ONLINE APPENDIX

For the paper:

Twin Birth and Maternal Condition
Sonia Bhalotra and Damian Clarke

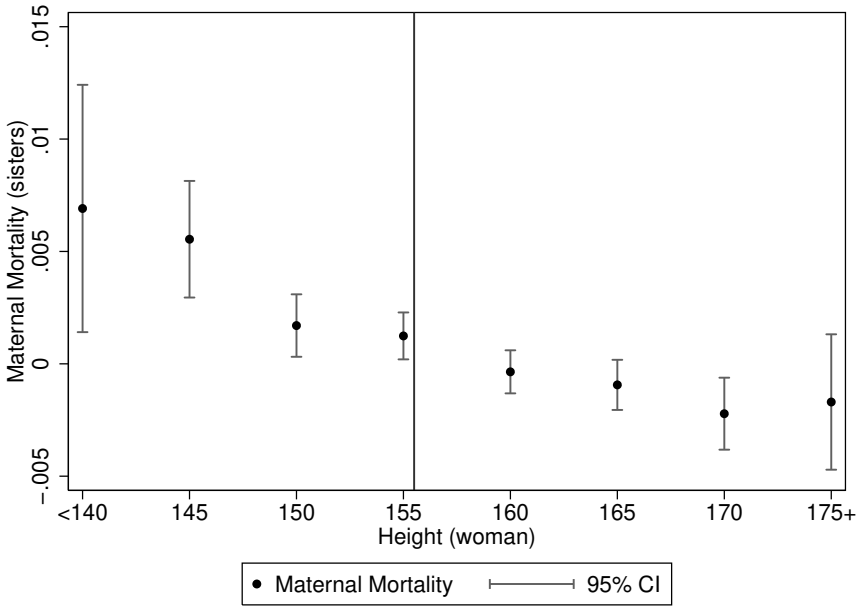
Contents

List of Appendix Figures

List of Appendix Tables

A Appendix Figures and Tables

Figure A1: Height and Selective Survival



Note to Figure A1: Data consists of all women in the DHS from countries where the maternal mortality module was applied in surveys. Heights are based on measures for all mothers at the time of the survey, and rates of maternal mortality are calculated based on the survival status of each sister of surveyed women. Each point represents the average rate of maternal death for sisters of women with heights in each group, concentrating out country and age FEs. The vertical line represents the mean height of 155.5 cm.

Table A1: Effects of Maternal Health on Twin Births (Unstandardised Variables)

Health Behaviours / Access			Health Conditions		
Variable	Estimate	[95% CI]	Variable	Estimate	[95% CI]
Panel A: United States					
Smoked Before Pregnancy	-0.336***	[-0.361,-0.311]	Height	0.084***	[0.082,0.086]
‡Smoked Trimester 1	-0.686***	[-0.715,-0.657]	Underweight	-0.666***	[-0.699,-0.633]
‡Smoked Trimester 2	-0.877***	[-0.908,-0.846]	Obese	0.106***	[0.084,0.128]
‡Smoked Trimester 3	-0.918***	[-0.949,-0.887]	Diabetes	-3.352***	[-3.464,-3.240]
Education	0.448***	[0.442,0.454]	Hypertension	-1.917***	[-2.007,-1.827]
Panel B: Sweden					
‡Smoked (12 weeks)	-0.704***	[-0.798,-0.610]	Height	0.099***	[0.095,0.103]
‡Smoked (30-32 weeks)	-1.030***	[-1.132,-0.928]	Underweight	-0.716***	[-0.887,-0.545]
			Obese	-0.411***	[-0.497,-0.325]
			Asthma	-0.085	[-0.189,0.019]
			Diabetes	-3.737***	[-4.113,-3.361]
			Kidney Disease	-1.359***	[-1.731,-0.987]
			Hypertension	-1.872***	[-2.286,-1.458]
Panel C: United Kingdom (Avon)					
‡Healthy Foods	1.333***	[0.635,2.031]	Height	0.059***	[0.016,0.102]
‡Fresh Fruit	0.039	[-0.594,0.672]	Underweight	-0.794	[-2.164,0.576]
‡Alcohol (Infrequently)	-0.259	[-0.976,0.458]	Obese	-0.218	[-1.512,1.076]
‡Alcohol (Frequently)	-1.567***	[-2.759,-0.375]	Diabetes	-0.951	[-5.573,3.671]
‡Passive Smoke	0.096	[-0.494,0.686]	Hypertension	-2.536***	[-3.975,-1.097]
‡Smoked during Pregnancy	-0.433	[-1.199,0.333]			
Education	0.225*	[0.000,0.450]			
Panel D: Chile					
‡Smoked during Pregnancy	-1.084***	[-1.894,-0.274]	Underweight	-0.753*	[-1.641,0.135]
‡Drugs (Infrequently)	0.021	[-3.439,3.481]	Obese	-1.734***	[-3.000,-0.468]
‡Drugs (Frequently)	-3.053***	[-3.717,-2.389]			
‡Alcohol (Infrequently)	-0.274	[-1.383,0.835]			
‡Alcohol (Frequently)	-2.783***	[-3.436,-2.130]			
Education	0.140***	[0.052,0.228]			
Panel E: Developing Countries					
Doctor Availability	0.319***	[0.203,0.435]	Height	0.039***	[0.035,0.043]
Nurse Availability	0.235***	[0.113,0.357]	Underweight	-0.282***	[-0.358,-0.206]
Prenatal Care Availability	0.514***	[0.377,0.651]	Obese	0.187***	[0.091,0.283]
Education	0.031***	[0.025,0.037]			

Regressions replicate Table 2, however all variables are unstandardised. Refer to additional notes to Table 2.

Table A2: Effects of Maternal Health on Twin Births (Conditional Results)

Health Behaviours / Access			Health Conditions		
Variable	Estimate	[95% CI]	Variable	Estimate	[95% CI]
Panel A: United States					
Smoked Before Pregnancy	0.186***	[0.170,0.202]	Height	0.545***	[0.535,0.555]
‡Smoked Trimester 1	-0.056***	[-0.080,-0.032]	Underweight	-0.185***	[-0.193,-0.177]
‡Smoked Trimester 2	-0.132***	[-0.163,-0.101]	Obese	0.108***	[0.098,0.118]
‡Smoked Trimester 3	-0.183***	[-0.210,-0.156]	Diabetes	-0.261***	[-0.271,-0.251]
Education	0.678***	[0.668,0.688]	Hypertension	-0.213***	[-0.223,-0.203]
Panel B: Sweden					
‡Smoked (12 weeks)	0.049*	[-0.008,0.106]	Height	0.612***	[0.587,0.637]
‡Smoked (30-32 weeks)	-0.307***	[-0.352,-0.262]	Underweight	-0.148***	[-0.181,-0.115]
			Obese	-0.082***	[-0.106,-0.058]
			Asthma	-0.005	[-0.023,0.013]
			Diabetes	-0.243***	[-0.268,-0.218]
			Kidney Disease	-0.066***	[-0.088,-0.044]
			Hypertension	-0.082***	[-0.104,-0.060]
Panel C: United Kingdom (Avon)					
‡Healthy Foods	0.537***	[0.253,0.821]	Height	0.408***	[0.122,0.694]
‡Fresh Fruit	-0.116	[-0.422,0.190]	Underweight	-0.191	[-0.469,0.087]
‡Alcohol (Infrequently)	0.069	[-0.251,0.389]	Obese	-0.008	[-0.286,0.270]
‡Alcohol (Frequently)	-0.398**	[-0.716,-0.080]	Diabetes	-0.065	[-0.337,0.207]
‡Passive Smoke	0.193	[-0.121,0.507]	Hypertension	-0.479***	[-0.751,-0.207]
‡Smoked during Pregnancy	-0.165	[-0.475,0.145]			
Education	0.384*	[-0.039,0.807]			
Panel D: Chile					
‡Smoked during Pregnancy	-0.256**	[-0.501,-0.011]	Underweight	-0.172	[-0.388,0.044]
‡Drugs (Infrequently)	0.018	[-0.243,0.279]	Obese	-0.256***	[-0.444,-0.068]
‡Drugs (Frequently)	-0.096***	[-0.141,-0.051]			
‡Alcohol (Infrequently)	-0.037	[-0.331,0.257]			
‡Alcohol (Frequently)	-0.115***	[-0.160,-0.070]			
Education	0.486***	[0.155,0.817]			
Panel E: Developing Countries					
Doctor Availability	0.043*	[-0.002,0.088]	Height	0.265***	[0.234,0.296]
Nurse Availability	0.045**	[0.006,0.084]	Underweight	-0.084***	[-0.109,-0.059]
Prenatal Care Availability	0.035*	[-0.004,0.074]	Obese	0.044***	[0.013,0.075]
Education	0.073***	[0.040,0.106]			

Regressions replicate Table 2, however all variables in each panel are included in one regression. Refer to additional notes to Table 2.

Table A3: Effect of Maternal Health on Twinning (Education Quadratic and Wealth Quintiles)

Health Behaviours / Access			Health Conditions		
Variable	Estimate	[95% CI]	Variable	Estimate	[95% CI]
Panel E: Developing Countries [N =2,050,795, % Twin = 2.07]					
Doctor Availability	0.049**	[0.004,0.094]	Height	0.265***	[0.234,0.296]
Nurse Availability	0.045**	[0.006,0.084]	Underweight	-0.085***	[-0.110,-0.060]
Prenatal Care Availability	0.016	[-0.023,0.055]	Obese	0.044***	[0.013,0.075]
Education	0.049	[-0.035,0.133]			
Education Squared	0.026	[-0.056,0.108]			

Results are reported following the specifications in table 2, for DHS only (where ART usage is observed for all births). The sample period and specification is identical to those in Table 2, however now additionally control for a quadratic in maternal education, and family wealth quintile fixed effects. A test of joint insignificance of the coefficients on maternal education is rejected ($F = 8.83$, $p < 0.01$.)

Table A4: Effects of Maternal Health on Twin Births Correcting for Large Sample Size

Health Behaviours / Access			Health Stocks and Conditions		
Variable	Estimate	[95% CI]	Variable	Estimate	[95% CI]
Panel A: United States [N =13,646,236, % Twin = 2.84]					
Smoked Before Pregnancy	-0.108#	[-0.116,-0.100]	Height	0.612#	[0.604,0.620]
‡Smoked Trimester 1	-0.195#	[-0.203,-0.187]	Underweight	-0.156#	[-0.164,-0.148]
‡Smoked Trimester 2	-0.232#	[-0.240,-0.224]	Obese	0.042#	[0.032,0.052]
‡Smoked Trimester 3	-0.238#	[-0.246,-0.230]	Diabetes	-0.286#	[-0.296,-0.276]
Education	0.800#	[0.790,0.810]	Hypertension	-0.223#	[-0.233,-0.213]
Panel B: Sweden [N =1,240,621, % Twin = 2.55]					
‡Smoked (12 weeks)	-0.266#	[-0.301,-0.231]	Height	0.617#	[0.592,0.642]
‡Smoked (30-32 weeks)	-0.285#	[-0.312,-0.258]	Underweight	-0.140#	[-0.173,-0.107]
			Obese	-0.113#	[-0.137,-0.089]
			Asthma	-0.015	[-0.033,0.003]
			Diabetes	-0.253#	[-0.278,-0.228]
			Kidney Disease	-0.079#	[-0.101,-0.057]
			Hypertension	-0.099#	[-0.121,-0.077]
Panel C: United Kingdom (Avon) [N =10,463, % Twin = 2.37]					
‡Healthy Foods	0.538#	[0.256,0.820]	Height	0.399	[0.115,0.683]
‡Fresh Fruit	0.019	[-0.281,0.319]	Underweight	-0.161	[-0.439,0.117]
‡Alcohol (Infrequently)	-0.099	[-0.373,0.175]	Obese	-0.046	[-0.322,0.230]
‡Alcohol (Frequently)	-0.358	[-0.630,-0.086]	Diabetes	-0.056	[-0.328,0.216]
‡Passive Smoke	0.047	[-0.243,0.337]	Hypertension	-0.480#	[-0.752,-0.208]
‡Smoked during Pregnancy	-0.162	[-0.448,0.124]			
Education	0.416	[-0.002,0.834]			
Panel D: Chile [N =14,050, % Twin = 2.55]					
‡Smoked during Pregnancy	-0.327	[-0.572,-0.082]	Underweight	-0.183	[-0.399,0.033]
‡Drugs (Infrequently)	0.002	[-0.253,0.257]	Obese	-0.258	[-0.446,-0.070]
‡Drugs (Frequently)	-0.161#	[-0.196,-0.126]			
‡Alcohol (Infrequently)	-0.072	[-0.362,0.218]			
‡Alcohol (Frequently)	-0.172#	[-0.213,-0.131]			
Education	0.529#	[0.200,0.858]			
Panel E: Developing Countries [N =2,050,795, % Twin = 2.07]					
Doctor Availability	0.092#	[0.059,0.125]	Height	0.276#	[0.245,0.307]
Nurse Availability	0.060	[0.029,0.091]	Underweight	-0.090#	[-0.115,-0.065]
Prenatal Care Availability	0.103#	[0.076,0.130]	Obese	0.059	[0.028,0.090]
Education	0.141#	[0.110,0.172]			

‡: Conditions which are measured during pregnancy, and so may be behavioural responses to twins.

This table replicatesTable 2, however reports significance based on the criterion laid out by Deaton (1997) and Leamer (1978). This corrects for the increased likelihood of rejecting the null hypothesis as the sample size grows and the null is not exactly true, by adjusting the significance criterion in line with sample size. As discussed by Deaton (1997), the Leamer (1978) criterion can be approximated by comparing t-statistics with $\sqrt{\log(N)}$. # Refers to variables which are significant based on this criterion.

Table A5: Smoking and birthweight

Dependent Variable: Birthweight	All Births	Non-Twin Births	Twin Births
Smokes 3 Months Prior to Pregnancy	-98.45***	-100.41***	-67.63***
	(0.376)	(0.380)	(1.825)
Smokes Trimester 1	-141.56***	-144.88***	-100.97***
	(0.421)	(0.426)	(2.080)
Smokes Trimester 2	-164.22***	-168.36***	-116.49***
	(0.445)	(0.450)	(2.227)
Smokes Trimester 3	-169.19***	-173.51***	-119.96***
	(0.454)	(0.459)	(2.311)
Average Birthweight	3286.8	3313.3	2371.9
Observations	13,595,271	13,212,574	382,697

Each cell represents a multivariate regression of smoking behaviour on birthweight using the sample of US birth data used in Table 2. All specifications follow those reported in Table 2. Smoking in each period is a binary measure, and birthweight is measured in grams.

Table A6: Twinning and Maternal Height with Height Trimmed of Outliers (USA)

	Full Sample		Trimmed Sample	
Twin×100	(1)	(2)	(3)	(4)
Height	0.084***		0.086***	
	(0.001)		(0.001)	
Height (Quintile 1)		-1.771***		-1.380***
		(0.016)		(0.022)
Height (Quintile 2)		-1.269***		-1.061***
		(0.018)		(0.023)
Height (Quintile 3)		-0.944***		-0.737***
		(0.016)		(0.022)
Height (Quintile 4)		-0.529***		-0.323***
		(0.017)		(0.023)
Observations	13,616,587	13,616,587	11,462,162	11,462,162

Regressions of twinning on maternal height are presented using US birth certificate data with and without trimming the sample to remove outliers. Columns 1 and 2 use the full sample, while columns 3 and 4 trim the sample at the 5th and 95th percentile. In the trimmed sample, heights vary from 152.4 to 175.26 cm. Columns 1 and 3 present height in centimetres, while columns 2 and 4 use categorical variables for height quintiles, with quintile 5 (tallest mothers) as the omitted base category.

Table A7: Effects of Maternal Health on Twin Births (ART only)

Health Behaviours / Access			Health Conditions		
Variable	Estimate	[95% CI]	Variable	Estimate	[95% CI]
United States: ART Only [N =217,703, % Twin = 35.39]					
Smoked Before Pregnancy	-1.113***	[-1.266,-0.960]	Height	1.880***	[1.719,2.041]
‡Smoked Trimester 1	-1.175***	[-1.324,-1.026]	Underweight	0.105	[-0.050,0.260]
‡Smoked Trimester 2	-1.123***	[-1.274,-0.972]	Obese	-2.204***	[-2.365,-2.043]
‡Smoked Trimester 3	-1.042***	[-1.193,-0.891]	Diabetes	-1.318***	[-1.489,-1.147]
Education	2.426***	[2.254,2.598]	Hypertension	-1.592***	[-1.766,-1.418]

Results are reported following the specifications in table 2, for USA only (where ART usage is observed for all births). The sample period and specification is identical to those in Table 2, however now only Artificial Reproductive Technology users are included in the regression.

Table A8: Probability of Giving Birth to Twins (Developing Countries by Income and Time Period)

Twin×100	All	Income		Time		Prenatal
		Low inc	Middle inc	1990-2013	1972-1989	
Mother's Age	0.540***	0.550***	0.517***	0.601***	0.314***	0.541***
	(0.027)	(0.033)	(0.047)	(0.031)	(0.058)	(0.027)
Mother's Age Squared	-0.007***	-0.007***	-0.007***	-0.008***	-0.003**	-0.007***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Age at First Birth	-0.050***	-0.082***	-0.001	-0.052***	-0.040***	-0.050***
	(0.008)	(0.010)	(0.013)	(0.010)	(0.015)	(0.008)
Mother's Education (years)	0.021***	0.019**	0.013	0.021***	0.019	0.018**
	(0.007)	(0.009)	(0.010)	(0.008)	(0.012)	(0.007)
Mother's Height (cm)	0.059***	0.058***	0.058***	0.063***	0.044***	0.058***
	(0.004)	(0.005)	(0.007)	(0.005)	(0.007)	(0.004)
Mother's BMI	0.047***	0.059***	0.038***	0.045***	0.050***	0.045***
	(0.006)	(0.009)	(0.009)	(0.007)	(0.010)	(0.006)
Prenatal Care (Doctor)						0.333**
						(0.142)
Prenatal Care (Nurse)						0.312**
						(0.142)
Prenatal Care (None)						0.008
						(0.181)
Observations	2,046,907	1,287,585	759,322	1,525,966	520,941	2,043,217
R-Squared	0.006	0.006	0.005	0.006	0.005	0.006

Notes: This table presents results for the developing country sample splitting by pre- and post-1990. Main specifications for the developing country sample are pooled for all years. All specifications include a full set of year of birth and country dummies, and are estimated as linear probability models. Twin is multiplied by 100 for presentation. Height is measured in cm and BMI is weight in kg divided by height in metres squared. Prenatal care variables refer to average levels of coverage in DHS clusters. These prenatal measures are only recorded for births in 5 years preceding each survey wave, and as such, a small number of (small) clusters do not have records available. Standard errors clustered by mothers are presented in parentheses. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A9: Prior Children's Birth Outcomes of Twin and Non-Twin Mothers (Developing Countries)

	Reported Birth Size (1)	Small Birth (2)
Panel A: 2+ Sample		
Treated	0.067*	-0.052***
	(0.036)	(0.016)
Mean Values	3.07	0.20
Observations	48,028	48,028
Panel B: 3+ Sample		
Treated	0.115***	-0.049***
	(0.037)	(0.015)
Mean Values	3.14	0.17
Observations	40,156	40,156
Panel C: 4+ Sample		
Treated	0.132***	-0.036**
	(0.046)	(0.017)
Mean Values	3.17	0.17
Observations	29,621	29,621

Refer to notes in Table 4. Identical regressions are estimated using observations in the developing country sample where birth sizes are recorded. Treated takes the values of one if the second, third or fourth birth (respecitvely in panels A, B and C) is a twin, and zero if a singleton. The estimation samples consists of siblings born *before* the indicator birth. Reported birth size is a categorical variable coded from 1 (very small) to 5 (very large) as reported by mothers, and small birth refers to births reported to be very small or smaller than average. Birth measures in the DHS are collected for any children born in the five years preceding the survey date. Standard errors are clustered by mother. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A10: Effects of Maternal Health on Twin Births (Same Sex Twins for DHS Only)

Health Behaviours / Access			Health Conditions		
Variable	Estimate	[95% CI]	Variable	Estimate	[95% CI]
Developing Countries					
Doctor Availability	0.084***	[0.041,0.127]	Height	0.259***	[0.220,0.298]
Nurse Availability	0.061***	[0.022,0.100]	Underweight	-0.066***	[-0.097,-0.035]
Prenatal Care Availability	0.092***	[0.057,0.127]	Obese	0.071***	[0.032,0.110]
Education	0.119***	[0.080,0.158]			

Regressions replicate panel E of Table 2, however now the outcome variable is equal to 100 only for same sex twins, and 0 for all singleton children. Refer to additional notes to Table 2. This specification is only estimated using DHS data, as in this data set we are able to match twins with their siblings.

Table A11: Fetal Deaths, Twinning, and Health Behaviours (Conditional)

	(1) Smokes	(2) Drinks	(3) No College	(4) Anemic	(5) N Cigs	(6) N Drinks	(7) Years Educ
Panel A: Uninteracted Twin – Non-Twin Difference							
Twin	9.704***	10.083***	10.105***	11.020***	9.697***	10.075***	10.105***
	(0.118)	(0.119)	(0.108)	(0.115)	(0.118)	(0.119)	(0.108)
Constant	4.775***	4.936***	4.635***	5.282***	4.776***	4.936***	4.635***
	(0.052)	(0.052)	(0.047)	(0.050)	(0.052)	(0.052)	(0.047)
Panel B: Health, Twin and Twin×Health Interaction							
Twin	9.648***	10.080***	8.621***	10.972***	9.679***	10.063***	19.150***
	(0.123)	(0.119)	(0.145)	(0.117)	(0.121)	(0.119)	(0.552)
Health (Dis)amenity	0.867***	4.219***	1.229***	0.337**	0.071***	0.525***	-0.122***
	(0.067)	(0.260)	(0.043)	(0.131)	(0.005)	(0.038)	(0.008)
Twin × Health	0.856**	3.322*	3.548***	1.307**	0.040	0.755***	-0.667***
	(0.416)	(1.753)	(0.218)	(0.640)	(0.032)	(0.206)	(0.040)
Constant	4.717***	4.909***	4.152***	5.276***	4.727***	4.926***	6.311***
	(0.052)	(0.052)	(0.050)	(0.050)	(0.052)	(0.052)	(0.115)
Observations	13,660,400	13,809,830	15,909,836	16,158,564	13,679,142	13,828,573	15,909,836

Refer to notes in Table 5 for full details. Identical regression results are presented here, however now each regression also controls for mother's age fixed effects, total number of mother's birth, and the year of birth. $^{*}p<0.10$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A12: Twin Births and Maternal Health: Differences by Child Gender

	Developing Country Data		US Vital Stats	
	(1)		(2)	
Twin×100	Characteristic	× Male	Characteristic	× Male
Male	-0.100***		-0.309***	
	(0.022)		(0.008)	
Height	0.274***	-0.017	0.525***	0.041***
	(0.020)	(0.022)	(0.006)	(0.009)
Underweight	-0.083***	-0.003	-0.182***	-0.007
	(0.017)	(0.020)	(0.006)	(0.008)
Obese	0.029	0.029	0.114***	-0.011
	(0.020)	(0.024)	(0.007)	(0.009)
Education	0.053**	0.037	0.685***	-0.003
	(0.021)	(0.024)	(0.007)	(0.009)
Doctor Availability	0.031	0.023		
	(0.027)	(0.029)		
Nurse Availability	0.013	0.062**		
	(0.025)	(0.030)		
Prenatal Care Availability	0.049*	-0.027		
	(0.025)	(0.029)		
Smoked Before Pregnancy			0.182***	0.006
			(0.012)	(0.016)
Smoked Trimester 1			-0.057***	0.002
			(0.017)	(0.023)
Smoked Trimester 2			-0.128***	-0.006
			(0.023)	(0.033)
Smoked Trimester 3			-0.194***	0.021
			(0.020)	(0.028)
Diabetes			-0.255***	-0.012
			(0.007)	(0.010)
Hypertension			-0.220***	0.011
			(0.008)	(0.011)
Observations	2,050,795			13,616,587

Notes: In each case (developing country data and US data) a single regression is estimated, with coefficients and standard errors reported in two columns. The first column displays the impact of each socioeconomic or health measure on twinning, while the second column displays the interaction between each variable and whether the child is a male. Developing country data refers to the pooled DHS sample.

Table A13: Full Survey Countries and Years (DHS)

		Survey Year						
Country	Income	1	2	3	4	5	6	7
Albania	Middle	2008						
Armenia	Low	2000	2005	2010				
Azerbaijan	Middle	2006						
Bangladesh	Low	1994	1997	2000	2004	2007	2011	
Benin	Low	1996	2001	2006				
Bolivia	Middle	1994	1998	2003	2008			
Brazil	Middle	1991	1996					
Burkina Faso	Low	1993	1999	2003	2010			
Burundi	Low	2010						
Cambodia	Low	2000	2005	2010				
Cameroon	Middle	1991	1998	2004	2011			
Central African Republic	Low	1994						
Chad	Low	1997	2004					
Colombia	Middle	1990	1995	2000	2005	2010		
Comoros	Low	1996						
Congo Brazzaville	Middle	2005	2011					
Congo Democratic Republic	Low	2007						
Cote d Ivoire	Low	1994	1998	2005	2012			
Dominican Republic	Middle	1991	1996	1999	2002	2007		
Egypt	Low	1992	1995	2000	2005	2008		
Ethiopia	Low	2000	2005	2011				
Gabon	Middle	2000	2012					
Ghana	Low	1993	1998	2003	2008			
Guatemala	Middle	1995						
Guinea	Low	1999	2005					
Guyana	Middle	2005	2009					
Haiti	Low	1994	2000	2006	2012			
Honduras	Middle	2005	2011					
India	Low	1993	1999	2006				
Indonesia	Low	1991	1994	1997	2003	2007	2012	
Jordan	Middle	1990	1997	2002	2007			
Kazakhstan	Middle	1995	1999					
Kenya	Low	1993	1998	2003	2008			
Kyrgyz Republic	Low	1997						
Lesotho	Low	2004	2009					
Liberia	Low	2007						
Madagascar	Low	1992	1997	2004	2008			
Malawi	Low	1992	2000	2004	2010			
Maldives	Middle	2009						
Mali	Low	1996	2001	2006				
Moldova	Middle	2005						
Morocco	Middle	1992	2003					
Mozambique	Low	1997	2003	2011				
Namibia	Middle	1992	2000	2006				
Nepal	Low	1996	2001	2006	2011			
Nicaragua	Low	1998	2001					

Niger	Low	1992	1998	2006				
Nigeria	Low	1990	1999	2003	2008			
Pakistan	Low	1991	2006					
Paraguay	Middle	1990						
Peru	Middle	1992	1996	2000				
Philippines	Middle	1993	1998	2003	2008			
Rwanda	Low	1992	2000	2005	2010			
Sao Tome and Principe	Middle	2008						
Senegal	Middle	1993	1997	2005	2010			
Sierra Leone	Low	2008						
South Africa	Middle	1998						
Swaziland	Middle	2006						
Tanzania	Low	1992	1996	1999	2004	2007	2010	2012
Togo	Low	1998						
Turkey	Middle	1993	1998	2003				
Uganda	Low	1995	2000	2006	2011			
Ukraine	Middle	2007						
Uzbekistan	Middle	1996						
Vietnam	Low	1997	2002					
Yemen	Low	1991						
Zambia	Low	1992	1996	2002	2007			
Zimbabwe	Low	1994	1999	2005	2010			

Notes: Country income status is based upon World Bank classifications described at http://data.worldbank.org/about/country-classifications and available for download at http://siteresources.worldbank.org/DATASTATISTICS/Resources/OGHIST.xls (consulted 1 April, 2014). Income status varies by country and time. Where a country's status changed between DHS waves only the most recent status is listed above. Middle refers to both lower-middle and upper-middle income countries, while low refers just to those considered to be low-income economies.

B Measurement Considerations and Data Appendix

B.1 Measurement Considerations

B.1.1 Maternal health data are relatively scarce

High quality registry data used to estimate the quantity–quality fertility model in the literature using twin births to instrument fertility has limited or no measures of maternal health. This includes census data from Israel used by Angrist et al. (2010) (see questionnaire here: http://www.cbs.gov.il/mifkad/q_census1995_e.pdf) and administrative Norwegian data used in Black et al. (2005). An exception is the sick leave register which captures spells off of work, but no measures of health stocks (Barth, 2012). Even in rich survey data collected expressly for the purposes of research into twins (Rosenzweig and Zhang, 2009), measures of health of twin *mothers* are scarce.

B.1.2 Unobserved Miscarriages in Vital Statistics Data

We examine fetal death data in the United States Vital Statistics to test mechanisms relating to twin-selection. These data record all fetal deaths occurring *after* 20 weeks of pregnancy, which is about 25,000 per year. Estimates from the National Center of Health Statistics suggest that there are about 1 million fetal losses per year, and 90% of these occur before the 20^{th} week of gestation (MacDorman and Kirmeyer, 2009). Only certain states report fetal deaths occurring earlier, so to ensure a consistent measurement across the country, we focus only on fetal losses occurring at after 20 weeks. Fetal loss earlier in pregnancy often goes unnoticed, resulting in measurement error. While there is some evidence of under-reporting of fetal deaths around 20 weeks in some states (Martin and Hoyert, 2002), the majority of fetal deaths at this point of gestation are recorded in the NVSS data.

Using the Vital Statistics threshold of 20 weeks should not create any selection problems for our analysis. For it to bias our results, mothers who were healthier would need to be more likely to miscarry twins in the first 20 weeks of pregnancies than less healthy mothers. This is the opposite of what is observed from week 20 onwards. Indeed, we can partially test this by including fetal deaths from the number of states which report deaths *prior* to 20 weeks, and in each case the same health gradient remains, while in 5 of the 7 cases reported in Table 5 the twin–health gradient of fetal deaths becomes even steeper, suggesting that if anything having the universe of fetal deaths would strengthen our results.

B.1.3 Selective Recall Bias

It is well documented that recall bias in retrospective survey data exists in a range of circumstances. Evidence from Beckett et al. (2001) provides discussion and analysis of survey data in a developing country context (Malaysia). Beckett et al. (2001) state that while events like pregnancy are rarely forgotten, details of the timing of these events may be mismeasured, and find evidence of this in Malaysian Family Life Survey (MFLS) data. In particular, concerns relating to heaping of birth dates and other life events exist. In the case of DHS, analysis on even the earliest round of surveys finds that heaping is not a major problem when considering child age, though some minor heaping is observed on ages ending in decades. For example, as stated in Arnold (1990):

> "In summary, while digit preference exists to some extent in most DHS surveys, it is not a major problem in the reporting of children's ages. Moreover, the impact of age heaping on fertility rates is quite small. Efforts have been made in all DHS surveys to obtain the exact calendar year and month of birth of children" (Arnold, 1990, p. 95).

In general we would be most concerned about selective recall bias if it affected the measurement of our dependent variables of interest (twin births), and the independent variable of interest (maternal health). In the case of administrative

records measuring birth outcomes and maternal conditions, these are captured at the time of birth, and retrospective measures are main life events (eg prior births, chronic health conditions) and so are unlikely to be affected by recall bias. In the case of household survey data, there is little support in the literature on recall bias to suggest that the number of births are misreported (Mathiowetz, 1999). The DHS data collection procedure puts significant emphasis on managing and examining data quality, and measures of fertility and missing responses are better than measures in other surveys such as the World Values Surveys (Arnold, 1990). At the stage of data elicitation, enumerators are given Age/Birth Date consistency check cards to provide an initial check of measurement. Additionally, there is a cross-check question about twin births available. When asking about education, our principal measure in regressions is based on years of education. Given concerns that years of education may be miscalculated, the DHS procedure asks for levels and courses of education, which is then converted into years of education based on the particular educational system in each country (ICF, 2017). In the case of anthropometric measures, these are physically captured by enumerators, and are so not subject to recall bias. In general, we do not expect recall bias to lead to a bias in the relationship between maternal characteristics and twinning.

B.2 Data Appendix

We analyse a number of datasets, which are:

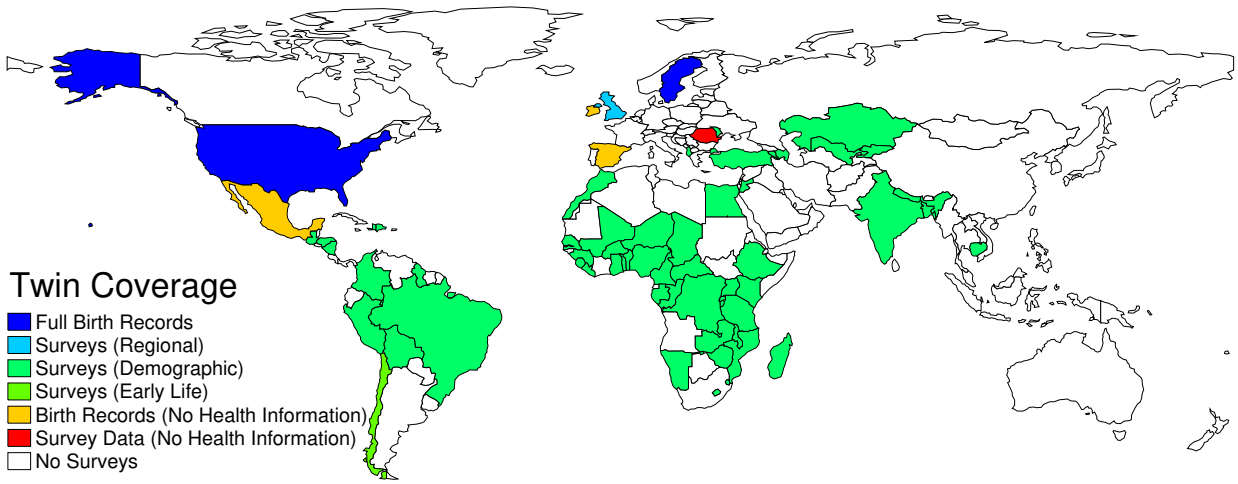
- United States National Vital Statistics Birth Data
- The Swedish Medical Birth Register
- Spanish Vital Statistics (available from Spain's INE)
- Longitudinal Early Life Survey, Chile (ELPI)
- The Avon Longitudinal Study of Parents and Children (ALSPAC, UK)
- The Demographic and Health Surveys (DHS)

The first three datasets are administrative records of all births, and and the remaining three data sets are large representative surveys. We always use the sample of mothers aged 18-49 and drop births which triplets and higher order multiple births. In the United States Vital Statistics data, from 2009 onwards we observe Artificial Reproductive Technology (ART) use status of birth, and remove the 1.6% of births which were conceived using ART from the estimation sample. The ELPI survey from Chile focuses on early childhood and records mother's behaviours before, during and after pregnancy, along with child birth outcomes. We use all index children from the first wave of this survey who meet the inclusion criteria discussed above. The ALSPAC survey follows prospectively-enrolled mothers and their children who were born in the early 1990's in the county of Avon, UK. We use all mothers from the original survey cohort. A small number of mothers who were later enrolled as a refreshment sample are not included, as a range of required pre-pregnancy measures are not available for these women. Finally, the Demographic and Health Surveys (DHS) are a set of nationally representative surveys which have been administered in low- and middle-income countries between 1985 and the present. A full list of the DHS countries and years of surveys which make up this sample is provided in Table A13. Women aged 15–49 in surveyed households respond to an in-depth series of questions reporting their full fertility history (listing all surviving and non-surviving children), their actual and desired contraceptive use and number of births, education level, marital status, and their height and body mass index are not self-reported but measured by surveyors using state of the art instruments. For all of a mother's births, a shorter series of responses are recorded, including their birth date, birth type (singleton, twin, triplet, etc.) and survival status. We pool all publicly available DHS data. The geographic coverage of datasets with measure of maternal health available is displayed in Figure A2. Full summary statistics corresponding to tests displayed in Table 2 is provided in Table A14.

When examining mechanisms of twin selection, we use data the United States National Vital Statistics Foetal Death Records, pooling all fetal deaths and births occurring in 1999-2002. This selection of years is made to ensure consistency

in measures of maternal characteristics in the birth and fetal death records. In 2003 there was an update to birth certificates and fetal death records.

Figure A2: Coverage of Data Containing Indicators of Twin Births and Maternal Health by Country and Data Type



Different colours represent different types of data (surveys, national vital statistics, or no data collected). Each data type is described in the figure legend.

)

Table A14: Summary Statistics (Twin Regressions)

	N	Mean	Std.Dev.	Min	Max
Panel A: United States					
Mother's height (cm)	13,646,236	163.00	7.26	86.36	198.12
Mother's education (years)	13,646,236	4.19	1.79	1.00	9.00
Mother Smoked Before Pregnancy	13,646,236	0.12	0.32	0.00	1.00
Mother Smoked in 1st Trimester	13,646,236	0.09	0.28	0.00	1.00
Mother Smoked in 2nd Trimester	13,646,236	0.08	0.27	0.00	1.00
Mother Smoked in 3rd Trimester	13,646,236	0.07	0.26	0.00	1.00
Mother had pre-pregnancy diabetes	13,646,236	0.01	0.09	0.00	1.00
Mother had pre-pregnancy hypertension	13,646,236	0.01	0.12	0.00	1.00
Mother is underweight (pre-pregnancy)	13,646,236	0.06	0.23	0.00	1.00
Mother is obese (pre-pregnancy)	13,646,236	0.20	0.40	0.00	1.00
Percent Twin Births	13,646,236	2.84	16.60	0.00	100.00
Mother's Age in years	13,646,236	28.09	5.78	18.00	49.00
Panel B: Sweden					
Pre-pregnancy asthma	1,240,621	0.07	0.25	0	1
Pre-pregnancy diabetes	1,240,621	0.01	0.07	0	1
Pre-pregnancy kidney disease	1,240,621	0.01	0.07	0	1
Pre-pregnancy hypertension	1,240,621	0.01	0.06	0	1
Smoked at 12 weeks gestation	1,240,621	0.09	0.29	0	1
Smoked at 30-32 weeks gestation	1,240,621	0.07	0.27	0	1
Height	1,240,621	166.38	6.35	100	200
Underweight (BMI $<$ 18.5) Prior to Pregnancy	1,240,621	0.02	0.15	0	1
Obese (BMI $\geq$ 30) Prior to Pregnancy	1,240,621	0.11	0.31	0	1
Percent Twin Births	1,240,621	2.55	15.77	0	100
Mother's Age in Years	1,240,621	29.96	5.11	18	49
Panel C: United Kingdom (Avon)					

Underweight (BMI $<$ 18.5) Prior to Pregnancy	10,463	0.04	0.21	0.00	1.00
Obese (BMI $\geq$ 30) Prior to Pregnancy	10,463	0.05	0.22	0.00	1.00
Mother's height (cm)	10,463	164.10	6.68	129.54	200.66
Pre-pregnancy diabetes	10,463	0.00	0.06	0.00	1.00
Pre-pregnancy hypertension	10,463	0.04	0.20	0.00	1.00
Pre-pregnancy infections (total)	10,463	3.26	1.24	0.00	7.00
Frequent Healthy Food in Pregnancy	10,463	0.21	0.40	0.00	1.00
Frequent Fresh Fruit in Pregnancy	10,463	0.30	0.46	0.00	1.00
Infrequent Alcohol Consumption in Pregnancy	10,463	0.19	0.39	0.00	1.00
Frequent Alcohol Consumption in Pregnancy	10,463	0.06	0.24	0.00	1.00
Exposed to Passive Smoke in Pregnancy	10,463	0.43	0.50	0.00	1.00
Smoked during Pregnancy	10,463	0.17	0.38	0.00	1.00
Mother's education (years)	10,463	12.29	1.83	10.00	16.00
Percent Twin Births	10,463	2.37	15.21	0.00	100.00
Mother's Age in years	10,463	28.67	4.61	18.00	45.00
Panel D: Chile					
Mother Smoked During Pregnancy	14,050	0.10	0.30	0.00	1.00
Drugs During Pregnancy (Sporadically)	14,050	0.01	0.07	0.00	1.00
Drugs During Pregnancy (Regularly)	14,050	0.00	0.05	0.00	1.00
Alcohol During Pregnancy (Sporadically)	14,050	0.07	0.25	0.00	1.00
Alcohol During Pregnancy (Regularly)	14,050	0.00	0.06	0.00	1.00
Mother Obese Prior to Pregnancy	14,050	0.02	0.14	0.00	1.00
Mother Low Weight Prior to Pregnancy	14,050	0.07	0.25	0.00	1.00
Mother's Education in Years	14,050	10.83	3.59	0.00	16.00
Percent Twin Births	14,050	2.55	15.76	0.00	100.00
Mother's Age in Years	14,050	27.89	6.61	18.00	49.00
Panel E: Developing Countries					
Mother's Height (cm)	2,050,795	155.82	7.11	84.30	230.50
Mother is underweight	2,050,795	0.12	0.32	0.00	1.00
Mother is obese	2,050,795	0.11	0.32	0.00	1.00
Mother's Education	2,050,795	4.22	4.56	0.00	24.00
Attended Births in Area (% Doctor)	2,050,795	0.35	0.29	0.00	1.00
Attended Births in Area (% Nurse)	2,050,795	0.41	0.26	0.00	1.00
Attended Births in Area (% Any)	2,050,795	0.79	0.20	0.00	1.00
Percent Twin Births	2,050,795	2.07	14.24	0.00	100.00
Mother's age in years	2,050,795	34.15	7.54	18.00	49.00

Each panel presents descriptive statistics of data from each sample analysed in Table 2 of the paper. Panel A comes from the United States Vital Statistics System for all non-ART users from 2009-2013, Panel B consists of all births from the Swedish Medical Birth Register from 1993-2012, Panel C comes Avon Longitudinal Study of Parents and Children, Panel D is the Early Life Longitudinal Survey from Chile, and Panel E consists of pooled DHS data. Further details on data are available in Section 2 and appendix B.2. All variables are either binary measures, or with units indicated in the variable name.

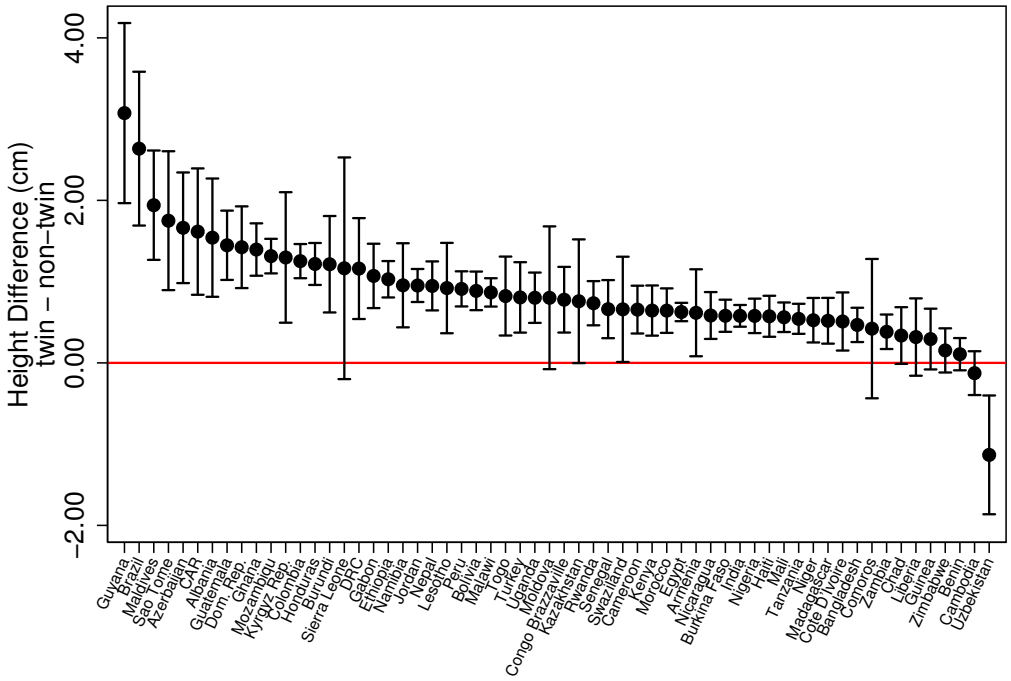
C Cross-country comparisons and the role of income

In this appendix we present results showing the comparability and consistency of twin selection across all the available estimation samples. We use mother's height as this is available matched to birth records in 70 of the countries in our sample, including richer and poorer countries. Figure A3 shows that in 68 of the 70 countries, twin mothers are on average significantly taller than non-twin mothers. Each estimate reflects the mean difference between twin and non-twin mothers, conditioning on age and parity fixed effects. As the comparison is within country, it nets out country differences including differences in the genetic pool (Deaton, 2007).

Since many women in poorer countries are under-nourished, it seems plausible that their resources are particularly challenged in carrying twins to term. As a result, we may expect that income growth and poverty reduction attenuate the association of mother's health and twin births. On the other hand, risky behaviours in pregnancy may be increasing in income, so the gradient will depend upon the health indicator that is analysed. To assess this, we need a comparable index of mother's health for countries that span a range of income levels. As height is widely available, we plot the point estimates from Figure A3 against GDP per capita in Figure A4. The estimates lie above the zero line, indicating that the relationship persists in high income countries. In fact, the coefficients in Table 2 show larger marginal effects of height on twinning in richer countries. Similarly, the mother's height has a significantly larger impact on the probability that boy twins are born than it does on the probability of girl twins (see Table A12).

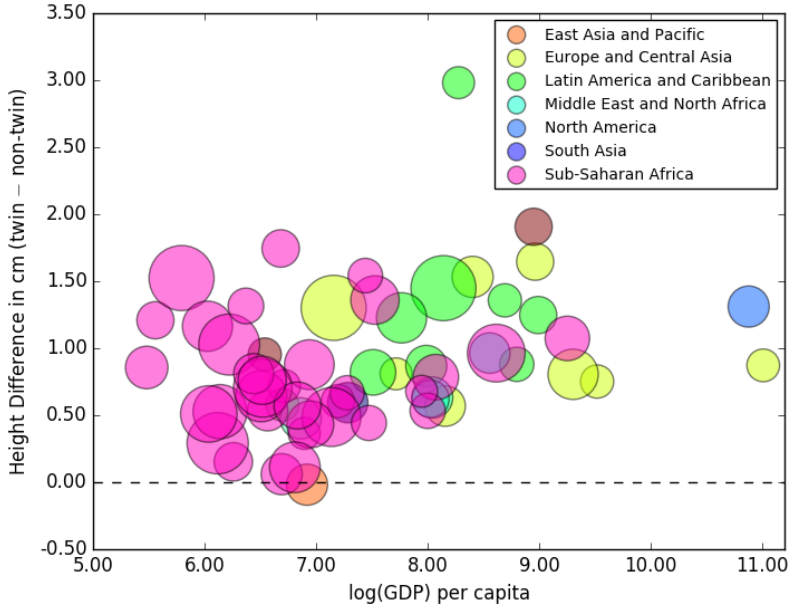
Since education is also widely available and we have seen it is predictive of twinning (both conditional and unconditional on maternal health), below we present plots displaying systematically positive education differences between twin and non-twin mothers in all countries in the sample (Figure A5) and just as for height this is true at high and low levels of GDP, if anything, there is a weakly positive correlation between country income and the education differential (Figure A6), which may reflect the finding cited earlier that the effects of education on health care access and uptake are most substantial in environments in which health-care technologies are changing rapidly (Lleras-Muney and Lichtenberg, 2005; Lleras-Muney and Cutler, 2010).

Figure A3: Height Differential By Twin and non-Twin Mothers by Country



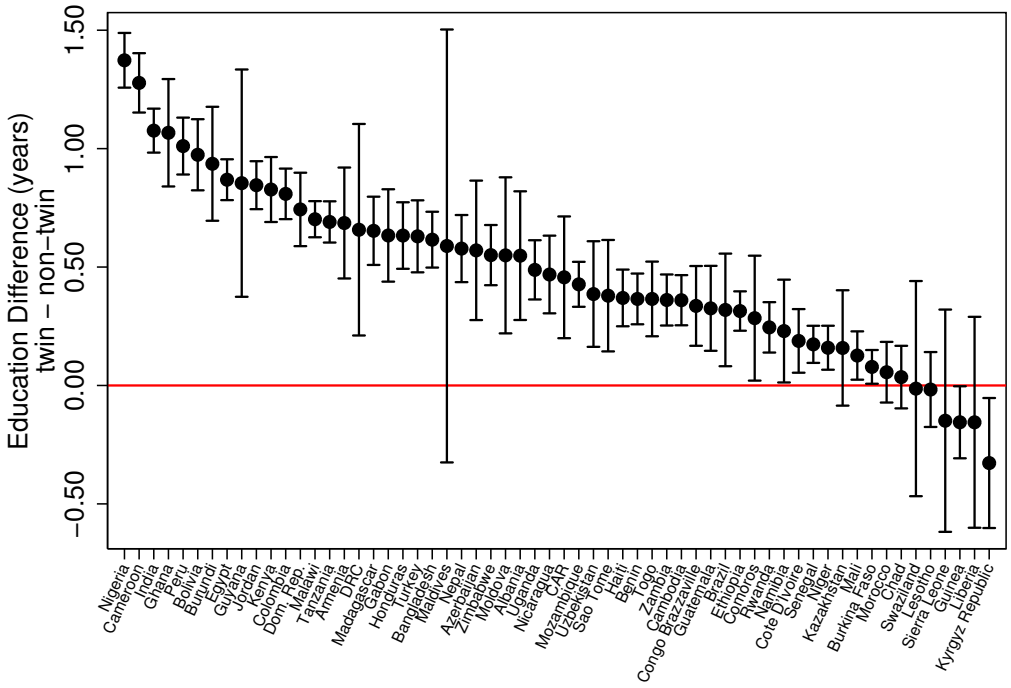
Note to Figure A3: Point estimates of the average difference in height between mothers of twin and singleton births are presented along with the 95% confidence intervals for each country for which the required microdata are available. Sources of data are described in section 2. When based on survey data, each point is weighted to be nationally representative, and if based on vital statistics data, the universe of births is included. The difference-in-mean estimates are conditioned upon total fertility, mother's age and child year of birth.

Figure A4: Height Differential By Twin and non-Twin Mothers by Country and GDP



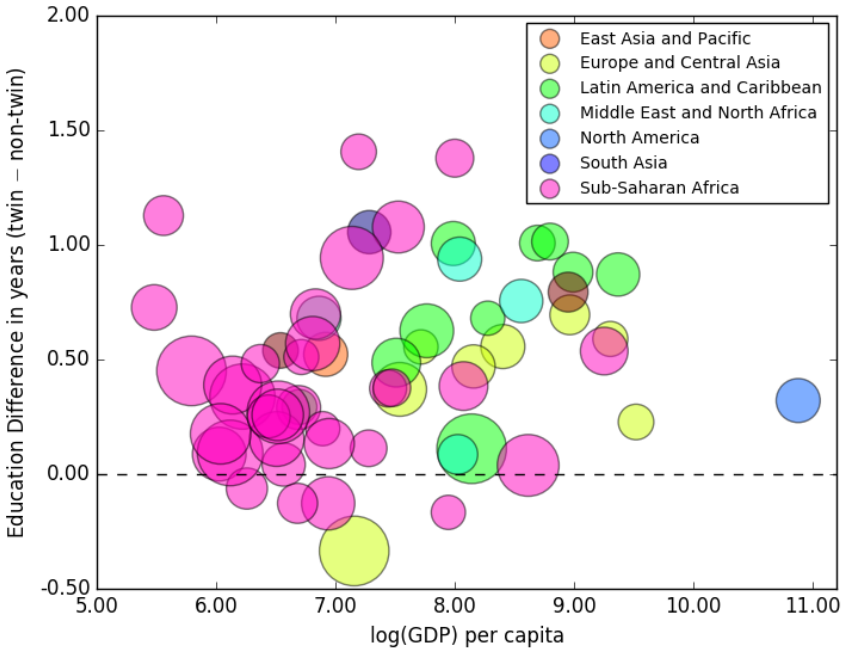
Note to Figure A4: The correlation of the average height differential between twin and singleton mothers in a country with the country's log GDP per capita is plotted. Estimates for the height differential are calculated using the same controls and methodology as in Figure A3. Each circle represents a country and the size of the circle indicates the proportion of births in the country that are twins. Circles above the horizontal dotted line imply that mothers of twins are taller on average. The global correlation between the height difference and GDP conditional on continent fixed effects is 0.259 (t-statistic 1.95).

Figure A5: Completed Education Differential By Twin and non-Twin Mothers by Country



Note to Figure A5: Refer to Figure A3.

Figure A6: Completed Education Differential By Twin and non-Twin Mothers by Country and GDP



Note to Figure A6: Refer to Figure A4

D A Latent-Health Index Measure of Twin versus Non-Twin Mothers

We presented results for various individual measures of maternal health and condition. So at to obtain a summary measure, we calculated a factor index of maternal health based on all available measures, appropriately re-scaled so that each variable measures a positive health improvement. For example, instead of using smoking we use not smoking, and instead of using chronic health conditions, we use no chronic health conditions.

Following Biroli (2016), we use the principal factor method to estimate factor loadings of all health measures available in each country, and based on these factor loadings and individual health measures, calculate each mother's unidimensional latent health score. See Table A15 below. In each case we observe that twin mothers score significantly higher on this index than non-twin mothers.

Table A15: Difference in Aggregate Health Between Twin and non-Twin Mothers

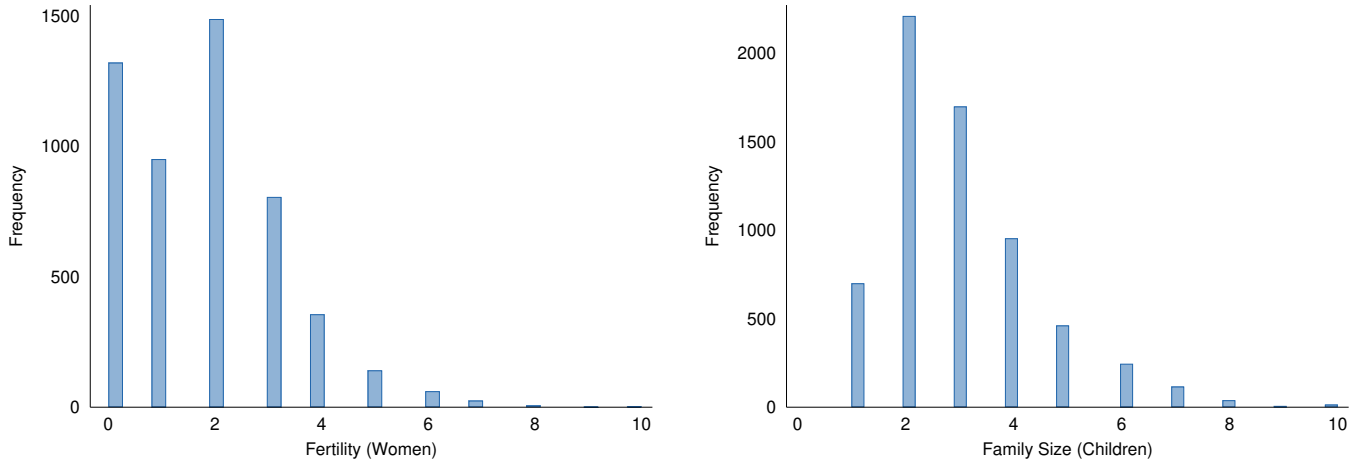
Aggregate Health	DHS	UK	USA	Chile
Twin−Non-twin	0.163	0.128	0.031	0.111
	(0.005)	(0.059)	(0.001)	(0.034)

Notes: We follow Biroli (2016) in using the principal factor method to estimate factor loadings for each positive health measure (for a particular country), and from these factor loadings calculate each mother's standardised latent health. Where a health measure is a negative variable (for example smoking) we multiply by minus one, so that all components are cast as positive effects. This latent health measure for each mother is regressed on whether her birth is a twin (1) or singleton (0). Regression results for each data source are displayed above, along with their standard errors in parentheses.

E Panel Data for Mothers: Robustness to Genetic Traits

As discussed in the paper we sought panel data for mothers that include information on their births and time-varying measures of their health, so that we can estimate the association of maternal health and twinning conditional on woman fixed effects. The NLSY allows us to do this. It has a sample of 5,159 women aged 14 and 24 in 1968 is followed until 2003 although, as discussed in the text, we stop in 1999 when the youngest are aged 45, and the most recent birth observed in the sample occurred in 1998. Of the 5,159 women first surveyed, 3,838 had at least 1 child at any point in their life. Of those, 368 had births prior to 1968 only when the panel survey was not yet implemented, 144 had births exclusively while aged under 18 years, and are excluded from the estimation sample, and an additional 28 have missing information on at least one covariate of interest. This results in an estimation sample of 3,298 mothers, who have a total of 6,439 children. This sample is smaller than most of the estimation samples used in the paper, but has the benefit of measuring health outcomes at various points of time.

Figure A7: Fertility Descriptives: NLSY Young Women Panel



(a) Total Births for All Women

(b) Family Size of All Children

The full distribution of fertility for all women surveyed and the family size of all children born to women with at least one child are displayed in Figure A7. When generating these data, we use the 19 survey waves implemented between 1968 and 1999. Surveys are typically implemented every year or every two years, and at each point any births since the previous survey are reported, along with their birth year. When a birth occurs in between years in which a survey was implemented, covariates are set equal to the value for the survey the year before birth, so that all values refer to *pre-gestational* measures. There are relatively few health measures which are recorded consistently from 1968 to 1999. Summary statistics for the health variables which are consistently recorded for the whole period under study are displayed in Table A16. Alternative health variables such as alcohol consumption and maternal weight were only recorded in most recent waves, once the majority of women had completed the reproductive period. In regressions, we use only health measures which are available consistently.

We display regression results in Table A17. Columns 1 and 2 capture maternal age using a quadratic term, while columns 3 and 4 include full maternal age at birth fixed effects. In column 2 and 4 maternal education is included as a control. Each of these variables are measured as a standardised Z-score, and so are interpreted as the impact of increasing the prevalence of the condition by 1 standard deviation. The estimates are larger than those reported in Table 2, although are also accompanied by a large standard error. The sample size available here is smaller than that of most of the data sets used in the main analysis.

Table A16: NLSY Summary Statistics for Births in Young Women Survey

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Mother's Age	6,439	25.37	4.87	18	49
Fertility	6,439	2.97	1.46	1	10
Twin	6,439	0.0332	0.1793	0	1
Birth Order	6,439	2.18	1.25	1	10
Health Limiting Work	6,439	0.08	0.27	0	1
Smoker	6,439	0.04	0.19	0	1
Had Prior Cancer Diagnosis	6,439	0.04	0.20	0	1

Table A17: Twinning and Maternal Characteristics: Mother Fixed Effects Results from NLSY

Twin$\times$100	(1)
Health Limits Work	-0.092
	(0.304)
Smoker	-1.754**
	(0.838)
Cancer Diagnosis	-1.131***
	(0.329)
Mean Dependent Variable	3.323
Number of Children	6,439
Number of Mothers	3,298

Notes: A panel of births is constructed of each child born to a mother aged 18-49 years between the years of 1968 and 1998 (based on NLSY waves 1968-1999). Each specification includes a quadratic in family income and fixed effects for mother, mother's age at birth, child birth order and child birth year. Mother sampling weights (fixed in 1968) are included in each specification, and standard errors are clustered by mother.

References

J. Angrist, V. Lavy, and A. Schlosser. Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, 28(4):pp. 773–824, 2010.

F. Arnold. Assessment of the Quality of Birth History Data in the Demographic and Health Surveys. In I. for Resource Development, editor, *An Assessment of DHS-I Data Quality*, DHS Methodological Reports, No. 1, chapter 3, pages 81–112. Macro Systems Inc., 1990.

E. Barth. Mutual Learning Programme: Peer Country Comments Paper – Norway. Administrative data for research and evaluation, ICF GHK Consulting, Oct. 2012.

M. Beckett, J. D. Vanzo, N. Sastry, C. Panis, and C. Peterson. The Quality of Retrospective Data: An Examination of Long-Term Recall in a Developing Country. *Journal of Human Resources*, 36(3):593–625, 2001.

P. Biroli. Health and Skill Formation in Early Childhood. Working Paper 17, University of Zurich, UBS International Center of Economics in Society, Nov. 2016.

S. E. Black, P. J. Devereux, and K. G. Salvanes. The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120(2):669–700, 2005.

A. Deaton. *The Analysis of Household Surveys – A Microeconometric Approach to Development Policy*. The Johns Hopkins University Press, 1997.

A. Deaton. Height, health, and development. *Proceedings of the National Academy of Sciences*, 104(33):13232–13237, August 2007.

ICF. *Demographic and Health Survey Interviewer's Manual*. Rockville, Maryland USA, June 2017.

E. E. Leamer. *Specification Searches – Ad Hoc Inference with Nonexperimental Data*. John Wiley & Sons, Inc., 1978.

A. Lleras-Muney and D. Cutler. Understanding Differences in Health behaviors by Education. *Journal of Health Economics*, 29(1):1–28, 2010.

A. Lleras-Muney and F. Lichtenberg. The Effect Of Education On Medical Technology Adoption: Are The More Educated More Likely To Use New Drugs? *Annales d'Economie et Statistique*, 79/80, 2005.

M. F. MacDorman and S. Kirmeyer. The Challenge of Fetal Mortality. NCHS Data Brief 16, National Center for Health Statistics, Apr. 2009.

J. A. Martin and D. L. Hoyert. The national fetal death file. *Seminars in Perinatology*, 26(1):3–11, 2002.

N. A. Mathiowetz. Measurement Error in Surveys of the Low Income Population. Presented at the Workshop on Data Collection for Low Income and Welfare Populations Chapter 6, National Academy of Sciences, Nov. 1999.

M. R. Rosenzweig and J. Zhang. Do population control policies induce more human capital investment? twins, birth weight and China's one-child policy. *Review of Economic Studies*, 76(3):1149–1174, 2009.