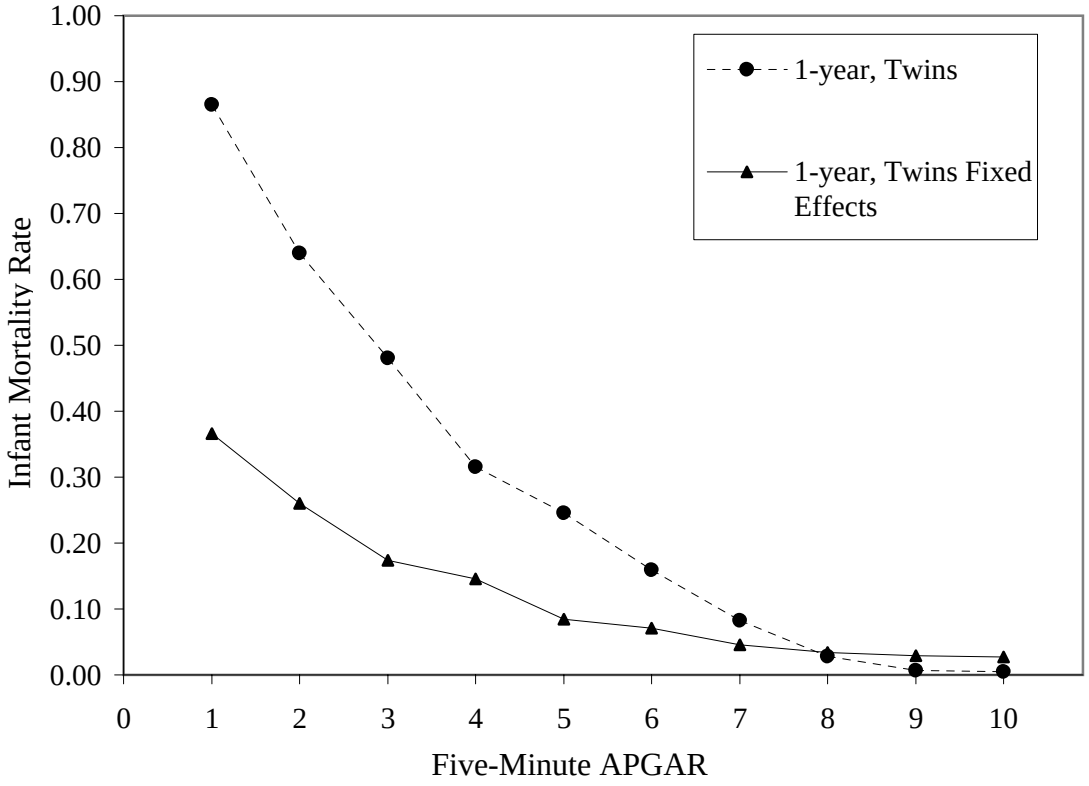


Figure 4B: Infant Mortality Rate (1-day) and APGAR

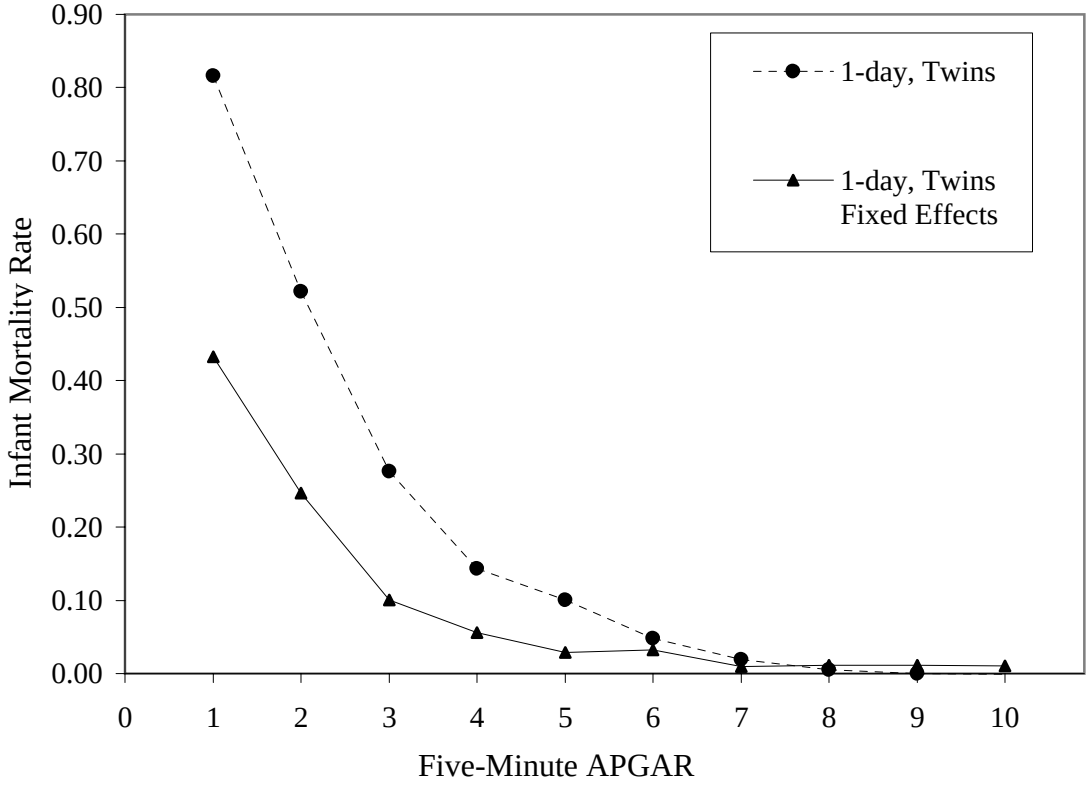
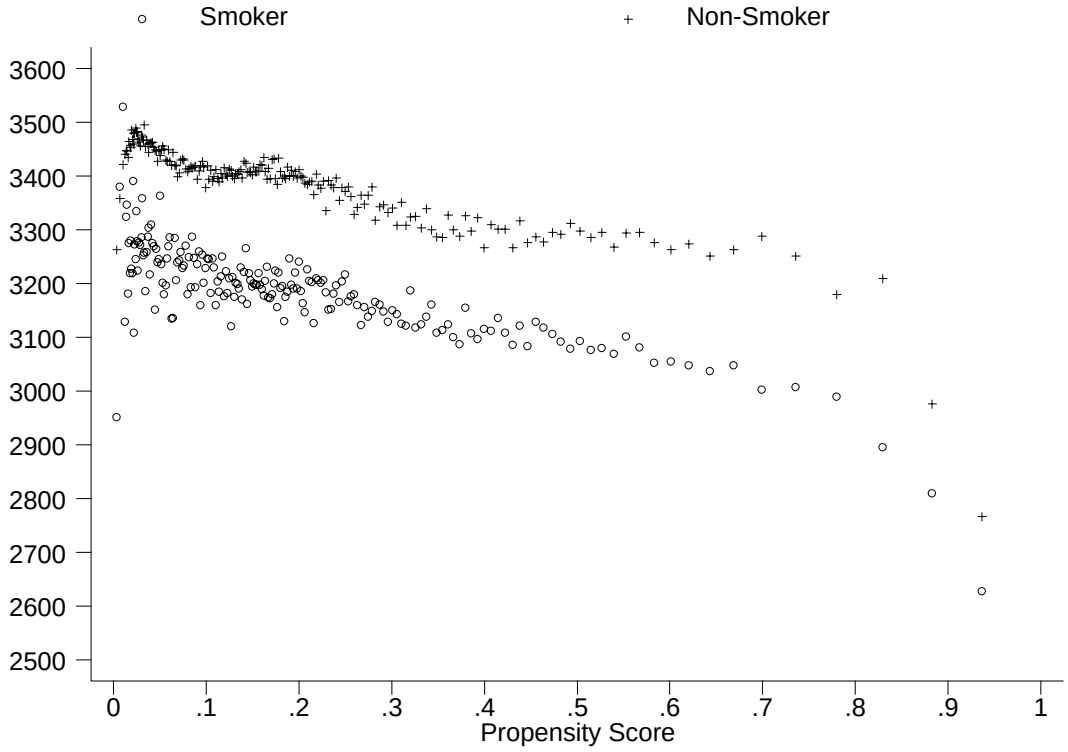
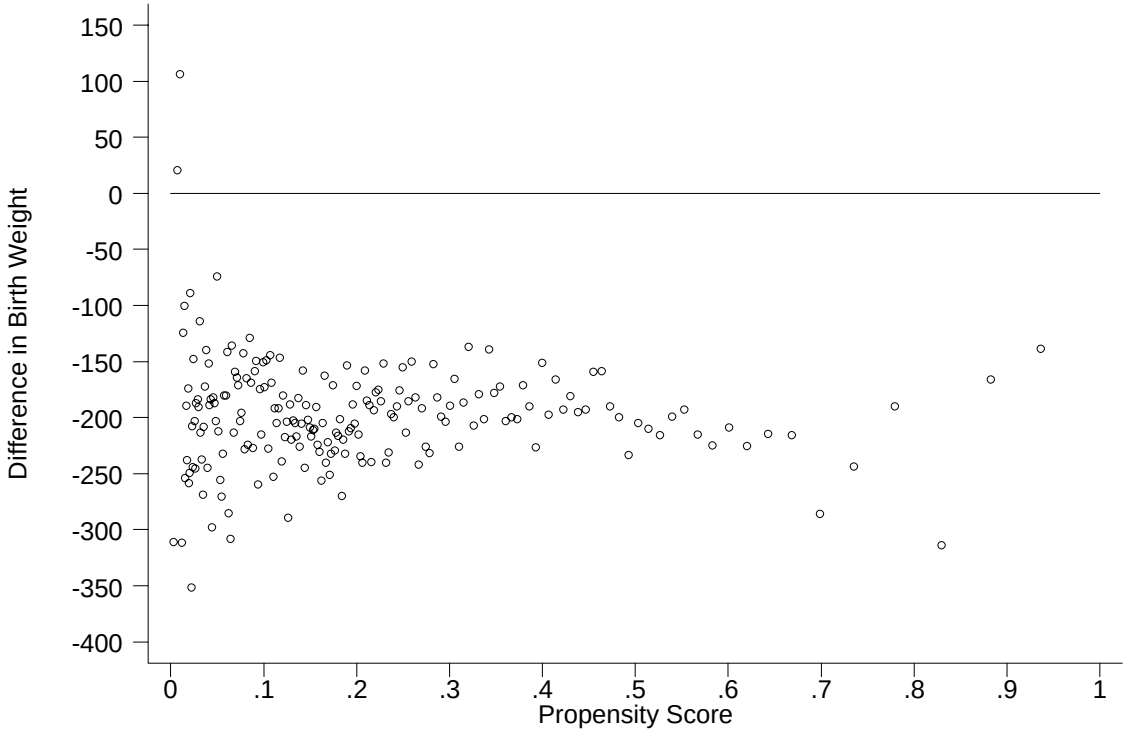


Figure 5: Estimated Birth Weight Effects of Maternal Smoking by the Propensity Score

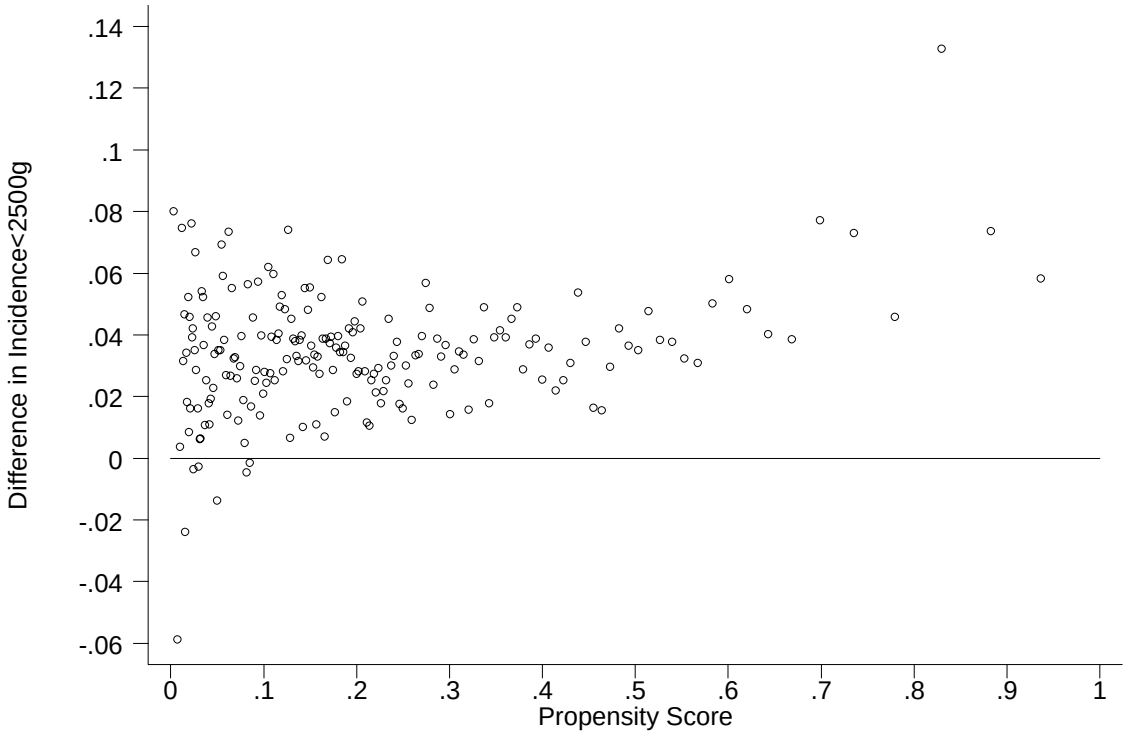
A. Average Birth Weight of Smoking and Non-Smoking Infants by the Propensity Score



B. Difference in Average Birth Weight of Smoking and Non-Smoking Infants by the Propensity Score



C. Difference between Smoking and Non-Smoking Infants in Low Birth Weight (<2500 grams) Incidence



D. Difference between Smoking and Non-Smoking Infants in Incidence less than 1000 grams

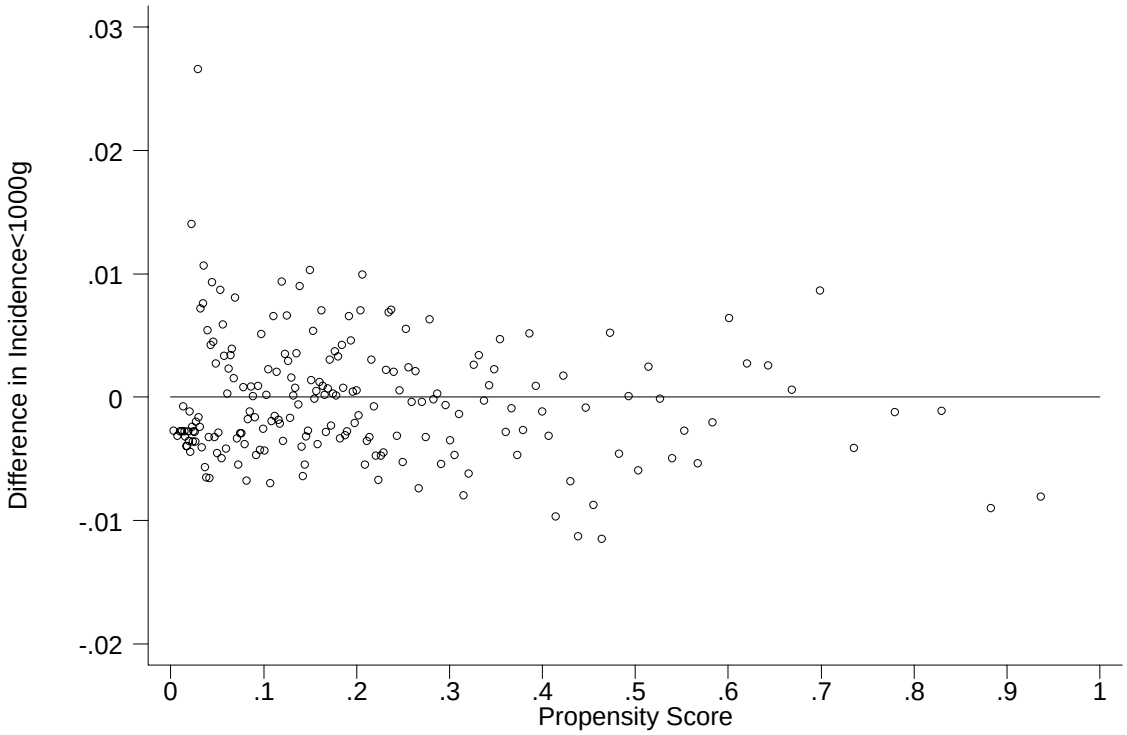
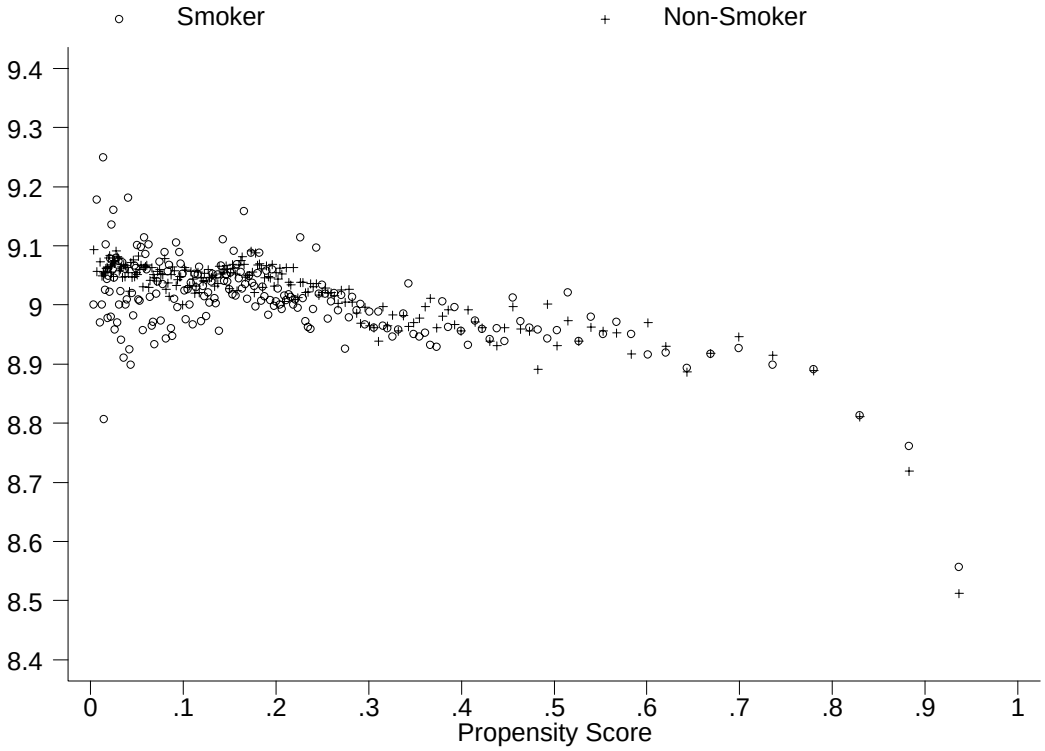


Figure 6: Estimated Effects of Maternal Smoking on the 5-Minute APGAR Score

A. Average 5-Minute APGAR Score of Smoking and Non-Smoking Infants by the Propensity Score



B. Difference in 5-Minute APGAR Score of Smoking and Non-Smoking Infants by the Propensity Score

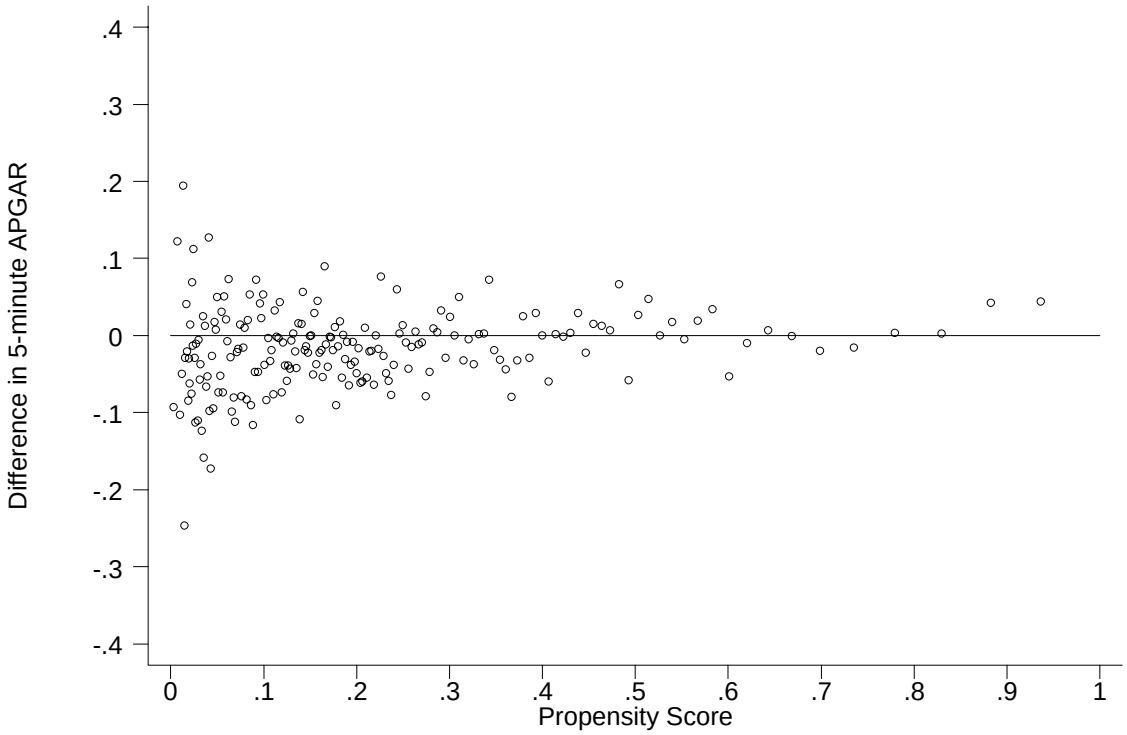
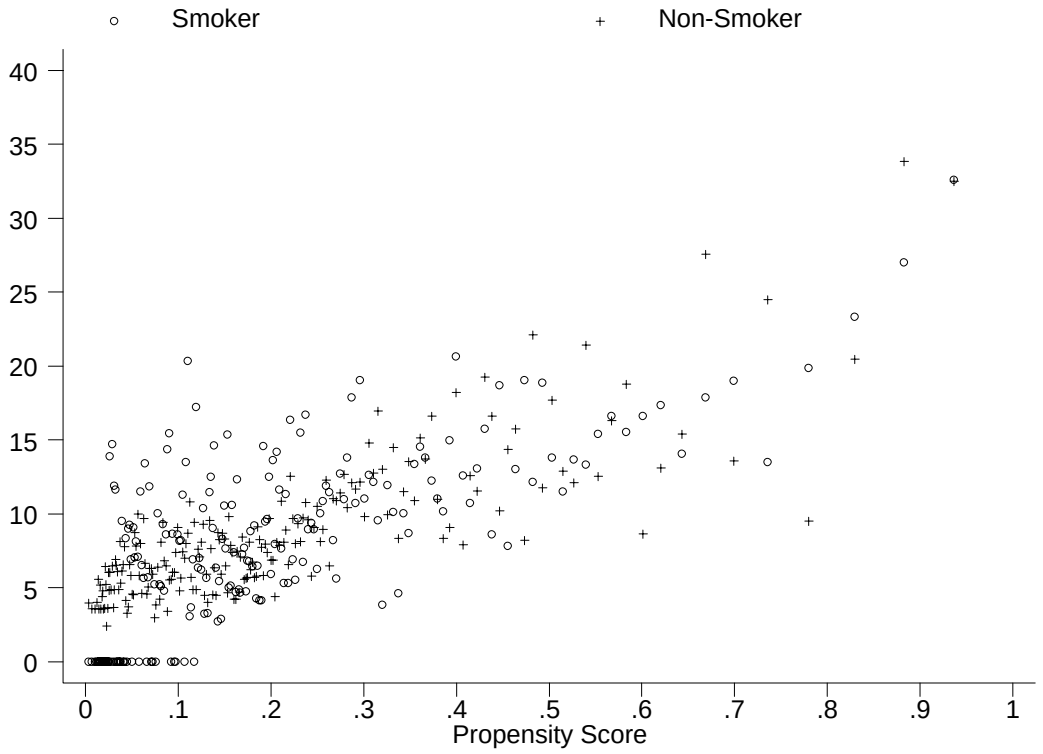
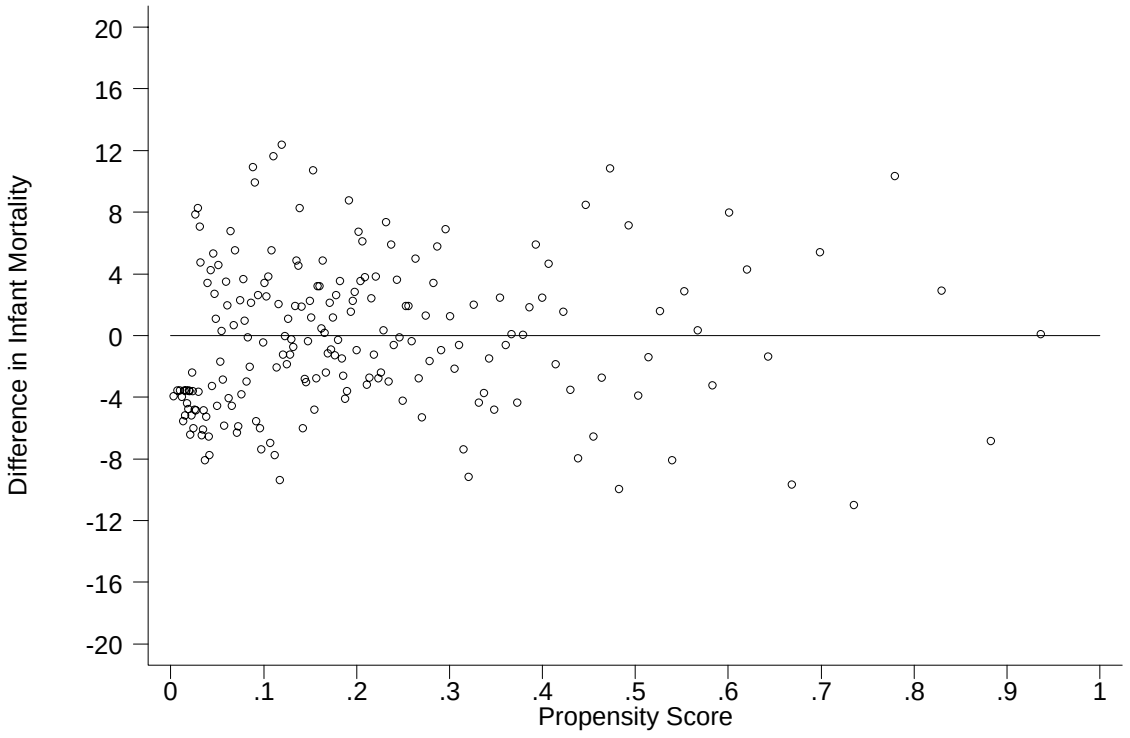


Figure 7: Estimated Effects of Maternal Smoking on Infant Mortality (per 1,000 live births)

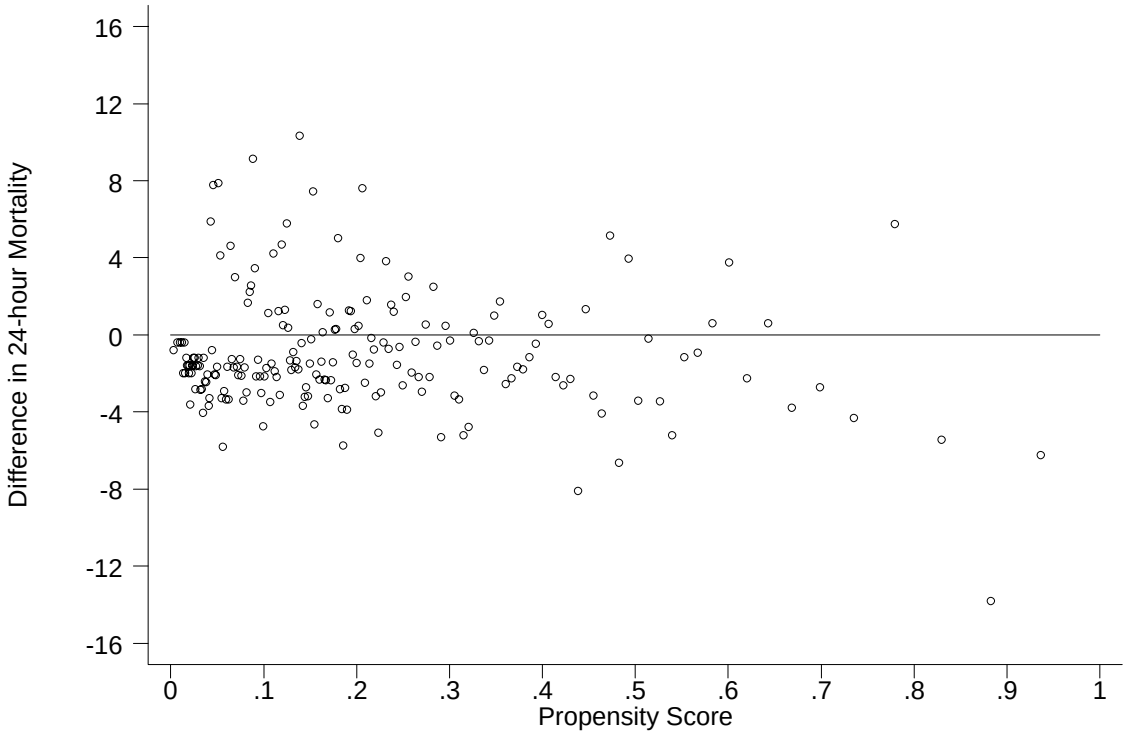
A. Smoking and Non-Smoking Infant Deaths within 1-Year of Birth by the Propensity Score



B. Difference in Smoking and Non-Smoking Infant Deaths within 1-Year of Birth



C. Difference in Smoking and Non-Smoking Infant Deaths within 24 Hours of Birth



D. Difference in Smoking and Non-Smoking Infant Deaths within 28 Days of Birth

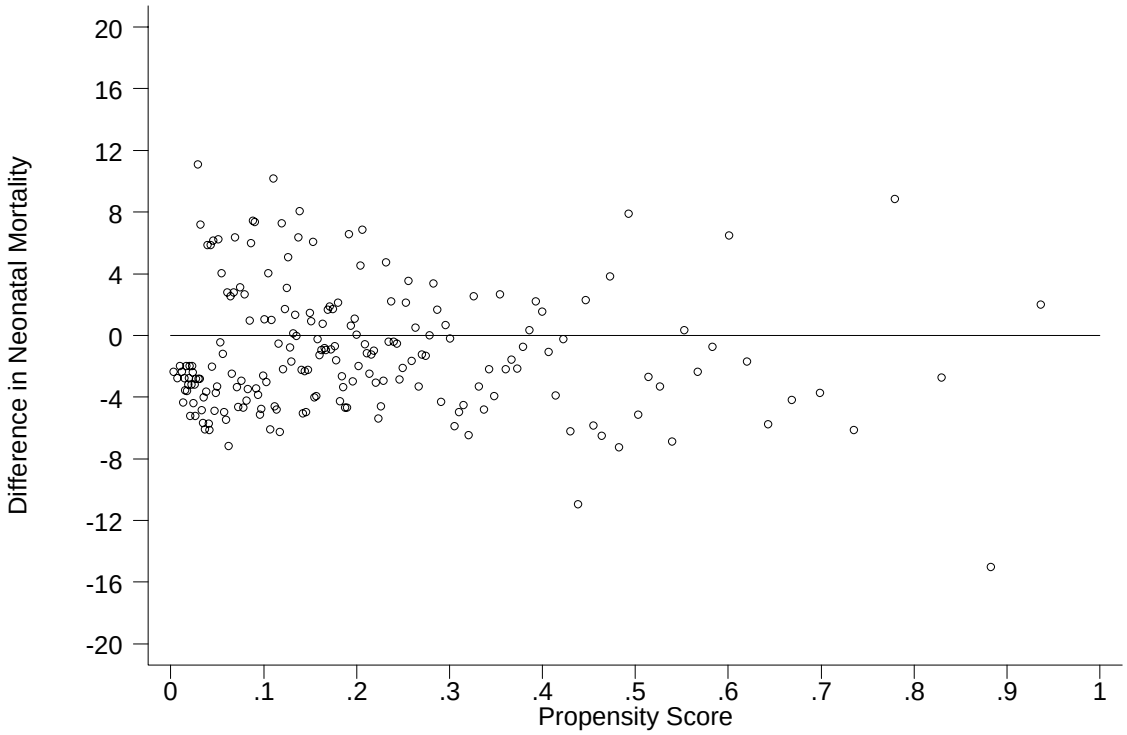


Table 1: Sample Means for Singletons and Twins from 1989 Linked Natality-Mortality Detail File
(standard deviation in parentheses)

	Sample Means		T-ratio for Diff
	Singletons	Twins	In Sample Means
Mother's Demographic Characteristics			
Fraction Black	0.195	0.219	14.1
Education	12.8	13.0	22.2
Fraction High School Dropout	0.184	0.157	17.3
Fraction College Graduate	0.187	0.212	15.2
Age	26.3	27.4	51.3
Fraction Teenager	0.129	0.078	45.8
Fraction 30 or Older	0.289	0.360	35.5
Fraction Married	0.736	0.746	5.1
Mother's Risk Factors			
Fraction Prenatal Care	0.982	0.982	0.4
Number of Prenatal Visits	11.2	12.4	49.7
Pregnancy Associated Hypertension	0.030	0.067	35.1
Anemia of Mother	0.018	0.035	21.0
Fraction Smoke during Pregnancy	0.212	0.201	5.4
Characteristics of Birth			
Fraction Male	0.512	0.504	4.1
Breech Birth	0.035	0.256	119.5
Abnormal Conditions of Newborn	0.055	0.148	62.9
Assisted Ventilation (<30 minutes)	0.011	0.024	20.4
Assisted Ventilation (>=30 minutes)	0.006	0.038	40.1
Congenital Anomaly	0.019	0.027	12.4
Sample Size	2,655,977	58,132	

Table 1 (cont'd)

	Sample Means		T-ratio for Diff
	Singletons	Twins	In Sample Means
Infant Birth Weight (grams)			
Mean	3,369	2,417	339.6
	(591)	(670)	
Median	3,402	2,495	
25th percentile	3,060	2,070	
10th percentile	2,693	1,503	
5th percentile	2,410	1,080	
1st percentile	1,430	503	
Fraction Low birth weight (<2500g)	0.061	0.504	
Gestation in weeks	39.3	36.0	201.2
	(2.65)	(3.85)	
1-minute APGAR score (1-10)	8.02	7.30	79.9
	(1.35)	(1.97)	
5-minute APGAR score (1-10)	8.97	8.52	70.2
	(0.81)	(1.40)	
Infant Deaths (per 1,000 live births)			
Within 1-year of birth (infant mortality)	8.46	38.71	37.7
Within 24 hours of birth	2.73	19.27	29.0
Within 7 days	3.99	26.92	34.1
Within 28 days (neonatal)	4.99	30.62	35.8
28 days to 1 year (postneonatal)	3.49	8.19	12.5
Fraction of dead with birth weight<2500 g			
Infant mortality	0.57	0.93	
Within 24-hour mortality	0.89	0.98	
Neonatal mortality	0.76	0.97	
Postneonatal mortality	0.30	0.79	
Sample Size	2,655,977	58,132	

Notes: Data come from the National Center of Health Statistics 1989 Linked Birth-Infant Death Detail File. Sample contains non-Hispanic, black and white mothers born in the United States.

Table 2: Sample Means for Twin Pairs in 1989

	Heavier versus Lighter Twin			First versus Second Born Twin		
	Sample Means		T-ratio for	Sample Means		T-ratio for
	Heavy	Light	Diff	First	Second	Diff
Birth Weight	2,572	2,270	54.7	2,431	2,398	5.76
Male	0.546	0.461	20.0	0.507	0.502	1.09
Breech Birth	0.253	0.260	1.83	0.218	0.297	20.3
Abnormal Conditions	0.145	0.149	1.41	0.142	0.156	4.53
Ventilation (<30 mins)	0.024	0.025	1.21	0.022	0.027	3.87
Ventilation (>=30 mins)	0.038	0.038	0.27	0.036	0.041	3.36
Congenital Anomaly	0.025	0.029	2.72	0.027	0.026	0.73
1-minute APGAR	7.37	7.26	6.12	7.52	7.07	24.1
5-minute APGAR	8.55	8.51	2.98	8.58	8.46	9.21
Infant Deaths						
Within 1-year of birth	34.4	40.3	3.70	36.9	40.0	1.86
Within 24 hours	17.1	19.4	2.00	18.2	18.7	0.42
Within 7 days	23.7	27.4	2.77	25.1	27.6	1.74
Within 28 days	27.2	31.4	2.97	29.1	31.3	1.48
28 days to 1 year	7.3	9.0	2.22	7.9	8.8	1.15
Sample Size	27,691	27,691		26,243	26,243	

Notes: See notes to Table 1.

Table 3: Pooled Twins Estimates of Association between Birth Weight and Infant Mortality Rates
(estimated standard errors in parentheses)

	Infant Deaths within 1-year of birth (per 1,000 live births)					
	(1)	(2)	(3)	(4)	(5)	(6)
Birth Weight	-0.121	-0.122	-0.119	-0.058	-0.056	-0.046
	(0.0030)	(0.0030)	(0.0030)	(0.0025)	(0.0025)	(0.0025)
Gestation Length				-14.70		
				(0.55)		
20-27 weeks					-296.7	
					(48.6)	
28-31 weeks					-738.6	
					(47.4)	
32-35 weeks					-766.4	
					(47.3)	
36 weeks					-755.1	
					(47.3)	
37-39 weeks					-743.7	
					(47.4)	
40 weeks					-732.0	
					(47.5)	
41 weeks					-739.7	
					(47.5)	
42 or more weeks					-735.7	
					(47.6)	
R-squared	0.18	0.18	0.18	0.22	0.36	0.42
Sample Size	58,132	58,132	57,156	56,836	56,836	56,836
Gender	N	Y	Y	Y	Y	Y
Mother's variables	N	Y	Y	Y	Y	Y
Prenatal care	N	N	Y	Y	Y	Y
Flexible form	N	N	Y	Y	Y	Y
Gestation Fixed Effects	N	N	N	N	N	Y

Notes: Linear probability regressions based on pooled twins data for 1989. Standard errors are corrected for heteroskedasticity and family-level clustering in the residuals. Mother's characteristics include race, education, age, and marital status. The flexible specification includes categories of the variables and interactions.

Table 4: Twins Pooled and Fixed Effects Estimates of Birth Weight Effect on Infant Mortality
(estimated standard errors in parentheses)

A. All Observations

	Infant Deaths (per 1,000 live births)									
	Within 1-Year		Within 24 Hours		Within 7 Days		Within 28 Days		Postneonatal Mortality	
	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.
Birth Weight	-0.121	-0.022	-0.075	-0.005	-0.100	-0.013	-0.109	-0.016	-0.012	-0.005
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)
Family Fixed Effects	N	Y	N	Y	N	Y	N	Y	N	Y
R-squared	0.18	0.80	0.13	0.86	0.17	0.86	0.18	0.84	0.01	0.52
Sample Size	58,132	58,132	58,132	58,132	58,132	58,132	58,132	58,132	58,132	58,132

B. Dropping Observations Born with or Dying from a Congenital Anomaly

	Infant Deaths (per 1,000 live births)									
	Within 1-Year		Within 24 Hours		Within 7 Days		Within 28 Days		Postneonatal Mortality	
	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.
Birth Weight	-0.109	-0.007	-0.067	0.002	-0.090	-0.003	-0.098	-0.004	-0.011	-0.003
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)	(0.001)	(0.001)	(0.002)
Family Fixed Effects	N	Y	N	Y	N	Y	N	Y	N	Y
R-squared	0.17	0.82	0.13	0.90	0.16	0.90	0.17	0.88	0.01	0.54
Sample Size	56,393	56,393	56,393	56,393	56,393	56,393	56,393	56,393	56,393	56,393

Table 5A: Infant Mortality (1-year) and Birthweight, Linear Specifications

	OLS	F.E.	Correlated R.E.		Logit	Cond. Logit
			1st born	2nd born		
	(1)	(2)	(3a)	(3b)	(4)	(5)
Birthweight (Own)	-0.120	-0.022	-0.078	-0.073	-0.319	-0.124
	(0.003)	(0.005)	(0.004)	(0.004)	(0.008)	(0.021)
Birthweight (Twin)	---	---	-0.048	-0.059	---	---
			(0.004)	(0.004)		
Firstborn	0.015	-0.020	---	---	-0.040	-0.123
	(0.012)	(0.017)			(0.051)	(0.081)
Constant	3.263	---	3.408	3.587	2.295	---
	(0.090)		(0.105)	(0.106)	(0.098)	
Family Effects	No	Yes	No	No	No	Yes
Test of equal effects (p-value)	---	---	0.0399	*	---	---
Test of symmetry (p-value)	---	---	0.0383	*	---	---
Number of Observations	43618	43618	21809	21809	43618	1328

Note: Standard errors (consistent with within-family correlation) in parentheses. Sample includes only twins with valid one-minute and five-minute apgar scores. Birth-weight enters linearly in all specifications. Birth-weight coefficients expressed as change in probability per 1000 grams. Column (1) reports least squares estimates, and Column (2) report family (mother) fixed effects estimates. Columns 3(a) and 3(b) report coefficients from separate regression for the first- and second-born twin, where mortality indicator is regressed on own and twin birth-weight. Column (4) and Column (5) coefficients are changes in log-odds per 100 grams.

Table 5B: Infant Mortality (1-year) and Birthweight, Piecewise Linear Spline Specifications

		OLS	F.E.	Correlated R.E.		Logit	Cond. Logit
				1st born	2nd born		
		(1)	(2)	(3a)	(3b)	(4)	(5)
Birthweight (Own)	1st seg.	-1.254	-0.324	-0.822	-0.812	-0.682	-0.165
		(0.027)	(0.153)	(0.086)	(0.087)	(0.294)	(0.074)
	2nd seg.	-0.152	-0.176	-0.168	-0.132	-0.354	-0.332
		(0.035)	(0.080)	(0.053)	(0.051)	(0.039)	(0.093)
	3rd seg.	-0.067	-0.043	-0.009	-0.112	-0.131	-0.083
		(0.034)	(0.071)	(0.047)	(0.047)	(0.015)	(0.024)
Birthweight (Twin)	1st seg.	---	---	-0.484	-0.511	---	---
				(0.085)	(0.089)		
	2nd seg.	---	---	0.050	-0.002	---	---
				(0.046)	(0.052)		
	3rd seg.	---	---	-0.069	0.035	---	---
				(0.048)	(0.047)		
Firstborn		-0.014	-0.019	---	---	-0.081	-0.118
		(0.012)	(0.017)			(0.065)	(0.081)
Constant		14.616	---	14.994	15.298	5.368	---
		(0.090)		(0.281)	(0.285)	(0.247)	
Family Effects		No	Yes	No	No	No	Yes
Test of equal effects (p-value)		---	---	0.7962	*	---	---
Test of symmetry (p-value)		---	---	0.3524	*	---	---
Number of Observations		43618	43618	21809	21809	43618	1328

Note: Standard errors (consistent with within-family correlation) in parentheses. Sample includes only twins with valid one-minute and five-minute apgar scores. Birth-weight enters as a twenty-segment piecewise linear spline, with knot points at birthweights corresponding to the 5th, 10th, 15th, etc. percentiles of the twins birthweight distribution. Coefficients on the slopes are reported only for the first three segements of the spline. Logit and conditional logit uses a three-segment piecewise linear spline, with knot points at birthweights corresponding to the 5th and 10th percentiles of the twins birthweight distribution. Birth-weight coefficients expressed as change in probability per 1000 grams. Column (1) reports least squares estimates, and Column (2) report family (mother) fixed effects estimates. Columns 3(a) and 3(b) report coefficients from separate regression for the first- and second-born twin, where mortality indicator is regressed on own and twin birth-weight. Column (4) and Column (5) coefficients are changes in log-odds per 100 grams.

Table 5C: Infant Mortality (1-day) and Birthweight, Linear Specifications

	OLS	F.E.	Correlated R.E.		Logit	Cond. Logit
			1st born	2nd born		
	(1)	(2)	(3a)	(3b)	(4)	(5)
Birthweight (Own)	-0.072	-0.005	-0.046	-0.038	-0.485	-0.051
	(0.003)	(0.004)	(0.003)	(0.003)	(0.027)	(0.026)
Birthweight (Twin)	---	---	-0.032	-0.041	---	---
			(0.003)	(0.003)		
Firstborn	0.023	-0.001	---	---	0.043	-0.026
	(0.007)	(0.010)			(0.069)	(0.135)
Constant	1.922	---	2.739	2.105	2.652	---
	(0.081)		(0.093)	(0.093)	(0.214)	
Family Effects	No	Yes	No	No	No	Yes
Test of equal effects (p-value)	---	---	0.6159	*	---	---
Test of symmetry (p-value)	---	---	0.1153	*	---	---
Number of Observations	43618	43618	21809	21809	43618	1328

Note: Standard errors (consistent with within-family correlation) in parentheses. Sample includes only twins with valid one-minute and five-minute apgar scores. Birth-weight enters linearly in all specifications. Birth-weight coefficients expressed as change in probability per 1000 grams. Column (1) reports least squares estimates, and Column (2) report family (mother) fixed effects estimates. Columns 3(a) and 3(b) report coefficients from separate regression for the first- and second-born twin, where mortality indicator is regressed on own and twin birth-weight. Column (4) and Column (5) coefficients are changes in log-odds per 100 grams.

Table 5D: Infant Mortality (1-day) and Birthweight, Piecewise Linear Spline Specifications

		OLS	F.E.	Correlated R.E.		Logit	Cond. Logit
				1st born	2nd born		
		(1)	(2)	(3a)	(3b)	(4)	(5)
Birthweight (Own)	1st seg.	-1.161	-0.027	-0.629	-0.651	-0.785	0.015
		(0.029)	(0.121)	(0.077)	(0.072)	(0.036)	(0.093)
	2nd seg.	0.208	-0.037	0.103	0.157	-0.403	-0.134
		(0.022)	(0.053)	(0.037)	(0.031)	(0.114)	(0.126)
	3rd seg.	-0.127	-0.027	-0.088	-0.138	-0.119	-0.046
		(0.018)	(0.036)	(0.026)	(0.026)	(0.058)	(0.038)
Birthweight (Twin)	1st seg.	---	---	-0.592	-0.645	---	---
				(0.077)	(0.074)		
	2nd seg.	---	---	0.182	0.161	---	---
				(0.034)	(0.031)		
	3rd seg.	---	---	-0.107	-0.067	---	---
				(0.025)	(0.021)		
Firstborn		0.002	0.000	---	---	0.039	-0.036
		(0.007)	(0.010)			(0.090)	(0.136)
Constant		12.032	---	12.483	13.139	4.655	---
		(0.286)		(0.328)	(0.320)	(0.254)	
Family Effects		No	Yes	No	No	No	Yes
Test of equal effects (p-value)		---	---	0.1092	*	---	---
Test of symmetry (p-value)		---	---	0.1651	*	---	---
Number of Observations		43618	43618	21809	21809	43618	456

Note: Standard errors (consistent with within-family correlation) in parentheses. Sample includes only twins with valid one-minute and five-minute apgar scores. Birth-weight enters as a twenty-segment piecewise linear spline, with knot points at birthweights corresponding to the 5th, 10th, 15th, etc. percentiles of the twins birthweight distribution. Logit and conditional logit uses a three-segment piecewise linear spline, with knot points at birthweights corresponding to the 5th and 10th percentiles of the twins birthweight distribution. Coefficients on the slopes are reported only for the first three segements of the spline. Birth-weight coefficients expressed as change in probability per 1000 grams. Column (1) reports least squares estimates, and Column (2) report family (mother) fixed effects estimates. Columns 3(a) and 3(b) report coefficients from separate regression for the first- and second-born twin, where mortality indicator is regressed on own and twin birth-weight. Column (4) and Column (5) coefficients are changes in log-odds per 100 grams.

Table 6: Twins Pooled and Fixed Effects Estimates by Mother's Race and Education
(estimated standard errors in parentheses)

	Infant Deaths (per 1,000 live births)							
	Within 1-Year		Within 24 Hours		Within 28 Days		Postneonatal	
Birth weight coeff.	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.
White	-0.108	-0.022	-0.066	-0.006	-0.100	-0.018	-0.008	-0.005
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.001)	(0.002)
Black	-0.161	-0.018	-0.102	-0.001	-0.141	-0.010	-0.020	-0.008
	(0.006)	(0.008)	(0.006)	(0.004)	(0.006)	(0.006)	(0.002)	(0.005)
White Education								
HS Dropout	-0.123	-0.045	-0.070	-0.011	-0.113	-0.030	-0.010	-0.016
	(0.010)	(0.013)	(0.008)	(0.007)	(0.010)	(0.009)	(0.003)	(0.010)
HS Graduate	-0.116	-0.023	-0.077	-0.012	-0.110	-0.021	-0.007	-0.003
	(0.006)	(0.005)	(0.005)	(0.004)	(0.006)	(0.005)	(0.001)	(0.003)
Some College	-0.099	-0.016	-0.062	-0.001	-0.094	-0.012	-0.005	-0.004
	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.005)	(0.002)	(0.003)
College Graduate	-0.095	-0.017	-0.055	-0.001	-0.086	-0.014	-0.010	-0.004
	(0.007)	(0.005)	(0.006)	(0.004)	(0.006)	(0.005)	(0.002)	(0.002)
Black Education								
HS Dropout	-0.176	-0.039	-0.109	-0.018	-0.152	-0.032	-0.025	-0.007
	(0.012)	(0.018)	(0.011)	(0.011)	(0.012)	(0.014)	(0.005)	(0.012)
HS Graduate	-0.155	-0.011	-0.100	0.003	-0.138	-0.003	-0.017	-0.008
	(0.009)	(0.011)	(0.009)	(0.007)	(0.009)	(0.008)	(0.003)	(0.008)
Some College	-0.156	-0.000	-0.107	0.009	-0.135	0.008	-0.021	-0.008
	(0.013)	(0.012)	(0.013)	(0.005)	(0.013)	(0.009)	(0.005)	(0.008)
College Graduate	-0.151	-0.042	-0.082	0.001	-0.140	-0.028	-0.010	-0.014
	(0.024)	(0.026)	(0.020)	(0.002)	(0.024)	(0.022)	(0.005)	(0.014)

Table 7: Fixed Effects Estimates of Birth Weight Effect for Same-Sex and Boy-Girl Twin Pairs
(estimated standard errors in parentheses)

	Infant Deaths (per 1,000 live births)							
	Same Gender Twins		Both Girl Twins		Both Boy Twins		Boy-Girl Twins	
Fixed Effects Coeff.	1-year	24 hour	1-year	24 hour	1-year	24 hour	1-year	24 hour
Birth Weight	-0.023	-0.006	-0.024	-0.004	-0.023	-0.007	-0.022	-0.006
	(0.004)	(0.002)	(0.005)	(0.003)	(0.006)	(0.004)	(0.005)	(0.005)
Boy							6.06	1.90
							(1.84)	(1.21)
Family Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
R-squared	0.80	0.86	0.79	0.86	0.80	0.86	0.79	0.86
Sample Size	39,836	39,836	19,692	19,692	20,144	20,144	18,296	18,296

Table 8: Twins Pooled and Fixed Effects Estimates by Cause of Death
(estimated standard errors in parentheses)

	Infant Deaths (per 100,000 live births)									
	LBW & Short Gestation		Respiratory Distress		Congenital Anomalies		Sudden Infant Death		Accident & Homicide	
	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.
Birth Weight	-0.753	-0.058	-4.53	-0.278	-0.773	-1.43	-0.170	-0.092	-0.049	0.018
	(0.094)	(0.056)	(0.201)	(0.166)	(0.079)	(0.236)	(0.032)	(0.130)	(0.017)	(0.060)
R-squared	0.01	0.80	0.07	0.78	0.01	0.56	0.00	0.50	0.00	0.57
Dependent var. mean	180.6	180.6	1,249	1,249	454.1	454.1	287.3	287.3	51.6	51.6
Sample Size	58,132	58,132	58,132	58,132	58,132	58,132	58,132	58,132	58,132	58,132

Table 9: Twins Pooled and Fixed Effects Estimates for 1985-1986, 1990-1991, and 1995
(estimated standard errors in parentheses)

	Infant Deaths (per 1,000 live births)							
	Within 1-Year		Within 24 Hours		Within 28 Days		Postneonatal	
	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.
1985-1986								
Birth Weight	-0.132	-0.013	-0.085	-0.004	-0.119	-0.008	-0.014	-0.006
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
Dep. var. mean	44.4	44.4	23.2	23.2	35.5	35.5	8.9	8.9
R-squared	0.19	0.82	0.15	0.88	0.19	0.86	0.01	0.55
Sample Size	99,782	99,782	99,782	99,782	99,782	99,782	99,782	99,782
1990-1991								
Birth Weight	-0.115	-0.023	-0.074	-0.008	-0.103	-0.015	-0.012	-0.008
	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Dep. var. mean	36.6	36.6	18.6	18.6	28.8	28.8	7.8	7.8
R-squared	0.17	0.79	0.13	0.88	0.17	0.84	0.01	0.53
Sample Size	130,904	130,904	130,904	130,904	130,904	130,904	130,904	130,904
1995								
Birth Weight	-0.090	-0.016						
	(0.003)	(0.002)						
Dep. var. mean	26.4	26.4						
R-squared	0.13	0.79						
Sample Size	68,764	68,764						

Table 10A: 5-minute APGAR Scores and Birthweight, Linear Specifications

	OLS	F.E.	Correlated R.E.	
			1st born	2nd born
	(1)	(2)	(3a)	(3b)
Birthweight (Own)	0.1030	0.0091	0.0667	0.0543
	(0.0021)	(0.0034)	(0.0024)	(0.0028)
Birthweight (Twin)	---	---	0.0420	0.0609
			(0.0023)	(0.0027)
Firstborn	0.0833	0.1163	---	---
	(0.0082)	(0.0110)		
Constant	5.994201	---	5.9534	5.6787
	(0.0545)		(0.0620)	(0.0609)
Family Effects	No	Yes	No	No
Test of equal effects (p-value)	---	---	0.0001	*
Test of symmetry (p-value)	---	---	0.0000	*
Number of Observations	43618	43618	21809	21809

Note: Standard errors (consistent with within-family correlation) in parentheses. Birth-weight enters linearly in all specifications. Birth-weight coefficients expressed as change in five-minute APGAR score per 100 grams. Column (1) reports least squares estimates, and Column (2) report family (mother) fixed effects estimates. Columns 3(a) and 3(b) report coefficients from separate regression for the first- and second-born twin, where mortality indicator is regressed on own and twin birth-weight.

Table 10B: 5-minute APGAR Scores and Birthweight, Piecewise Linear Spline Specifications

		OLS	F.E.	Correlated R.E.	
				1st born	2nd born
		(1)	(2)	(3a)	(3b)
Birthweight (Own)	1st seg.	0.7243	0.1910	0.5125	0.4332
		(0.0182)	(0.0749)	(0.0482)	(0.0479)
	2nd seg.	0.0460	0.0324	0.0477	-0.0113
		(0.0195)	(0.0456)	(0.0318)	(0.0302)
	3rd seg.	0.1823	0.0141	0.0902	0.1561
		(0.0251)	(0.0428)	(0.0349)	(0.0344)
Birthweight (Twin)	1st seg.	---	---	0.2323	0.3338
				(0.0471)	(0.0488)
	2nd seg.	---	---	-0.00887	-0.0189
				(0.0306)	(0.0307)
	3rd seg.	---	---	0.079654	0.1474
				(0.0317)	(0.0363)
Firstborn		0.0996	0.1153	---	---
		(0.0079)	(0.0110)		
Constant		-0.2307	---	-0.3293	-0.6356
		(0.1667)		(0.1922)	(0.1833)
Family Effects		No	Yes	No	No
Test of equal effects (p-value)		---	---	0.0000	*
Test of symmetry (p-value)		---	---	0.0000	*
Number of Observations		43618	43618	21809	21809

Note: Standard errors (consistent with within-family correlation) in parentheses. Birth-weight enters as a twenty-segment piecewise linear spline, with knot points at birthweights corresponding to the 5th, 10th, 15th, etc. percentiles of the twins birthweight distribution. Coefficients on the slopes are reported only for the first three segements of the spline. Birth-weight coefficients expressed as change in APGAR score per 100 grams. Column (1) reports least squares estimates, and Column (2) report family (mother) fixed effects estimates. Columns 3(a) and 3(b) report coefficients from separate regression for the first- and second-born twin, where mortality indicator is regressed on own and twin birth-weight.

Table 11: Estimates of Effects of Birth Weight and 5-Minute APGAR on Assisted Ventilation after Birth

	Effect of Birth Weight (1/1000)				Effect of 5-Minute APGAR			
	On Ventilator		Ventilator>30 mins		On Ventilator		Ventilator>30 mins	
	OLS	F.E.	OLS	F.E.	OLS	F.E.	OLS	F.E.
Birth Weight	-0.0857	-0.0054	-0.0739	-0.0018				
	(0.0027)	(0.0032)	(0.0024)	(0.0025)				
5-Minute APGAR					-0.0434	-0.0277	-0.0339	-0.0171
					(0.0017)	(0.0021)	(0.0015)	(0.0018)
Family Fixed Effects	N	Y	N	Y	N	Y	N	Y
R-squared	0.056	0.853	0.067	0.857	0.060	0.857	0.058	0.859
Sample Size	58,132	58,132	58,132	58,132	48,776	48,776	48,776	48,776

Table 12: Estimated Effects of Maternal Smoking on Proxies for Infant Health at Birth and Infant Mortality,
All Mothers/Infants in Pennsylvania, 1989-1991
(estimated standard errors in parentheses)

A. Proxies for Infant Health at Birth

	Effects of Maternal Smoking on the Distribution of Birth Weight								Effects on Gestation Duration (wks)				Effects on	
	Birth Weight		Incidence<2500g		Incidence<1500g		Incidence<1000g		Gestation Length		Incidence<28 wk		5-min. APGAR	
	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust
Smoking coeff.	-281.3	-199.4	65.0	35.2	9.33	0.39	4.23	-0.79	-0.32	-0.07	5.27	-1.24	-0.07	-0.01
	(2.11)	(2.15)	(1.08)	(1.07)	(0.48)	(0.47)	(0.33)	(0.34)	(0.01)	(0.01)	(0.36)	(0.36)	(0.003)	(0.003)
Constant	3398.3		58.3		11.90		6.08		39.2		6.86		9.03	
	(0.94)		(0.37)		(0.17)		(0.12)		(0.004)		(0.13)		(0.001)	
Adjusted with control variables	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
R-squared	0.035	0.190	0.010	0.128	0.001	0.110	0.000	0.084	0.002	0.174	0.001	0.096	0.001	0.074
Sample Size	509,382	509,382	509,382	509,382	509,382	509,382	509,382	509,382	509,195	509,195	509,195	509,195	507,722	507,722

B. Infant Deaths (per 1,000 live births)

	Infant Deaths (per 1,000 live births)									
	Within One-Year		Within 24 Hours		Within 28 Days		Low Birth Weight		Respiratory Distress	
	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust
Smoking coeff.	4.68	-0.20	1.09	-1.25	1.78	-1.49	0.66	-0.31	0.44	-0.46
	(0.37)	(0.39)	(0.23)	(0.25)	(0.29)	(0.31)	(0.14)	(0.15)	(0.14)	(0.16)
Constant	8.14		3.54		5.75		1.10		1.37	
	(0.14)		(0.09)		(0.12)		(0.05)		(0.06)	
Adjusted with control variables	N	Y	N	Y	N	Y	N	Y	N	Y
R-squared	0.000	0.041	0.000	0.048	0.000	0.048	0.000	0.024	0.000	0.016
Sample Size	509,830	509,830	509,830	509,830	509,830	509,830	509,830	509,830	509,830	509,830

Table 13: Estimated Effects of Maternal Smoking on Proxies for Infant Health at Birth and Infant Mortality,
Across All Twin Pairs in 1989
(estimated standard errors in parentheses)

A. Proxies for Infant Health at Birth

	Effects of Maternal Smoking on the Distribution of Birth Weight						Effects on Gestation Duration (wks)				Effects on	
	Birth Weight		Incidence<1500g		Incidence<1000g		Gestation Length		Incidence<28 wks		5-min. APGAR	
	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust	no adj.	adjust
Smoking coeff.	-202.8	-160.1	37.5	25.8	8.26	1.68	-0.11	-0.03	9.18	3.10	-0.07	-0.02
	(10.01)	(7.25)	(4.99)	(3.92)	(3.36)	(2.91)	(0.07)	(0.07)	(3.49)	(3.66)	(0.02)	(0.02)
Constant	2454.7		90.1		41.9		36.0		41.2		8.54	
	(4.72)		(1.99)		(1.41)		(0.03)		(1.45)		(0.01)	
Adjusted with control variables	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
R-squared	0.015	0.553	0.003	0.419	0.000	0.336	0.000	0.032	0.000	0.029	0.000	0.274
Sample Size	46,318	45,584	46,318	45,584	46,318	45,584	46,224	45,584	46,224	45,584	43,601	42,926

B. Infant Deaths (per 1,000 live births)

	OLS Estimates of Effect of Maternal Smoking on Infant Death					
	Within One-Year		Within 24 Hours		Within 28 Days	
	no adj.	adjust	no adj.	adjust	no adj.	adjust
Smoking coeff. and Birth Weight coeff.	5.88	0.06	-1.95	-4.98	0.10	-3.93
	(2.88)	(2.74)	(2.05)	(2.03)	(2.59)	(2.43)
Constant	37.4		20.6		31.0	
	(1.25)		(0.97)		(1.17)	
Adjusted with control variables	N	Y	N	Y	N	Y
R-squared	0.000	0.215	0.000	0.186	0.000	0.228
Sample Size	46,318	45,584	46,318	45,584	46,318	45,584

Notes: Sampling errors corrected for twin clustering and heteroskedasticty.

Appendix Table 1A: 1-minute APGAR Scores and Birthweight, Linear Specifications

	OLS	F.E.	Correlated R.E.	
			1st born	2nd born
	(1)	(2)	(3a)	(3b)
Birthweight (Own)	0.1363	0.0304	0.0933	0.0835
	(0.0019)	(0.0050)	(0.0031)	(0.0037)
Birthweight (Twin)	---	---	0.0481	0.0680
			(0.0030)	(0.0037)
Firstborn	0.3989	0.4362	---	---
	(0.0134)	(0.0185)		
Constant	3.80838	---	4.1002	3.4206
	(0.0504)		(0.0595)	(0.0587)
Family Effects	No	Yes	No	No
Test of equal effects (p-value)	---	---	0.0000	*
Test of symmetry (p-value)	---	---	0.0000	*
Number of Observations	43618	43618	21809	21809

Appendix Table 1B: 1-minute APGAR Scores and Birthweight, Piecewise Linear Spline Specifications

		OLS	F.E.	Correlated R.E.	
				1st born	2nd born
		(1)	(2)	(3a)	(3b)
Birthweight (Own)	1st seg.	0.4575	0.1169	0.3677	0.2186
		(0.0163)	(0.0688)	(0.0456)	(0.0450)
	2nd seg.	0.2882	0.1545	0.2514	0.1660
		(0.0258)	(0.0577)	(0.0396)	(0.0414)
	3rd seg.	0.208355	-0.0124	0.1276	0.1321
		(0.0389)	(0.0664)	(0.0515)	(0.0550)
Birthweight (Twin)	1st seg.	---	---	0.1213	0.2190
				(0.0456)	(0.0459)
	2nd seg.	---	---	0.05058	0.0660
				(0.0384)	(0.0421)
	3rd seg.	---	---	0.12416	0.1635
				(0.0477)	(0.0567)
Firstborn		0.4135	0.4337	---	---
		(0.0132)	(0.0185)		
Constant		-0.0795	---	-0.0566	0.0579
		(0.1284)		(0.1601)	(0.1579)
Family Effects		No	Yes	No	No
Test of equal effects (p-value)		---	---	0.0000	*
Test of symmetry (p-value)		---	---	0.0000	*
Number of Observations		43618	43618	21809	21809