REAP Rede de Economia Aplicada

Low Birth Weight and Infant Mortality: Lessons from Brazil

Bladimir Carrillo
José G. Feres

Junho, 2017
Working Paper 109

Low Birth Weight and Infant Mortality: Lessons from Brazil

Bladimir Carrillo
José G. Feres

Low Birth Weight and Infant Mortality: Lessons from Brazil[•]

Bladimir Carrillo José G. Feres

June 2017

Abstract

Governments devote considerable resources on reducing the incidence of low birth weight with the reasoning that low birth weight is the cause of poor infant health. Much of what we know on the causal link between these variables comes from developed countries. However, these estimates may have limited external validity to the developing world if families with more resources are better able to remediate the effects of poor neonatal health. This study uses unique, rich data and a within-twin identification strategy to estimate the effect of low birth weight on infant mortality in Brazil. We document that the effects of low birth weight are much larger than those derived from the US and Norwegian context. This finding suggests that applying from these economies to poorer countries may be misleading for cost-benefit assessments of policy.

Keywords: Health human capital; health endowments at birth; Brazil; Twins.
JEL Classification: H51, I12, I18

[•] We thank Prashant Bharadwaj, Emily Oster, Miguel Foguel, Danyelle Branco, Cristiana Tristao, Micheliana Costa, Wilman Iglesias, Carlos Charris, Raul Velilla, and Marcelo José Braga for useful comments and suggestions. This study was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior (CAPES), Brazil. Bladimir Carrillo: Department of Rural Economics, Universidade Federal de Viçosa, email: bladimir.bermudez@ufv.br. José G. Ferés IPEA, email: jose.feres@ipea.gov.br jose.feres@ipea.gov.br

1. Introduction

It has been widely believed that malnutrition *in utero*, commonly proxied by low birth weight, is an important contributor to poor infant health.[1] As a result, governments and international agencies have devoted considerable resources on preventing low birth weight. In India, for example, the World Bank allocated over US$100 million for a program aimed at cutting in half the incidence of low birth weight.[2] Preventing low birth weight has been also a major motivation for nutritional programs and maternal smoking campaigns worldwide.[3] The strong and well-documented association between low birth weight and infant health has led to the position that the social returns of these investments are large. Numerous studies indicate that low birth weight babies have increased risk of death within one year of birth, and that who survive infancy are likely to suffer from a number of health and developmental difficulties, some of which are known to negatively affect acquisition of human capital.[4] Understanding whether low birth weight is in fact the cause of poor infant health and not simply a correlate of such problems is crucial for guiding the targeting of policies intended to reduce inequalities by improving early life health.

The challenges with uncovering the causal effect of birth weight are well known in the literature. A strong correlation between low birth weight and infant health may be the product of unobserved factors because the determinants of nutrition during pregnancy, including socioeconomic status and parent's knowledge about health care, are also likely determinants of infant health. So any attempt to ascertain the importance of birth weight for infant health by simply looking at their correlation, or equivalently estimating a simple ordinary least square (OLS) regression, is unlikely to provide convincing evidence. Studies within economics have overcome these challenges using rich data from the United States and Norway, and within-twin identification strategies (Almond et al., 2005; Black et al., 2007; Oreopoulos et al., 2008). These studies suggest that low birth weight leads to increased risk

[1] Low birth weight is conventionally defined as a birth weight less than 2,500 grams.
[2]This is the Second Tamil Nadu Integrated Nutrition Project. The program also had other goals, such as improving nutrition and health status of children0-72 months (The World Bank, 1998).
[3] In the United States, for example, a motivation for the Medicaid expansion to pregnant women was the reduction of the incidence of low birth weight (Currie and Gruber, 1996).
[4] Previous studies have shown, for example, that low birth weight is associated with health problems such as cerebral palsy, deafness, epilepsy, blindness, asthma, and lung disease (Brooks et al., 2001; Kaelber and Pugh, 1969; Lucas et al., 1998; Matte et al., 2001).

of mortality, although the effects are much smaller than those derived from cross-sectional regressions. This body of research even suggests that low birth can have long-lasting effects on human capital accumulation, which in turn has been interpreted as evidence consistent with the literature emphasizing that early health conditions are a major determinant of individual capabilities.[5] For example, Figlio et al. (2014) illustrate that birth weight has negative effects on cognitive development, while that Black, Devereux, and Salvanes (2007) show that low birth weight babies exhibit reduced earnings, lower educational attainment, and worse health outcomes as adults.

While these studies have undoubtedly advanced our understanding of the effect of birth weight on infant welfare, we know fairly little about this relationship in developing countries. As emphasized by Currie and Vogl (2013), research on the consequences of early health insults has much policy relevance in poorer countries, but precisely measured birth weight data are rare in large sample surveys from these countries. Thus, it is very little known about whether the effects of birth weight vary at different economic development contexts. In the absence of a well-functioning public health system and the presence financial constraints, the capacity to remediate health shocks may be simply more limited in poor countries, which would imply that birth weight might have a larger overall health impact in these economies. Moreover, one may observe different effects if there are non-linearities in the production function for child health or if there are interactions between birth weight and environmental factors (Almond and Mazumder, 2013; Yi et al., 2015). In consequence, estimates derived from rich countries may not be externally valid to the developing country context.

Many of the existing studies for developing countries are in the epidemiological literature. These studies has relied on cross-sectional estimates while controlling for parents' background characteristics. However, this empirical strategy might be subject to omitted variable bias from unobserved factors that can affect both birth weight and infant health. Furthermore, these studies are generally based on small and non-representative samples, making it the results difficult to generalize and limiting the development of clear stylized facts. Remarkably, research in the economic literature that aims to have a more causal and

[5] See Conti and Heckman (2010), Cunha, Heckman and Schennach (2010), Cunha, and Heckman (2007, 2008, 2009) for a theoretical discussion about the role of early health conditions in the accumulation of human capital.

general interpretation of the relationship between birth weight and infant health in a developing country context is rare. To the best of our knowledge, only McGovern (2014) investigates the effects of birth weight on infant health in developing countries. He uses data from the Demographic and Health Surveys (DHS), which is conducted in more than 90 countries worldwide. However, the use of self-reported information on birth weight is likely to suffer from measurement error that may not be random. Most people in developing country rural areas, especially in sub-Saharan Africa, do not give births in hospitals, so birth weight is likely to be badly measured. Moreover, the use of these surveys does not allow excluding twin pairs with congenital defects and Almond, Chay and Lee (2005) show that it can lead to severe bias.

In this paper, we provide estimates of the effect of birth weight on infant mortality using administrative data on the universe of births linked to death records in Brazil. As we describe in more detail in section 2.2, these matched data provide comprehensive information on birth weight, congenital defects, date and cause of death, and mother's background characteristics. With these rich data, we follow 19 million singletons and 300,000 pairs of twins from birth through the first year of life. The enormous sample size from this dataset gives us a strong statistical power to discern patterns. For identification, we take advantage of quasi-random variation in birth weight within twin pairs, as described in section 2.1. Using precisely measured birth weight data in a large nationally representative sample and a within-twin identification strategy, we provide what we believe is the most credible evidence on the causal effect of birth weight on infant health in a developing country context.

We document that lower birth weight babies exhibit higher rates of mortality within one year of birth. Our estimates imply that very low birth weight babies have 4 percentage points higher risk of death within one year. The mortality effects are concentrated on conditions originating in the perinatal period, which include respiratory and cardiovascular disorders specific to the perinatal period, and hematological disorders of fetus and newborn. In line with earlier studies for developed countries, the cross-sectional estimates tend be substantially larger in magnitude than the ones derived from the twin-fixed effects estimator. This confirms that policy designs based on cross-sectional estimates may exaggerate the benefits of reducing the incidence of low birth weight.

We then compare our estimates to those derived in the US and Norway. Specifically, we compare our estimates to Almond, Chay, and Lee (2005) and Black, Devereux and Salvanes (2007). In general, our estimates are larger in magnitude than those derived from these studies. The differences are substantial. For example, our estimates are about two times larger than those reported by Almond, Chay, and Lee (2005) for the United States. We argue that these results cannot be explained by specific features of our empirical setting, such as measurement error and the possibility of selection bias induced by miscarriage or stillbirth. A more plausible interpretation of these results is that developing and developed countries have a very different causal relationship between birth weight and infant mortality. Although it is difficult to make causal claims on the specific reasons behind these differences, we assess whether related explanations that are more common to developing countries, such as low parental education, might be plausible candidates. Our results indicate that the effects of birth weight are stronger for infants born to mothers who have low educational attainment and are unmarried. The effects generally increase by 5 to 71 percent relative to infants born to more advantaged families. We also find that the effects of birth weight are smaller for families residing in municipalities with sanitation coverage over 85 percent. For these families, the impacts falls by 41 to 83 percent, which suggests that birth weight may be interacting with environmental factors. Taken together, we conclude tentatively that applying estimates that are derived from the US or Norway to developing countries may be misleading for cost-benefit assessments of policy.

The rest of the paper is organized as follows. In the next section, we describe our estimation strategy and the data used. Section 3 presents our main results, including robustness checks and a comparison of our estimates to those derived in the US and Norwegian setting. Section 4 explores different forms of heterogeneity in the impacts of birth weight on infant mortality. Finally, section 5 concludes.

2. Empirical Approach and Data

2.1. Identification strategy

The goal of the empirical analysis is to estimate the effect of birth weight on infant death. Following Almond and Lee (2005) and Black, Devereux, and Salvanes (2007), let:

$$Death_{ijt} = \alpha + \beta bw_{ijt} + x'_{jt}\delta + \mu_{jt} + \varepsilon_{ijt} \qquad (1)$$

The variable $Death$ is the probability of death within one year of life of the infant i born to mother j in year t. The variable bw is birth weight; x is a vector of mother's characteristics, including education, age at birth and marital status; μ_{jt} is a set of unobservable that are mother- and birth-specific, such as family background, the quality of prenatal, genetic factors, and mother's knowledge or awareness about health care; and ε_{ijt} is an idiosyncratic error term assumed orthogonal to other terms in the equation.

The parameter of interest for policy is β. If it is negative and large in magnitude, then targeting interventions during the in utero period to prevent low birth weight may yield high returns. OLS estimates of the equation (1) that ignore μ_{jt} will be likely biased because many factors in μ_{jt} are also determinants of birth weight. For example, the quality of parent's education is likely to affect both prenatal and postnatal investments. Therefore, any OLS estimate of β would need to be a combination of omitted variable bias and the causal effect of birth weight. To isolate the effect of birth weight from unobservable factors, we use a twin-fixed effects estimator. This approach compares the probability of death of twin *i* to twin *k*, who were born to the same mother but had different levels of birth weight. Including twin-fixed effects is equivalent to estimating the following equation:

$$Death_{1jt} - Death_{2jt} = \beta\left(bw_{1jt} - bw_{2jt}\right) + (\varepsilon_{1jt} - \varepsilon_{2jt}) \qquad (2)$$

Where "1" refers to the first-born twin and "2" refers to the second-born twin. Note that the use of the equation (2) will produce consistent estimates since the mother- and birth specific component is differenced out and ε_{ijt} is assumed independent of birth weight. To see this, it is important to understand the mechanisms by which nearly all twin pairs differ in the birth weights. As discussed by Almond and Lee (2005) and Black, Devereux, and Salvanes (2007), the differences in birth weight could be the result of differences in nutritional intake induced by different umbilical cord insertion points within the placenta and different positions in the womb (Phillips, 1993; Zhang et al., 2001). As parental control over these factors is limited, it becomes plausible the identifying assumption that within-twin differences in birth weight are exogenous.

Our main results are based on the equation (2). To account for gender differences in birth weight, we also include an indicator variable of infant's sex as a control in all regressions. There is no a priori justification for using a determined functional form. We compare the explanatory power of specifications that use either birth weight, log(birth weight), and a set of dummy variables for discrete birth weight categories (i.e, 1,500, 1,500-2,500 and 2,500-3,000 grams). This exercise reveals that the specification using dummy variables provides the best fit. Therefore, we use these birth weight categories as the independent variables of interest in our analysis.

2.2.Data

Our empirical analysis requires data on birth weight and infant mortality. We use microdata from the Brazilian National System of Information on Birth Records (SINASC) and the National System of Mortality Records (SIM). The SINASC provides information on all births in Brazil since 1994, although it did not cover most of the municipalities before 1998. The data includes the exact date of birth, weeks of gestation, sex, birth weight, and maternal characteristics such as marital status, age and education. The death certificates from the SIM provide comprehensive information on date and cause of death, birthdate, race, and gender, and mother's characteristics (education, marital status and age) are also provided for individuals who were under one year of life at death. Municipal governments are responsible for collecting all death certificates and sending them to the Health Ministry of the Government of Brazil, which consolidates finally the information in the SIM database. The laws governing the collection of the death certificates are national and no burial can be performed without a death certificate. The SIM covers over 96 percent of all annual deaths inferred from demographic census.[6]

We linked death certificate information for the infants who die in their first year of life to the SINASC database by using unique personal identifiers provided by the Health Ministry. The unique personal identifiers are available for births occurring from 2006 through 2012. During this period, there were 20,265,131 births, of which 1.9 percent were twins. The matching

[6] Information on coverage of the deaths from SIM are available at http://tabnet.datasus.gov.br/cgi/sim/dados/cid10_indice.htm.

rates are nearly constant across time and States. About 70 percent of the infant death records are matched to one of the birth records. The matching rate is not 100 percent because the unique personal identifiers are missing for some infants in the death records. The matching rate is notably higher for twin births, approximately 80 percent. This is reassuring because our main analysis relies on the sample composed by twins. The fact that some infant deaths are not matched to the birth certificate records will introduce measurement error in our dependent variable. We discuss the implications of this issue in section 3.2.

Since the unique personal identifiers allow only identifying individual births, infants who were born to the same mother cannot be directly inferred. We exploit the fact that multiple birth records are generally located next to each other in birth certificate files to construct twin-pairs codes. First, we identified all these adjacent twin pairs in our data. Second, we consider that a given adjacent twin pair is part of the same twin set if pregnancy characteristics are identical. Specifically, adjacent twin pairs are considered as born in the same mother if they have identical information for the following set of covariates: hospital of birth, exact date of birth, gestational age, mother's age, mother's education, mother's marital status, and municipality of residence. Approximately 90 percent of the 308,010 adjacent twin pair records have identical information for these variables. We restrict the sample to these twin births, although the results are extremely similar when used all adjacent twin pair records.

In total, there are 276,268 twin births in our sample. We exclude twin pairs where either twin was born with a congenital defect (about 7 percent), as differences in birth weight that are driven by this condition may introduce bias. Twin pairs where either twin had missing information about sex or birth weight are also excluded from the analysis. This restriction results in dropping about 0.1 percent of the sample. Our final sample consists of 255,362 twin births. While our main analysis focuses on the twin sample, we also present results for singletons. There are 18,929,949 singletons with non-missing information for birth weight and infant's sex.

Table 1 presents basic descriptive statistics splitting the sample between twins and singletons. It is apparent that there is differences between the two populations. Indeed, twins are more likely to be born with a weight less than 2,500 grams, have higher rates of prematurity, and are more likely to die within one year of birth than singletons. The differences are large. For example, the probability of dying in the first year is 4.5 times higher for twins than for singletons. These differences also suggest a negative relationship between birth weight and infant mortality. Since low birth weight is also the result of prematurity, it is difficult to establish in principle from these cross-sectional comparisons either whether birth weight or prematurity is the responsible for the increased rates of infant mortality among twins. As Table 1 shows, there are also substantive differences in mother's characteristics between the two groups. In general, twinning probabilities seem to be higher among advantaged families. Indeed, mother of twins are more likely to be older, more educated and more likely to be married. It is well known that the use of fertility treatments, such as in vitro fertilization pre-embryo transfer, can increase the likelihood of multiple births.[7] Since these treatments are costly or provided by private health insurance, families with more resources may be more likely to use them and consequently parents' background characteristics could be systematically related to the incidence of twin births.[8] This fact calls into question the external validity of the analysis from twins. Despite these dissimilarities between the two populations, we provide suggestive evidence that the results from twins may be generalizable to the general population. In particular, we show that the pooled cross-sectional estimates for the twin population are remarkably similar with that for the singleton population.

Because our statistical approach relies on within-twin variation, we confirm that there is substantial within-twin variability in birth weight and mortality outcomes. Table 2 and Figure 2 show the distribution of the twin birth weight-difference. The mean birth weight difference is 276 grams, or 11 percent of the average twin's birth weight. The data also indicate that 60 percent of twin pairs exhibit a birth weight difference higher than 260 grams, and 10 percent have a birth weight difference higher than 600 grams. In Table 3, which reports mean squared errors from regressions with either birth weight or mortality outcomes as dependent variables, we explore in more detail the sources of the variation in both outcomes. Columns (2) reveals that gestational age explains over half of the overall variance in birth weight. This is

[7] See http://www.ivf.comhttp://www.ivf.com.

[8] Ponczek and Souza (2012) provide a comprehensive discussion about the relationship between fertility treatments, twinning probabilities and parents' background characteristics in Brazil. They also show that mother of twins are more educated.

consistent with prior literature indicating that gestation length plays a critical role in intrauterine growth (Kramer, 1987). Despite the significant contribution of gestation length to variation in birth weight, there are great deal of variation that is due to within-twin differences. Indeed, column (3) shows that 20 percent of differences of the birth weight variation due to differential fetal growth rates is due to within-twin differences. This wide variation is the basis of our identification strategy.

3. Results

3.1. Main results and robustness checks

Table 4 shows estimates of the effect of birth weight on four mortality outcomes: one-year mortality, neonatal mortality, seven-day mortality and one-day mortality. Each panel reports results for different sample and estimation techniques. In Panel A, we use the sample of singletons and estimate OLS regressions. Pooled OLS estimates for the sample of twins are presented in Panel B. Finally, our preferred results are presented in Panel C, where we use the twin-fixed effects strategy. All regressions include control for infant's sex. The sample sizes and R^2's of the regressions are shown at the bottom of each panel. All standard errors are robust against arbitrary heteroscedasticity, and allow for clustering at the twin-pair level when the sample of twins is used.

The pooled OLS estimates using the singleton sample suggest strong and negative effects of birth weight on infant mortality. The probability of dying within one year of birth of a baby weighting less than 1,500 grams is as much as 31 percentage points higher than that of babies weighting 3,000 grams or more. Relative to a sample average probability of one-year mortality of 34 percent for infants with birth weight less than 1,500, the effect is substantial at 91 percent. For the 1,500-2,500 grams birth weight category the effect falls to 2 percentage points, but remains significant. These estimated coefficients are quite similar in magnitude for neonatal mortality. This suggests that the cross-sectional relationship between birth weight and infant mortality is driven largely by deaths that occur within 28 days of birth. When we use pooled OLS in the sample of twins, we find a similar pattern. Importantly, the estimated coefficients are also similar in magnitude for twins than for singletons, which suggests that both populations are subject to the same relationship between birth weight and infant mortality. This provides reassuring evidence that the results from twins may be generalizable to the general population. Overall, these findings confirm the strong cross-sectional relationship between birth weight and infant mortality found in earlier studies for developed countries.

In direct contrast to the cross-sectional estimates, the twin-fixed effects estimator suggests much smaller impacts. For infants born with very low weight (less than 1,500 grams), there are about 4 percentage points higher risk of death within one year. This estimated effect is only one sixth the size of the OLS coefficient. Similarly, when we look at the other mortality measures, we find much smaller impacts. The effect of very low birth weight falls from 23 to 3 percentage points for neonatal mortality, from 19 to 1.8 percentage points for seven mortality, and falls from 12 to 0.3 percentage points for one-day mortality. In sum, the estimated effect of very low birth weight falls by a factor of 6 to 40. Despite the estimates fall notably when family unobserved characteristics are accounted for, they remain statistically significant, with exception of one-day mortality. The fact that twin-fixed effects estimates are much smaller suggests that there is a prima facie case for a severe omitted variable bias in the cross-sectional regressions.

Since our sample includes both fraternal and monozygotic twins, one might be even worried if discordance in birth weight between twins are related to these genetic conditions. This type of caveat is recurrently mentioned in the literature that uses within-twin identification strategies. Still, these studies tend to find quite similar results when their samples are restricted to same-sex twin pairs, which clearly contain a larger fraction of identical twin births. In fact, Black, Devereux, and Salvanes (2007) are able to observe directly zigosity in a sub-set of twins and find identical results. The robustness of the findings in such sub-samples suggests that the bias generated by zygosity is, at best, small in practice. We perform the same robustness check by estimating the birth weight effects in a sub-sample that includes only same-sex twin pairs. As shown in Table 5, the estimates are similar to the baseline ones, indicating that zygosity is not affecting our estimates.

The enormous sample size we have at our disposal allows us to explore the relationship between birth weight and infant mortality by cause. We group our sample into four categories: conditions originating in the perinatal period, infectious and parasitic diseases, diseases of the respiratory system, and all other diagnoses. These results are presented in

Table 6 using our preferred estimation technique, namely the twin-fixed effects estimator. The results show no robust evidence of a birth weight effect on mortality by infectious, parasitic, or respiratory causes. Furthermore, we find small estimates that are tightly bound around zero, indicating that birth weight does not have any discernible effects on these causes of death. In contrast, the results indicate that the birth weight effects are driven by deaths from conditions originating in the perinatal period, which include respiratory and cardiovascular disorders specific to the perinatal period, hematological disorders of fetus and newborn, and disorders related to low birth weight. This is perhaps unsurprisingly because some of these disorders are directly diagnosed based on baby's birth weight.

While the functional form used allows for non-linear effects, it may not completely capture the relationship between birth weight and infant mortality if there is specific effects at other birth weight categories. For example, the effects of birth weight may be particularly higher among infants with a birth weight less than 1,000 grams. Next, we use a specification that allows the effects of birth weight to be more flexible. Specifically, we estimate models given by:

$$Death_{ijt} = \alpha + \sum_k D_{ijt}^k \beta^k + x_{jt}' \delta + \mu_{jt} + \varepsilon_{ijt} \qquad (3)$$

where D_{ijt}^k is a dummy variable that indicates if the birth weight of an infant is in the *k*th bin. We use 27 dummy variables corresponding to 100 gram-wide birth weight bins of the distribution of birth weight below 3,000 grams. The bins range from a low of 300-400 grams to a high of 2,900-3,000 grams. The omitted category is birth weight of 3,000 grams or more. We estimate these regressions using both OLS and twin-fixed effects. The results from this more flexible functional form are presented in Figure 3, which plots the coefficients from these weight-bins. In general, there appears to be a concave relationship between birth weight and infant mortality, indicating that reductions in birth weight are more detrimental at lower levels of birth weight. The effects tend to disappear when the birth weight is over 1,800 grams. The results also make clearer the severe omitted variable bias in the OLS regressions. Consider, for example, the cross-sectional results for one-day mortality. They indicate that infants who are born with a weight below 300 grams are 40 percentage points more likely to die within one day of birth. In contrast, the twin-fixed effects results suggest a statistically insignificant impact of about 1 percentage point. In general, the twin-fixed effects estimates are never significant for one-day mortality, with estimated coefficients tightly bound around zero.

Table 7 explores further alternative specifications of the relationship between birth weight and infant mortality. Panel A shows results using birth weight (in grams) as the primary variable of interest, while that Panel B uses the log of birth weight. In general, our results are qualitatively similar using these variables. Our estimates from Panel A imply that a 50 grams increase in birth weight would reduce one-year mortality by one death per 1,000 births. Since infant mortality is a rare event, estimates may be sensitive to functional form. Panels C and D estimate logit models with twin-fixed effects. Using this functional form, we find results qualitatively similar, but the marginal effects tend to be higher. For example, the coefficient of -0.002 in Panel C implies that a 50 grams increase in birth weight would reduce one-year mortality by three deaths per 1,000 births. This is perhaps unsurprising given that logit models only includes cases in which one twin lives and one twin dies, which may change the composition of the sample.

Next, we replace our measure of infant health with other common measure of infant welfare, namely APGAR scores. This is a clinical test that is given to the newborn in which five parameters are assessed. These include muscle tone, respiratory effort, heart rate, reflexes and skin color. The test provides a total score between 0 and 10, where a higher score means “healthier”. The results in Table 8 suggests one strong cross-sectional relationship between birth weight and APGAR scores. Very low birth weight babies are 33 percentage points more likely to have a low 5-minute APGAR score (less than 8). This estimated effect falls by a factor of 13 when twin-fixed effects are included, although it remains statistically significant.

3.2. Selective mortality and measurement error

As our analysis is based on live births, a bias could arise if a disproportioned number of the marginal fetus that survive are in the low end of the birth weight distribution. That is, if weak fetuses with potentially low birth weight are less likely to be born alive, then our results would be based on a select sample of surviving (and presumably stronger) births. However, note that the use of this select sample most likely will bias our estimates of the effect of birth weight on infant mortality towards zero. Therefore, we are less concerned about selection bias from selective miscarriage or stillbirth. As such, in the presence of this bias, our estimates should be viewed as lower bounds of the true effect of birth weight on infant mortality.

Another potential concern with our results is measurement error in health outcomes. As we mentioned earlier, some certificate death records were not matched to one of the birth records, which implies that infant mortality is measured with error for some infants. If the measurement error is random then the consistency and unbiasedness of our estimates would be unaffected. Alternatively, if birth weight covaries with the measurement error, then our estimates would be biased. In Appendix A, we describe in full detail a simple test that measures the extent to which this measurement error may affect our estimates. In particular, the test takes advantage of the fact that the measurement error is observable in the certificate death records and we have information about birth weight for all these births. Thus, the within-twin correlation between the measurement error (or equivalently the likelihood of being matched to birth files) and the birth weight of the infants would be a simple test for this potential bias. Since we are unable to identify twin pairs who were born to the same mother in the death records, the within-twin correlation cannot estimated, but the overall correlation would provide useful information if it goes in the same direction and magnitude. The data indicate that a 200 grams increase in birth weight is associated with a decrease of 0.8 percentage point in the likelihood of being matched. While significant at the 5 percent level, this estimate is small in magnitude, with an implied elasticity of only -0.05. Assuming that the within-twin covariance is smaller than overall covariance between the probability of being matched to certificate birth records and birth weight, then the resulting bias is unlikely to be relevant in practice. [9]

3.3.Comparison to existing studies for developed countries

A natural question is whether estimates derived from developed countries are externally valid to the developing world. If the access to medical care is more limited in developing countries

[9] This seems a plausible assumption since it is difficult to think of reasons why, in a given twin pair, one twin has a non-missing unique personal identifier and not the other. In this case, the twin-specific component of the measurement error will tend to zero.

or if there are interactions between birth weight and infant mortality, the estimates derived from the US or Norway may not valid in conducting cost-benefit analysis of public health policies in developing countries. As we discussed in the Introduction, the presence of these potential factors likely imply an underestimating of the benefits of such policies.

We compare our estimates to those from Almond and Lee (2005) and Black, Devereux, and Salvanes (2007) in Table 9. Panel A presents our estimates for infant mortality. We present estimates that use either birth weight or log of birth weight as independent variables in order to make our results comparable to these previous studies. Panel B provides twin-fixed effects estimates for the papers in the US and Norway setting. The means of infant mortality and birth weight are also provided for ease of interpretation.

We find that a 50 grams increase in birth weight would reduce one-year mortality by one death per 1,000 births. Given the mean rate of 36.88, this implies that a 1 percent increases in birth weight leads to a 1.6 percent reduction in infant mortality. Thus, we find a much larger effect on the infant mortality rate than either Almond and Lee (2005) or Black, Devereux, and Salvanes (2007). The estimated elasticity for the US is -0.51, and for Norway is about -0.83. Given the discussion in section 3.2, it is clear that these results cannot be explained by bias from selective mortality or measurement error in mortality outcomes. Thus, a tentative conclusion is that estimates derived from the developed world are not generalizable to poor countries.

4. Heterogeneity

While it is beyond the scope of this study to understand why the causal effect of birth weight differ between developing and developed countries, we can assess whether the effects vary heterogeneously across different dimensions to provide tentative evidence of possible explanations. Furthermore, learning whether there are significant interactions also offers evidence about specific channels linking birth weight and mortality outcomes, as well as about possible policy prescriptions that may act to mitigate the consequences of low birth weight. Brazil provides a compelling setting for these purposes because it has a large, demographically heterogeneous, and socio-economically diverse population.

We look at two potential factors. First, the effects of low birth weight may depend on household behavior and it in turn might vary with family disadvantage. Parents with more resources may be simply better able to remediate the health consequences of low birth weight. As poorer families are more likely to be credit constraint, the use of important health services may be more limited. Moreover, neonatal health and parental inputs may be complements in the production function for child quality either because richer families are more likely to adopt compensating health investments or because the investments richer families make have higher returns. Second, one might expect the effects of birth weight to vary with economic development due to differences in access to public health infrastructure. For instance, it is well-known that widespread open defecation that does not make use of a toilet is one leading cause of infant mortality in developing countries. Thus, low birth weight babies in poor regions are potentially at higher risk of death partly because they are more exposed to unhealthy environments.

We begin by exploring whether the impacts of birth weight vary with family disadvantage. As family disadvantage is unobservable, we proxy it by maternal education and marital status at the time of birth. In this case, family disadvantage should be view as differences in the quality and quantity of available household resources, including child-rearing inputs and parental attention (Autor et al., 2016). In Table 10, we estimate our preferred model separately for less- and more-educated mothers, and for married and unmarried mothers. The results for these separate regressions replicate qualitatively the pattern found before. The coefficients for infants born to married and more-educated mothers tend to be smaller. The decline in the estimates ranges from 5 percent to nearly 71 percent. The cross-equation tests of coefficients reject that the coefficients are the same. These results are consistent with the notion that the health effects of low birth weight vary proportionally with family disadvantage.

In Table 11, we assess whether the birth weights impacts vary with economic development. To that end, we first divide the sample according to the quintile of the municipal GDP and then we estimate regressions separately for each group. In general, the effects tend to be smaller for infants born in municipalities with higher economic development. The falls in the estimates are striking, ranging from 41 to 83 percent. The tests for equality of coefficients generally reject the null hypothesis that they are the same. Analogously, in panel B, we divide the sample by level of sanitation coverage and estimate regressions separately for each group. The results indicate that higher access to sanitation is associated with smaller health impacts of low birth weight. Again, the differences in the estimates tend to be large and statistically significant.

Overall, these results confirm that the effects of birth weight interact with family disadvantage and economic development. To disentangle the relative importance of both dimensions, we estimate our basic regression with twin fixed effects including interactions between the discrete birth weight categories and mother's education, marital status, sanitation coverage and GDP. We do so in Table 12. Column (1) replicates our baseline estimates, while the remaining columns add progressively the interactions. The first thing to note is that there are significant interactions when considered each dimension individually, indicating that the effects of low birth weight vary inversely with family advantage and economic developing. This is consistent with the patterns found in Tables 10 and 11. The second thing is that the interactions tend to be larger and statistically significant for very low birth weight category. This is perhaps unsurprisingly because the effects of birth weight are detrimental at lower levels of birth weight. When all interactions are simultaneously added in column (6), the magnitude and significance of the interactions for education and GDP falls notably. In contrast, the interactions with sanitation and married continue to be large and statistically significant.

5. Conclusion

Despite the important reductions in infant mortality rate worldwide during the last 20 years, it continues to be high today in many developing countries. While a variety of factors are likely determinant of poor infant health, the understanding of specific causes is necessary for the most efficient design of policies. Previous studies suggest that low birth weight is a major cause of infant mortality, but much of what we know on the causal link between these variables is derived from developed countries and there is no a priori reason to believe that the results are generalizable to poorer countries. Previous studies for developing countries rely on self-reported survey data and do not have access to comprehensive birth record data, which makes it complicated to estimate the magnitude of the effect of birth weight on infant mortality.

In this article, we address these limitations by using rich administrative data on the universe of births in Brazil and shed light on the importance of birth weight for infant health in a developing country context. Using a within-twin identification strategy, we find that lower babies have increased risk of death within one year. Our estimates imply that very low birth babies have 4 percentage points higher risk of death within one year. Deaths from conditions originating in the perinatal period account for much of these effects. Our results are generally larger than those estimated with data from the US and Norway. This finding illustrate that there may be differences in estimates for developed and developing countries, which suggests that using estimates derived from rich countries may understate the benefits from interventions aimed at decreasing infant mortality by increasing birth weight in developing countries.

A natural question is why the effect of birth weight in developed and developing countries is different. Our findings suggest that financial constraints and parental attention may be an important explanation. If financial constraints hamper the use of important health services or if parental time is a powerful determinant of infant health, we should see even larger health impacts of birth weight in poorer regions. Indeed, we find the strongest birth weight effects for infants born to unmarried and less-educated mothers. In addition, we find that the effects are reduced when local sanitation coverage is high, suggesting that access to public health infrastructure may mitigate the consequences of low birth weight. Overall, these findings suggest that poverty is a likely driver behind the differences we observe in the effects of birth weight between developing and developed countries. Further research on the topic is needed to clarify these relationships.

References

Almond, D., Chay, K.Y., Lee, D.S., 2005. The Costs of Low Birth Weight. Q. J. Econ. 120, 1031–1083.

Almond, D., Mazumder, B., 2013. Fetal Origins and Parental Responses. Annu. Rev.

Econom. 5, 37–56.

Autor, D., Figlio, D., Karbownik, K., Roth, J., Wasserman, M., 2016. Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes (No. 22267), NBER Working Papers. Washington, DC.

Black, S.E., Devereux, P.J., Salvanes, K.G., 2007. From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. Q. J. Econ. 122, 409–439.

Brooks, A., RS, B., Weitzman, M., Auinger, P., JT, M., 2001. IMpact of low birth weight on early childhood asthma in the united states. Arch. Pediatr. Adolesc. Med. 155, 401–406.

Conti, G., Heckman, J.J., 2010. Understanding the Early Origins of the Education–Health Gradient: A Framework That Can Also Be Applied to Analyze Gene–Environment Interactions. Perspect. Psychol. Sci. 5, 585–605.

Cunha, F., Heckman, J., 2007. The Technology of Skill Formation. Am. Econ. Rev. 97, 31–47.

Cunha, F., Heckman, J.J., 2009. THE ECONOMICS AND PSYCHOLOGY OF INEQUALITY AND HUMAN DEVELOPMENT. J. Eur. Econ. Assoc. 7, 320–364.

Cunha, F., Heckman, J.J., 2008. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. J. Hum. Resour. 43, 738–782.

Cunha, F., Heckman, J.J., Schennach, S.M., 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation. Econometrica 78, 883–931.

Currie, J., Gruber, J., 1996. Health Insurance Eligibility, Utilization of Medical Care, and Child Health. Q. J. Econ. 111, 431–466.

Currie, J., Vogl, T., 2013. Early-Life Health and Adult Circumstance in Developing Countries. Annu. Rev. Econom. 5, 1–36.

Figlio, D., Guryan, J., Karbownik, K., Roth, J., 2014. The Effects of Poor Neonatal Health on Children's Cognitive Development. Am. Econ. Rev. 104, 3921–3955.

Kaelber, C.T., Pugh, T.F., 1969. Influence of Intrauterine Relations on the Intelligence of Twins. N. Engl. J. Med. 280, 1030–1034.

Kramer, M.S., 1987. Intrauterine Growth and Gestational Duration Determinants. Pediatrics 80, 502–511.

Lucas, A., Morley, R., Cole, T.J., 1998. Randomised trial of early diet in preterm babies and later intelligence quotient. BMJ 317, 1481–1487.

Matte, T.D., Bresnahan, M., Begg, M.D., Susser, E., 2001. Influence of variation in birth weight within normal range and within sibships on IQ at age 7 years: cohort study. BMJ 323, 310–314.

McGovern, M., 2014. How Much Does Birth Weight Matter for Child Health in Developing Countries? Estimates from Siblings and Twins, Unpublished manuscript.

Oreopoulos, P., Stabile, M., Walld, R., Roos, L.L., 2008. Short-, Medium-, and Long-Term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins. J. Hum. Resour. 43, 88–138.

Phillips, D., 1993. Twin studies in medical research: can they tell us whether diseases are genetically determined? Lancet 341, 1008–1009.

Ponczek, V., Souza, A.P., 2012. New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country. J. Hum. Resour. 47, 64–106.

The World Bank, 1998. Second Tamil Nadu Integrated Nutrition Project.

Yi, J., Heckman, J.J., Zhang, J., Conti, G., 2015. Early Health Shocks, Intra-household Resource Allocation and Child Outcomes. Econ. J. 125, F347--F371.

Zhang, J., Brenner, R.A., Klebanoff, M.A., 2001. Differences in Birth Weight and Blood Pressure at Age 7 Years among Twins. Am. J. Epidemiol. 153, 779–782.

Appendix A: *A simple test for non-random measurement error*

Consider the twin-fixed effects estimator:

$$Death^{*}_{1jt} - Death^{*}_{2jt} = \beta\left(bw_{1jt} - bw_{2jt}\right) + (\varepsilon_{1jt} - \varepsilon_{2jt}) \qquad (4)$$

where $Death^{*}_{ijt}$ is the true probability of death, but we only observe:

$$Death_{ijt} = Death^{*}_{ijt} + \eta_{ijt}$$

$$\eta_{ijt} = \zeta_{jt} + \phi_{ijt}$$

The measurement error is η_{ijt}, with family-and birth-specific (ζ_{jt}) and twin-specific (ϕ_{ijt}) components. Since the mortality outcomes can only take two values (0 or 1), the measurement error will be equal to 0 if the mortality outcome is measured without error and -1 otherwise. That is, the measurement error will be equal -1 if an infant in the death records is not matched in one of the births. Thus, the omitted variable formula implies that the within-twin estimator of β in the equation (4) is given by

$$\beta_{FE} = \beta + \frac{cov(bw_{1jt} - bw_{2jt}, \phi_{1jt} - \phi_{2jt})}{var(bw_{1jt} - bw_{2jt})}$$

The second term of the right-hand side is the resulting bias from the measurement error. Note that only a significant correlation between birth weight and the twin-specific component of the measurement error would lead to bias in our estimates. As the measurement error is a deterministic function of the probability of being matched, the following regression may be used to determine the importance of bias induced by measurement error:

$$Birthweight_{ijt} = \delta Matched_{ijt} + \varphi_{jt} + \xi_{ijt} \qquad (5)$$

where $Matched_{ijt}$ is a dummy variable indicating whether the infant death record i was matched to one of the birth records. The twin-fixed effects are represented by φ_{jt}, while that ξ_{ijt} is an idiosyncratic error term. The parameter δ measures the importance of the bias induced by measurement error. If we are unable to reject the hypothesis that $\delta = 0$, then we would conclude that the measurement error is unlikely to bias our estimates of the effect of birth weight on infant mortality.

Table 1. Summary Statistics

	Singletons	Twins	Same-sex male Twins	Same-sex female twins
Characteristics of birth				
Birth weight (in grams)	3,202.44	2,322.16	2,334.93	2,270.65
	(533.40)	(573.91)	(593.68)	(555.82)
Fraction low birth weight (<2,500 grams)	0.07	0.59	0.56	0.63
	(0.26)	(0.49)	(0.50)	(0.48)
Fraction preterm births (<37 weeks)	0.07	0.47	0.48	0.47
	(0.26)	(0.50)	(0.50)	(0.50)
1 minute APGAR score	8.25	7.77	7.71	7.78
	(1.25)	(1.61)	(1.68)	(1.59)
5 minute APGAR score	9.31	8.98	8.94	8.98
	(0.90)	(1.19)	(1.26)	(1.18)
Fraction C-section	0.50	0.79	0.78	0.78
	(0.50)	(0.41)	(0.41)	(0.41)
Fraction male	0.51	0.49	1.00	0.00
	(0.50)	(0.50)	(0.00)	(0.00)
Mother's characteristic				
Fraction high education (12 or more years of schooling)	0.16	0.21	0.20	0.20
	(0.37)	(0.41)	(0.40)	(0.40)
Fraction married	0.34	0.40	0.40	0.40
	(0.48)	(0.49)	(0.49)	(0.49)
age	25.47	27.61	27.21	27.20
	(6.47)	(6.41)	(6.42)	(6.44)
Mortality outcomes				
one-year mortality rate (per 1,000 births)	8.13	36.88	43.89	34.35
	(89.79)	(188.47)	(204.85)	(182.14)
Neonatal mortality rate (per 1,000 births)	6.13	31.29	37.61	29.10
	(78.04)	(174.09)	(190.26)	(168.10)
Seven-day mortality rate (per 1,000 births)	4.78	24.46	29.92	22.27
	(68.95)	(154.47)	(170.37)	(147.55)
one-day mortality rate (per 1,000 births)	2.86	14.11	17.51	12.92
	(53.39)	(117.94)	(131.16)	(112.92)
N	18,929,949	255,362	90,976	93,904

Source. Authors.

Note. Standard deviations are given in parentheses.

Table 2. Summary statistics: heavier versus lighter twins

	Heavier (1)	Lighter (2)
Birth weight:		
Mean	2456.5	2182.9
	(575.47)	(537.91)
Median	2,535	2,250
Twenty-fifth percentile	2,180	1,905
Tenth percentile	1,720	1,475
Fifth percentile	1,320	1,130
First percentile	665	575
Fraction low birth weight (<2,500 grams)	0.46	0.70
Mortality outcomes:		
one-year mortality rate (per 1,000 births)	34.19	39.66
	(181.72)	(195.17)
Neonatal mortality rate (per 1,000 births)	29.50	33.13
	(169.21)	(178.97)
Seven-day mortality rate (per 1,000 births)	23.58	25.35
	(151.76)	(157.21)
one-day mortality rate (per 1,000 births)	14.28	13.93
	(118.64)	(117.20)

Sources. Authors.
Note. Standard deviations are given in parentheses.

Table 3. Components of variance for birth weight and outcomes among twins

	Mean squared error in OLS regressions			Ratio
	(1)	(2)	(3)	(3)/(2)
Birth weight	32.93	16.8	3.52	0.20
One year mortality	0.035	0.024	0.008	0.35
Neonatal mortality	0.030	0.020	0.006	0.34
Seven-day mortality	0.023	0.016	0.005	0.31
one-day mortality	0.013	0.010	0.002	0.28
Controls for:				
Gestation length dummies	No	Yes	-	
Twin-fixed effects	No	No	Yes	

Source. Authors.

Notes. Columns (1)–(3) provide the means squared error from OLS regressions that include no controls, dummies for gestation length (less than 22 weeks, 22-27 weeks, 28-31 weeks, 32-36 weeks, and 37-41 weeks), and twin-fixed effects, respectively. The final column provides the ratio of column (3) to column (2). Birth weight is measured in 100s of grams. The sample size is 255,362.

Table 4. OLS and Twin-Fixed effects of the relationship between birth weight and infant mortality

	One-year mortality	Neonatal mortality	Seven-day mortality	One-day mortality
	(1)	(2)	(3)	(4)
	Panel A: OLS - singleton sample			
Birth weight < 1,500	0.319	0.286	0.228	0.142
	[0.001]***	[0.001]***	[0.001]***	[0.001]***
Birth weight 1,500-2,500	0.021	0.016	0.012	0.007
	[0.000]***	[0.000]***	[0.000]***	[0.000]***
Birth weight 2,500-3,000	0.002	0.001	0.001	0.001
	[0.000]***	[0.000]***	[0.000]***	[0.000]***
R^2	0.132	0.141	0.115	0.074
N	18,929,949	18,929,949	18,929,949	18,929,949
	Panel B: OLS - Twins sample			
Birth weight < 1,500	0.328	0.296	0.238	0.145
	[0.004]***	[0.004]***	[0.004]***	[0.003]***
Birth weight 1,500-2,500	0.013	0.01	0.007	0.003
	[0.001]***	[0.000]***	[0.000]***	[0.000]***
Birth weight 2,500-3,000	0.001	0.001	0.001	0.000
	[0.000]***	[0.000]***	[0.000]**	[0.000]
R^2	0.227	0.219	0.181	0.115
N	255,362	255,362	255,362	255,362
	Panel C: FE - Twins sample			
Birth weight < 1,500	0.057	0.04	0.024	0.004
	[0.007]***	[0.006]***	[0.005]***	[0.003]
Birth weight 1,500-2,500	0.007	0.004	0.003	0.002
	[0.001]***	[0.001]***	[0.001]***	[0.001]**
Birth weight 2,500-3,000	0.003	0.002	0.001	0.000
	[0.001]**	[0.001]*	[0.001]	[0.001]
R^2	0.755	0.77	0.777	0.779
N	255,362	255,362	255,362	255,362

Source. Authors.

Notes. The standard errors are in parentheses and are corrected for heteroskedasticity. In addition, Panels B and C use standard errors corrected for within-twin-pair correlation in the residuals. All regressions control for infant's sex. In addition, regressions in Panels C controls for twin-fixed effects. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.1.

Table 5. Twin-Fixed effects of the relationship between birth weight and infant mortality
(The role of Zigosity)

	One-year mortality (1)	Neonatal mortality (2)	Seven-day mortality (3)	One-day mortality (4)
	Panel A: Male same-sex twins			
Birth weight < 1,500	0.069	0.047	0.029	0.004
	[0.013]***	[0.012]***	[0.010]***	[0.006]
Birth weight 1,500-2,500	0.008	0.006	0.004	0.001
	[0.002]***	[0.002]***	[0.002]**	[0.001]
Birth weight 2,500-3,000	0.002	0.001	0.000	0.000
	[0.002]	[0.002]	[0.001]	[0.001]
R^2	0.761	0.774	0.782	0.788
N	90,976	90,976	90,976	90,976
	Panel B: Female same-sex twins			
Birth weight < 1,500	0.052	0.036	0.020	-0.000
	[0.011]***	[0.009]***	[0.008]***	[0.005]
Birth weight 1,500-2,500	0.006	0.003	0.002	0.000
	[0.002]**	[0.002]*	[0.002]	[0.001]
Birth weight 2,500-3,000	0.003	0.001	0.000	-0.000
	[0.002]	[0.001]	[0.001]	[0.001]
R^2	0.759	0.773	0.777	0.777
N	93,904	93,904	93,904	93,904

Source. Authors.
Notes. The standard errors are in parentheses and are corrected for heteroskedasticity and within-twin-pair correlation in the residuals. All regressions control for infant's sex and twin-fixed effects. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.1.

Table 6. Twin-Fixed effects of the relationship between birth weight and one-year mortality (by cause of death)

	Conditions originating in the perinatal period	Infectious and parasitic diseases	Respiratory diseases	Other Diagnoses
	(1)	(2)	(3)	(4)
Birth weight < 1,500	0.037	0.004	0.002	0.013
	[0.006]***	[0.002]**	[0.001]*	[0.003]***
Birth weight 1,500-2,500	0.004	0.000	0.001	0.002
	[0.001]***	[0.000]	[0.000]	[0.001]**
Birth weight 2,500-3,000	0.002	0.000	0.000	0.000
	[0.001]***	[0.000]	[0.000]	[0.001]
R^2	0.775	0.536	0.513	0.523
N	255,362	255,362	255,362	255,362
Mean of dependent variable	0.031	0.001	0.0008	0.0035

Source. Authors.

Notes. The standard errors are in parentheses and are corrected for heteroskedasticity and within-twin-pair correlation in the residuals. All regressions control for infant's sex and twin-fixed effects. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.1.

Table 7. Twin-Fixed effects of the relationship between birth weight and infant mortality (Alternative specifications)

	One-year Mortality	Neonatal Mortality	Seven-day mortality	One-day mortality
	(1)	(2)	(3)	(4)
	Panel A: Twin FE- linear specification			
Birth weight (in grams)	-0.026	-0.019	-0.012	-0.003
	[0.002]***	[0.002]***	[0.002]***	[0.001]***
	Panel B: Twin FE- linear-log specification			
Log(Birth weight)	-76.015	-59.351	-36.562	-10.969
	[6.069]***	[5.596]***	[4.877]***	[3.686]***
	Panel C: Logit model with Twin FE- linear-log specification			
Birth weight	-0.002	-0.002	-0.001	-0.001
	[0.000]***	[0.000]***	[0.000]***	[0.000]***
	Panel D: Logit model with Twin FE- linear-log specification			
Log(Birth weight)	-2.566	-2.317	-1.801	-0.878
	[0.153]***	[0.163]***	[0.176]***	[0.210]***
	Panel E: Logit model with Twin FE- discrete birth weight categories			
Birth weight < 1,500	2.067	1.942	1.545	1.04
	[0.228]***	[0.294]***	[0.332]***	[0.439]**
Birth weight 1,500-2,500	1.22	1.222	0.948	0.918
	[0.212]***	[0.281]***	[0.317]***	[0.421]**
Birth weight 2,500-3,000	0.549	0.491	0.217	0.195
	[0.205]***	[0.273]*	[0.318]	[0.422]

Source. Authors.

Notes. The standard errors are in parentheses and are corrected for heteroskedasticity and within-twin-pair correlation in the residuals. All regressions control for infant's sex and twin-fixed effects. Panels C, D, and E show estimated coefficients from logit models with twin-fixed effects. These logit models only includes cases in which one twin lives and one twin dies, implying a sample size of 8,904 for one-year mortality, 7,130 for neonatal mortality, 5,444 for seven-day mortality, and 3,142 for one-day mortality. In Panels A and B, we have multiplied the coefficients and standard errors by 100 to make them easier to read. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.1.

Table 8. OLS and Twin-Fixed effects of the relationship between birth weight and APGAR scores

	1 minute APGAR score (1)	low 1 minute APGAR score (<8) (2)	5 minute APGAR score (3)	low 5 minute APGAR score (<8) (4)
		Panel A: OLS - Twins sample		
Birth weight < 1,500	-2.489	0.532	-1.843	0.332
	[0.023]***	[0.005]***	[0.021]***	[0.004]***
Birth weight 1,500-2,500	-0.407	0.117	-0.264	0.032
	[0.010]***	[0.003]***	[0.007]***	[0.001]***
Birth weight 2,500-3,000	-0.067	0.019	-0.044	0.004
	[0.009]***	[0.003]***	[0.007]***	[0.001]***
R^2	0.161	0.096	0.164	0.128
N	249,701	249,701	249,440	249,440
		Panel B: FE - Twins sample		
Birth weight < 1,500	-0.36	0.09	-0.138	0.024
	[0.045]***	[0.013]***	[0.030]***	[0.008]***
Birth weight 1,500-2,500	-0.11	0.032	-0.038	0.005
	[0.019]***	[0.006]***	[0.012]***	[0.003]**
Birth weight 2,500-3,000	-0.043	0.009	-0.013	0.003
	[0.016]***	[0.005]*	[0.010]	[0.002]
R^2	0.784	0.737	0.833	0.754
N	249,701	249,701	249,440	249,440

Source. Authors.

Notes. The standard errors are in parentheses and are corrected for heteroskedasticity. In addition, Panel B uses standard errors corrected for within-twin-pair correlation in the residuals. All regressions control for infant's sex. In addition, regressions in Panel B control for twin-fixed effects. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.1.

Table 9. Comparison with Literature about developed countries

	Infant mortality rate (per 1000 births)	Mean of birth weight	Specification using birth weight (grams)		Specification using log of Birth weight	
			Effect size	Elasticity	Effect size	Elasticity
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Brazil						
Infant mortality	36.88	2,322	-0.026	-1.63	-76.01	-2.06
			[0.002]***		[6.06]***	
Panel B: Estimates from the US and Norway:						
Almond, Chay and Lee (2005)	38.71	2,417	-0.008	-0.51	-	-
			[0.001]***			
Black, Devereux and Salvanes (2007)	31.11	2,598	-0.010	-0.83	-41.1	-1.32
			[0.003]***		[7.74]***	

Source. Authors.

Notes. In Panel A, each column presents the results of a specification that use birth weight (in grams) and log of birth weight as the primary independent variable of interest. The dependent variable is mortality within one year of birth (per 1,000 births). All regressions control for infant's sex and twin-fixed effects. Panel B presents the corresponding estimates previous studies for developed countries. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.10.

Table 10. Twin-Fixed effects of the relationship between birth weight and one-year mortality (by mother's education and marital status)

	Less-educated mothers (1)	More-educated mothers (2)	Unmarried (3)	Married (4)
Birth weight < 1,500	0.063	0.034	0.072	0.036
	[0.008]***	[0.012]***	[0.009]***	[0.010]***
Birth weight 1,500-2,500	0.008	0.003	0.007	0.007
	[0.002]***	[0.002]	[0.002]***	[0.002]***
Birth weight 2,500-3,000	0.003	0.001	0.003	0.003
	[0.001]**	[0.002]	[0.002]*	[0.002]
Test of equality of coefficients:				
$\chi 2$		61.998		76.098
p-value		0.000		0.000
R^2	0.753	0.767	0.758	0.749
N	200,870	53,542	151,856	103,506

Source. Authors.

Notes. The standard errors are in parentheses and are corrected for heteroskedasticity and within-twin-pair correlation in the residuals. All regressions control for infant's sex and twin-fixed effects. Less-educated mothers refer to mothers who have 11 years of schooling or less. More-educated mothers refer to mothers who have 12 years of schooling or more. The dependent variable is mortality within one year of birth. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.1.

Table 11. Twin-Fixed effects of the relationship between birth weight and infant mortality (by GDP and sanitation coverage)

	Municipality GDP at the:					*% sanitation coverage at the municipality:*			
	1st quintile (1)	2nd quintile (2)	3rd quintile (3)	4th quintile (4)	5th quintile (5)	<20 (6)	20-50 (7)	50-85 (8)	>85 (9)
Birth weight < 1,500	0.067	0.056	0.084	0.068	0.045	0.063	0.062	0.077	0.037
	[0.030]**	[0.026]**	[0.020]***	[0.016]***	[0.009]***	[0.02]***	[0.02]***	[0.01]***	[0.01]***
Birth weight 1,500-2,500	0.022	0.004	0.008	0.006	0.005	0.012	0.010	0.005	0.005
	[0.006]***	[0.004]	[0.004]**	[0.003]*	[0.002]***	[0.004]***	[0.004]**	[0.003]**	[0.002]**
Birth weight 2,500-3,000	0.009	-0.002	0.003	0.003	0.002	0.004	0.004	0.002	0.002
	[0.005]**	[0.003]	[0.003]	[0.003]	[0.002]	[0.003]	[0.003]	[0.002]	[0.002]
Test of equality of coefficients									
$\chi 2$		11.342	9.414	23.470	43.428		3.574	35.429	81.371
p-value		0.010	0.024	0.000	0.000		0.311	0.000	0.000
R^2	0.734	0.754	0.755	0.759	0.759	0.739	0.736	0.766	0.761
N	21,004	27,602	33,788	47,418	125,536	40,738	34,878	81,878	97,868

Source. Authors.

Notes. The standard errors are in parentheses and are corrected for heteroskedasticity and within-twin-pair correlation in the residuals. All regressions control for infant's sex and twin-fixed effects. Per capita GDP is expressed in constant 2000 prices. The data on GDP and sanitation coverage are obtained from the Brazilian Institute of Geography and Statistics (IBGE). For the analysis by GDP quintile, The test of equality of coefficients compares the results from column (1) to those from columns (2)-(5). For the analysis by sanitation coverage, the test of equality of coefficients compares the results from column (6) to those from columns (7)-(9). The dependent variable is mortality within one year of birth. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.1.

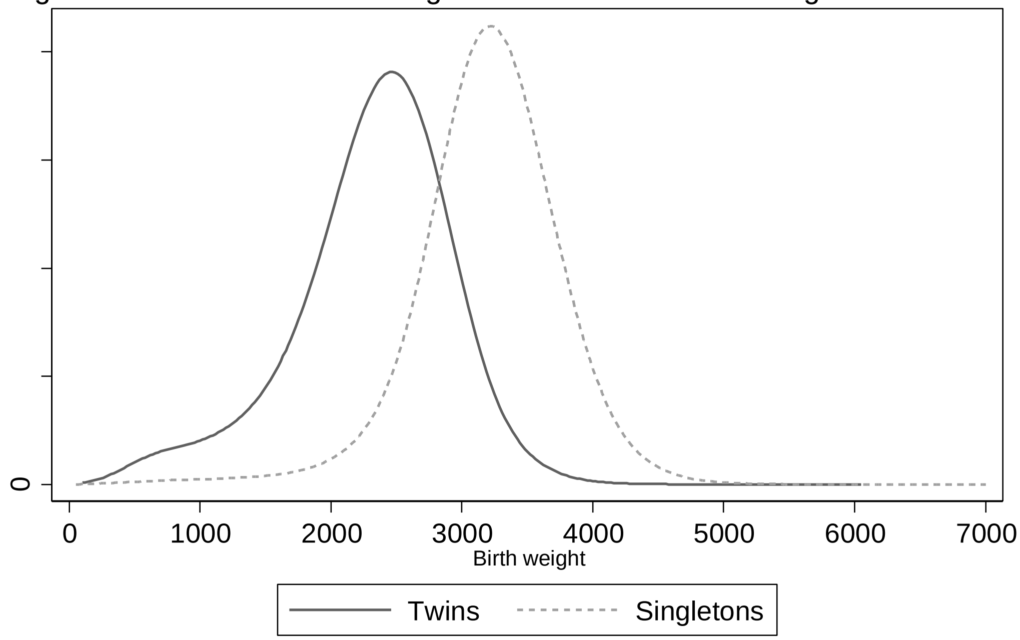
Table 12. Twin-Fixed effects of the relationship between birth weight and one-year mortality (Heterogeneous effects)

	(1)	(2)	(3)	(4)	(5)	(6)
Birth weight < 1,500	0.057 [0.005]***	0.064 [0.006]***	0.072 [0.007]***	0.070 [0.007]***	0.070 [0.008]***	0.086 [0.009]***
Birth weight 1,500-2,500	0.007 [0.001]***	0.008 [0.001]***	0.007 [0.001]***	0.008 [0.001]***	0.009 [0.001]***	0.009 [0.002]***
Birth weight 2,500-3,000	0.003 [0.001]***	0.003 [0.001]***	0.003 [0.001]***	0.003 [0.001]***	0.003 [0.001]**	0.003 [0.001]**
(Birth weight < 1,500) interacted with:						
More-educated mothers		-0.030 [0.010]***				-0.013 [0.011]
Married			-0.036 [0.010]***			-0.030 [0.010]***
Sanitation coverage (>85 %)				-0.033 [0.010]***		-0.024 [0.011]**
GDP at 5th quintile					-0.025 [0.010]**	-0.008 [0.012]
(Birth weight 1,500-2,500) interacted with:						
More-educated mothers		-0.005 [0.002]***				-0.005 [0.002]**
Married			-0.000 [0.002]			0.001 [0.002]
Sanitation coverage (>85 %)				-0.002 [0.002]		-0.000 [0.002]
GDP at 5th quintile					-0.003 [0.002]*	-0.003 [0.002]
(Birth weight 1,500-2,500) interacted with:						
More-educated mothers		-0.002 [0.001]				-0.002 [0.002]
Married			-0.000 [0.002]			0.000 [0.002]
Sanitation coverage (>85 %)				-0.001 [0.002]		-0.001 [0.002]
GDP at 5th quintile					-0.001 [0.002]	-0.000 [0.002]
R^2	0.755	0.755	0.755	0.755	0.755	0.755
N	255,362	255,362	255,362	255,362	255,348	255,348

Source. Authors.

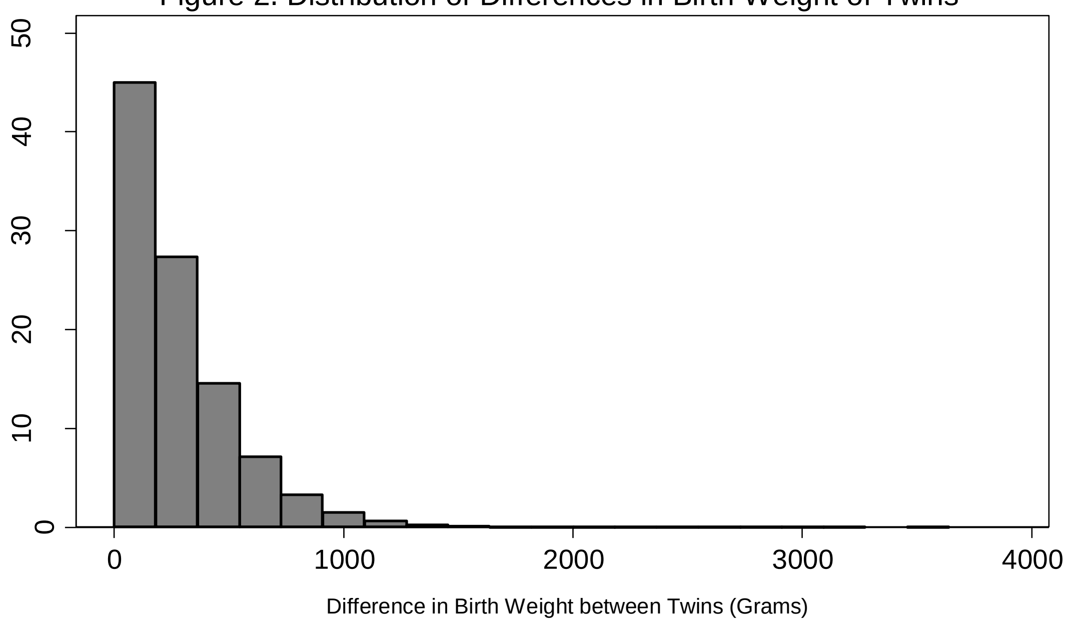
Notes. The standard errors are in parentheses and are corrected for heteroskedasticity and within-twin-pair correlation in the residuals. All regressions control for infant's sex and twin-fixed effects. More-educated mothers refer to mothers who have 12 years of schooling or more. The dependent variable is mortality within one year of birth. Statistical significance is denoted by: ***p < 0.01, **p < 0.05, *p < 0.1.

Figure 1. Difference in birth weight distributions between singletons and twins



Notes. Figure 1 plots kernel density distributions of infant birth weight for twins (solid line) and singletons (dashed line) in our sample.

Figure 2. Distribution of Differences in Birth Weight of Twins



Notes. Each bar represents the percentage of twins whose birth weight difference falls within the specified range. The mean birth weight difference among twins in our sample is 276 grams.

Figure 3. Relationship between infant mortality and birth weight

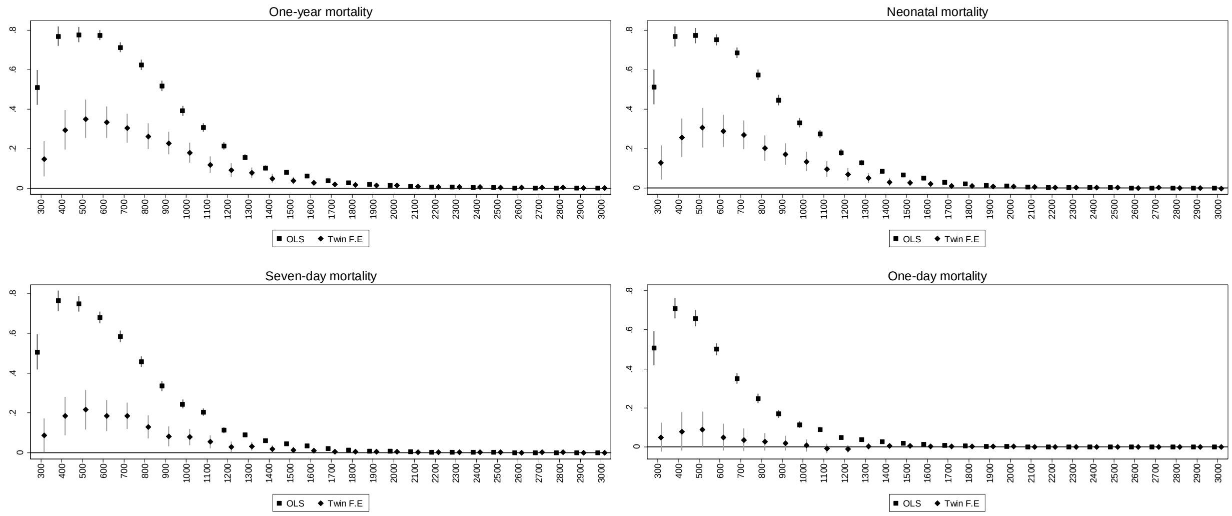


Figure 3 plots the coefficients from the equation (3), which is estimated using either OLS or Twin-fixed effects. We use 27 dummy variables corresponding to 100 gram-wide birth weight bins of the distribution of birth weight below 3,000 grams. The bins range from a low of 300-400 gr to a high of 2,900-3,000 grams.