Online Appendix

Stacey H. Chen, Yen-Chien Chen and Jin-Tan Liu

Appendix 1: Sample Analogs of Average Direct and Indirect Effects

Under the assumption of randomized sibling gender, we find two useful properties to proceed:

Property (1): ϵ_{1M} and ϵ_{0M} have the same distribution, conditional on sibsize M.

This implies $E[\epsilon_{1M} - \epsilon_{0M}|M = m] = 0$ for any sibsize m.

Property (2): Observed and potential sibsizes have the same distribution conditional on sibling gender.

This implies $E[M_D] = E[M|D = d]$ for $d = 0$ or 1.

Using these two properties, we can construct the sample analogs of DE and IE by averaging over all possible values of potential sibsize in $\mathcal{M}$. We start with ADE, standing for the sample analogs of DE:

$$
\begin{aligned}
ADE &= \Sigma_{m \in \mathcal{M}}\{E[Y|M_1, D = 1] - E[Y|M_1, D = 0]\}Pr\{M_1 = m|D = 1\} \\
&= \Sigma_{m \in \mathcal{M}}E[(\beta_0 + \beta_1 + \beta_2 M_1 + \beta_3 M_1 + \epsilon_{1M}) - (\beta_0 + \beta_2 M_1 + \epsilon_{0M})]Pr\{M_1 = m|D = 1\} \\
&= \beta_1 + \beta_3 E[M|D = 1] + \Sigma_{m \in \mathcal{M}}E[\epsilon_{1M} - \epsilon_{0M}|M = m]Pr\{M_1 = m|D = 1\} \\
&= \beta_1 + \beta_3 E[M|D = 1].
\end{aligned}
$$

The second equality is achieved by applying equation (**??**). The third and fourth equalities are achieved by using Property (2) and Property (1) respectively. The sample analog of IE (denoted by AIE) can be derived similarly or simply by subtracting ADE from ATE.

Table A1: Outcome Equations and Decomposed Effects of Sibling Gender on High School Completion (with *Morethan2* as Mediating Variable), Using Restricted Sample with First Two Births Before 1985

Dependent variable = High school completion	Firstborn females					Firstborn males				
	OLS	2SLS	2SLS 1st-stage interact	OLS interact	2SLS interact	OLS	2SLS	2SLS 1st-stage interact	OLS interact	2SLS interact
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A Next Brother (D)	-0.0031	-0.0067	-0.0039	0.0024	-0.0814	-0.0055	-0.0049	-0.0049	-0.0017	-0.0067
	(0.0015)*	(0.0056)	(0.0054)	(0.0026)	(0.0440)*	(0.0014)*	(0.0020)*	(0.0020)*	(0.0021)	(0.0203)
*Morethan*2 (M)	-0.0166	-0.0330	-0.0201	-0.0117	-0.1042	-0.0271	-0.0180	-0.0179	-0.0233	-0.0200
	(0.0017)*	(0.0250)	(0.0240)	(0.0025)*	(0.0532)	(0.0015)*	(0.0185)	(0.0184)	(0.0021)*	(0.0291)
$D \times M$				-0.0082	0.1079				-0.0074	0.0036
				(0.0032)	(0.0609)				(0.0029)*	(0.0387)
Decomposition										
Average Total Effect	0.0005	0.0005	0.0005	0.0005	0.0005	-0.0035	-0.0035	-0.0035	-0.0035	-0.0035
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0014)*	(0.0014)*	(0.0014)*	(0.0014)*	(0.0014)*
(a) *Average Indirect Effect*	0.0036	0.0072	0.0044	0.0026	0.0228	0.0021	0.0014	0.0014	0.0018	0.0015
	(0.0004)*	(0.0055)	(0.0052)	(0.0006)*	(0.0116)	(0.0001)*	(0.0014)	(0.0014)	(0.0002)*	(0.0022)
(b) *Average Direct Effect*	-0.0031	-0.0067	-0.0039	-0.0021	-0.0223	-0.0055	-0.0048	-0.0049	-0.0053	-0.0050
	(0.0015)*	(0.0056)	(0.0054)	(0.0016)	(0.0117)	(0.0014)*	(0.0020)*	(0.0020)*	(0.0014)*	(0.0026)
Controlled Direct Effect	-0.0031	-0.0067	-0.0039	-0.0029	-0.0109	-0.0055	-0.0048	-0.0049	-0.0055	-0.0049
	(0.0015)*	(0.0056)	(0.0054)	(0.0015)	(0.0067)	(0.0014)*	(0.0020)*	(0.0020)*	(0.0014)*	(0.0020)*
Difference $CDE - ADE$				-0.0009	0.0114				-0.0003	0.0001
				(0.0003)*	(0.0064)				(0.0001)*	(0.0014)
Sample Mean			0.207					0.210		

Note: This table reports OLS and IV estimates for the educational outcome equation (with high school completion as dependent variable) and decomposed effects of sibling gender on the outcome, with fertility choice being measured by *Morethan2*. The corresponding results for fertility choice being measured by *Sibsize* are reported in Table A3. We exclude the interaction between the twins instrument and sibling gender from the first stage for columns (2) and (7), while we include it for columns (3) and (8) in order to be comparable to columns (5) and (10). The decomposed effects are measured according to equations (**??**) to (**??**) as summarized below:

$$AIE = \beta_2\{E[M|D=1] - E[M|D=0]\}, \qquad ADE = \beta_1 + \beta_3 E[M|D=1], \qquad CDE = \beta_1 + \beta_3 E[M].$$

AIE and ADE are evaluated at the conditional mean of $M = Sibsize$: $(E[M|D=1], E[M|D=0]) = (0.5475, 0.7662)$ for firstborn females and $(0.4791, 0.5546)$ for firstborn males, while CDE is evaluated at the unconditional mean $E[M] = 0.6532$ for firstborn females and 0.5158 for firstborn males. We control for the full set of dummies for urban, the subject's age and district of birth, parents' education and years of birth, and mother's age at the first birth. Robust standard errors are reported in parentheses.

Table A2: Outcome Equations and Decomposed Effects of Sibling Gender on University Admission (with *Morethan2* as Mediating Variable), Using Restricted Sample with First Two Births Before 1985

Dependent variable = University admission	Firstborn females					Firstborn males				
	OLS	2SLS	2SLS 1st-stage interact	OLS interact	2SLS interact	OLS	2SLS	2SLS 1st-stage interact	OLS interact	2SLS interact
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A Next Brother (D)	-0.0012	-0.0051	-0.0024	0.0042	-0.0779	-0.0018	-0.0017	-0.0017	0.0009	-0.0058
	(0.0014)	(0.0051)	(0.0049)	(0.0024)	(0.0399)	(0.0012)	(0.0017)	(0.0017)	(0.0018)	(0.0176)
Morethan2 (M)	-0.0147	-0.0328	-0.0203	-0.0099	-0.1022	-0.0183	-0.0174	-0.0171	-0.0156	-0.0216
	(0.0015)*	(0.0227)	(0.0217)	(0.0023)*	(0.0482)*	(0.0013)*	(0.0160)	(0.0159)	(0.0018)*	(0.0252)
$D \times M$				-0.0080	0.1052				-0.0053	0.0078
				(0.0029)*	(0.0552)				(0.0025)*	(0.0335)
Decomposition										
Average Total Effect	0.0020	0.0021	0.0020	0.0020	0.0021	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
(a) *Average Indirect Effect*	0.0032	0.0072	0.0044	0.0022	0.0224	0.0014	0.0013	0.0013	0.0012	0.0016
	(0.0003)*	(0.0050)	(0.0047)	(0.0005)*	(0.0105)*	(0.0001)*	(0.0012)	(0.0012)	(0.0001)	(0.0019)
(b) *Average Direct Effect*	-0.0012	-0.0052	-0.0024	-0.0001	-0.0203	-0.0018	-0.0017	-0.0017	-0.0016	-0.0020
	(0.0014)	(0.0051)	(0.0049)	(0.0014)	(0.0106)	(0.0012)	(0.0017)	(0.0017)	(0.0012)	(0.0023)
Controlled Direct Effect	-0.0012	-0.0052	-0.0024	-0.0010	-0.0092	-0.0018	-0.0017	-0.0017	-0.0018	-0.0018
	(0.0014)	(0.0051)	(0.0049)	(0.0014)	(0.0061)	(0.0012)	(0.0017)	(0.0017)	(0.0012)	(0.0017)
Difference $CDE - ADE$				-0.0008	0.0111				-0.00019	0.0003
				(0.0003)*	(0.0058)				(0.00009)*	(0.0012)
Sample Mean			0.155					0.138		

Note: This table reports OLS and IV estimates for the educational outcome equation (with university admission as dependent variable) and decomposed effects of sibling gender on the outcome, with fertility choice being measured by *Morethan2*. The corresponding results for fertility choice being measured by *Sibsize* are reported in Table A4. We exclude the interaction between the twins instrument and sibling gender from the first stage for columns (2) and (7), while we include it for columns (3) and (8) in order to be comparable to columns (5) and (10). The decomposed effects are measured according to equations (??) to (??) as summarized below:

$$AIE = \beta_2\{E[M|D=1] - E[M|D=0]\}, \qquad ADE = \beta_1 + \beta_3 E[M|D=1], \qquad CDE = \beta_1 + \beta_3 E[M].$$

AIE and *ADE* are evaluated at the conditional mean of $M = Sibsize$: $(E[M|D=1], E[M|D=0]) = (0.5475, 0.7662)$ for firstborn females and $(0.4791, 0.5546)$ for firstborn males, while *CDE* is evaluated at the unconditional mean $E[M] = 0.6532$ for firstborn females and 0.5158 for firstborn males. We control for the full set of dummies for urban, the subject's age and district of birth, parents' education and years of birth, and mother's age at the first birth. Robust standard errors are reported in parentheses.

Table A3: Outcome Equations and Decomposed Effects of Sibling Gender on High School Completion (with *Sibsize* as Mediating Variable), Using Restricted Sample with First Two Births Before 1985

Dependent variable = High school completion	Firstborn females					Firstborn males				
	OLS	2SLS	2SLS 1st-stage interact	OLS interact	2SLS interact	OLS	2SLS	2SLS 1st-stage interact	OLS interact	2SLS interact
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A Next Brother (D)	-0.0064	-0.0093	-0.0096	0.0056	-0.1450	-0.0064	-0.0052	-0.0052	0.0101	-0.0069
	(0.0015)*	(0.0075)	(0.0075)	(0.0053)	(0.0961)	(0.0014)*	(0.0022)*	(0.0022)*	(0.0054)	(0.0766)
Sibsize (M)	-0.0145	-0.0206	-0.0212	-0.0128	-0.0433	-0.0232	-0.0136	-0.0136	-0.0202	-0.0139
	(0.0009)*	(0.0156)	(0.0156)	(0.0012)*	(0.0221)	(0.0011)*	(0.0139)	(0.0139)	(0.0014)*	(0.0203)
$D \times M$				-0.0042	0.0464				-0.0062	0.0006
				(0.0018)*	(0.0328)				(0.0020)*	(0.0290)
Decomposition										
Average Total Effect	0.0005	0.0005	0.0005	0.0005	0.0005	-0.0035	-0.0035	-0.0035	-0.0035	-0.0035
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0014)*	(0.0014)*	(0.0014)*	(0.0014)*	(0.0014)*
(a) *Average Indirect Effect*	0.0069	0.0098	0.0101	0.0061	0.0206	0.0029	0.0014	0.0017	0.0025	0.0017
	(0.0004)*	(0.0074)	(0.0075)	(0.0006)*	(0.0105)	(0.0001)*	(0.0017)	(0.0017)	(0.0002)*	(0.0025)
(b) *Average Direct Effect*	-0.0064	-0.0093	-0.0096	-0.0056	-0.0201	-0.0064	-0.0052	-0.0052	-0.0060	-0.0052
	(0.0015)*	(0.0075)	(0.0075)	(0.0016)*	(0.0106)	(0.0014)*	(0.0022)*	(0.0022)*	(0.0014)*	(0.0029)
Controlled Direct Effect	-0.0064	-0.0093	-0.0096	-0.0065	-0.0094	-0.0064	-0.0052	-0.0052	-0.0064	-0.0052
	(0.0015)*	(0.0075)	(0.0075)	(0.0015)*	(0.0076)	(0.0014)*	(0.0022)*	(0.0022)*	(0.0014)*	(0.0022)*
Difference $CDE - ADE$				-0.0010	0.0107				-0.0004	0.0004
				(0.0004)*	(0.0076)				(0.0001)*	(0.0018)
Sample Mean			0.207					0.210		

Note: This table reports OLS and IV estimates for the educational outcome equation (with high school completion as dependent variable) and decomposed effects of sibling gender on the outcome, with fertility choice being measured by *Sibsize*. The corresponding results for fertility choice being measured by *Morethan2* are reported in Table A1. We exclude the interaction between the twins instrument and sibling gender from the first stage for columns (2) and (7), while we include it for columns (3) and (8) in order to be comparable to columns (5) and (10). The decomposed effects are measured according to equations (**??**) to (**??**) as summarized below:

$$AIE = \beta_2\{E[M|D=1] - E[M|D=0]\}, \qquad ADE = \beta_1 + \beta_3 E[M|D=1], \qquad CDE = \beta_1 + \beta_3 E[M].$$

AIE and *ADE* are evaluated at the conditional mean of $M = Sibsize$: $(E[M|D=1], E[M|D=0]) = (2.6930, 3.1696)$ for firstborn females and $(2.5758, 2.7003)$ for firstborn males, while *CDE* is evaluated at the unconditional mean $E[M] = 2.9234$ for firstborn females and 2.6364 for firstborn males. We control for the full set of dummies for urban, the subject's age and district of birth, parents' education and years of birth, and mother's age at the first birth. Robust standard errors are reported in parentheses.

Table A4: Outcome Equations and Decomposed Effects of Sibling Gender on University Admission (with *Sibsize* as Mediating Variable), Using Restricted Sample with First Two Births Before 1985

	Firstborn females					Firstborn males				
	OLS	2SLS	2SLS 1st-stage interact	OLS interact	2SLS interact	OLS	2SLS	2SLS 1st-stage interact	OLS interact	2SLS interact
Dependent variable = University admission	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A Next Brother (D)	-0.0039	-0.0077	-0.0080	0.0092	-0.1392	-0.0024	-0.0020	-0.0020	0.0093	-0.0123
	(0.0014)*	(0.0068)	(0.0068)	(0.0048)	(0.0870)	(0.0012)*	(0.0019)	(0.0019)	(0.0047)	(0.0663)
Sibsize (M)	-0.0124	-0.0205	-0.0211	-0.0105	-0.0425	-0.0156	-0.0131	-0.0131	-0.0135	-0.0151
	(0.0008)*	(0.0141)*	(0.0141)	(0.0011)*	(0.0200)*	(0.0009)*	(0.0120)	(0.0120)	(0.0012)*	(0.0176)
$D \times M$				-0.0045	0.0449				-0.0044	0.0039
				(0.0016)*	(0.0297)				(0.0017)*	(0.0251)
Decomposition										
Average Total Effect	0.0020	0.0021	0.0021	0.0020	0.0021	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
(a) *Average Indirect Effect*	0.0059	0.0097	0.0103	0.0050	0.0202	0.0019	0.0016	0.0016	0.0017	0.0019
	(0.0004)*	(0.0067)	(0.0067)	(0.0005)*	(0.0095)*	(0.0001)*	(0.0015)	(0.0015)	(0.0002)*	(0.0022)
(b) *Average Direct Effect*	-0.0039	-0.0077	-0.0080	-0.0030	-0.0182	-0.0024	-0.0020	-0.0020	-0.0021	-0.0023
	(0.0014)*	(0.0068)	(0.0068)	(0.0014)*	(0.0096)	(0.0011)*	(0.0019)	(0.0019)	(0.0012)	(0.0025)
Controlled Direct Effect	-0.0039	-0.0077	-0.0080	-0.0040	-0.0078	-0.0024	-0.0020	-0.0020	-0.0024	-0.0021
	(0.0014)*	(0.0068)	(0.0068)	(0.0014)*	(0.0068)	(0.0011)*	(0.0019)	(0.0019)	(0.0012)*	(0.0019)
Difference $CDE - ADE$				-0.0010	0.0104				-0.0003	0.0002
				(0.0004)*	(0.0068)				(0.00010)	(0.0015)
Sample Mean			0.155					0.138		

Note: This table reports OLS and IV estimates for the educational outcome equation (with university admission as dependent variable) and decomposed effects of sibling gender on the outcome, with fertility choice being measured by *Sibsize*. The corresponding results for fertility choice being measured by *Morethan2* are reported in Table A2. We exclude the interaction between the twins instrument and sibling gender from the first stage for columns (2) and (7), while we include it for columns (3) and (8) in order to be comparable to columns (5) and (10). The decomposed effects are measured according to equations (**??**) to (**??**) as summarized below:

$$AIE = \beta_2\{E[M|D = 1] - E[M|D = 0]\}, \qquad ADE = \beta_1 + \beta_3 E[M|D = 1], \qquad CDE = \beta_1 + \beta_3 E[M].$$

AIE and *ADE* are evaluated at the conditional mean of $M = Sibsize$: $(E[M|D = 1], E[M|D = 0]) = (2.6930, 3.1696)$ for firstborn females and $(2.5758, 2.7003)$ for firstborn males, while *CDE* is evaluated at the unconditional mean $E[M] = 2.9234$ for firstborn females and 2.6364 for firstborn males. We control for the full set of dummies for urban, the subject's age and district of birth, parents' education and years of birth, and mother's age at the first birth. Robust standard errors are reported in parentheses.

Table A5: Family Compositional Effects on the First Child's Education, OLS

	Firstborn females						Firstborn males					
	No controls			Controls			No controls			Controls		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Y = High school completion		*0.246*			*0.246*			*0.239*			*0.239*	
A next brother (D)	0.0037	-0.0297	0.0135	0.0018	-0.0044	0.0133	-0.0018	-0.0106	0.0082	-0.0028	-0.0051	0.0099
	(0.0013)*	(0.0014)*	(0.0047)*	(0.0012)	(0.0013)*	(0.0043)*	(0.0013)	(0.0013)*	(0.0046)	(0.0012)*	(0.0012)*	(0.0043)*
Sibsize (M)		-0.0775	-0.0711		-0.0146	-0.0120		-0.0869	-0.0834		-0.0233	-0.0206
		(0.0007)*	(0.0010)*		(0.0008)*	(0.0010)*		(0.0008)*	(0.0011)*		(0.0008)*	(0.0011)*
$D \times M$			-0.0156			-0.0064			-0.0074			-0.0059
			(0.0015)*			(0.0014)*			(0.0016)*			(0.0015)*
Y= University admission		*0.177*			*0.177*			*0.153*			*0.153*	
A next brother (D)	0.0038	-0.0214	0.0104	0.0023	-0.0028	0.0104	0.0013	-0.0046	0.0098	0.0006	-0.0010	0.0103
	(0.0012)*	(0.0012)*	(0.0041)*	(0.0011)*	(0.0012)*	(0.0039)*	(0.0011)	(0.0011)*	(0.0039)*	(0.0010)	(0.0010)	(0.0037)*
Sibsize (M)		-0.0584	-0.0537		-0.0119	-0.0100		-0.0576	-0.0549		-0.0156	-0.0135
		(0.0007)*	(0.0009)*		(0.0007)*	(0.0008)*		(0.0007)*	(0.0009)*		(0.0007)*	(0.0009)*
$D \times M$			-0.0114			-0.0048			-0.0057			-0.0044
			(0.0013)*			(0.0013)*			(0.0014)*			(0.0013)*

Note: This table reports the OLS results of how the firstborn's education is associated with the sex of the secondborn (D), sibsize (M) and their interaction. We also report in Table ?? the corresponding results with M being measured by *Morethan2*. We include 416,315 firstborn females and 434,729 firstborn males born between 1978 and 1984 who have at least one sibling. We control for the full set of dummies for urban, the subject's age and district of birth, parents' education and years of birth, and mother's age at the first birth. The sample mean of the educational outcome (Y) is in italics, and robust standard errors are in parentheses.

Table A6: Heterogeneous Decomposed Effects on First Child's Education

	Firstborn females				Firstborn males			
	HS (1)	Univ (2)	HS (3)	Univ (4)	HS (5)	Univ (6)	HS (7)	Univ (8)
	Father HS+		Father HS-		Father HS+		Father HS-	
Average total effect / Mean	0.010	0.012	0.005	0.017	-0.008	0.007	-0.019	-0.003
	(0.006)	(0.007)	(0.009)	(0.012)	(0.006)	(0.008)	(0.009)	(0.012)
(a) Average indirect effect / Mean	0.098	0.100	0.013	-0.007	0.005	0.009	0.010	0.006
	(0.028)	(0.036)	(0.072)	(0.090)	(0.004)	(0.006)	(0.014)	(0.018)
(b) Average direct effect / Mean	-0.088	-0.088	-0.009	0.024	-0.013	-0.002	-0.029	-0.009
	(0.029)	(0.036)	(0.073)	(0.091)	(0.007)	(0.010)	(0.017)	(0.022)
Sample mean	0.377	0.277	0.152	0.105	0.369	0.240	0.146	0.091
	Mother HS+		Mother HS-		Mother HS+		Mother HS-	
Average total effect / Mean	0.003	0.002	0.013	0.027	0.008	0.002	-0.016	0.007
	(0.006)	(0.008)	(0.008)	(0.010)	(0.006)	(0.009)	(0.008)	(0.011)
(a) Average indirect effect / Mean	0.082	0.072	0.043	0.064	0.001	0.000	0.026	0.033
	(0.026)	(0.033)	(0.066)	(0.082)	(0.003)	(0.005)	(0.012)	(0.016)
(b) Average direct effect / Mean	-0.079	-0.070	-0.030	-0.037	0.007	0.001	-0.043	-0.026
	(0.027)	(0.034)	(0.066)	(0.083)	(0.007)	(0.010)	(0.015)	(0.019)
Sample mean	0.415	0.307	0.164	0.114	0.401	0.265	0.160	0.099
	Urban		Rural		Urban		Rural	
Average total effect / Mean	0.009	0.014	0.007	0.012	-0.014	-0.006	-0.010	0.010
	(0.008)	(0.010)	(0.007)	(0.008)	(0.008)	(0.011)	(0.007)	(0.009)
(a) Average indirect effect / Mean	0.076	0.106	0.095	0.051	0.002	0.003	0.016	0.020
	(0.034)	(0.043)	(0.050)	(0.063)	(0.005)	(0.007)	(0.009)	(0.013)
(b) Average direct effect / Mean	-0.067	-0.092	-0.088	-0.039	-0.016	-0.009	-0.026	-0.010
	(0.035)	(0.044)	(0.051)	(0.064)	(0.009)	(0.013)	(0.011)	(0.015)
Sample mean	0.298	0.216	0.219	0.157	0.281	0.181	0.217	0.139

Note: This table reports the IV estimated decomposed effects of having a secondborn brother on the first child's high school completion (HS) and university admission (Univ), by place of birth (urban versus rural) and by maternal/paternal education ((academic or vocational) high school diploma or more (HS+) versus less high school (HS-)). $M = Morethan2$; results using $Sibsize$ are almost identical. Robust standard errors are in parentheses. Other covariates not listed in the table are the subject's age and district of birth, parents' years of birth, and maternal age at the first birth. The decomposed effects are evaluated at the conditional means of $M = Morethan2$, i.e. $E[M|D = 1], E[M|D = 0]$ in the corresponding subgroup.

Table A7: Robustness Checks, Using Data from Firstborn Males

	High school completion				University admission			
	Baseline estimates (1)	Add birthweight percentile (2)	Add gestation length (3)	Add birth spacing (4)	Baseline estimates (5)	Add birthweight percentile (6)	Add gestation length (7)	Add birth spacing (8)
A next brother (D)	-0.0128	-0.0100	-0.0123	-0.0128	-0.0122	-0.0102	-0.0118	-0.0122
	(0.0126)	(0.0126)	(0.0131)	(0.0126)	(0.0109)	(0.0109)	(0.0113)	(0.0109)
Morethan2 (M)	-0.0307	-0.0155	-0.0303	-0.0290	-0.0261	-0.0150	-0.0287	-0.0254
	(0.0203)	(0.0206)	(0.0210)	(0.0210)	(0.0176)	(0.0178)	(0.0182)	(0.0174)
$D \times M$	0.0193	0.0158	0.0173	0.0194	0.0266	0.0241	0.0246	0.0269
	(0.0275)	(0.0275)	(0.0286)	(0.0275)	(0.0238)	(0.0239)	(0.0248)	(0.0238)
Mean birthweight percentile of 2nd birth		0.0177				0.0131		
		(0.0022)*				(0.0019)*		
Gestational length of 2nd birth			-0.0010				-0.0003	
			(0.0006)				(0.0005)	
Ln (birth spacing between first 2 births)				-0.0093				-0.0093
				(0.0030)*				(0.0026)*
Decomposition								
Average Total Effect	-0.0028	-0.0024	-0.0032	-0.0027	0.0006	0.0084	0.0004	0.0006
	(0.0012)*	(0.0012)*	(0.0012)*	(0.0012)*	(0.0011)	(0.0011)*	(0.0011)	(0.0011)
(a) *Average Indirect Effect*	0.0020	0.0010	0.0019	0.0019	0.0017	0.0010	0.0018	0.0016
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0011)	(0.0011)	(0.0012)	(0.0011)
(b) *Average Direct Effect*	-0.0048	-0.0034	-0.0051	-0.0046	-0.0011	-0.0001	-0.0014	-0.0011
	(0.0018)*	(0.0018)	(0.0018)*	(0.0018)*	(0.0015)	(0.0015)	(0.0016)	(0.0015)
Sample mean	0.239	0.240	0.240	0.239	0.153	0.154	0.154	0.153
Sample size	434,729	428,493	421,586	434,301	434,729	428,493	421,586	434,301

Note: This table reports robustness checks on the 2SLS results of the educational outcome equation and the decomposed effects of sibling gender on firstborn males. Though not reported here, we find no evidence of nonlinearity in gestation length. Since boys are heavier than girls at birth, we use gender-specific birthweight percentiles and calculate the mean percentile of secondborn twins. We control implicitly for the full set of dummies for urban, the subject's age and district of birth on average, parents' education and years of birth, and maternal age at the first birth. Robust standard errors are reported in parentheses, and * indicates significant at the 5% level. See Table **??** for the results using data for firstborn females.

Table A8: Family Size and Initial Health Condition of the Secondborn

Family size	Mean gestation length	Mean birthweight	Mean birthweight percentile	Number of families
2	39.571	3.301	0.515	406,255
3	39.669	3.267	0.501	328,111
4	39.675	3.233	0.490	84,418
5	39.672	3.210	0.482	15,812
6	39.669	3.203	0.478	3,198
7	39.697	3.183	0.462	674
8	39.667	3.196	0.473	168
9 or more	39.820	3.187	0.461	60
All	39.623	3.279	0.506	838,696

Note: This table reports the relationship between family size and the initial health condition of the secondborn whose oldest sibling was born between 1978 and 1984. Birthweight is in kilograms and birthweight percentiles are calculated given the child's gender. We calculate the mean if the secondborn is twins. This table uses fewer observations than the firstborn population in Table ?? since some secondborn children have no information about initial health condition.