Web Appendix A: Data Creation

I use two steps to create the California twins dataset. In the first step, I identify the birth of twins in the 1960-1988 California birth files using the plurality indicator variable. I retain all twins for which there are exactly two observations for each unique child birth date and child first and last name combination.[67] Since I focus on differences in adult outcomes for twins later observed giving birth, I discard all twin pairs in which at least one of the twins is male.

In the second step, I match the birth record information of the twins' births to that of the birth of their offspring. That is, for each twin, I determine whether I later observe her giving birth in California. I perform this matching procedure for all 1989 to 2002 California births. The matching is based on the twin's first and last name and her date of birth.[68] This matching algorithm will miss all births to twins occurring outside of California or happening before 1989 or after 2002. The final dataset consists of same sex female twins born between 1960 and 1988; for those twins whom I observe having a California birth between 1989 and 2002, I have longer-run outcomes (e.g., health, education) measured when they give birth. Hence, because birthweight may be a predictor of whether I observe a twin as a mother, the estimation sample used to estimate the effect of birthweight on adult outcomes may be selective, an issue I address directly in the main text.

Appendix Table A1 provides further details about the creation of the estimation sample (e.g., the dropping of observations). Each row in this table represents a different birth year. The columns represent different samples. For instance, in 1979, there were 380,271 births in California. Of those births, there were 7,407 twin births and 2,490 same sex female twin births.[69] My primary estimation sample consists of twins for whom I observe both twins having their first birth or both twins having their second birth (Selection 7). As the median number of children is two, this sample is likely reflective of the typical mother. Moreover, qualitatively, the results are the same when only using twins for whom I observe the first birth. The foreseeable advantage to including the second-birth mothers is the increase in statistical precision.

In addition to imposing the first and second birth restriction, I also limit the twins sample to women born between 1960 and 1982. As seen in Appendix Table A1, some of the 1960-1988 twins were born too recently to have had births yet. Since the mean age at first birth is twenty, the earlier cohorts in the estimation sample are not so selective.

Appendix Figure A1 provides a sense of the performance of the matching algorithm, which links the twin's own birth to that of her children. The solid line displays the percent of same-sex female births born in California having an observed first birth in California in the data. To indicate how large this percentage could be, the long dashed line represents the fraction of all women in California born between 1960 and 1988 who later gave birth in California. Here I am not performing the one-to-one matching I did for the twins. Instead, based on the California birth records for 1989-2002, I calculate the number of women who were born in each year and

[67] This selection criterion will delete all twins born on different days and those with different last names. However, the number of twins discarded is small.

[68] Some women born between 1960 and 1988 will have births before 1989. I could match these pre-1989 births to the birth certificates of their twin mothers born between 1960 and 1988. However, I use the twin's exact date of birth and her first and last name to match her to the births of her children, and mother's exact date of birth was first reported in 1989.

[69] In these calculations, each twin pair represents two births. However, the number of twin births is odd. Errors in the plurality variable, which I use to identify twin pairs, cause this discrepancy. When I create the twins' dataset, I confirm that both twins are present in the data. Otherwise, the twin pair (or lack thereof) is dropped.

had births between 1989 and 2002. Then, I divide these counts by the number of females born in those years. The long-dashed line is this ratio by year of birth. Therefore, if the fertility patterns of twins exactly followed that of the overall female population and if matching were perfect, the solid line and the long-dashed line would coincide. These lines differ in most years, but imperfect matching likely explains the difference. The best guess at the match rate is 80 percent, which is consistent with the gap between the two curves. It does not explain, however, why the gap varies across years.

From the national Detailed Natality data, I also estimate the fraction of women who were born in California but gave birth outside of California. While this is not a trivial fraction, women born in California usually give birth in California. To reiterate, while Appendix Figure A1 effectively summarizes how well the matching works, one should not be so concerned with these measures. The twin estimates will only be biased by selection to the extent that amongst twins, there is a differential probability of giving birth that is correlated with the twins' differences in birthweight. One cannot ascertain this from Appendix Figure A1.

Web Appendix B: A Test of Sample Selection Bias

There are several methods resting on different assumptions that one could use to test or quantify the effects of sample selection. Here I develop a "non-parametric" test for sample selection bias motivated by a simple model of sample selection. To understand this approach, first consider the traditional selection model applied to the fixed-effects model:

(A1) $$s_{ij}^{*} = \boldsymbol{z}_{ij}\theta + \varepsilon_{ij}$$

and

(A2) $$y_{ij} = \begin{cases} bw_{ij}\beta + \gamma_i + u_{ij} & \text{if } s_{ij} = 1(s_{ij}^{*} > 0) = 1 \\ \text{unobserved} & \text{otherwise} \end{cases}$$

where s_{ij}^{*} is the selection index, $\boldsymbol{z}_{ij}$ is a vector of characteristics affecting sample selection, and γ_i is the twin fixed effect. Therefore, differencing across twins,

(A3) $$y_{i2} - y_{i1} = \begin{cases} (bw_{i2} - bw_{i1})\beta + u_{i2} - u_{i1} & \text{if } s_{i1} = 1 \text{ and } s_{i2} = 1 \\ \text{unobserved} & \text{otherwise} \end{cases}$$

Now suppose we assume the standard assumptions of the James Heckman (1979) selection model including the joint bivariate normality of ε_{ij} and u_{ij}, then

(A4) $$y_{i2} - y_{i1} = \begin{cases} (bw_{i2} - bw_{i1})\beta + \left(\dfrac{\phi(\boldsymbol{z}_{i2}\theta)}{\Phi(\boldsymbol{z}_{i2}\theta)} - \dfrac{\phi(\boldsymbol{z}_{i1}\theta)}{\Phi(\boldsymbol{z}_{i1}\theta)} \right)\rho\sigma + \xi_{i2} - \xi_{i1} & \text{if } s_{i1} = 1 \text{ and } s_{i2} = 1 \\ \text{unobserved} & \text{otherwise} \end{cases}$$

where ϕ is the standard normal pdf, Φ is the standard normal cdf, ρ is the correlation between u and ε, σ is the standard deviation of u, and ξ_{ij} is an idiosyncratic error term. If $\boldsymbol{z}_{ij}$ is twin-invariant (i.e., factors contributing to selection are the same for each twin), then sample selection is not an issue because the selection term is zero. However, it is highly-probable that $\boldsymbol{z}_{ij}$ varies within twin pairs. In particular, it might be a function of birthweight. If so, we might be inclined to include powers of birthweight (e.g., birthweight-squared) in the fixed-effects regressions. However, the simple equation above is not informative about how many powers of birthweight to include and moreover, the joint normality assumption makes the selection bias term additive separable whereas under other distributional assumptions, this would not be necessarily true.

The reasoning behind the non-parametric test becomes clearer if I rewrite equation (A4) as follows (John DiNardo, McCrary, and Lisa Sanbonmatsu (2006)):

(A5)

$$y_{i2} - y_{i1} = \begin{cases} (bw_{i2} - bw_{i1})\beta + \left(\frac{\phi(\Phi^{-1}(p_{i2}))}{p_{i2}} - \frac{\phi(\Phi^{-1}(p_{i1}))}{p_{i1}} \right) \rho\sigma + \xi_{i2} - \xi_{i1} & \text{if } s_{i1} = 1 \text{ and } s_{i2} = 1 \\ \text{unobserved} & \text{otherwise} \end{cases}$$

where p_{ij} is the fraction observed beyond birth (i.e., at time of motherhood).

A priori, we might believe that $\hat{\beta}$, when not taking into account sample selection, is downward-biased. This is because ρ is likely negative and the inverse Mill's ratio, a decreasing function of the degree of sample selection, are probably positively correlated with birthweight. However, this intuition does not give us any sense of the magnitude of the sample selection bias. But equation (A5) implies that for samples in which the twin birthweight difference strongly predicts the within twin-pair difference in the probability of later observation, the usual fixed-effects estimate will be the most biased. Therefore, because the degree of bias is related to the correlation between $\frac{\phi(\Phi^{-1}(p_{i2}))}{p_{i2}} - \frac{\phi(\Phi^{-1}(p_{i1}))}{p_{i1}}$ and $bw_{i2} - bw_{i1}$, if selection bias is severe, we should expect the fixed effects estimates to vary across groups classified by the degree of correlation between $\frac{\phi(\Phi^{-1}(p_{i2}))}{p_{i2}} - \frac{\phi(\Phi^{-1}(p_{i1}))}{p_{i1}}$ and $bw_{i2} - bw_{i1}$.

This intuition provides the basis for the "non-parametric" test. First, suppose I arrange the twin pairs into k groups, which I believe *a priori* may differ by the degree of correlation between $\frac{\phi(\Phi^{-1}(p_{i2}))}{p_{i2}} - \frac{\phi(\Phi^{-1}(p_{i1}))}{p_{i1}}$ and $bw_{i2} - bw_{i1}$. Then, I estimate the following regression:

(A6) $$s_{i2} - s_{i1} = \sum_k \lambda_k D_i^k (bw_{i2} - bw_{i1}) + \psi_{i2} - \psi_{i1}$$

where D_i^k is an indicator for whether the twin pair i belongs to the kth group and Ψ_{ij} is an idiosyncratic error term. Then, I test whether $\lambda_k = \lambda_l \ \forall k, l$. If I reject this null hypothesis that the effect of birthweight on selection is the same across groups, then when I estimate the main regression equation but allow the effect of birthweight to differ across these k groups (i.e., estimate $y_{i2} - y_{i1} = \sum_k \eta_k D_i^k (bw_{i2} - bw_{i1}) + \phi_{i2} - \phi_{i1}$), I should be able to reject the hypothesis that $\eta_k = \eta_l \ \forall k, l$ in the presence of strong sample selection bias. If I am unable to reject this hypothesis, then I would suspect that although selection into the sample is related to within-twin-pair birthweight differences, such selection does not severely bias the fixed-effects estimates. I perform this test for all outcomes where each group is a birth-year cohort.

Web Appendix C: The Inconsistency of the Twin Fixed Effect Estimator Using Fetal Growth

While there are several appealing reasons to use fetal growth (birthweight/gestation) rather than birthweight to look at the long-run effects of infant health, measurement error in models using fetal growth as an independent variable is likely quite problematic. In particular, under classical measurement error, one can show that measurement error in fetal growth due to mismeasurement of gestational length can potentially lead to *upward*-biased estimates of the effect of fetal growth.

To demonstrate this, consider a simplified version of equation (3) where I replace the within-twin difference in birthweight by the within-twin difference in fetal growth[70]:

(A7) $$y_{i2} - y_{i1} = \left[\frac{bw_{i2} - bw_{i1}}{g_i^*}\right]\theta + \varepsilon_{i2} - \varepsilon_{i1}$$

where g_i^* is the true gestational length of the twins, assumed to be identical for the twins. Equation (A7) is the true model. Now suppose that birthweight is measured without error but gestational length is imprecisely measured. That is,

(A8) $$g_i = g_i^* + \eta_i$$

where g_i is the mismeasured gestational length and the measurement error η_i is classical (i.e., uncorrelated with the true gestational length and with an expected value of 0). Then, the twin fixed effect estimate of the effect of fetal growth using the mismeasured gestational length will be:

(A9) $$\hat{\theta} = \frac{\sum\left[\frac{bw_{i2} - bw_{i1}}{g_i}\right][y_{i2} - y_{i1}]}{\sum\left[\frac{bw_{i2} - bw_{i1}}{g_i}\right]^2}.$$

or, equivalently, given the assumptions,

(A10) $$\hat{\theta} = \frac{\theta\sum\left[\frac{(bw_{i2} - bw_{i1})^2}{(g_i^* + \eta_i)g_i^*}\right] + \sum\left[\frac{bw_{i2} - bw_{i1}}{g_i^* + \eta_i}\right][\varepsilon_{i2} - \varepsilon_{i1}]}{\sum\left[\frac{bw_{i2} - bw_{i1}}{g_i^* + \eta_i}\right]^2}$$

Taking the plim of both sides,

(A11) $$\text{plim}\,\hat{\theta} = \theta\,\text{plim}\frac{\frac{1}{N}\sum\left[\frac{(bw_{i2} - bw_{i1})^2}{(g_i^* + \eta_i)g_i^*}\right]}{\frac{1}{N}\sum\left[\frac{bw_{i2} - bw_{i1}}{g_i^* + \eta_i}\right]^2}$$

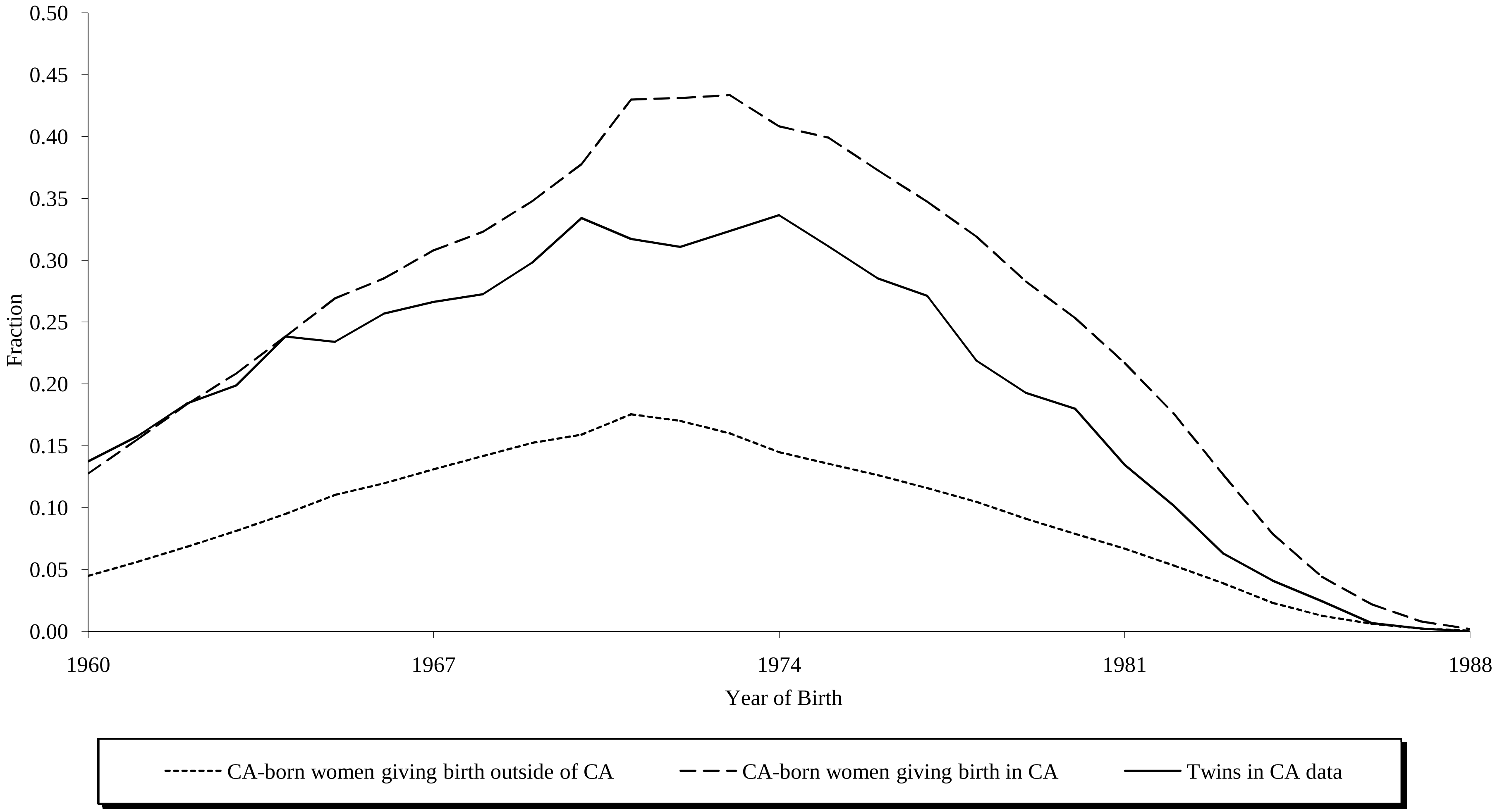
Therefore, only in the unusual case that

[70] I am ignoring covariates in this equation as the estimating equation excluding covariates more closely resembles the main estimating equation.

(A12) $$\text{plim}\frac{1}{N}\sum\left[\frac{(bw_{i2}-bw_{i1})^2}{(g_i^*+\eta_i)g_i^*}\right]=\text{plim }\frac{1}{N}\sum\left[\frac{bw_{i2}-bw_{i1}}{g_i^*+\eta_i}\right]^2$$

will the probability limit of $\hat{\theta}$ equal θ. Unlike the classical error-in-variables model, from the theory, the direction of bias is unclear. In particular, it is unclear whether $\text{plim}\frac{1}{N}\sum\left[\frac{(bw_{i2}-bw_{i1})^2}{(g_i^*+\eta_i)g_i^*}\right]$ is larger or smaller than $\text{plim }\frac{1}{N}\sum\left[\frac{bw_{i2}-bw_{i1}}{g_i^*+\eta_i}\right]^2$ without further assumptions.

Appendix Figure A1: Fraction of Women Having a First Birth Between 1989 and 2002 by Their Own Year of Birth



Notes: The fraction of CA-born women having a 1st birth between 1989 and 2001 outside of CA is calculated using the national Natality Detail Files. The other fractions are computed using the California Birth Statistical Master Files.

Appendix Figure A2: Fraction of Births that are Twins By Year

Notes: The sample includes all California births.

Appendix Table A1: Selection of the Twins Sample

	Number of Observations									
Year of Birth	Births	All Twins Births	Selection 1	Selection 2	Selection 3	Selection 4	Selection 5	Selection 6	Selection 7
1960	372,994	7,616	6,761	6,430	2,184	970	296	104	158
1961	382,420	7,722	6,731	6,442	2,130	996	298	122	186
1962	379,562	7,615	6,768	6,470	2,172	1,120	368	140	210
1963	381,756	7,745	6,781	6,454	2,162	1,092	418	186	270
1964	375,762	7,335	6,506	6,214	2,098	1,192	470	214	312
1965	356,268	7,339	6,469	6,060	2,038	1,110	402	180	256
1966	338,915	7,046	6,246	5,658	2,004	1,082	464	216	332
1967	337,952	7,397	6,624	5,776	1,944	1,044	458	242	330
1968	340,507	6,946	6,224	5,856	2,048	1,092	440	240	320
1969	354,069	7,060	6,351	6,042	2,076	1,120	474	282	386
1970	363,799	6,998	6,329	6,042	2,122	1,154	542	402	468
1971	330,796	6,084	5,511	5,262	1,834	926	412	326	374
1972	307,400	5,617	5,137	4,906	1,798	882	376	304	342
1973	298,874	5,526	5,048	4,806	1,736	840	396	326	366
1974	312,802	5,964	5,456	5,180	1,952	978	424	370	400
1975	318,390	6,051	5,501	5,242	1,924	880	370	340	354
1976	333,130	6,364	5,812	5,516	2,068	848	392	352	370
1977	348,467	6,441	5,936	5,652	2,012	838	300	268	276
1978	357,135	6,919	6,722	6,414	2,416	802	306	276	288
1979	380,271	7,407	7,231	6,914	2,490	724	286	264	270
1980	403,980	7,668	7,497	7,194	2,712	762	242	230	236
1981	421,771	7,925	7,754	7,462	2,704	560	188	178	178
1982	430,905	8,211	8,126	7,830	2,968	498	112	108	110
1983	436,984	8,532	8,440	8,082	3,072	324	68	62	64
1984	448,694	8,766	8,630	8,300	2,832	200	36	36	36
1985	472,190	9,338	9,214	8,868	3,300	146	18	16	16
1986	483,381	9,893	9,780	9,436	3,360	42	4	4	4
1987	504,853	10,406	10,289	9,900	3,590	16	0	0	0
1988	534,174	10,908	10,806	10,398	3,782	0	0	0	0
Total	11,108,201	218,839	204,680	194,806	69,528	22,238	8,560	5,788	6,912

Selection criteria:
Selection 1: Non-missing name, non-missing date of birth, non-missing birthweight
Selection 2: Selection 1 + Twins with only two births per unique date of birth and last name combination + twins with each twin with a unique date of birth, first and last
Selection 3: Selection 2 + Same sex female
Selection 4: Selection 3 + Observe at least one twin giving birth
Selection 5: Selection 3 + Observe both twins giving birth
Selection 6: Selection 3 + Observe first birth for both twins
Selection 7: Selection 3 + Observe first or second birth for both twins

Appendix Table A2: 2003 American Community Survey Comparison

Descriptive Statistics - Fraction of Female Population with Certain Characteristics

	Born in CA between 1960 and 1982		*Born outside of CA between 1960 and 1982*
	Meets twin sample criteria	Does not meet twin sample	
Less than high school degree	0.11	0.07	0.09
High school degree	0.29	0.19	0.28
Some college	0.40	0.38	0.35
College	0.15	0.26	0.21
More than college	0.05	0.10	0.07
Married	0.68	0.38	0.54
Separated, divorced, or widowed	0.16	0.13	0.15
Never married	0.16	0.50	0.31
Employed	0.92	0.92	0.92
Not in labor force	0.69	0.78	0.77
Disability	0.04	0.05	0.04
Mobility limitation	0.02	0.03	0.02
Personal care limitation	0.01	0.02	0.01
Physical difficulty	0.05	0.05	0.04
White	0.72	0.74	0.79
Number of observations	4,020	7,691	144,377

Notes: To meet the twin criteria, a woman must (a) have given birth between 1989 and 2002 and still be living with that child and (2) have been living in her state of birth.

Appendix Table A3: Pooled OLS and Fixed Effect Estimates: Effects of Birthweight in Kilograms Female Twins with First Birth Observed

Education

Maximum Education		Mean Education		Education at First Birth	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
0.13	0.17	0.14	0.21	0.15	0.21
(0.06)	(0.08)	(0.06)	(0.08)	(0.06)	(0.09)
Mean: 13.12		Mean: 12.89		Mean: 12.77	

Birth and Adult Health Outcomes

Child's Birthweight (in Grams)		Gestational Length (in Days)		Hypertension		Diabetes	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
171.23	60.48	2.08	0.99	-0.005	-0.02	-0.005	-0.01
(16.24)	(33.43)	(0.54)	(1.34)	(0.005)	(0.01)	(0.003)	(0.01)
Mean: 3363.34		Mean: 277.71		Mean: 0.03		Mean: 0.01	

Anemia		# of Pregnancy Complications		# of Labor Complications		Neonatal Intensive Care Unit Transfer	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-0.003	-0.01	-0.02	-0.08	-0.01	-0.04	-0.008	-0.01
(0.003)	(0.01)	(0.01)	(0.03)	(0.03)	(0.06)	(0.005)	(0.01)
Mean: 0.01		Mean: 0.12		Mean: 0.58		0.0259	

Birth Delivery / **Residential Location**

C-Section Delivery		Public Payment for Delivery		Median Household Income (1999) of Zipcode		Poverty Rate (1999) of Zipcode	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-0.03	-0.05	-0.03	0.0003	918.05	-542.85	-0.001	0.001
(0.01)	(0.03)	(0.01)	(0.0245)	(525.62)	(844.84)	(0.003)	(0.004)
Mean: 0.24		Mean: 0.39		Mean: 47363.1		Mean: 0.15	

Maternal and Paternal Characteristics

Maternal Age (in Days)		Father Present		Paternal Age (in Days)		Paternal Education	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-34.15	75.29	0.02	0.003	-21.25	72.78	0.19	0.24
(37.36)	(68.71)	(0.01)	(0.014)	(59.37)	(120.07)	(0.09)	(0.16)
Mean: 8621.99		Mean: 0.95		Mean: 9785.37		Mean: 12.80	

Notes: Robust standard errors adjusted for within-twin-pair correlation are shown in parentheses. All regressions are based on the sample of twins observed having a first birth in California between 1989 and 2002 (2,835 twin pairs). The reported mean is the mean of the dependent variable. The pooled OLS regressions include controls for year of birth, birth order, and race.

Appendix Table A4: Pooled OLS and Fixed Effect Estimates: Effects of Birthweight in Kilograms
Female Twins Born Between with First or Second Birth Observed
Sample with Congenital Anomaly Information

Education

Maximum Education		Mean Education		Education at Birth	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
0.16	0.13	0.17	0.16	0.20	0.22
(0.08)	(0.11)	(0.08)	(0.11)	(0.08)	(0.11)
Mean: 13.40		Mean: 13.19		Mean: 13.2	

Birth and Adult Health Outcomes

Child's Birthweight (in Grams)		Gestational Length (in Days)		Hypertension		Diabetes	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
187.15	50.75	3.36	0.48	-0.01	-0.03	-0.01	-0.01
(20.91)	(44.32)	(0.69)	(1.60)	(0.01)	(0.01)	(0.01)	(0.01)
Mean: 3411.03		Mean: 276.87		Mean: 0.03		Mean: 0.02	

Anemia		# of Pregnancy Complications		# of Labor Complications		Neonatal Intensive Care Unit Transfer	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-0.003	0.01	-0.02	-0.05	0.02	0.02	-0.01	0.01
(0.003)	(0.01)	(0.02)	(0.03)	(0.03)	(0.07)	(0.01)	(0.01)
Mean: 0.01		Mean: 0.12		Mean: 0.50		Mean: 0.03	

Birth Delivery				Residential Location			
C-Section Delivery		Public Payment for Delivery		Median Household Income (1999) of Zipcode		Poverty Rate (1999) of Zipcode	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-0.03	0.02	-0.03	-0.01	1686.30	1131.53	-0.003	-0.01
(0.02)	(0.04)	(0.02)	(0.03)	(734.06)	(1176.89)	(0.004)	(0.01)
Mean: 0.26		Mean: 0.31		Mean: 49682.62		Mean: 0.14	

Maternal and Paternal Characteristics

Maternal Age (in Days)		Father Present		Paternal Age (in Days)		Paternal Education	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-26.62	63.66	0.03	0.01	-23.74	98.26	0.23	0.46
(45.74)	(94.10)	(0.01)	(0.02)	(74.03)	(157.75)	(0.10)	(0.18)
Mean: 9710.46		Mean: 0.95		Mean: 10918.62		Mean: 13.37	

Notes: Robust standard errors adjusted for within-twin-pair correlations are shown in parentheses. All regressions are based on the sample of twins born between either 1960 and 1967 or 1978 and 1982 and observed having a first or second birth in California between 1989 and 2002 (1,568 twin pairs). For those twins whose first and second births are both observed, only the second birth is included in the regressions. The reported mean is the mean of the dependent variable. The pooled OLS regressions include controls for year of birth, birth order, and race.

Appendix Table A5: Pooled OLS and Fixed Effect Estimates: Effects of Birthweight in Kilograms
Female Twins with First or Second Birth Observed
Sample with No Congenital Anomalies

Education

Maximum Education		Mean Education		Education at Birth	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
0.15	0.15	0.17	0.19	0.19	0.25
(0.08)	(0.11)	(0.08)	(0.11)	(0.08)	(0.12)
Mean: 13.41		Mean: 13.20		Mean: 13.21	

Birth and Adult Health Outcomes

Child's Birthweight (in Grams)		Gestational Length (in Days)		Hypertension		Diabetes	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
188.06	67.82	3.29	0.66	-0.01	-0.03	-0.01	-0.02
(21.44)	(45.09)	(0.68)	(1.58)	(0.01)	(0.01)	(0.01)	(0.01)
Mean: 3411.75		Mean: 277.01		Mean: 0.03		Mean: 0.02	

Anemia		# of Pregnancy Complications		# of Labor Complications		Neonatal Intensive Care Unit Transfer	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-0.003	0.01	-0.02	-0.05	0.03	0.02	-0.01	0.02
(0.003)	(0.01)	(0.02)	(0.03)	(0.03)	(0.07)	(0.01)	(0.01)
Mean: 0.01		Mean: 0.12		Mean: 0.50		Mean: 0.03	

Birth Delivery / **Residential Location**

C-Section Delivery		Public Payment for Delivery		Median Household Income (1999) of Zipcode		Poverty Rate (1999) of Zipcode	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-0.02	0.02	-0.03	-0.005	2001.15	1220.59	-0.004	-0.01
(0.02)	(0.04)	(0.02)	(0.030)	(745.51)	(1210.32)	(0.004)	(0.01)
Mean: 0.26		Mean: 0.30		Mean: 49794.08		Mean: 0.14	

Maternal and Paternal Characteristics

Maternal Age (in Days)		Father Present		Paternal Age (in Days)		Paternal Education	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-27.62	74.26	0.03	0.01	-23.90	85.33	0.25	0.47
(46.94)	(96.64)	(0.01)	(0.02)	(76.16)	(161.67)	(0.10)	(0.19)
Mean: 9738.61		Mean: 0.95		Mean: 10940.85		Mean: 13.38	

Notes: Robust standard errors adjusted for within-twin-pair correlation are shown in parentheses. All regressions are based on the sample of twin pairs without congenital anomalies who were born between either 1960 and 1967 or 1978 and 1982, observed having a first or second birth in California between 1989 and 2002 (1,530 twin pairs). For those twins whose first and second births are both observed, only the second birth is included in the regressions. The reported mean is the mean of the dependent variable. The pooled OLS regressions include controls for year of birth, birth order, and race.

Appendix Table A6: Fixed Effect Linear Spline Estimates of the Effect of Birthweight in Kilograms
Female Twins with First or Second Birth Observed

	Infant Mortality	Education		
	Death within First Year	Maximum Education	Mean Education	Education at Birth
<1500 g	-0.23	-0.02	0.09	0.49
	(0.03)	(1.47)	(1.40)	(1.54)
1500-2500 g	-0.02	-0.07	-0.03	-0.10
	(0.01)	(0.14)	(0.13)	(0.15)
2500-3000 g	-0.01	0.29	0.31	0.30
	(0.01)	(0.18)	(0.17)	(0.19)
3000g+	0.02	0.30	0.30	0.49
	(0.01)	(0.23)	(0.22)	(0.24)
F-stat of equal slopes	27.69	1.08	0.99	1.86

	Birth and Adult Health Outcomes							
	Child's Birthweight (in Grams)	Gestational Length (in Days)	Hypertension	Diabetes	Anemia	# of Pregnancy Complications	# of Labor Complications	Neonatal Intensive Care Unit Transfer
<1500 g	-517.45	3.11	-0.04	0.12	0.00	0.18	-0.06	-0.05
	(596.60)	(22.09)	(0.18)	(0.14)	(0.11)	(0.44)	(0.94)	(0.19)
1500-2500 g	92.04	-0.22	-0.03	-0.02	0.00	-0.12	-0.05	0.00
	(57.09)	(2.21)	(0.02)	(0.01)	(0.01)	(0.04)	(0.09)	(0.02)
2500-3000 g	50.34	1.59	-0.01	0.01	0.01	0.01	0.13	-0.02
	(72.89)	(2.78)	(0.02)	(0.02)	(0.01)	(0.05)	(0.11)	(0.02)
3000g+	81.74	1.93	0.01	0.00	-0.03	-0.02	-0.13	0.03
	(92.85)	(3.56)	(0.03)	(0.02)	(0.02)	(0.07)	(0.15)	(0.03)
F-stat of equal slopes	0.35	0.13	0.61	0.87	1.09	1.20	0.55	0.37

Appendix Table 6 Con't: Fixed Effect Linear Spline Estimates of the Effect of Birthweight in Kilograms

	Birth Delivery		Residential Location		Maternal and Paternal Characteristics			
	C-Section Delivery	Public Payment for Delivery	Median Household Income (1999) of Zipcode	Poverty Rate (1999) of Zipcode	Maternal Age (in Days)	Father Present	Paternal Age (in Days)	Paternal Education
<1500 g	0.384	-0.98	-316.13	0.033	361.06	0.80	1598.14	-3.77
	(0.469)	(0.44)	(14690.10)	(0.073)	(1289.23)	(0.24)	(3479.60)	(4.31)
1500-2500 g	-0.047	-0.03	2609.66	-0.010	-21.88	0.02	-290.91	0.22
	(0.045)	(0.04)	(1431.88)	(0.007)	(123.37)	(0.02)	(207.16)	(0.25)
2500-3000 g	0.002	-0.04	-957.36	-0.001	91.04	-0.04	5.10	0.40
	(0.057)	(0.05)	(1825.39)	(0.009)	(157.50)	(0.03)	(261.17)	(0.32)
3000g+	-0.081	-0.02	-3159.56	0.009	265.60	-0.01	1005.01	-0.19
	(0.073)	(0.07)	(2349.80)	(0.012)	(200.63)	(0.04)	(332.24)	(0.41)
F-stat of equal slopes	0.45	1.54	1.86	0.74	0.54	4.79	3.83	0.64

Notes: The point estimates represent the estimated slope within the relevant birthweight interval (e.g., 0-2500g). Robust standard errors adjusted for within- twin-pair correlation are reported in parentheses. The reported F-stat tests whether the four segments of the linear spline have equal slopes. All regressions are based on the sample of twins having first or second births in California between 1989 and 2002 (3,396 twin pairs). For those twins whose first and second births are both observed, only the second birth is included in the regressions.

Appendix Table A7: Pooled OLS and Fixed Effect Estimates: Effect of Low Birthweight Female Twins with First or Second Birth Observed

Education

Maximum Education		Mean Education		Education at Birth	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-0.17	-0.14	-0.18	-0.13	-0.17	-0.12
(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.06)
Mean: 13.09		Mean: 12.87		Mean: 12.87	

Birth and Adult Health Outcomes

Child's Birthweight (in Grams)		Gestational Length (in Days)		Hypertension		Diabetes	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-132.78	-42.74	-2.18	-0.53	0.004	0.01	0.005	-1.360E-19
(14.14)	(23.87)	(0.48)	(0.92)	(0.004)	(0.01)	(0.003)	(0.005)
Mean: 3399.89		Mean: 277.19		Mean: 0.02		Mean: 0.01	

Anemia		# of Pregnancy Complications		# of Labor Complications		Neonatal Intensive Care Unit Transfer	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
-0.002	-0.004	0.01	0.02	-0.03	-0.05	0.002	-0.001
(0.002)	(0.004)	(0.01)	(0.02)	(0.02)	(0.04)	(0.004)	(0.008)
Mean: 0.01		Mean: 0.11		Mean: 0.49		Mean: 0.02	

Birth Delivery				Residential Location			
C-Section Delivery		Public Payment for Delivery		Median Household Income (1999) of Zipcode		Poverty Rate (1999) of Zipcode	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
0.003	-0.02	0.02	0.03	-663.43	-416.93	0.001	0.003
(0.010)	(0.02)	(0.01)	(0.02)	(463.03)	(596.18)	(0.002)	(0.003)
Mean: 0.24		Mean: 0.37		Mean: 47418.04		Mean: 0.15	

Maternal and Paternal Characteristics

Maternal Age (in Days)		Father Present		Paternal Age (in Days)		Paternal Education	
Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
16.47	8.52	-0.01	0.003	3.49	37.76	-0.21	-0.07
(32.69)	(51.58)	(0.01)	(0.010)	(50.99)	(85.65)	(0.07)	(0.11)
Mean: 9125.28		Mean: 0.96		Mean: 10283.04		Mean: 12.89	

Notes: The number reported in the first row of each set of estimates is the low birthweight (birthweight < 2500 grams) coefficient. Robust standard errors adjusted for within-twin-pair correlation are shown in parentheses. The reported mean is the mean of the dependent variable. All regressions are based on the sample of twins and observed having a first or second birth in California between 1989 and 2002 (3,396 twin pairs). For those twins whose first and second births are both observed, only the second birth is included in the regressions. The pooled OLS regressions include controls for year of birth, birth order, and race.