Short-, Medium-, and Long-Term Consequences of Poor Infant Health

An Analysis Using Siblings and Twins

Philip Oreopoulos
Mark Stabile
Randy Walld
Leslie L. Roos

ABSTRACT

We use administrative data on a sample of births between 1978 and 1985 to investigate the short-, medium-, and long-term consequences of poor infant health. Our findings offer several advances to the existing literature on the effects of early infant health on subsequent health, education, and labor force attachment. First, we use a large sample of both siblings and twins, second, we use a variety of measures of infant health, and finally, we track children through their schooling years and into the labor force. Our findings suggest that poor infant health predicts both mortality within one year, and mortality up to age 17. We also find that infant health is a strong predictor of educational and labor force outcomes. In particular, infant health is found to predict both high school completion and welfare takeup and length.

Philip Oreopoulos is an associate professor of economics at the University of Toronto. Mark Stabile is an associate professor of business economics and public policy at the University of Toronto. Randy Walld is an information technologist at the University of Manitoba. Leslie Roos is a professor of community health science at the University of Manitoba. *The authors gratefully acknowledge those individuals who helped make this research possible. These include the Manitoba Ministry of Education, Citizenship and Youth—John Van Wallenghem, Richard Perrault, Carol Crera, Jean Britton, Ken Clark, and Shirley McLellan; from the Ministry of Family Services and Housing—Harvey Stevens, Grant Doak, Gudrun Fritz, and Jan Forster; and from the Ministry of Health—Louis Barre. Janet Currie and Sandra Black provided very helpful comments. Stabile and Roos thank the Canadian Institutes for Health Research for financial support. Oreopoulos thanks the Social Sciences and Humanities Research Council of Canada (MH:2004/2005–27). Additional support was provided by the Canadian Institute for Advanced Research and the RBC Financial Group (MH:2003/2004-32). The authors also thank Florian Hoffman for excellent research assistance. The data used in this article are housed at the Manitoba Centre for Health Policy. Information on access can be obtained by contacting Leslie Roos, Manitoba Centre for Health Policy, 4th Floor, 727 McDermot Avenue, Winnipeg, MB R3E 3P5 Canada; e-mail:Leslie_Roos@cpe.umanitoba.ca Beginning from August 2008 through July 2011, All codes to replicate the tables in this paper will be kept on file during this time*
ISSN 022-166X E-ISSN 1548-8004

I. Introduction

Infants born in poor health, as measured by low and very low birth weights and low Apgar scores, have lower chances of survival, and also may experience further health and social difficulties later in life (Conley 2003). Low-birth-weight babies are also increasingly expensive to treat in hospital. Almond et al. (2005) calculate that among babies weighing 2000 grams, an additional 450 grams is associated with a $10,000 savings in hospital charges for inpatient services. As such, understanding the causes and consequences of poor infant health has been a primary concern of both the medical and health policy literature for some time.

Medical advice to expecting mothers on how to prevent low birth weight, including refraining from smoking and seeking prenatal care, is centered around the notion that preventing low birth weight will improve both the life chances of the child and chances of future success. Researchers also have noted the potential to reduce hospital costs significantly through inexpensive prenatal interventions aimed at reducing low birth weight in particular (Almond et al. 2005). Program evaluations on both Medicaid expansions in the United States (Currie and Gruber 1996) and the introduction of national health insurance in Canada (Hanratty 1996) have examined improved prenatal treatment and its potential effects on infant health, providing further evidence of the policy importance of, and potential benefits associated with, improving infant health.

As noted in Almond et al. (2005), interventions aimed particularly at reducing low birth weight are premised on the notion that low birth weight in particular is the cause of poor health and related outcomes in the future, and not simply a marker and correlate of such problems. While interventions and public policy aimed at improving overall infant health, including reducing the incidence of low birth weight, are likely to have both short- and long-term benefits, a clearer understanding of the causes and consequences of poor infant health only can help to improve the efficacy of both healthcare and public policy. An analysis of the long-term impact of infant health also may uncover important relationships not realized from focusing on earlier outcomes. Infants born lower than average birth weight but not considered at risk of early death, for example, may in fact benefit from prenatal care. Or, the majority of low-birth-weight infants that survive past one year may face few subsequent risks.

A considerable body of research attempts to quantify the effects of early infant health on both early childhood survival and on future health, education, and social outcomes. Conley, Strully, and Bennett (2003), for example, examined the effects of birth weight for both fraternal and identical twins on both neonatal and post-neonatal mortality. They conclude that birth weight differences between twins affects infant mortality and that this effect is stronger for fraternal than identical twins. Almond et al. (2005) examine the relationship between low birth weight, low Apgar scores, and mortality in the first year of life. Using a large sample of twin births from the National Center for Health Statistics, they show that, while both birth weight and Apgar scores are strongly related to infant mortality across families, the relationship between birth weight and infant mortality significantly decreases when differences between twins are examined. In contrast, the relationship between Apgar scores and infant mortality remains strong both across families and within twin pairs. Both papers note that, while twin samples can be extremely helpful in eliminating unobserved heterogeneity across families, the resulting sample is somewhat unique in that twins tend to be of lower weight than the average in the singleton infant population.

A second stream of social science literature has used twin studies to examine the longer-term effects of birth weight on health and education. Behrman and Rosenzweig (2004) use twin data from the Minnesota Twins Registry to examine the effects of low birth weight on the educational attainment and adult health of women. They find that increasing birth weight increases schooling attainment by about one-third of a year and that this effect is stronger within twins than across children of different families. Conley, Strully, and Bennett (2003) examine the effects of low birth weight on high school graduation and placement in special education using the Panel Study of Income Dynamics. They find that the effects of low birth weight on timely high school graduation are more pronounced among siblings than across families. The study does not look at other measures of infant health (Apgar and gestation) nor does it explore the potential nonlinear effects of low birth weight on infant health. Black et al. (2007) use a sample of Norwegian twins to examine the long-run consequences of low birth weight. Their evidence confirms that low birth weight is not a good predictor of infant death within twin pairs. However, they do find long-term effects of low birth weight on cognitive outcomes, educational outcomes, and on earnings.

Our paper expands on the results of previous work in the following ways: First, it uses an administrative sample of both siblings and twins to examine the effects of infant health on mortality within one year. Comparing sibling findings and twin findings allows us to overcome concerns that twins are a select sample of the population and that inference from this sample is not, therefore, generalizable to the broader population. Second, tracking both siblings and twins through school and into their early experiences in the work force provides longer-term evidence for both groups, including educational outcomes, healthcare costs, and social assistance receipt. Third, a variety of infant health measures, including birth weight, Apgar scores, and gestational length, are used to contrast the effects of these measures on outcomes and to reconcile and expand the findings of other research using multiple measures of infant health. Gestational length is an important determinant of low birth weight, one which twin only studies are unable to examine. Finally, using a sample of children with uniform access to health insurance further corrects for any potential unobserved heterogeneity *within* families across siblings that might not be captured in sibling-fixed-effects models and offers an interesting comparison with a U.S. sample lacking universal coverage.

Our findings offer several advances to the existing literature on the effects of early infant health on subsequent health, education, and labor force attachment. First, we confirm earlier results by Almond et al., which show that the effect of infant health as measured by birth weight less than 2,500 grams largely disappears when looking at within twin variation. The five-minute Apgar score and measures of very low birth weight (less than 1,500 grams) are stronger predictors of infant mortality within one year than birth weight for twin samples. However, we find that within sibling pairs Apgar, low birth weight, and gestational age predict infant mortality within one year, even though we continue to account for unobserved heterogeneity across families. Second, infant health is found to predict both high school completion and social assistance (welfare) takeup and length. We find evidence of longer-term consequences of infant health both across families, within siblings, and within twin pairs, although different measures of infant health predict outcomes differently. The results suggest strong effects of infant health on death between ages one and 17, grade completion, and months on social assistance after age 18, even for ranges not considered overtly concerning (for example, birth weights between 2,500 and 3,500 grams and Apgar scores of seven or eight). The results are similar comparing families living in more or less disadvantaged neighborhoods. Interestingly, we find weaker evidence of the longer-term effects of infant health on either cognitive ability as measured by language arts test scores or longer-term physician visits and costs. Overall, we conclude that there are indeed long-term consequences of poor infant health, and that a better understanding of these consequences can be determined by examining a variety of infant health measures and by examining the variation both within families and within twin pairs. The implication of these findings is that reductions in poor infant health will lead to lower mortality, greater human capital accumulation, and lower welfare usage.

II. Data

The data are from the Population Health Research Data Repository at the Manitoba Centre for Health Policy (MCHP). Our main data match hospital records at birth to other administrative records on education, physician visits, and social-assistance takeup. We also match socioeconomic characteristics at the postal code level using the 2001 Canadian Census. The sample includes more than 96 percent of all children born in Manitoba in 1978–82 and 1984–85 and more than 99 percent of this group remaining in the province up to June of their 18th year.[1] Health, educational, and social assistance outcomes are tracked up to 2004.[2] Further details on the construction of the data set are available in the data appendix (see the *JHR* website, www.ssc.wisc.edu/JHR/).

Table 1 presents descriptive statistics of the infant health measures recorded on the hospital records and used in our study: birth weight (in grams), gestation (in weeks), and five-minute Apgar score (on a ten-point scale). Means, standard deviations, and percentiles for these measures are shown for the full sample of births between 1979 and 1985. These statistics are also shown for the subset sample of births with at least two siblings identified within this cohort range and the subset sample of twins. The sibling sample excludes twins.

The frequency distributions of these variables compare similarly with those generated from nationally representative samples of Canada or the United States. The mean birth weight among the full sample is about 3,500 grams. Twins weigh about 950 grams less and are born about three weeks earlier. About 7 percent of the full sample is born low birth weight, defined as weighing less than 2,500 grams. In the analysis below, we explore not only the effects of being born less than 2,500 grams and less than 1,500 grams, but also the effects of being born below average birth weight, between 2,500 to 3,000 grams and between 3,001 to 3,500 grams. Gestation

1. The cohort born in 1983 was not included because grade 12 provincial tests were not given in the school year 2000/2001 (when the 1983 birth cohort would be expected to be in grade 12).

2. In Canada, welfare is more commonly referred to as social assistance. For consistency, we maintain this terminology.

Table 1
Descriptive Statistics of Infant Health Measures

				5 Minute APGAR Score (0–10)						
			Within Family	Percentile						
Sample	Mean	s.d.	s.d.	1st	5th	10th	25th	50th	75th	N
All births 1979–85	9.094	0.928	NA	6	8	8	9	9	10	108,800
Siblings only	9.131	0.919	0.646	6	8	8	9	9	10	54,123
Twins only	8.594	1.397	0.625	1	6	7	9	9	9	1,742
				Birth Weight (in Grams)						
			Within Family	Percentile						
Sample	Mean	s.d.	s.d.	1st	5th	10th	25th	50th	75th	N
All births 1979–85	3,424.5	570.3	NA	1,680	2,500	2,760	3,100	3,450	3,780	109,125
Siblings only	3,458.3	556.6	314.1	1,780	2,580	2,820	3,140	3,480	3,800	54,986
Twins only	2,517.4	610.6	201.9	790	1,360	1,725	21,81.5	2,580	2,930	1,752
				Gestation (in Weeks)						
			Within Family	Percentile						
Sample	Mean	s.d.	s.d.	1st	5th	10th	25th	50th	75th	N
All births 1979–85	39.4	2.0	NA	32	36	37	39	40	40	90,135
Siblings only	39.4	1.9	1.1	33	37	38	39	40	40	45,583
Twins only	36.5	3.3	0.0	25	30	32	35	37	39	1,492

before birth typically takes about 40 weeks. Preterm births are often defined as births before 37 weeks gestation; there are 7 percent preterm births in our full sample. Late births occur after 41 weeks.

The Apgar score summarizes five vital sign conditions at birth. Heath care providers assess an infant's heart rate, respiration, muscle tone, reflex, and color; they assign values of zero, one, or two for each category, with the best possible total score equaling ten. A score less than seven often triggers additional action to stabilize conditions. A score of seven to ten is considered normal. As shown below, lower five-minute Apgar scores even within this normal range affect subsequent educational outcomes and social assistance takeup.

The typical variation in these infant health measures between a pair of siblings or a pair of twins is about 55 to 70 percent of the typical variation between any randomly chosen infant pair. Column 2 of Table 1 lists standard deviation for each variable, across all individuals. Column 3 shows standard deviations in these infant health measures within families, among siblings and twins. These amounts are the standard deviations of the residuals generated after regressing the health measures on a set of family fixed effects. The standard deviation for Apgar scores is about 0.92 over the full sample and 0.65 within families. The standard deviation for gestation is about two weeks and one week between siblings. The within family standard deviation of birth weight is still 314 grams between siblings, and 202 grams between twins. In perspective, Almond, Chay, and Lee (2005) report that the average difference in birth weight between a newborn with a mother who smokes and one with a mother who does not is 285 grams.[3] The average difference in gestation is 0.3 weeks, while the average difference in five-minute Apgar score is 0.07. Our main analysis uses within family variation in infant health to explore short- and long-run differences in socioeconomic outcomes. Column 3 indicates considerable variation within families to work with in exploring later outcome differences.

We observe differences in infant health across both siblings and twins for several reasons. Between siblings, birth weight can differ due to both gestational length and differences in intrauterine growth retardation (IUGR). Between twins, differences in birth weight are mainly attributable to differences in IUGR. Twin studies have emphasized that much of the literature has focused on differences in IUGR, despite the fact that gestational length accounts for a significant percent of the low-birth-weight infants. One possible reason for this, as noted by Almond et al. (2005) and reported in Goldenberg and Rouse (1998), is that there is little medical evidence on how to effectively increase gestational length, whereas there are widely accepted policy interventions aimed at IUGR (reducing smoking and ensuring appropriate nutrition during pregnancy, are the most common of these). Apgar scores differ between both siblings and twins. The test was initially designed to measure whether infants required immediate medical care and has been shown to be highly correlated with infant mortality (Almond et al. 2005).[4] After testing whether the infant health measures

3. These figures are for the sample of Pennsylvania singletons born between 1989 and 1991. The means, standard deviation, and percentiles reported in Table 1 are similar to those reported by Almond, Chay, and Lee for their Pennsylvania sample.

4. The definition and purpose of the Apgar were obtained from the NIH web site at http://www.nlm.nih.gov/medlineplus/ency/article/003402.htm

Table 2
Descriptive Statistics of Outcome Measures (Sibling Sample) 1979–85 Manitoba Births

	Mean	Standard Deviation	Age of Individual	N
Infant mortality	0.011	0.105	To 365 days	54,310
Death between ages 1 and 17	0.006	0.080	17	53,700
Moved from Manitoba	0.208	0.406	17	53,750
Total physician visits	14.358	12.610	Age 12–17	40,203
Language score (standardized scaled logit)	−0.016	1.013	Grade 12	40,203
Reached grade 12 by age 17	0.694	0.461	17	40,203
Ever on social assistance[a]	0.080	0.271	Age 18 to Mar-04	22,870
Months on social assistance[a]	1.372	5.985	Age 18 to Mar-04	22,870

a. Includes only the cohorts born between 1978 and 1982 to maximize a consistent exposure window for social assistance.

presented here are good predictors of death in the first year, we then consider, conditional on survival, whether they are also predictors of poor health later in life and potentially measures of cognitive ability and human capital as well.

Table 2 lists the health and socioeconomic outcomes explored in our paper. The infant mortality variable comes from matching births and deaths from the Manitoba Vital Statistics over the first year of life. The variable takes on the value of one if a birth is matched to a death in the first year, and zero otherwise. A death between ages one and 17 is similarly recorded.

The other outcome variables came from administrative data on physician costs, education, and social assistance. These data are available only for Manitoba residents. The analysis of the effects of infant health on these longer-term outcomes, therefore, is conditional on survival and conditional on remaining a resident in the province. We focus on estimating the long-term effects of infant health for those born in Manitoba and living in the province at least until they reach 17.5 years old. Table 2 indicates that 24 percent of our original sample of births in Manitoba between 1979 and 1985 either died or left the province before this age. We shall document that health at birth does indeed affect mortality before age 17, even after one year, but it does not affect mobility. For all outcome measures except mortality and mobility, we condition on the sample of those remaining in Manitoba at least until age 17.

The Manitoba Repository data record hospital discharge abstracts and physician claims extending back to 1970. Physician claims include diagnostic information and are primarily reimbursed on a fee-for-service system. We summarize adolescent health by summing the number of ambulatory physician visits recorded between ages 12 and 17.[5] An ambulatory physician visit is any contact with a physician that is

5. We chose to only use ambulatory visits after age 11, as many children have many routine visits before age 12 for immunizations, which are not due to poor health.

billable by the physician to Manitoba Health and occurs while the patient is not a hospital in-patient. This includes physician services received in hospital emergency rooms and outpatient departments, contacts with physicians in salaried positions, consultative and nonconsultative care, and physician visits to residents of personal care homes. Excluded from ambulatory physician visits are all claims for optometrist, oral surgery, dental, periodontal, and chiropractor contacts; inpatient visits (that is, contact with a physician while admitted to a hospital); and laboratory tests, radiology, and similar services. More than 90 percent of this population contacts a physician over a two-year period and the average visit rate is more than four visits per year.

We link education enrollment records with the provincial registry to determine whether a student has attained grade 12 by age 17. Not attaining grade 12 by this age could indicate that a student entered school late for age, has been held back in a grade at least once, or has dropped out. This measure is available for all seven birth cohorts used. Students may not have reached grade 12 because they have been held back. On the other hand, many students held back are more likely to drop out. Our measure proxies as an overall indicator for being at risk of ending up with a low level of education attainment.[6]

We also have information from provincial language arts standards tests taken in grade 12. These tests contribute 30 percent to the students' final course grade. Individuals pass the language arts test by scoring 50 percent or more on a comprehensive exam. The test focuses on reading comprehension, exploring and expanding on ideas from texts, the management of ideas and information, and writing and editing skills. For each birth cohort, we record the test score in five percentage point categories (13 in total, with a residual 14th for students scoring between 0 and 35 percent) in the year that most students write the test. Within each birth cohort, approximately 35 percent of test scores are missing. For these students we impute test scores based on the reason for missing information (ranking them below the lowest scoring category among those who wrote the test) and estimate models both including and excluding imputed values. Details on the imputation methods are available in the data appendix (see the *JHR* website, www.ssc.wisc.edu/jhr/)

Finally, the sample of Manitoba residents is matched to monthly social assistance records up to March 2004. Our youngest birth cohort only can be followed for about a year after the age of 18. The oldest cohorts are followed from age 18 to age 25. Eight percent of our sample received some social assistance before April 2004. In order to avoid censoring issues we define our social assistance exposure window two ways. First, using the cohorts born between 1978 and 1982 we use the maximum exposure to social assistance eligibility possible in our data such that all siblings are observed for the same length of time. This produces an exposure window of 3.25 years. Second, we use the total possible exposure for all birth cohorts. We report only the former here but the results are not sensitive to selecting the sample this way. In case infant health also may affect the length of time on social assistance, we focus on the number of months individuals in our sample used these services. The average number of months on social assistance over our selected sample is 1.4.

6. Approximately 3 percent of the children start school (kindergarten) a year late. Analysis by month of birth shows these children to be disproportionately those born in November and December.

III. Empirical Methods

We estimate models of the effects of early infant health on mortality, health expenditures, educational performance, and social assistance receipt as follows:

$$(1)\quad y_{ijt} = \alpha + \beta \text{inf}\mathit{health}_{ij} + X_{ij} + \delta_j + \tau_t + \varepsilon_{ijt}$$

Where y_{ijt} represents the outcomes for individual i in family j, at time t. X measures family or individual specific controls such as marital status, sex of the child, and mother's age at birth. We also include a set of dummies for the birth order of the child within each family size to completely control for any effects of both birth order and family size.[7] The τ_t are year of birth fixed effects to account for any differences by year of birth of the child. The δ_j are family fixed effects, which, as we outline in greater detail below, are included in some specifications.

Our primary parameter of interest is β, which is the coefficient associated with our estimate of the effect of early infant health. As discussed above, we use three different measures: birth weight, five-minute Apgar score, and gestational length in weeks. Our main analysis classifies these infant health measures into categories and uses dummy variables to estimate possible nonlinear effects of infant health. This approach helps uncover more detailed relationships between infant health and our outcome measures. For example, education attainment may differ by birth weight only for the small fraction born weighing less than 2,500 grams and surviving. In this case, a linear regression model would not adequately capture this relationship. For Apgar score, we estimate effects at birth by comparing scores of six or less, seven to eight, or nine, to a score of ten. For birth weight, we group infants by whether they weigh 1,000 grams or less, 1,001 to 1,500 grams, 1,501 to 2,500 grams, 2,501 grams to 3,000 grams, 3,001 to 3,500 grams, and 3,501 grams or more. For gestation, we compare normal gestation length, between 40 and 41 weeks, to infants born with less than 37 weeks gestation, with 37, 38, or 39 weeks of gestation, and with 42 weeks or more.[8]

For each measure of infant health, we estimate five models: OLS using our entire sample, OLS using the sample of children with siblings, OLS using the sample of twins in the data, the sibling sample including family fixed effects (δ_j) and finally the twin sample including family fixed effects.

One of the advantages of our study is that we are able to examine how the relationship between infant health and our outcome measures change when we use twin and sibling fixed effects with the same data. We are able to compare OLS coefficients to the within-twin and within-sibling estimates, and within the sibling estimates we are able to compare the coefficients when we allow the source of birth weight variation to vary due to

7. Several studies point out that family size correlates with education and other socioeconomic outcomes. Black, Deveraux, and Salvanes (2005), also find important differences in outcomes depending on birth order. We control for these differences with family-size and birth-order fixed effects, in case these variables also relate to infant health. Family size and birth order are based on the final family size and birth order in the family (using all siblings in the data up to 2004). Excluding such controls does not change our baseline results in significant ways.

8. To check how sensitive our results are to functional form, and to allow comparison with studies that use different specifications, we reestimate our models using linear specifications. The results are qualitatively similar and are available in Oreopoulos et al. (2006).

both IUGR and gestational length, to those using IUGR alone (for example, controlling for gestation). Comparing twin and sibling estimates in this way exploits the benefits of both identification strategies. Using the siblings we are able to use a more representative sample of children, but we risk that our estimates are biased due to both the potential change in parental investment following the birth of the first child (noted in Rosenzweig and Wolpin 1995) as well as the potential change in socioeconomic status between births (more on this below). Using the twins, we are able to eliminate these potential biases, but we estimate coefficients for a limited and unrepresentative sample. It also is possible that the patterns of postnatal investment as a function of infant health are different for siblings than for twins and this could also lead to differences in the coefficient estimate.[9] Finding common patterns across these estimation strategies allows for more confidence in the inference. Differences across these methods also may be informative as to the nature of the infant health-future outcome relationship.

Our estimates of Equation 1 serve two purposes. First, we are able to replicate the results found in Almond et al. (2005), contrasting OLS and twin models with family fixed effects, using a smaller sample of Canadian children. The differences using between-family and within-family variation found in that research are shown to hold for this sample of Canadian children as well. Second, we are able to expand on the Almond et al. analysis of the effects of infant health on one-year mortality by estimating fixed effects models using variation in infant health across siblings instead of across twins.

In addition, we estimate alternate versions of Equation 1 using the other outcome measures described above, including: whether the child was held back a grade, the child's language arts test scores measured in grade 12, whether the child dropped out of high school before graduation, and whether the child was on social assistance. Thus, we are able to apply the same OLS, twin, and sibling analyses to a variety of longer-term measures of child health and social outcomes.

IV. Results

Table 3a shows the effects of our measures of infant health on infant mortality (death before age one). The subsequent tables presenting results with different outcome variables have a similar structure. Column 1 displays the coefficients on the infant health categories for the full sample of singletons and siblings, without family fixed effects. These results are from the linear probability model for whether an infant died in the first year regressed on the infant health dummy variables, plus controls for mother's marital status, gender of child, and a complete set of dummy variables for all family size and birth order combinations. The second column shows the same regression, but for the subset sample of births with at least one other sibling identified within the birth cohorts 1979–85 (but excluding twins). In the third

9. As noted in the empirical specification section below, we include a full set of birth-order fixed effects to partially control for this, although differences in the interaction between birth order and infant health could still lead to differences in the results. To explore this possibility we perform a further specification check of interacting birth order and poor infant health on reaching grade 12 by age 17 to explore whether parents respond differently to an early child with poor infant health compared to a later child. We find no significant interaction terms between infant health and birth order.

Table 3a

Estimated Effects of Infant Health at Birth on Infant Mortality (Death within One Year of Birth)With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample With Family F. E.	Twins Sample No Family F. E.	Twins Sample With Family F. E.
	APGAR Score (Omitted Category APGAR=10)				
APGAR≤6	0.2588 ***	0.3092 ***	0.3197 ***	0.3123 ***	0.0957 ***
	(0.0021)	(0.0034)	(0.0047)	(0.0171)	(0.0237)
APGAR=7–8	0.0613 ***	0.0189 ***	0.0198 ***	0.0186 *	−0.0045
	(0.0010)	(0.0015)	(0.0022)	(0.0111)	(0.0155)
APGAR=9	0.0017 ***	0.0025 ***	0.0020	0.0024	−0.0014
	(0.0006)	(0.0009)	(0.0014)	(0.0097)	(0.0147)
F-test: No information about health effects	5,036.06 ***	2,887.21 ***	1,588.14 ***	137.49 ***	8.93 ***
Sample size	{108,893}	{54,091}	{54,091}	{1,740}	{1,740}
R-squared	0.12	0.13	0.50	0.34	0.84
	Birth Weight (Omitted Category BW = 3,500+grams)				
BW<1,000	0.8120 ***	0.8572 ***	0.8723 ***	0.7249 ***	0.2532 ***
	(0.0046)	(0.0071)	(0.0099)	(0.0292)	(0.0701)
BW 1,000–1,500	0.2657 ***	0.3622 ***	0.3912 ***	0.2099 ***	0.0848 **
	(0.0039)	(0.0066)	(0.0093)	(0.0220)	(0.0384)
BW 1,500–2,500	0.0320 ***	0.0479 ***	0.0630 ***	0.0051	0.0048
	(0.0013)	(0.0022)	(0.0034)	(0.0174)	(0.0246)
BW 2,500–3,000	0.0058 ***	0.0065 ***	0.0130 ***	0.0063	−0.0043
	(0.0008)	(0.0012)	(0.0021)	(0.0174)	(0.0231)

BW 3,000–3,500	0.0015 ***	0.0018 ***	0.0042 ***	−0.0006	−0.0017
	(0.0006)	(0.0009)	(0.0014)	(0.0182)	(0.0218)
F-test: no information about health effects	7,099.73 ***	3,586.36 ***	1,899.91 ***	213.85 ***	3.25 ***
Sample size	{109,125}	{54,310}	{54,310}	{1,752}	{1,752}
R-squared	0.25	0.25	0.57	0.48	0.84
	Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	0.0837 ***	0.1060 ***	0.1187 ***	0.0804 ***	NA
	(0.0014)	(0.0023)	(0.0039)	(0.0104)	
Gestation 37 weeks	0.0056 ***	0.0084 ***	0.0142 ***	0.0060	NA
	(0.0016)	(0.0025)	(0.0041)	(0.0141)	
Gestation 38 weeks	0.0035 ***	0.0052 ***	0.0073 ***	0.0044	NA
	(0.0010)	(0.0015)	(0.0026)	(0.0123)	
Gestation 39 weeks	0.0010	0.0005	−0.0003	0.0089	NA
	(0.0008)	(0.0013)	(0.0021)	(0.0132)	
Gestation≥42 weeks	0.0005	−0.0003	0.0005	0.0001	NA
	(0.0013)	(0.0019)	(0.0031)	(0.0237)	
F-test: no information about health effects	722.48 ***	435.15 ***	198.70 ***	7.55 ***	
Sample size	{90,135}	{45,583}	{45,583}	{1,492}	NA
R-squared	0.05	0.05	0.51	0.19	NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

column, the coefficients presented correspond to the regression model that now includes family fixed effects. The fourth and fifth columns show the results among twins, without and with family fixed effects respectively.

The first panel shows the results defining infant health by five-minute Apgar score. Infants born assessed with an Apgar score below seven are about 26 percentage points more likely to die within one year than those with Apgar scores of ten, and 31 percentage points more likely to die among the sibling sample. This relationship holds when we use only differences between siblings in Column 3. The coefficient remains virtually unchanged. However, after adding family fixed effects in the twin sample, the coefficient falls by about two-thirds. The relatively higher association between Apgar and early death is far less severe for those with scores of seven or eight. While such assignments are not normally considered indicators of critical need, nontwin siblings in this category are about 1.9 percentage points more likely to die within a year than other siblings with scores of ten. For twins, however, this relationship drops by a third, and is measured less precisely because of the smaller sample size. The results also suggest only a minute difference in infant mortality between infants with Apgar scores of nine versus ten.

The second panel presents the same set of results, but using birth weight instead of Apgar categories. Interestingly, the same contrast in results between the sibling and the twins samples arises when we compare the effects of very low levels of birth weight on infant mortality with and without fixed effects. The estimated effects of low birth weight slightly increase after adding the family fixed effects for the sibling sample. Even for infants born between 2,500 and 3,500 grams—below average weight but not typically considered low birth weight—there is about a one percentage point higher risk of death within one year. The estimated effect associated with weighing less than 1,500 grams falls by about two-thirds when comparing twins from the same family compared to using cross-variation of the nontwin sibling sample. Similar results were found by Almond, Chay, and Lee (2005), who focus on twins exclusively.

Variation in infant health between twins cannot be due to changes in socioeconomic circumstances of the parents between births. Changes in such circumstances over the one- to seven-year period in our sibling sample do not seem large enough to explain the different estimates, especially since the coefficients do not fall at all after adding the fixed effects (the twins results suggest a downward omitted variable bias). Another explanation is that twin birth-weight variation cannot result from differences in gestation, but it certainly can with the sibling sample. The last panel indicates siblings born premature are significantly more likely to die in one year than another sibling not born premature. A sibling born 37 weeks since conception faces a 1.4 percentage point higher chance of infant mortality than another sibling born between 40 and 41 weeks since conception. We also find slightly higher chances of infant mortality from 39 weeks gestation. Thus, one possibility to explain the different estimated effects from low birth weight and low Apgar score between nontwin siblings and twins is gestation, an important source of variation correlated with these measures of infant health but left out from the between twins analysis.[10]

10. Approximately 14 percent of the births in our sample are from C-sections. C-sections can alter the gestational length in ways that may affect our estimates. We reestimated the gestational models excluding C-sections and the results are very similar.

In order to compare the sibling and twin fixed-effects more closely we also estimate models of the effects of birth weight on infant death conditional on gestation. Interestingly, there is little change in the coefficients on birth weight (they are, on average, 90 percent of the size of the coefficients excluding controls for gestation) while most of the coefficients on gestation become insignificant. This suggests that most of the relationship between birth weight and infant health is indeed due to IGUR and not differences in gestational length, supporting the inference from estimates of twin samples. These results are reported in Table 3b.

There are sufficient numbers of infant deaths in our sample to explore the relationship between infant health and death by cause if we group our sample into four classifications: congenital anomalies (24 percent), symptoms, sign and ill-defined conditions (11 percent) (including SIDS, although 76 percent of these deaths are coded as "sudden death, cause unknown"), conditions originating in the perinatal period (41 percent), and all other diagnoses (20 percent, including injuries which make up 3 percent of the 20 percent). Although the twin samples get too small to infer much, the sibling-fixed-effects results suggest that at lowest birth weights and Apgar scores, the largest effect is from conditions originating in the perinatal period. These include respiratory distress syndrome, disorders relating to low birth weight and short gestation, and hematological disorders of fetus and newborn. Because many of the examples given for these disorders are exactly our infant health measures, the strength of this relation should not be surprising. At low, but not very low birth weights, and mid-level Apgar scores, congenital anomalies are also reasonably large predictors (and conditions originating in the perinatal period remain large). These results are presented in Appendix Tables A1 through A4.[11]

To explore the extent of possible adjustments in fertility decisions in response to having a sick child we reestimate our model splitting the sibling sample into (a) the sample for which the first child has poor health and (b) a subsequent child has poor health.[12] These results are reported in Appendix Table A5. Using both birth weight and Apgar scores as the measure of infant health we note that the effects of poor infant health on death within one year are fairly consistent between the samples where the first child experienced poor infant health and where a subsequent child experienced poor infant health. While the coefficients on infant health for the sample with a subsequent child having poor health are larger in some cases, the coefficients in both samples continue to be large and significant.

Extending the analysis on child death out to the first 17 years of life, we continue to find an effect of both low Apgar scores and low birth weight on survival when we examine children across families and between siblings within families (Table 4). Indeed, the coefficients on birth weight between 1,000 and 2,500 grams actually increase once we include family fixed effects, and the coefficients on the lowest Apgar scores remain relatively stable. Using the twin sample, however, we find no evidence of a negative

11. A main contribution of this paper is to explore robustly whether the consequences of infant health continue over the longer term. Extrapolating from cause of death to the cause of longer-run relationships, while extremely important, is an area for further exploration.

12. This subsample includes only those siblings who fit the following criteria: one group has first-born <3,000g (or Apgar <eight) with no later sibling <3,000g, and the second group is first-born >3,000g and some later sibling <3,000g.

Table 3b

Estimated Effects of Infant Health at Birth on Infant Mortality (Death within One Year of Birth) With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample With Family F. E.	Twins Sample No Family F. E.	Twins Sample With Family F. E.
	Birth Weight (Omitted Category BW = 3,500+ grams)				
BW<1,000	0.8069 ***	0.8494 ***	0.8683 ***	0.7296 ***	0.2532 ***
	(0.0048)	(0.0074)	(0.0103)	(0.0301)	(0.0701)
BW 1,000–1,500	0.2609 ***	0.3546 ***	0.3872 ***	0.2145 ***	0.0848 **
	(0.0041)	(0.0069)	(0.0097)	(0.0230)	(0.0384)
BW 1,500–2,500	0.0292 ***	0.0435 ***	0.0604 ***	0.0091	0.0048
	(0.0015)	(0.0024)	(0.0037)	(0.0181)	(0.0246)
BW 2,500–3,000	0.0051 ***	0.0053 ***	0.0120 ***	0.0093	−0.0043
	(0.0008)	(0.0013)	(0.0021)	(0.0177)	(0.0231)
BW 3,000–3,500	0.0013 **	0.0014 ***	0.0038 ***	0.0006	−0.0017
	(0.0006)	(0.0009)	(0.0014)	(0.0182)	(0.0218)
F-test: No infant health effects	6,192.65 ***	3,043.72 ***	1,640.12 ***	200.80 ***	3.25 ***

Sample size	{109,125}	{54,310}	{54,310}	{1,752}	{1,752}
R-squared	0.25	0.25	0.56	0.48	0.84
	Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	0.0061 ***	0.0095 ***	0.0053	−0.0091	NA
	(0.0015)	(0.0024)	(0.0035)	(0.0117)	
Gestation 37 weeks	0.0001	0.0016	0.0041	−0.0095	NA
	(0.0014)	(0.0022)	(0.0032)	(0.0133)	
Gestation 38 weeks	0.0015	0.0026 *	0.0024	−0.0113	NA
	(0.0009)	(0.0014)	(0.0020)	(0.0123)	
Gestation 39 weeks	0.0000	0.0000	−0.0015	−0.0013	NA
	(0.0007)	(0.0011)	(0.0016)	(0.0128)	
Gestation≥42 weeks	0.0009	0.0006	0.0005	−0.0093	NA
	(0.0011)	(0.0017)	(0.0024)	(0.0296)	
Gestation missing	−0.0004	−0.0009	−0.0018	−0.0124	
	(0.0016)	(0.0025)	(0.0035)	(0.0199)	
F-test:	3.31 ***	3.14 ***	1.14	0.24	

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

relationship between infant health and death up to age 17. In fact, the coefficient on Apgar scores between seven and eight is the wrong sign and marginally significant. Given the drop in predictive power within twins for our one-year mortality rate estimates, it is perhaps not surprising that we do not find results with twins for mortality 17 years out. On the whole, however, we take these results as evidence that infant health continues to be a strong predictor of mortality both across and within families even up to age 17.

Table 5 indicates little correlation between infant health and Manitoba emigration. The OLS results with the full sample indicate a small correlation between poorer infant health and moving away from the province before age 17 (particularly for Apgar ≤ six, although the confidence interval here is quite wide) suggesting some potential sample selection bias here. However, the point estimates gravitate toward zero after including the family fixed effects regardless of which measure of infant health is used. Our results suggest that once we control for family fixed effects, our estimates on the impact of infant health on later outcomes among Manitoba residents do not appear to be biased from some fraction of our sample leaving the province.[13]

Another important source of selection bias is mortality. Our results beyond infant mortality are conditional on being alive past age one (and for those later results that extend beyond age 17, the results are similarly conditional on being alive past age 17). If we assume that those in the worst health died and that these children would also have had poor outcomes later in life then our estimates can be seem as underestimates of the true effects of poor infant health on longer term outcomes.

Conditional on survival until age 17, we find little evidence of significant effects of infant health on physician utilization between ages 12 and 17. Table 6 displays the estimates of the effects of Apgar score, birth weight, and gestation on total physician visits between these ages, with and without including family fixed effects. The dependent variable here is number of physician visits between the ages of 12 and 17. We find little consistent evidence here to support a relationship between infant health and physician visits 12 to 17 years later. Although we do find some evidence of a greater number of visits within twin families for those children with birth weights between 1,000 and 1,500 grams, the majority of the coefficient estimates are insignificant and some are the wrong sign.

Table 7 shows the results of the language arts standards test for the sample of Manitoban residents at age 17. Recall that for the approximately 30 percent of residents who did not write the test, the score is imputed by ranking these individuals lower than those writing the test and categorizing them by enrollment and attainment categories (for example, withdrawn from school or held back). A score is given to each associated test score and education attainment category using a standardized logit transformation weighted by the population size in each group.

Columns 1 and 2 indicate a clear positive correlation between infant health and the language arts test measure. For example, a low-birth-weight child averages a score about 0.23 standard deviations below a child born weighing above 3,500 grams. Apgar scores less than eight and gestation lengths less than 38 weeks are also associated with significantly lower grade 12 test scores. The relationship weakens notably after adding the

13. Being born 1,500-2,500g or having an Apgar score of seven to eight in the sibling-fixed-effects model is associated with a 0.2 percent (and insignificant) increase in the probability of moving out of the province off a base of 20[0] percent.

Table 4
Estimated Effects of Infant Health of Birth on Mortality between Ages 1 to 17 With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample With Family F. E.	Twins Sample No Family F. E.	Twins Sample With Family F. E.
	APGAR Score (Omitted Category APGAR=10)				
APGAR≤6	0.0111 ***	0.0147 ***	0.0131 ***	0.0100	−0.0159
	(0.0020)	(0.0033)	(0.0045)	(0.0076)	(0.0136)
APGAR=7–8	0.0026 ***	0.0038 ***	0.0034 *	−0.0071 *	−0.0150 *
	(0.0008)	(0.0013)	(0.0018)	(0.0043)	(0.0087)
APGAR=9	0.0001	−0.0002	0.0004	−0.0040	−0.0092
	(0.0005)	(0.0008)	(0.0011)	(0.0038)	(0.0082)
F-test: No infant health effects	14.1 ***	10.4 ***	3.88 ***	2.31 *	1.13
Sample size	{107,772}	{53,527}	{53,527}	{1,693}	{1,693}
R-squared	0.01	0.01	0.48	0.07	0.55
	Birth Weight (Omitted Category BW = 3,500+ grams)				
BW<1,000	0.0444 ***	−0.0058	−0.0384 *	0.0004	0.0170
	(0.0094)	(0.0169)	(0.0226)	(0.0216)	(0.0488)
BW 1,000–1,500	0.0124 ***	0.0193 ***	0.0291 ***	0.0161	0.0169
	(0.0040)	(0.0073)	(0.0101)	(0.0100)	(0.0211)
BW 1,500–2,500	0.0025 **	0.0043 **	0.0068 **	0.0002	0.0013
	(0.0011)	(0.0020)	(0.0030)	(0.0075)	(0.0132)
BW 2,500–3,000	0.0003	0.0006	0.0019	0.0010	0.0028
	(0.0007)	(0.0011)	(0.0018)	(0.0075)	(0.0124)

(*continued*)

Table 4 (*continued*)

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample With Family F. E.	Twins Sample No Family F. E.	Twins Sample With Family F. E.
BW 3,000–3,500	0.0006	0.0009	0.0034 ***	0.0043	−0.0013
	(0.0005)	(0.0008)	(0.0012)	(0.0078)	(0.0117)
F-test: No infant health effects	7.36 ***	2.44 **	4.24 **	1.15	0.26
Sample size	{108,073}	{53,700}	{53,700}	{1,700}	{1,700}
R-squared	0.01	0.01	0.48	0.06	0.55
	Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	0.0028 **	0.0024	0.0039	0.0023	NA
	(0.0011)	(0.0019)	(0.0030)	(0.0049)	
Gestation 37 weeks	0.0010	0.0009	0.0012	−0.0027	NA
	(0.0012)	0.00192	(0.0030)	(0.0060)	
Gestation 38 weeks	0.0004	0.0014	0.0027	−0.0029	NA
	(0.0008)	(0.0012)	(0.0019)	(0.0056)	
Gestation 39 weeks	0.0003	0.0014	0.0008	0.0020	NA
	(0.0006)	(0.0012)	(0.0015)	(0.0059)	
Gestation≥42 weeks	0.0008	−0.0004	−0.0002	−0.0068	NA
	(0.0010)	(0.0010)	(0.0023)	(0.0137)	
F-test: No infant health effects	1.38	0.71	0.63	0.47	
Sample size	{89,276}	{45,078}	{45,078}	{1,447}	
R-squared	0.01	0.01	0.54	0.08	

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Table 5
Estimated Effects of Infant Health at Birth on Mobility out of Manitoba Before Age 18 With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample With Family F. E.	Twins Sample No Family F. E.	Twins Sample With Family F. E.
	APGAR Score (Omitted Category APGAR=10)				
APGAR≤6	0.0241 **	0.0169	−0.0195 *	0.0465	0.0169
	(0.0115)	(0.0165)	(0.0110)	(0.0548)	(0.0178)
APGAR=7–8	0.0120 ***	0.0072	0.0023	0.0470	0.0111
	(0.0046)	(0.0062)	(0.0043)	(0.0308)	(0.0113)
APGAR=9	0.0042	0.0067 *	−0.0002	0.0174	0.0148
	(0.0030)	(0.0038)	(0.0028)	(0.0270)	(0.0107)
F-test: No infant health effects	3.38 **	1.29	1.24	0.94	0.70
Sample size	{107,772}	{53,527}	{53,527}	{1,693}	{1,693}
R-squared	0.07	0.05	0.88	0.06	0.99
	Birth Weight (Omitted Category BW = 3,500+ grams)				
BW<1,000	−0.0016	−0.0161	−0.0031	0.2391	−0.0348
	(0.0545)	(0.0836)	(0.0552)	(0.1548)	(0.0634)
BW 1,000–1,500	−0.0175	−0.00908	0.0041	0.0564	−0.0253
	(0.0230)	(0.0361)	(0.0245)	(0.0716)	(0.0274)
BW 1,500–2,500	0.0126 *	0.0271 ***	0.0027	0.1148 **	−0.0249
	(0.0066)	(0.0098)	(0.0073)	(0.0536)	(0.0171)
BW 2,500–3,000	0.0099 **	0.0079	−0.0045	0.0988 *	−0.0314 *
	(0.0039)	(0.0053)	(0.0043)	(0.0537)	(0.0161)

(*continued*)

Table 5 (*continued*)

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample With Family F. E.	Twins Sample No Family F. E.	Twins Sample With Family F. E.
BW 3,000–3,500	0.0058 **	0.0061	0.0027	0.1307 **	−0.0188
	(0.0029)	(0.0038)	(0.0029)	(0.0560)	(0.0152)
F-test: No infant health effects	2.11 *	1.97 *	0.80	1.58	0.96
Sample size	{108,073}	{53,700}	{53,700}	{1,700}	{1,700}
R-squared	0.07	0.05	0.88	0.06	0.99
	Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	0.0076	0.0157 *	0.0102	0.0511	NA
	(0.0063)	(0.0092)	(0.0069)	(0.0318)	
Gestation 37 weeks	0.0122 *	0.0172 *	−0.0059	0.0092	NA
	(0.0070)	(0.0095)	(0.0070)	(0.0392)	
Gestation 38 weeks	−0.0090 **	−0.0022	0.0055	0.0644 *	NA
	(0.0045)	(0.0059)	(0.0045)	(0.0367)	
Gestation 39 weeks	−0.0008	−0.0010	−0.0003	0.0589	NA
	(0.0037)	(0.0048)	(0.0036)	(0.0386)	
Gestation≥42 weeks	0.0044	0.0042	−0.0033	−0.1367	NA
	(0.0056)	(0.0073)	(0.0052)	(0.0894)	
F-test: No infant health effects	2.12 *	1.36	1.12	1.93 *	
Sample size	{89,276}	{45,078}	{45,078}	{1,447}	
R-squared	0.06	0.04	0.90	0.06	

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Table 6

Estimated Effects of Infant Health at Birth on Total Physician Visits Between Ages 12 and 17 With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample With Family F. E.	Twins Sample No Family F. E.	Twins Sample Family F. E.
	APGAR Score (Omitted Category APGAR=10)				
APGAR≤6	1.9767 ***	1.5734 **	1.5356 **	1.9612	0.7656
	(0.4213)	(0.6186)	(0.6768)	(1.7188)	(2.2051)
APGAR=7–8	0.5790 ***	0.2692	−0.5566 **	−1.0923	−1.2239
	(0.1629)	(0.2243)	(0.2561)	(0.9297)	(1.3159)
APGAR=9	0.3954 ***	0.3173 **	−0.1826	0.0717	−0.4005
	(0.1043)	(0.1376)	(0.1628)	(0.8024)	(1.2520)
F-test: No infant health effects	11.76 ***	3.47 **	3.81 ***	1.58	0.74
Sample size	{79,143}	{40,078}	{40,078}	{1,348}	{1,348}
R-squared	0.04	0.05	0.68	0.09	0.78
	Birth Weight (Omitted Category BW = 3,500+ grams)				
BW<1,000	−0.9530	−0.8204	2.2148	2.6851	4.9343
	(1.9892)	(3.1370)	(3.2498)	(4.8230)	(10.6648)
BW 1,000–1,500	1.0137	−0.5904	−1.6869	9.3137 ***	7.3252 **
	(0.8185)	(1.3344)	(1.4809)	(2.1409)	(2.9635)
BW 1,500–2,500	0.4262 *	−0.1540	0.5461	3.3341 **	0.1926
	(0.2359)	(0.3662)	(0.4425)	(1.5216)	(1.8861)
BW 2,500–3,000	0.2481 *	0.3470 *	0.2464	3.5727 **	0.1849
	(0.1402)	(0.1941)	(0.2537)	(1.5280)	(1.7688)

(*continued*)

Table 6 (*continued*)

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample With Family F. E.	Twins Sample No Family F. E.	Twins Sample Family F. E.
BW 3,000–3,500	0.2125 **	0.3026 **	0.3215 *	2.9624 *	1.9898
	(0.1027)	(0.1375)	(0.1704)	(1.6084)	(1.6555)
F-test: No infant health effects	1.78	1.47	1.25	4.03 ***	2.77 **
Sample size	{79,363}	{40,203}	{40,203}	{1,354}	{1,354}
R-squared	0.04	0.05	0.68	0.10	0.78
	Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	0.2814	−0.2208	−0.5687	0.8215	NA
	(0.2271)	(0.3395)	(0.4291)	(0.9276)	
Gestation 37 weeks	0.3913	0.1131	−0.1167	3.1315 ***	NA
	(0.2491)	(0.3487)	(0.4345)	(1.1326)	
Gestation 38 weeks	0.4832 ***	0.2030	0.0325	−0.1308	NA
	(0.1582)	(0.2119)	(0.2693)	(1.0700)	
Gestation 39 weeks	0.0867	−0.1401	−0.1719	−0.0701	NA
	(0.1305)	(0.1731)	(0.2151)	(1.1335)	
Gestation≥42 weeks	0.0874	−0.0581	−0.0985	−1.3445	NA
	(0.1973)	(0.2636)	(0.3201)	(2.3796)	
F-test: No infant health effects	2.29 **	0.53	0.48	2.59 **	
Sample size	{66,504}	{33,921}	{33,921}	{1,166}	NA
R-squared	0.04	0.05	0.72	0.10	NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Table 7

Estimated Effects of Infant Health at Birth on Language Arts Score With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample Family F. E.	Twins Sample No Family F. E.	Twins Sample Family F. E.
	APGAR Score (Omitted Category APGAR=10)				
APGAR≤6	−0.1541 ***	−0.1730 ***	−0.0913 *	0.1254	0.0756
	(0.0297)	(0.0450)	(0.0481)	(0.1457)	(0.1568)
APGAR=7–8	−0.0503 ***	−0.0436 ***	−0.0245	0.0326	−0.1021
	(0.0116)	(0.0164)	(0.0183)	(0.0786)	(0.0936)
APGAR=9	0.0069	0.0184 *	−0.0012	0.0444	−0.0952
	(0.0074)	(0.0101)	(0.0116)	(0.0676)	(0.0890)
F-test: No infant health effects	18.52 ***	10.97 ***	1.82	0.30	0.99
Sample size	{79,194}	{40,514}	{40,514}	{1,364}	{1,364}
R-squared	0.17	0.20	0.75	0.20	0.87
	Birth Weight (Omitted Category BW = 3,500+grams)				
BW<1,000	−0.3142 **	−0.2576	−0.0491	−0.5415	0.01819
	0.1411	0.23089	0.23167	0.40812	0.7616
BW 1,000–1,500	−0.26 ***	−0.2302 **	−0.0812	−0.4608 **	−0.185
	0.05806	0.09822	0.10557	0.18116	0.21163
BW 1,500–2,500	−0.1519 ***	−0.2188 ***	−0.0492	−0.1826	−0.0987
	0.01673	0.02695	0.03155	0.12875	0.13469
BW 2,500–3,000	−0.1172 ***	−0.1303 ***	−0.0477 **	−0.2063	−0.2139 *
	(0.0099)	(0.0143)	(0.0181)	(0.1293)	(0.1263)

(*continued*)

Table 7 (*continued*)

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample Family F. E.	Twins Sample No Family F. E.	Twins Sample Family F. E.
BW 3,000–3,500	−0.0339 ***	−0.0459 ***	−0.0154	−0.0969	−0.0672
	(0.0073)	(0.0101)	(0.0121)	(0.1361)	(0.1182)
F-test: No infant health effects	42.74 ***	27.73 ***	1.57	1.95 *	1.54
Sample size	{79,363}	{40,203}	{40,203}	{1,354}	{1,354}
R-squared	0.17	0.20	0.74	0.21	0.86
	Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	−0.0797 ***	−0.1232 ***	0.0255	0.0061	NA
	(0.0161)	(0.0251)	(0.0303)	(0.0800)	
Gestation 37 weeks	−0.0678 ***	−0.0999 ***	−0.0076	0.1317	NA
	(0.0177)	(0.0258)	(0.0307)	(0.0976)	
Gestation 38 weeks	−0.0114	−0.0166	0.0263	−0.0708	NA
	(0.0112)	(0.0157)	(0.0190)	(0.0922)	
Gestation 39 weeks	0.0083	0.0003	0.0037	0.0543	NA
	(0.0093)	(0.0128)	(0.0152)	(0.0977)	
Gestation≥42 weeks	−0.0204	−0.0231	−0.0089	−0.1789	NA
	(0.0140)	(0.0195)	(0.0226)	(0.2051)	
F-test: No infant health effects	8.44 ***	7.70 ***	0.60	1.28	
Sample size	{66,504}	{33,921}	{33,921}	{1,166}	NA
R-squared	0.17	0.20	0.79	0.21	NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance.

family fixed effects. The point estimates for birth weight and Apgar score are all still negative, but many are no longer significant. Siblings given an Apgar of six or less receive a test score about one-tenth of a standard deviation lower than a sibling with a ten. We also find indications of small, but in some cases significant, lasting effects on test scores from being born low birth weight, even among youths born weighing between 2,500 and 3,500 grams (only slightly below the average). The negative coefficients associated with gestation of less than 37 weeks without fixed effects become close to zero, once sibling fixed effects are added. All the point estimates for the gestation results are very small and insignificant.

Because of the large number of imputed test scores in the language arts test, we also rerun all our specifications dropping the 35 percent of students with imputed scores to check whether these students are driving our results. We continue to find little effect of infant health on this particular test score.

While our results for language arts test scores are mixed, our results examining high school attainment suggest long-lasting effects of infant health. Table 8 shows the estimates for the effects of infant health on reaching grade 12 by age 17. An individual may fail to reach this grade because she either dropped out or repeated at least one earlier grade. The sibling-fixed-effects analysis in Column 3 indicates a substantial impact on grade 12 attainment from infant health. Moving from the results with no family fixed effects to those that include them show only a small fall in the coefficients.

A newborn assessed with an Apgar score of six or less, for example, has a 7.4 percentage point lower probability of reaching grade 12 by age 17 compared to a youth born with an Apgar of ten (Column 2). When family fixed effects are added (Column 3), the estimated effect drops to a 4.1 percentage point difference in the probability of reaching grade 12 by this age. Siblings with Apgar scores still considered normal but below average (seven or eight) are two percentage points more likely to drop out or repeat a grade.

A similar story holds when looking at birth weight. Students born with low birth weights between 1,500 and 2,500 grams, and who survive until age 17 are about eight percentage points less likely to be enrolled in grade 12 than those born weighing 3,500 grams or more. The chances of infants attaining grade 12 by age 17 are severely affected by being born weighing less than 1,500 grams. The gestation results are also strong and significant. Premature siblings born in 36 weeks or less are 4.0 percentage points less likely to have reached grade 12 by age 17 than those born in 40–41 weeks. Negative effects on this measure of educational attainment among those born in 38 or 39 weeks are also detected. A sibling born after 38 weeks gestation is 2.5 percentage points less likely to be in grade 12 at age 17 than another sibling born less than 40 weeks gestation.[14,15,16]

14. We have estimated the main results separately by sex. The effects in most cases do not differ significantly by sex. Interestingly, however, the effects of infant health on reaching grade 12 by age 17 appear to be somewhat stronger for females than males.

15. We note that for this outcome, if we omit C-section births the effects of having a gestational period of 38 weeks for the sibling-fixed-effects model is no longer significant.

16. We again test the extent of possible adjustments in fertility decisions in response to having a sick child by reestimating our model splitting the sibling sample into a) the sample for which the first child has poor health and b) a subsequent child has poor health. These results are reported in Appendix Table A6. Our results remain large and significant across both samples, with slightly larger coefficients in some cases for the second subsample.

Table 8
Estimated Effects of Infant Health at Birth on Reaching Grade 12 by Age 17 With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample Family F. E.	Twins Sample No Family F. E.	Twins Sample Family F. E.
	APGAR Score (Omitted Category APGAR=10)				
APGAR ≤6	−0.0697 ***	−0.0739 ***	−0.0410 *	0.0710	0.0317
	(0.0137)	(0.0206)	(0.0239)	(0.0701)	(0.0706)
APGAR=7–8	−0.0240 ***	−0.0302 ***	−0.203 **	0.0179	−0.0210
	(0.0053)	(0.0075)	(0.0091)	(0.0378)	(0.0421)
APGAR=9	0.0058 *	0.0111 **	0.0051	0.0428	−0.0156
	(0.0034)	(0.0046)	(0.0058)	(0.0325)	(0.0401)
F-test: No infant health effects	21.3 ***	16.3 ***	4.12 ***	0.81	0.32
Sample size	{79,194}	{40,514}	{40,514}	{1,364}	{1,364}
R-squared	0.15	0.18	0.70	0.18	0.88
	Birth Weight (Omitted Category BW = 3,500+ grams)				
BW<1000	−0.1842 ***	−0.1308	0.0546	−0.3346 *	−0.0837
	(0.0649)	(0.1060)	(0.1153)	(0.1968)	(0.3434)
BW 1000–1500	−0.1353 ***	−0.1271 ***	−0.1410 ***	−0.2708 ***	−0.2315 **
	(0.0267)	(0.0451)	(0.0525)	(0.0874)	(0.0954)
BW 1500–2500	−0.0793 ***	−0.1168 ***	−0.0815 ***	−0.1284 **	−0.0685
	(0.0077)	(0.0124)	(0.0157)	(0.0621)	(0.0607)
BW 2500–3000	−0.0398 ***	−0.0496 ***	−0.0430 ***	−0.1094 *	−0.0921
	(0.0046)	(0.0066)	(0.0090)	(0.0624)	(0.0570)

BW 3000–3500	−0.0125 ***	−0.0189 ***	−0.0243 **	−0.0639	−0.0488
	(0.0034)	(0.0046)	(0.0060)	(0.0656)	(0.0533)
F-test: No infant health effects	37.47 ***	27.74 ***	8.98 ***	2.91 **	1.72
Sample size	{79,363}	{40,203}	{40,203}	{1,354}	{1,354}
R-squared	0.16	0.18	0.69	0.19	0.88
	Gestation (Omitted Category 40–41 Weeks)				
Gestation ≤36 weeks	−0.0602 ***	−0.0843 ***	−0.0403 ***	−0.0498	NA
	(0.0074)	(0.0115)	(0.0150)	(0.0387)	
Gestation 37 weeks	−0.0246 ***	−0.0426 ***	−0.0345 **	0.0473	NA
	(0.0081)	(0.0118)	(0.0152)	(0.0472)	
Gestation 38 weeks	−0.0064	−0.0139 *	−0.0252 ***	0.0419	NA
	(0.0051)	(0.0072)	(0.0094)	(0.0446)	
Gestation 39 weeks	0.0076 *	0.0062	−0.0014	0.0580	NA
	(0.0042)	(0.0059)	(0.0075)	(0.0473)	
Gestation ≥42 weeks	−0.0067	−0.0099	−0.0116	−0.1740 *	NA
	(0.0064)	(0.0089)	(0.0112)	(0.0992)	
F-test: No infant health effects	17.14 ***	14.37 ***	3.19 ***	3.23 ***	
Sample size	{66,504}	{33,921}	{33,921}	{1,166}	NA
R-squared	0.16	0.19	0.75	0.19	NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Finally, we estimate models including both birth weight and gestational length. As with infant death, the coefficients on birth weight remain significant while the coefficients on gestation are generally no longer significant. Results are shown in Appendix Table A7.

Table 9 shows the estimated effects of infant health on the probability of receiving any social assistance between the age of 18 and 21.25 years of age. Our sample includes only individuals born between 1979 and 1982 (we exclude the younger cohorts in this analysis as noted above). Given the relatively short window of social assistance eligibility, only 8 percent of our sample received social assistance over this period. The least-squares results without family fixed effects indicate a substantial relationship between poor infant health and receiving social assistance. Apgar scores strongly predict takeup. A young adult given an Apgar of six or less at birth has a 6.6 percentage point higher probability of receiving welfare than a contemporary given an Apgar of ten. Individuals with Apgar scores between seven and nine also have a higher likelihood of receiving social assistance. These effects are closer to zero once the family fixed effects are added in Column 3. The point estimates are somewhat noisy, but they suggest that the causal effects of Apgar score on welfare use, independent of family circumstances that correlate with this score, are lower than predicted by the strong effects estimated using the cross-family variation. Interestingly, the estimates from the smaller sample of twins do suggest a big effect. From Column 5, a twin who is given an Apgar of nine stands an 8.1 percentage point higher probability of receiving social assistance than does her other twin with an Apgar of ten.

The effects of birth weight on social assistance takeup are also strong in the OLS, weaker in the sibling fixed effects, but very strong in the twin fixed effects. The least-squares results in Column 2 indicate a strong escalating relationship between being born low birth weight and receiving social assistance. The point estimates fall substantially in Column 3 from including the family fixed effects. While, once again, the estimates for this outcome variable are imprecise, the standard errors are small enough to rule out that the estimates are equal to those without the fixed effects, while not ruling out the effects are zero. The twins results, like the case with Apgar scores, are strongly positive and suggestive of a long-lasting effect from infant health.

The gestation results without fixed effects indicate a significant relationship among youths born after less than 38 weeks gestation. These coefficients, however, all drop close to zero (and are imprecisely measured) after adding in family fixed effects.

In an attempt to increase the variance of social assistance use in our sample, we also look at months on social assistance between age 18 and 21.25. Table 10 presents these results. The mean number of months on social assistance in our sample of siblings is 1.4. The results are also somewhat noisy, but generally suggest a continued link between our infant health measures and social assistance use. Birth weight appears to affect not only takeup but also duration of social assistance. The coefficients for the effects of low birth weight on months on social assistance using family fixed effects are about one-half to three-quarters the size of those without the fixed effects. The estimates for the average association between being born 1,500 to 3,000 grams and social assistance use are significant, and we can reject the hypothesis that all the estimated effects are zero or less. The coefficients from the twins

Table 9
Estimated Effects of Infant Health at Birth on Social Assistance Take-up With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample Family F. E.	Twins Sample No Family F. E.	Twins Sample Family F. E.
	APGAR Score (Omitted Category APGAR=10)				
APGAR≤6	0.0505 ***	0.0660 ***	0.0302	0.0105	0.0309
	(0.0108)	(0.0174)	(0.0211)	(0.0518)	(0.0670)
APGAR=7–8	0.0198 ***	0.0217 ***	0.0146 *	0.0087	0.0765 *
	(0.0042)	(0.0064)	(0.0081)	(0.0258)	(0.0390)
APGAR=9	0.0121 ***	0.0092 **	−0.0049	0.0171	0.0813 **
	(0.0025)	(0.0037)	(0.0049)	(0.0212)	(0.0355)
F-test: No infant health effects	16.17 ***	8.42 ***	3.01 **	0.23	2.03
Sample size	{54,254}	{22,802}	{22,802}	{874}	{874}
R-squared	0.08	0.09	0.65	0.11	0.74
	Birth Weight (Omitted Category BW = 3500+grams)				
BW<1000	−0.0381	−0.0251	−0.2155 **	−0.0635	0.1412
	(0.0541)	(0.0804)	(0.0955)	(0.2428)	(0.2802)
BW 1000–1500	0.0408 *	0.0253	0.0439	0.0707	0.2430 **
	(0.0211)	(0.0369)	(0.0448)	(0.0595)	(0.0943)
BW 1500–2500	0.0377 ***	0.0436 ***	0.0186	0.0565	0.1062 *
	(0.0060)	(0.0103)	(0.0138)	(0.0417)	(0.0568)
BW 2500–3000	0.0219 ***	0.0248 ***	0.0062	0.0403	0.1058 **
	(0.0035)	(0.0054)	(0.0078)	(0.0419)	(0.0533)

(*continued*)

Table 9 (*continued*)

	Full Sample	Sibling Sample No Family F. E.	Sibling Sample Family F. E.	Twins Sample No Family F. E.	Twins Sample Family F. E.
BW 3000–3500	0.0096 ***	0.0127 ***	0.0078	0.0807 *	0.1178 **
	(0.0026)	(0.0038)	(0.0053)	(0.0441)	(0.0491)
F-test: No infant health effects	14.43 ***	7.47 ***	1.85	1.02	1.88 *
Sample size	{54,412}	{22,870}	{22,870}	{880}	{880}
R-squared	0.08	0.09	0.65	0.12	0.74
	Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	0.0309 ***	0.0373 ***	−0.0042	0.0160	NA
	(0.0063)	(0.0106)	(0.0160)	(0.0260)	
Gestation 37 weeks	0.0214 ***	0.0317 ***	−0.0132	−0.0124	NA
	(0.0069)	(0.0108)	(0.0163)	(0.0348)	
Gestation 38 weeks	0.0113 ***	0.0155 ***	−0.0033	−0.0050	NA
	(0.0043)	(0.0065)	(0.0101)	(0.0322)	
Gestation 39 weeks	−0.0035	−0.0048	−0.0085	−0.0471	NA
	(0.0036)	(0.0053)	(0.0080)	(0.0307)	
Gestation≥42 weeks	0.0047	0.0040	−0.0137	−0.0187	NA
	(0.0052)	(0.0079)	(0.0116)	(0.0573)	
F-test: No infant health effects	8.13 ***	5.47 ***	0.51	1.09	
Sample size	{41,593}	{17,175}	{17,175}	{694}	NA
R-squared	0.08	0.09	0.76	0.15	NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Table 10
Estimated Effects of Infant Health at Birth on Months on Social Assistance With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample Family F.E.	Twins Sample No Family F.E.	Twins Sample Family F.E.
	APGAR Score (Omitted Category APGAR=10)				
APGAR≤6	1.3555 ***	1.7468 ***	0.8219 *	1.5889	1.6621
	(0.2406)	(0.3866)	(0.4889)	(1.1879)	(1.6354)
APGAR=7–8	0.5543 ***	0.6349 ***	0.5418 ***	0.6216	1.3031
	(0.0930)	(0.1419)	(0.1863)	(0.5914)	(0.9516)
APGAR=9	0.1954 ***	0.0986	−0.1994 *	0.3314	0.8966
	(0.0561)	(0.0824)	(0.1128)	(0.4865)	(0.8659)
F-test: No infant health effects	20.85 ***	12.78 ***	7.22 ***	0.78	0.68
Sample size	{54,254}	{22,802}	{22,802}	{874}	{874}
R-squared	0.06	0.06	0.61	0.17	0.73
	Birth Weight (Omitted Category BW = 3500+ grams)				
BW <1000	−0.4577	−1.7721	−1.0878	−1.1759	−1.2839
	(1.2061)	(1.7860)	(2.2111)	(5.5353)	(6.7116)
BW 1000–1500	1.6866 ***	0.5559	1.3630	1.4057	1.7435
	(0.4698)	(0.8195)	(1.0360)	(1.3570)	(2.2587)
BW 1500–2500	1.0086 ***	1.2861 ***	0.9042 ***	0.8977	2.0004
	(0.1331)	(0.2285)	(0.3190)	(0.9504)	(1.3600)
BW 2500–3000	0.5272 ***	0.6356 ***	0.3548 *	0.6650	3.0202 **
	(0.0787)	(0.1198)	(0.1817)	(0.9545)	(1.2773)

(*continued*)

Table 10 (*continued*)

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample Family F.E.	Twins Sample No Family F.E.	Twins Sample Family F.E.
BW 3000–3500	0.2113 ***	0.3160 ***	0.2653 **	1.8979 *	3.4056 ***
	(0.0579)	(0.0853)	(0.1230)	(1.0059)	(1.1771)
F-test: No infant health effects	20.02 ***	11.38 ***	2.36 **	1.37	2.25 **
Sample size	{54,412}	{22,870}	{22,870}	{880}	{880}
R-squared	0.06	0.06	0.61	0.17	0.73
	Gestation (Omitted Category 40-41 Weeks)				
Gestation≤36 weeks	0.7880 ***	0.9217 ***	0.2576	0.8108	NA
	(0.1410)	(0.2360)	(0.3712)	(0.5329)	
Gestation 37 weeks	0.6180 ***	0.7582 ***	0.3541	−0.3809	NA
	(0.1435)	(0.2395)	(0.3784)	(0.7141)	
Gestation 38 weeks	0.2100 **	0.3178 **	0.0279	0.1203	NA
	(0.0966)	(0.1444)	(0.2342)	(0.6602)	
Gestation 39 weeks	−0.1390 *	−0.2428 **	−0.1531	−0.3572	NA
	(0.0802)	(0.1178)	(0.1869)	(0.6304)	
Gestation ≥42 weeks	0.1287	0.1467	−0.1461	0.6918	NA
	(0.1153)	(0.1764)	(0.2699)	(1.1741)	
F-test: No infant health effects	11.44 ***	7.55 ***	0.55	1.30	
Sample size	{41,593}	{17,175}	{17,175}	{694}	NA
R-squared	0.06	0.07	0.73	0.28	NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

sample with family fixed effects included, all suggest a long-lasting and large effect of poor Apgar score or birth weight on months on social assistance. The Apgar score results are not significant, but the implied effects of birth weight are large, as they are when looking at only social assistance takeup (in the previous table).

We reexamined our main sibling results for a subsample of siblings less than two years and three years apart. (Tables showing these results are available in Oreopoulos et al. 2006.) A threat to validity in the siblings analysis comes from changes to family or environmental contexts in between births that could account for differences in socioeconomic outcomes. By looking at a subset of siblings closer in age, fewer changes in family circumstances that may affect these outcomes are likely to occur. However, length between births is potentially endogenous (families that experience a bad first birth may wait longer to have a second child) and in this case, it might be that the effects of infant healthcare muted. The coefficients on the estimated effects of Apgar scores, birth weight, and gestation on infant mortality largely remain intact after looking only at siblings less than two years apart. The estimates with family fixed effects remain quite similar to the estimates without them, indicate a strong relationship, and suggest a significant causal relationship between these measures of health at birth and one-year mortality. It is worth pointing out again that this contrasts with the twins results, where the estimated effects fall by as much as two-thirds when accounting for family factors common between twins. We perform a similar analysis for grade attainment outcomes between the sample with all siblings and the one with siblings less than two years apart. The subsample is one-fourth the size of the full sample. Yet, for birth weight, the results are remarkably stable. Low-birth-weight siblings are approximately 10 percentage points less likely to attain grade 12 by age 17. Even those born between 2,500 and 3,500 grams are 2-5 percentage points less likely to attain grade 12 compared to those born weighing more than 3,500 grams. The estimated effects from being born premature or with a low Apgar are measured less precisely with the subsample of siblings close apart, but the results generally point to the same conclusions about impact of these measures on grade attainment, with and without including family fixed effects. Finally, we also repeat the analysis for social assistance takeup outcomes. The full sample results suggest significant effects of low birth weight on months receiving social assistance after age 19, as does the sample of siblings less than three years apart. The standard errors around the point estimates for the smaller sample of siblings less than two years apart prevent definitive conclusions. The Apgar results include a nonintuitive results that those born with an Apgar score of nine instead of ten are slightly less likely to end up on social assistance, while those born with an Apgar score of seven or eight are more likely. None of the gestation results are significant.

We end our analysis by considering whether our estimated effects of infant health on subsequent outcomes differ by family background using census enumeration area-level income data from the 2001 Canadian Census to stratify our sample by income quintile. The results among the bottom fifth group of families still reveal quite similar point estimates compared to those from the full sample (tables of these results are available in Oreopoulos et al. 2006). For example, among children with parents from the lowest residential income quintile, an Apgar score of seven or eight for one sibling is associated with a 2.3 percentage point higher likelihood of death within one year than another sibling given a five-minute Apgar score of ten. Using the bottom two-income quintiles increases the sample, while still focusing on families from poor socioeconomic backgrounds. In general, we find no substantial differences in the estimated effects of infant health on mortality comparing these more disadvantaged groups with the entire population of births. This evidence is consistent with previous research (Currie and Hyson 1999) using an older cohort of births.

We perform similar analyses by residential income quintile to examine the held back and months on social assistance outcome. In general, we conclude that there is no strong evidence that the effects of infant health on high school attainment and social assistance takeup are any worse among families from lower income backgrounds. Our results by quintile suggest that the short-, medium-, and longer-term effects of infant health are not confined to a single quintile, but rather are uniform across the population.

V. Conclusions

We use a cohort of births from a single Canadian province to examine the short-, medium-, and long-term effects of poor infant health. Our results both confirm and extend recent work on the effects of infant health on survival and future measures of health, human capital, and labor force attachment.

Using three measures of infant health: birth weight, Apgar scores, and gestational length, we find that poor infant health predicts both mortality within one year, and mortality up to age 17. These results hold both across families and between siblings within families. Consistent with results in Almond et al. (2005), differences in infant health within families but between twin pairs lead to much smaller differences in both one-year and 17-year mortality rates. This drop in the estimated effects occurs for twins but not for siblings.

We also find that infant health is a strong predictor of educational and labor force outcomes. In particular, infant health is found to predict both high school completion and social assistance (welfare) takeup and length. We find evidence of longer-term consequences of infant health both across families, within siblings, and within twin pairs, although different measures of infant health predict outcomes differently. Birth weight and Apgar scores appear to be stronger predictors than gestation length in these areas when family fixed effects are included, and in models with both birth weight and gestation included as regressors, the coefficients on birth weight are stronger predictors of the outcomes of study. Interestingly, we find less evidence of the longer-term effects of infant health on either cognitive ability as measured by language arts test scores or longer-term physician visits.

In Conley's review of this literature (Conley et al. 2003), he notes that low-birth-weight babies are associated with a number of biological conditions that could affect later outcomes. Neurological handicaps, respiratory infections, and gastrointestinal problems, and higher rates of hypertension some of these. These problems all could manifest themselves in difficulty learning, leading to slower completion rates, and eventually higher takeup rates such as those found in our study for both low birth weight and low Apgar score babies. It is somewhat surprising, however, that we do not also find results in doctors' visits and language test scores, where we would also expect these conditions to manifest themselves.

Our evidence, along with a growing body of literature in this area, confirms the importance of early childhood health as a predictor of future outcomes. Examining differences across families, between siblings, and between twin pairs can help inform both the medical literature and public policy with regards to understanding mechanisms of child development and effective ways to improve childhood health and hence future outcomes.

References

Almond, D., K. Chay, and David Lee. 2005. "The Costs of Low Birth Weight." *Quarterly Journal of Economics* 120(3):1031–83.

Behrman, J., and Mark Rosenzweig. 2001. "The Returns to Increasing Body Weight." Penn Institute for Economic Research Working Paper No.01-052.

———. 2004. "The Returns to Birth Weight." *Review of Economics and Statistics* 86(2):586–601.

Black, S., P. Devereux, and Kjell Salvanes. 2007. "From the Cradle to the Labor Market: The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics* 122(1): 409–39.

———. 2005. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *Quarterly Journal of Economics* 120(2):669–700.

Conley, D., K. Strullly, and Neil Bennett. 2003. "A Pound of Flesh or Just Proxy? Using Twin Differences to Estimate the Effect of Birth Weight on Life Chances." NBER Working Paper No.9901.

———. 2003. *The Starting Gate: Birth Weight and Life Chances*. Berkeley: University of California Press.

Currie, Janet, and Jonathan Gruber. 1996. "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *Quarterly Journal of Economics* 111(2):431–66.

Currie, Janet, and Enrico Moretti. 2005. "Biology as Destiny? Short and Long-Run Determinants of Intergenerational Transmission of Birth Weight." NBER Working Paper No.11567.

Currie, Janet, and Rosemary Hyson. 1999. "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Birth Weight." *American Economic Review* 89(2):245–50.

Goldenberg, R., and Dwight Rouse. 1998. "Prevention of Premature Birth." *New England Journal of Medicine* 339(5):313–20.

Hanratty, M. 1996. "Canadian National Health Insurance and Infant Health." *American Economic Review* 86(1):276–84.

Kramer, M. 1987. "Intrauterine Growth and Gestational Duration Determinants." *Pediatrics* 80(4):502–11.

Mosteller, F., and J.W. Tukey. 1977. *Data Analysis and Regression*: *A Second Course in Statistics*. Reading: Addison-Wesley.

Mustard, C., S. Derksen, J. M. Berthelot, and M. Wolfson. 1999. "Assessing Ecologic Proxies for Household Income: A Comparison of Household and Neighbourhood Level Income Measures in the Study of Population Health Status." *Health and Place* 5(2):157–71.

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

Oreopoulos, P., M. Stabile, L. L. Roos, and R. Walld. 2006. "Short, Medium, and Long Term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins." NBER Working Paper No.11998.

Roos, L. L., R. Walld, J. Uhanova, and R. Bond. 2005. ‘‘Physician Visits, Hospitalizations and Socioeconomic Status: Ambulatory Care Sensitive Conditions in a Canadian Setting.’’ *Health Services Research* 40(4):1167–85.

Roos L. L., and J. P. Nicol. 1999. ‘‘A Research Registry: Uses, Development, and Accuracy.’’ *Journal of Clinical Epidemiology* 52(1):39–47.

Rosenzweig, M., and K. Wolpin. 1995. ‘‘Sisters, Siblings, and Mothers: The Effects of Teen-Age Childbearing on Birth Outcomes in a Dynamic Family Context.’’ *Econometrica* 63(2):303–26.

Watson, D. E., A. Katz, R. J. Reid, B. Bogdanovic, N. Roos, and P. Heppner. 2004. ‘‘Family Physician Workloads and Access to Care in Winnipeg: 1991 to 2001.’’ *Canadian Medical Association Journal* 171(4):339–42.

Willms, J. D. 1986. ‘‘Social Class Segregation and its Relationship to Pupils’ Examination Results in Scotland.’’ *American Sociological Review* 51(2):224–41.

Appendix Table A1

Estimated Effects of Infant Health at Birth on Infant Mortality (Death within One Year of Birth) Cause: Conditions Originating in the Perinatal Period (ICD-9 760-779) With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Twins Sample No Family F.E.	Twins Sample With Family F.E.
N cases	447	243	243	35	35
APGAR Score (Omitted Category APGAR=10)					
APGAR≤6	0.1613 ***	0.1907 ***	0.1922 ***	0.2464 ***	0.0364 **
	(0.0014)	(0.0021)	(0.0029)	(0.0144)	(0.0167)
APGAR=7–8	0.0076 ***	0.0077 ***	0.0069 ***	0.0169 *	−0.0035
	(0.0006)	(0.0010)	(0.0014)	(0.0093)	(0.0110)
APGAR=9	0.0007 *	0.0006	−0.0004	0.0005	−0.0001
	(0.0004)	(0.0006)	(0.0009)	(0.0082)	(0.0104)
F-test: No infant health effects	4582.74 ***	2769.32 ***	1484.64 ***	121.61 ***	2.81 **
Sample size	{108,893}	{54,091}	{54,091}	{1,740}	{1,740}
R-squared	0.11	0.14	0.51	0.31	0.89
Birth Weight (Omitted Category BW = 3500+ grams)					
BW <1000	0.7070 ***	0.7500 ***	0.7551 ***	0.6016 ***	−0.0199
	(0.0027)	(0.0039)	(0.0055)	(0.0238)	(0.0497)
BW 1000–1500	0.1884 ***	0.2298 ***	0.2372 ***	0.1661 ***	0.0421
	(0.0023)	(0.0037)	(0.0051)	(0.0179)	(0.0272)
BW 1500–2500	0.0116 ***	0.0188 ***	0.0227 ***	−0.0011	0.0027
	(0.0007)	(0.0012)	(0.0019)	(0.0141)	(0.0174)

(*continued*)

Table A1 (*continued*)

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Twins Sample No Family F.E.	Twins Sample With Family F.E.
BW 2500–3000	0.0013 ***	0.0013 **	0.0033 ***	−0.0012	0.0017
	(0.0005)	(0.0007)	(0.0011)	(0.0142)	(0.0164)
BW 3000–3500	0.0003	0.0000	0.0006	−0.0027	−0.0010
	(0.0003)	(0.0005)	(0.0008)	(0.0148)	(0.0164)
F-test: No infant health effects	15248.10 **	8113.15 ***	4207.34 ***	224.48 ***	1.20
Sample size	{109,125}	{54,310}	{54,310}	{1,752}	{1,752}
R-squared	0.41	0.43	0.67	0.50	0.88
		Gestation (Omitted Category 40–41 Weeks)			
Gestation≤36 weeks	0.0574 ***	0.0721 ***	0.0753 ***	0.0451 ***	NA
	(0.0009)	(0.0014)	(0.0024)	(0.0107)	
Gestation 37 weeks	0.0013	0.0014	−0.0018	0.0078	NA
	(0.0010)	(0.0015)	(0.0025)	(0.0133)	
Gestation 38 weeks	0.0010	0.0019 **	0.0016	0.0023	NA
	(0.0007)	(0.0009)	(0.0016)	(0.0125)	
Gestation 39 weeks	0.0004	0.0004	0.0000	0.0090	NA
	(0.0005)	(0.0008)	(0.0013)	(0.0131)	
Gestation ≥42 weeks	0.0005	0.0001	0.0002	0.0104	NA
	(0.0008)	(0.0012)	(0.0019)	(0.0304)	
F-test: No infant health effects	866.88 ***	550.53***	217.51 ***	6.81 ***	
Sample size	{90,135}	{45,583}	{45,583}	{1,492}	NA
R-squared	0.05	0.06	0.50	0.23	NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Appendix Table A2

Estimated Effects of Infant Health at Birth on Infant Mortality (Death within One Year of Birth) Cause: Congenital Anomalies (ICD-9 740–759) With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample With Family F.E	Twins Sample No Family F.E.	Twins Sample With Family F.E.
N cases	297	170	170	4	4
APGAR Score (Omitted Category APGAR=10)					
APGAR≤6	0.0634 ***	0.0751 ***	0.0791 ***	N too low	N too low
	(0.0013)	(0.0019)	(0.0027)	(4 cases)	(4 cases)
APGAR=7–8	0.0063 ***	0.0076 ***	0.0091 ***		
	(0.0006)	(0.0009)	(0.0013)		
APGAR=9	0.0010 ***	0.0014 ***	0.0019 **		
	(0.0004)	(0.0005)	(0.0008)		
F-test: No infant health effects	934.48 ***	528.99 ***	303.68 ***		
Sample size	{108,893}	{54,091}	{54,091}	{1,740}	{1,740}
R-squared	0.03	0.03	0.45		
Birth Weight (Omitted Category BW = 3500+ grams)					
BW <1000	0.0451 ***	0.0489 ***	0.0518 ***	N too low	N too low
	(0.0028)	(0.0043)	(0.0059)	(4 cases)	(4 cases)
BW 1000–1500	0.0476 ***	0.0878 ***	0.1062 ***		
	(0.0024)	(0.0040)	(0.0056)		
BW 1500–2500	0.0136 ***	0.0156 ***	0.0252 ***		
	(0.0008)	(0.0013)	(0.0021)		

(*continued*)

Appendix Table A2 (*continued*)

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Twins Sample No Family F.E.	Twins Sample With Family F.E.
BW 2500–3000	0.0033 ***	0.0048 ***	0.0097 ***		
	(0.0005)	(0.0007)	(0.0012)		
BW 3000–3500	0.0004	0.0006	0.0029 ***		
	(0.0004)	(0.0005)	(0.0008)		
F-test: No infant health effects	188.91 ***	150.19 ***	109.44 ***		
Sample size	{109,125}	{54,310}	{54,310}	{1,752}	{1752}
R-squared	0.01	0.02	0.44		
Gestation (Omitted Category 40–41 Weeks)					
Gestation≤36 weeks	0.0150 ***	0.0188 ***	0.0254 ***	N too low	NA
	(0.0008)	(0.0012)	(0.0021)	(4 cases)	
Gestation 37 weeks	0.0025 ***	0.0038 ***	0.0070 ***		NA
	(0.0009)	(0.0013)	(0.0022)		
Gestation 38 weeks	0.0015 ***	0.0019 **	0.0025 *		NA
	(0.0006)	(0.0008)	(0.0014)		
Gestation 39 weeks	0.0005	−0.0001	0.0003		NA
	(0.0005)	(0.0007)	(0.0011)		
Gestation ≥42 weeks	0.0008	0.0010	0.0022		NA
	(0.0007)	(0.0010)	(0.0016)		
F-test: No infant health effects	80.90 ***	49.00 ***	32.04 ***		
Sample size	{90,135}	{45,583}	{45,583}	{1,492}	NA
R-squared	0.01	0.02	0.50		NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Appendix Table A3

Estimated Effects of Infant Health at Birth on Infant Mortality (Death within One Year of Birth) Cause: Symptoms, Signs and Ill-Defined Conditions (includes SIDS) (ICD-9 780–799) With and Without Family Fixed Effects

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Twins Sample No Family F.E.	Twins Sample With Family F.E.
N cases	138	92	92	4	4
APGAR Score (Omitted Category APGAR=10)					
APGAR≤6	0.0050 ***	0.0053 ***	0.0051 **	N too low	N too low
	(0.0008)	(0.0014)	(0.0020)	(4 cases)	(4 cases)
APGAR=7–8	0.0011 ***	0.0022 ***	0.0024 **		
	(0.0004)	(0.0006)	(0.0010)		
APGAR=9	0.0001 *	0.0005	0.0002		
	(0.0002)	(0.0004)	(0.0006)		
F-test: No infant health effects	14.01 ***	7.92 ***	4.25 ***		
Sample size	{108,893}	{54,091}	{54,091}	{1,740}	{1,740}
R-squared	0.00	0.00	0.41		
Birth Weight (Omitted Category BW = 3500+ grams)					
BW <1000	0.0073 ***	0.0039	0.0039	N too low	N too low
	(0.0019)	(0.0032)	(0.0045)	(4 cases)	(4 cases)
BW 1000–1500	0.0052 ***	0.0090 ***	0.0072 *		
	(0.0016)	(0.0030)	(0.0042)		
BW 1500–2500	0.0015 ***	0.0030 ***	0.0021		
	(0.0005)	(0.0010)	(0.0016)		

(*continued*)

Appendix Table A3 (*continued*)

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Twins Sample No Family F.E.	Twins Sample With Family F.E.
BW 2500–3000	0.0003	0.0003	−0.0006		
	(0.0003)	(0.0006)	(0.0009)		
BW 3000–3500	0.0004	0.0008 **	−0.0001		
	(0.0002)	(0.0004)	(0.0006)		
F-test: No infant health effects	6.08 ***	4.17 ***	1.29		
Sample size	{109,125}	{54,310}	{54,310}	{1,752}	{1,752}
R-squared	0.00	0.00	0.41		
Gestation (Omitted Category 40–41 Weeks)					
Gestation≤36 weeks	0.0025 ***	0.0035 ***	0.0033 **	N too low	NA
	(0.0005)	(0.0009)	(0.0015)	(4 cases)	
Gestation 37 weeks	0.0014 **	0.0015	0.0041 ***		NA
	(0.0006)	(0.0010)	(0.0016)		
Gestation 38 weeks	0.0008 **	0.0014 **	0.0016		NA
	(0.0004)	(0.0006)	(0.0010)		
Gestation 39 weeks	0.0001	0.0003	−0.0007		NA
	(0.0003)	(0.0005)	(0.0008)		
Gestation ≥42 weeks	0.0001	0.0001	−0.0021 *		NA
	(0.0005)	(0.0008)	(0.0012)		
F-test: No infant health effects	6.55 ***	3.98 ***	3.71 ***		
Sample size	{90,135}	{45,583}	{45,583}	{1,492}	NA
R-squared	0.00	0.00	0.52		NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Appendix Table A4

Estimated Effects of Infant Health at Birth on Infant Mortality (Death within One Year of Birth) Cause: All Other Diagnoses With and Without Family Family Effects

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Twins Sample No Family F.E.	Twins Sample With Family F.E.
N cases	693	407	407	10	10
APGAR Score (Omitted Category APGAR=10)					
APGAR≤6	0.0209 ***	0.0267 ***	0.0281 ***	N too low	N too low
	(0.0019)	(0.0030)	(0.0041)	(10 cases)	(10 cases)
APGAR=7–8	0.0035 ***	0.0056 ***	0.0055 ***		
	(0.0009)	(0.0013)	(0.0019)		
APGAR=9	0.0000	−0.0004	0.0009		
	(0.0006)	(0.0008)	(0.0012)		
F-test: No infant health effects	48.47 ***	34.74 ***	18.19 ***		
Sample size	{108,893}	{54,091}	{54,091}	{1,740}	{1,740}
R-squared	0.01	0.01	0.46		
Birth Weight (Omitted Category BW = 3500+ grams)					
BW <1000	0.0157 ***	0.0077	0.0066	N too low	N too low
	(0.0043)	(0.0067)	(0.0091)	(10 cases)	(10 cases)
BW 1000–1500	0.0173 ***	0.0191 ***	0.0244***		
	(0.0037)	(0.0062)	(0.0085)		
BW 1500–2500	0.0069 ***	0.0124 ***	0.0128 ***		
	(0.0012)	(0.0021)	(0.0031)		

(*continued*)

Appendix Table A4 (*continued*)

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Twins Sample No Family F.E.	Twins Sample With Family F.E.
BW 2500–3000	0.0010	0.0005	0.0016		
	(0.0007)	(0.0012)	(0.0019)		
BW 3000–3500	0.0006	0.0003	0.0028 **		
	(0.0005)	(0.0008)	(0.0013)		
F-test: No infant health effects	13.16 ***	9.23 ***	5.08 ***		
Sample size	{109,125}	{54,310}	{54,310}	{1,752}	{1,752}
R-squared	0.01	0.01	0.46		
Gestation (Omitted Category 40–41 Weeks)					
Gestation≤36 weeks	0.0055 ***	0.0053 ***	0.0081 ***	N too low	NA
	(0.0011)	(0.0019)	(0.0030)	(10 cases)	
Gestation 37 weeks	0.0005	0.0012	0.0041		NA
	(0.0013)	(0.0020)	(0.0032)		
Gestation 38 weeks	0.0006	0.0014	0.0028		NA
	(0.0008)	(0.0013)	(0.0020)		
Gestation 39 weeks	0.0003	−0.0007	0.0002		NA
	(0.0007)	(0.0010)	(0.0016)		
Gestation ≥42 weeks	0.0003	−0.0004	−0.0002		NA
	(0.0010)	(0.0016)	(0.0024)		
F-test: No infant health effects	4.86 ***	2.16 *	1.80		
Sample size	{90,135}	{45,583}	{45,583}	{1,492}	NA
R-squared	0.01	0.01	0.52		NA

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Appendix Table A5

Estimated Effects of Infant Health at Birth on Infant Mortality (Death within One Year of Birth) With and Without Family Fixed Effects

	First-born <3000g, Later sibs >3000g		First-born >3000g, one later sib <3000g	
	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.
APGAR Score (Omitted Category APGAR=10)				
APGAR≤6	0.4979 ***	0.4929 ***	0.4047 ***	0.4199 ***
	(0.0104)	(0.0144)	(0.0111)	(0.0150)
APGAR=7–8	0.0329 ***	0.0366 ***	0.0455 ***	0.0524 ***
	(0.0056)	(0.0083)	(0.0065)	(0.0092)
APGAR=9	0.0068 *	0.0029	0.0032	0.0025
	(0.0038)	(0.0059)	(0.0043)	(0.0062)
F-test: No infant health effects	795.80 ***	417.97 ***	472.66 ***	288.10 ***
Sample size	{6,265}	{6,265}	{5,433}	{5,433}
R-squared	0.30	0.61	0.22	0.54
Birth Weight (Omitted Category BW = 3500+ grams)				
BW <1000	0.9345 ***	0.9323 ***	0.8021 ***	0.8263 ***
	(0.0172)	(0.0235)	(0.0200)	(0.0263)
BW 1000–1500	0.5444 ***	0.5674 ***	0.3679 ***	0.4226 ***
	(0.0170)	(0.0235)	(0.0197)	(0.0260)
BW 1500–2500	0.0648 ***	0.0738 ***	0.0826 ***	0.0942 ***
	(0.0080)	(0.0117)	(0.0082)	(0.0104)
BW 2500–3000	0.0107 *	0.0149	0.0161 ***	0.0304 ***
	(0.0064)	(0.0096)	(0.0061)	(0.0076)
BW 3000–3500	−0.0005	0.0139 **	−0.0013	0.0089
	(0.0042)	(0.0059)	(0.0053)	(0.0071)

(continued)

Appendix Table A5 (*continued*)

	First-born <3000g, Later sibs >3000g		First-born >3000g, one later sib <3000g	
	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.	Sibling Sample No Family F.E.	Sibling Sample With Family F.E.
F-test: No infant health effects	852.77 ***	461.34	408.29 ***	255.52 ***
Sample size	{6,280}	{6,280}	{5,451}	{5,451}
R-squared	0.42	0.68	0.29	0.58
Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	0.1344 ***	0.1532 ***	0.1340 ***	0.1592 ***
	(0.0074)	(0.0118)	(0.0076)	(0.0116)
Gestation 37 weeks	0.0074	0.0154	0.0037	−0.0017
	(0.0082)	(0.0136)	(0.0095)	(0.0147)
Gestation 38 weeks	0.0034	0.0011	−0.0010	0.0039
	(0.0056)	(0.0097)	(0.0071)	(0.0110)
Gestation 39 weeks	0.0029	−0.0037	0.0076	0.0086
	(0.0049)	(0.0082)	(0.0066)	(0.0097)
Gestation ≥42 weeks	0.0069	0.0149	−0.0017	−0.0104
	(0.0088)	(0.0145)	(0.0116)	(0.0164)
F-test: No infant health effects	71.99 ***	41.04 ***	71.58 ***	43.56 ***
Sample size	{5,300}	{5,300}	{4,566}	{4,566}
R-squared	0.10	0.52	0.09	0.53

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively. Sample includes only those siblings where at least one child had a birth weight of less than 3000g.

Appendix Table A6

Estimated Effects of Infant Health at Birth on Reaching Grade 12 by Age 17 With and Without Family Fixed Effects

	First-born <3000g, Later sibs >3000g		First-born >3000g, one later sib <3000g	
	Sibling Sample No Family F.E.	Sibling Sample Family F.E.	Sibling Sample No Family F.E.	Sibling Sample Family F.E.
	APGAR Score (Omitted Category APGAR=10)			
APGAR≤6	−0.1817 ***	−0.1726 ***	−0.0013	0.0381
	(0.0535)	(0.0619)	(0.0550)	(0.0622)
APGAR=7–8	−0.0484 **	−0.0416 *	−0.0035	−0.0002
	(0.0207)	(0.0253)	(0.0248)	(0.0288)
APGAR=9	−0.0101	0.0191	0.0453 ***	0.0503 ***
	(0.0138)	(0.0174)	(0.0160)	(0.0189)
F-test: No infant health effects	5.29 ***	3.07 **	3.46 **	3.09 **
Sample size	{4,668}	{4,668}	{3,942}	{3,942}
R-squared	0.21	0.73	0.18	0.70
	Birth Weight (Omitted Category BW = 3500+ grams)			
BW <1000	0.3327	0.6136 **	−0.3588 **	−0.1453
	(0.2388)	(0.2623)	(0.1585)	(0.1720)
BW 1000–1500	−0.0559	−0.0202	−0.1054	−0.2668 **
	(0.0971)	(0.1097)	(0.0934)	(0.1054)
BW 1500–2500	−0.0751 **	−0.0945 **	−0.1329 ***	−0.0880 **
	(0.0341)	(0.0405)	(0.0347)	(0.0366)
BW 2500–3000	−0.0282	−0.0573 *	−0.0154	−0.0487 *
	(0.0268)	(0.0328)	(0.0250)	(0.0257)

(*continued*)

Appendix Table A6 (*continued*)

	First-born <3000g, Later sibs >3000g		First-born >3000g, one later sib <3000g	
	Sibling Sample No Family F.E.	Sibling Sample Family F.E.	Sibling Sample No Family F.E.	Sibling Sample Family F.E.
BW 3000–3500	−0.0233	−0.0424 **	0.0107	0.0040
	(0.0176)	(0.0199)	(0.0216)	(0.0242)
F-test: No infant health effects	1.7	2.81 **	5.09 ***	2.94 ***
Sample size	{4,427}	{4,427}	{3,708}	{3,708}
R-squared	0.19	0.71	0.18	0.68
Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	−0.0492 *	−0.0032	−0.0408	−0.0417
	(0.0272)	(0.0330)	(0.0270)	(0.0338)
Gestation 37 weeks	−0.0174	−0.0023	−0.0395	−0.0712 *
	(0.0285)	(0.0366)	(0.0334)	(0.0423)
Gestation 38 weeks	0.0055	−0.0026	0.0038	−0.0133
	(0.0194)	(0.0256)	(0.0235)	(0.0299)
Gestation 39 weeks	0.0394 **	0.0216	0.0287	−0.0009
	(0.0173)	(0.0217)	(0.0222)	(0.0268)
Gestation ≥42 weeks	0.0552 *	0.0276	−0.0447	−0.0383
	(0.0311)	(0.0390)	(0.0385)	(0.0436)
F-test: No infant health effects	2.86 **	0.35	1.72	0.91
Sample size	{3,770}	{3,770}	{3,101}	{3,101}
R-squared	0.21	0.76	0.19	0.74

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.

Appendix Table A7
Estimated Effects of Infant Health at Birth on Reaching Grade 12 by Age 17 With and Without Family Fixed Effects

Unlike in other models, these coefficients were generated by a model including both birthweight and gestation

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample Family F.E.	Twins Sample No Family F.E.	Twins Sample Family F.E.
	Birth Weight (Omitted Category BW = 3500+ grams)				
BW <1000	−0.1707 ***	−0.1039	0.0603	−0.3029	−0.0837
	(0.0653)	(0.1066)	(0.1161)	(0.1979)	(0.3434)
BW 1000–1500	−0.1224 ***	−0.0984 **	−0.1362 **	−0.2446 ***	−0.2315 **
	(0.0277)	(0.0465)	(0.0542)	(0.0906)	(0.0954)
BW 1500–2500	−0.0722 ***	−0.1012 ***	−0.0768 ***	−0.1166	−0.0685 *
	(0.0088)	(0.0137)	(0.0171)	(0.0645)	(0.0607)
BW 2500–3000	−0.0392 ***	−0.0459 ***	−0.0395 ***	−0.1156 *	−0.0921
	(0.0048)	(0.0068)	(0.0093)	(0.0630)	(0.0570)
BW 3000–3500	−0.0131 ***	−0.0186 ***	−0.0227 ***	−0.0682	−0.0488
	(0.0034)	(0.0047)	(0.0061)	(0.0656)	(0.0533)
F-test: No infant health effects	24.17 ***	17.02 ***	6.73 ***	1.93 *	1.72
Sample size	{79,363}	{40,203}	{40,203}	{1,354}	{1,354}
R-squared	0.16	0.18	0.69	0.20	0.88
	Gestation (Omitted Category 40–41 Weeks)				
Gestation≤36 weeks	−0.0139	−0.0320 **	−0.0074	−0.0107	NA
	(0.0087)	(0.0132)	(0.0156)	(0.0441)	

(*continued*)

Appendix Table A7 (*continued*)

Unlike in other models, these coefficients were generated by a model including both birthweight and gestation

	Full Sample	Sibling Sample No Family F.E.	Sibling Sample Family F.E.	Twins Sample No Family F.E.	Twins Sample Family F.E.
Gestation 37 weeks	−0.0061	−0.0218 *	−0.0225	0.0648	NA
	(0.0083)	(0.0121)	(0.0142)	(0.0490)	
Gestation 38 weeks	0.0032	−0.0030	−0.0146 *	0.0593	NA
	(0.0053)	(0.0073)	(0.0088)	(0.0457)	
Gestation 39 weeks	−0.0118 ***	0.0110 *	0.0007	0.0650	NA
	(0.0043)	(0.0059)	(0.0070)	(0.0481)	
Gestation ≥42 weeks	−0.0088	−0.0121	−0.0027	−0.1764 *	NA
	(0.0065)	(0.0090)	(0.0105)	(0.1002)	NA
Gestation Missing	0.0077	0.0141	−0.0074	0.0354	
	(0.0095)	(0.0135)	(0.0153)	(0.0760)	
F-test (6 df):	2.74 **	2.98 ***	0.86	1.89 *	

All regression models include additional fixed effects for mother's marital status, gender of child, and family sibling size dummies for the birth order of the child within each family size. One, two, and three asteriks indicate statistical significance at the 10, 5, and 1 percent levels respectively.