Health at Birth, Parental Investments and Academic Outcomes

July 2016

Online Appendix

Prashant Bharadwaj
Department of Economics
UC San Diego
9500 Gilman Drive, La Jolla, CA 92093
prbharadwaj@ucsd.edu

Christopher A. Neilson
Department of Economics
Princeton University
A 16 J1 Firestone Library,
Princeton, NJ 08544
cneilson@princeton.edu

Juan Eberhard
Diagonal Las Torres 2640
Santiago 7941169,
Santiago Metropolitan Region, Chile
juan.eberhard@gmail.com

For Online Publication

A Appendix: Model

A.1 Deciding Between Educational and Non-Educational Inputs

In order to better understand what happens to effective time endowments in the case with and without public goods in parental investments, we consider a problem where education is not the only activity in the household, and other competing activities may be also important for raising a child. The second part of the problem is related to the possibility that parents can strategically use investment time to reinforce the difference between siblings for efficiency motives, or they can compensate the less endowed child for inequality aversion motives. In our model, we explore the implications of both cases.

We start assuming that parents allocate time among different activities to raise their children. Specifically, parents can allocate time between educational activities T_E or non educational activities T_{NE}. We can think of the parents' problem as

$$\begin{aligned} \max_{T_E, T_{NE}} \quad & V(T_E, T_{NE}) \\ \text{s.t.} \quad & T_E + T_{NE} \leq T \end{aligned} \tag{1}$$

Where V is the utility coming from educational and non educational activities[1]. T is the total time allocated to raise the children in the household. Note that, if there is more than one child in the household, parents use the aggregate educational and non educational times and utilities to make the allocation decision.

[1] An alternative formulation consists on assuming that parents maximize the production of "children quality," which uses time in both educational and non educational inputs. Thus, the allocation of time is related to the marginal productivity, instead of the marginal utility which is the key concept in the formulation presented in the main text.

We denote T_E^* and T_{NE}^* the optimal allocation of time, coming from the solution of the maximization problem in Equation 1. Note that the optimal allocation depends on the marginal utilities associated with the educational and non-educational activities. In the main text, for expositional ease, we refer to T_E^* as T_E.

A.2 Allocations in the Presence of Public Goods

$$
\begin{aligned}
\max_{T_E, T_{NE}} \quad & V(T_E, T_{NE}) \\
\text{s.t.} \quad & T_E + T_{NE} \leq T
\end{aligned}
\tag{2}
$$

If T_E^* and T_{NE}^* are the optimal allocation, they satisfied the first order conditions:

$$
\begin{aligned}
\frac{\partial V(T_E^*, T_{NE}^*)}{\partial T_E} &= \lambda \\
\frac{\partial V(T_E^*, T_{NE}^*)}{\partial T_{NE}} &= \lambda \\
\text{combined:} \quad \frac{\partial V(T_E^*, T_{NE}^*)}{\partial T_E} &= \frac{\partial V(T_E^*, T_{NE}^*)}{\partial T_{NE}}
\end{aligned}
\tag{3}
$$

When parents know the public good dimension of parental investment, they realize that their effort T_E effectively converts into $\hat{T}_E = (1 + \delta(i, i'))T_E$. Therefore, they solve the problem

$$
\begin{aligned}
\max_{\hat{T}_E, T_{NE}} \quad & V(\hat{T}_E, T_{NE}) \\
\text{s.t.} \quad & \frac{\hat{T}_E}{1 + \delta(i, i')} + T_{NE} \leq T
\end{aligned}
\tag{4}
$$

In a similar way to the absence of public good, we combined the first order equations to obtain

$$
\frac{\partial V(T_E^{**}, T_{NE}^{**})}{\partial T_E}(1 + \delta(i, i')) = \frac{\partial V(T_E^{**}, T_{NE}^{**})}{\partial T_{NE}}
\tag{5}
$$

where $T_E^{**}(1 + \delta(i, i') = \hat{T}_E^*$. Proposition: *If V is a Leontief utility function, the optimal allocation for* *educational activities when there is a degree of public good dimension in parental investments is smaller than in the case without public goods.*

Proof 1

Let's denote T_E^* and T_{NE}^* the optimal allocations in the absence of public good dimension on parental investment, and T_E^{**} and T_{NE}^{**} the optimal allocations when public good dimension on parental investment is present. Finally, $\hat{T}$ represents the effective time when the public good dimension feature is present.

V is a Leontief production function, expressed as

$$V(T_E, T_{NE}) = \min\{a_1 T_E, a_2 T_{NE}\}$$

It is well known that the solution for the optimal allocation for the Leontief utility function is that

$$T_E^* = \frac{a_2}{a_1 + a_2} T \quad \wedge \quad T_{NE}^* = \frac{a_1}{a_1 + a_2} T \tag{6}$$

The public good dimension of parental investment effectively increases the parameter a_1 from its original value to $a_1(1 + \delta(i, i'))$. Therefore, the new optimal allocations are

$$T_E^{**} = \frac{a_2}{a_1(1 + \delta(i, i')) + a_2} T \quad \wedge \quad T_{NE}^* = \frac{a_1(1 + \delta(i, i'))}{a_1 + a_2} T \tag{7}$$

Comparing the allocation assigned to educational activities in Equation 6 with the one displayed in Equation 7, it is easy to see that the public good dimension of parental investment induce a decrease in the time assigned to educational activities.

A.3 Proof of Proposition 1 from Section 4

If compensating (reinforcing) parents can fully differentiate the educational inputs allocated to each child, the test score gap between siblings will decrease (increase) over time. If there is only partial parental investment differentiation, the test score gap may decrease (increase), but this decrease (increase) will be less than in the case of full differentiation.

Proof 2

For the case where parents can fully differentiate across siblings, Equation 4 and Equation 5 in the main text indicate that, for given cognitive endowments θ_{1jg} and $\theta_{2jg'}$, the allocation for child 1 is just a factor of allocation for child 2.

In particular, the factor is

$$C(\gamma, \rho, \theta_{1jg}, \theta_{2jg'}) = \left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}$$

Without loss of generality, let's assume that child 1 has a higher cognitive endowment that child 2. Thus, $\frac{\theta_{2jg'}}{\theta_{1jg}} < 1$. Additionally, if $\rho < 0$, or when parents present a compensating behavior, the exponent $\frac{\gamma\rho}{(1-\gamma)\rho-1} > 0$ because numerator and denominator are both negative. We conclude that $C(\gamma, \rho, \theta_{1jg}, \theta_{2jg'}) < 1$, and, therefore, the parental investment allocation for child 2 is bigger than for child 1, which is consistent with the compensating behavior. Note that, if $\rho > 0$, $\frac{\gamma\rho}{(1-\gamma)\rho-1} < 0$, and, therefore $C(\gamma, \rho, \theta_{1jg}, \theta_{2jg'}) > 1$. If child 2 has lower cognitive endowment that child 1, he or she will receive higher educational inputs. Equation 6 from the main text captures the evolution of cognitive endowments, and it shows that higher values of educational inputs for child 2 will reduce the gap between the cognitive endowments[2]. As $\theta_{2jg'} \longrightarrow \theta_{1jg}$, the factor $C(\gamma, \rho, \theta_{1jg}, \theta_{2jg'}) \longrightarrow 1$, producing the convergency of cognitive endowments, optimal educational inputs, and test scores. In the case of partial differentiation, we can assume without loss of generality that the actual parental invest-

[2] In order to rule out a *overshooting* behavior from the parents, and to make the evolution of cognitive endowment a relatively persistent process, we assume a specific region for the parameters β_X, $\beta_{X\theta}$, and T.

ment received by the children is a weighted average of the optimal parental investment expressed in Equation 4 and Equation 5 in the main text. In other words,

$$\begin{aligned}\tilde{X}_{1jg} &= \alpha_1 X^*_{1jg} + (1-\alpha_1) X^*_{2jg'} \\ \tilde{X}_{2jg} &= \alpha_2 X^*_{1jg} + (1-\alpha_2) X^*_{2jg'}\end{aligned}$$

where the tilde represents the actual educational input received by each child. Partial differentiation implies that α_1 and α_2 are in the interval (0,1). From the previous discussion, we know that if child 2 has a lower endowment, $X_{1jg} < X_{2jg'}$, and, therefore

$$X^*_{1jg} < \tilde{X}_{1jg} < X^*_{2jg'} \quad \wedge \quad X^*_{1jg} < \tilde{X}_{2jg'} < X^*_{2jg'}$$

It is easy to conclude that the compensating effort in the partial differentiation case will reduce the gap in the cognitive endowment dimension less than in the case of full differentiation. This is because $X^*_{1jg} < \tilde{X}_{1jg}$, or the high endowed child receives more parental investment in the partial differentiation case, and $\tilde{X}_{2jg'} < X^*_{2jg'}$ implies that the low endowed child receives less parental investment in the partial differentiation case.

Corollary 2 *The public good dimension of parental investment implies partial differentiation across children. Thus, the compensating (reinforcing) behavior will take longer to reduce (increase) the test score gap than in the absence of public good dimension.*

Proof 3

According to our model, the public good dimension feature of parental investment implies that the optimal allocation for child 1 (denoted by double stars) satisfies

$$\begin{aligned}
\hat{X}^{*}_{1jg} &= X^{**}_{1jg} + \delta(1,2)X^{**}_{2jg'} \\
&= \frac{T^{**}_{E}}{(1-\delta(1,2))\left[1+\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right]}\left[\left(\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}} - \delta(1,2)\right) + \delta(1,2)\left(1-\delta(1,2)\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right)\right] \\
&= \frac{T^{**}_{E}}{(1-\delta(1,2))\left[1+\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right]}\left[(1-\delta(1,2)^{2})\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right] \\
&= \frac{T^{**}_{E}}{\left[1+\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right]}\left[(1+\delta(1,2))\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right] \quad \text{but} \quad T^{**}_{E} = \frac{\hat{T}^{*}_{E}}{1+\delta(1,2)} \\
&= \frac{\hat{T}_{E}^{\;*}}{\left[1+\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right]}\left(\frac{\theta_{2jg'}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}
\end{aligned}$$

Which is exactly the same expression asthan in the original case but with $\hat{T}^{*}_{E}$ instead of T^{*}_{E}. Furthermore, because $\hat{T}^{*}_{E} < T^{*}_{E}$, it is easy to show that there is α_1 such that $\hat{X}^{*}_{1jg}$ can be written as

$$\hat{X}^{*}_{1jg} = \alpha_1 X^{*}_{1jg} + (1-\alpha_1)X^{*}_{2jg'}$$

Similarly for $\hat{X}^{*}_{2jg'}$. Therefore, the public good dimension is a particular case of partial differentiation, and the results of the proposition can be applied for this case.

A.4 Simulation Details

All the figures in the main text where constructed using the solutions simulated in Matlab 7.12. Code is available from the authors.

The solutions for the optimal allocations are presented in Equation 4 and Equation 5 from the main text. We simulate the solutions with the following parameters:

Table A.1: Parameters used in Simulation

Optimal Allocation Parameters	
ρ	40 equidistant points in the interval $[-0.9, 0.9]$
γ	0.5
θ_{1j1}	1.5
θ_{2j1}	1.0
T_E	0.5
Evolution of Endowments	
β	3
η	1.01
ζ	1.25
Public Good Parameters	
$\delta(i, i')$	$\delta^{(\text{age difference})}$
δ	0.8
Age Difference	1.5

Starting with the initial values of θ presented in the table above and the solution for optimal allocation of parental investment X^*, we constructed the evolution of θ over time for each child.

Once we have the sequence of optimal X and the implied θ, we calculate the test scores, using the equation

$$T_{ijg} = \theta_{ijg}^{\gamma} \cdot X_{ijg}^{(1-\gamma)}$$

A.5 Additional Extensions to the Model

A.5.1 CES Test Score Production Function

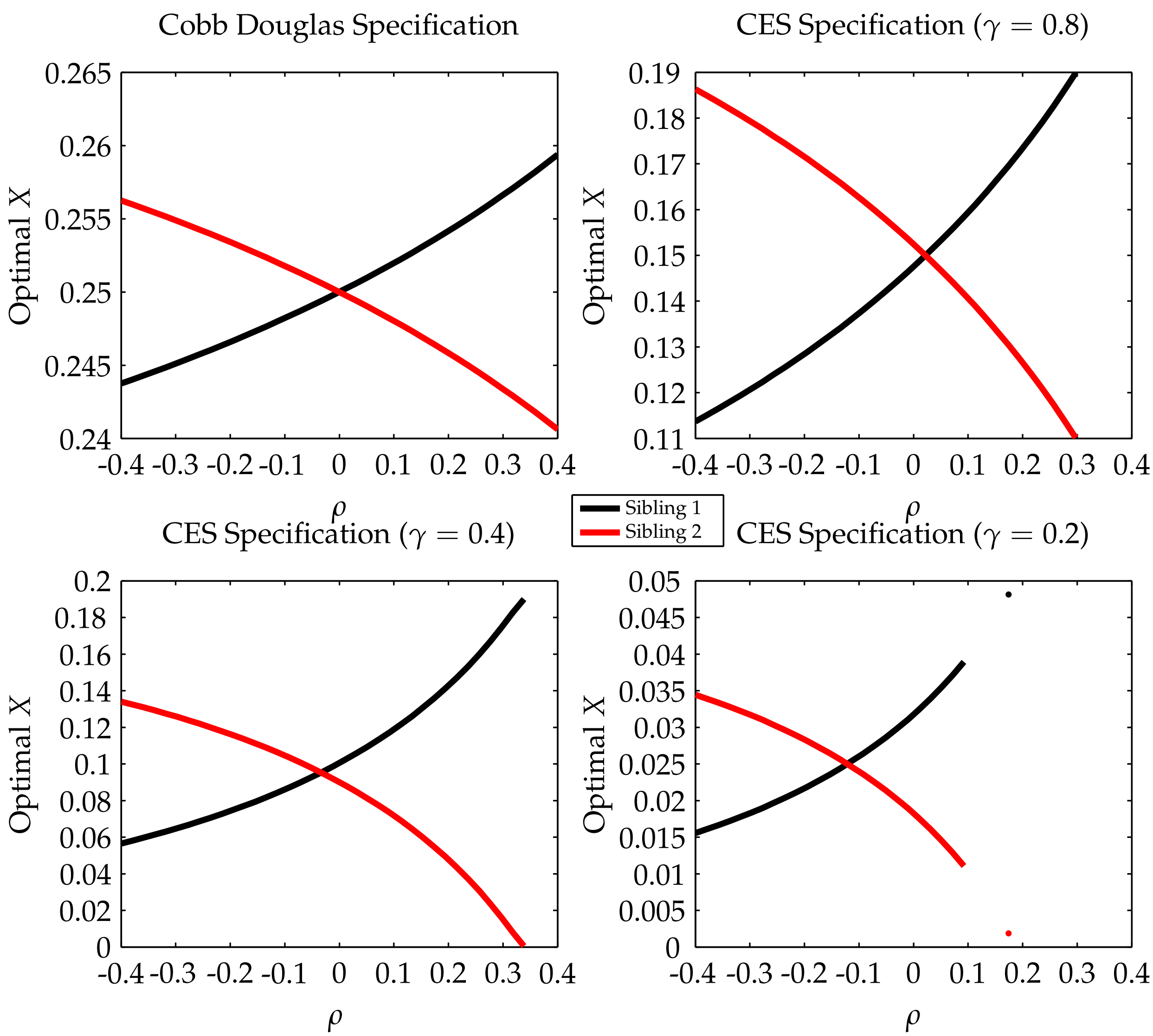
We assume that test scores take as input current cognitive endowment and current parental investment. In order to obtain a closed form solution, we use a Cobb-Douglas production function. However, a more general production function can be used.

Figure A.2 displays the optimal parental investment when test score production follows

$$T_{ig} = (X_{ig}^{\gamma} + \theta_{ig}^{\gamma})^{1/\gamma}$$

The figure shows, for different γ, we observe that the optimal parental investment crosses their paths as ρ increases. In other words, for a general test score production function, parents with inequality aversion invest more in the less endowed child.

Figure A.1: CES Case



Note: Note that all coefficients are significant at least at the 10% level with the exception of 7th and 8th grade, in which the number of observations is smaller and the siblings +3 and OLS is no longer significant.

A.5.2 Solution of the Model in a Dynamic Setup

One alternative approach to modeling the parents problem is a dynamic problem. At time 0, parents decide the optimal **path** of parental investments in order to maximize the discounted present value of the utilities at each grade. This problem is formalized in the following way

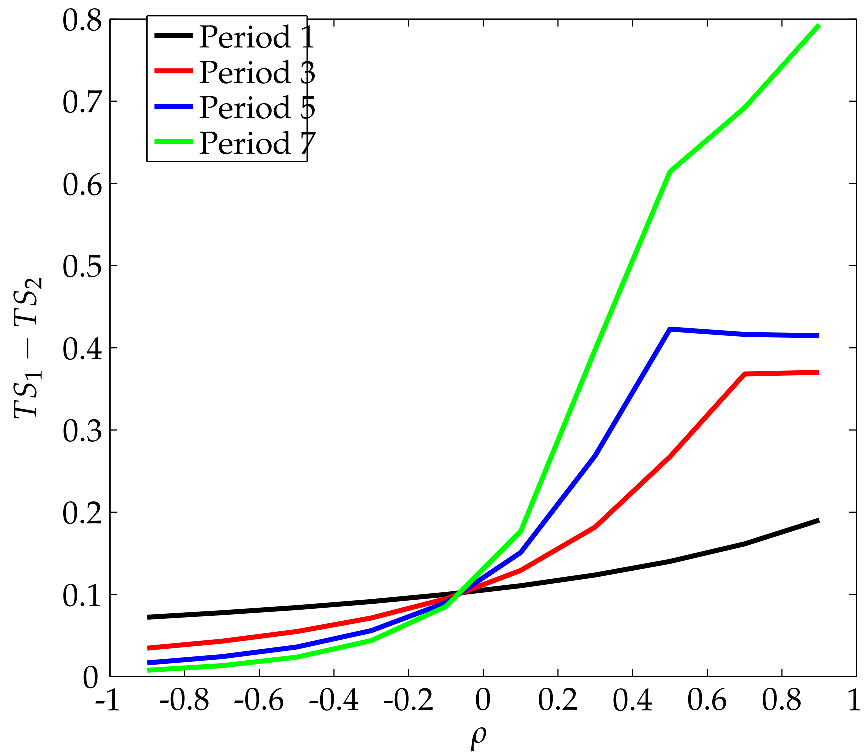
$$
\begin{aligned}
V_g(\theta_{1g}, \theta_{2g}) = \max_{X_{1g}, X_{2g}} \quad & \left(\theta_{1g}^{\gamma\rho}(X_{1g}^{1-\gamma})^{\rho} + \theta_{2g}^{\gamma\rho}(X_{2g}^{1-\gamma})^{\rho}\right)^{\frac{1}{\rho}} + V_{g+1}((\theta_{1(g+1)}, \theta_{2(g+1)}) \\
\text{s.t.} \quad & X_{1g} + X_{2g} \leq T_E \\
& \theta_{i(g+1)} = \beta\theta_{ig}^{\eta} X_{ig}^{\zeta} \qquad i \in \{1, 2\}
\end{aligned} \tag{8}
$$

Assuming that parents have perfect information, and assuming that $V_9 = constant$, we can solve the problem using backward induction. This means that we solve the problem for grade 8, which is equivalent to the one presented in the text, for each possible $(\theta_{1,8}; \theta_{2,8})$. That is, we solve the optimal parental investment at grade 8 for each possible cognitive endowment observed at that time. With that information we construct $V_8(\cdot, \cdot)$.

With the values of $V_8(\cdot, \cdot)$, we can solve the problem for grade 7. Notice that now, parental investment affects current utility through the current test score and **future** utility, through the effect on $(\theta_{1,8}; \theta_{2,8})$. Once we obtain the values of $V_7(\cdot, \cdot)$, we can keep iterating backwards.

Figure A.2 displays the evolution of the gap in test scores, when parents solve at time 0 the dynamic problem. As we can see, the pattern of the gap between test scores has the same features as the one presented in Figure 4 in the main text.

Figure A.2: Gap Between Test Scores ($\delta = 0$)



Note: Note that all coefficients are significant at least at the 10% level with the exception of 7th and 8th grade where the number of observations is smaller and the siblings +3 and OLS is no longer significant.

Table A.2: Twins Estimates: Log Birth Weight and Math Achievement in Grades 1-8 by types of Twin Pairs

Dependent Variable:	Grade							
Standardized Math Scores	1	2	3	4	5	6	7	8
Panel A - Same Sex Twin Pairs								
Log Birth Weight	0.446	0.533	0.502	0.632	0.559	0.615	0.551	0.546
	[0.065]***	[0.060]***	[0.061]***	[0.060]***	[0.064]***	[0.071]***	[0.081]***	[0.097]***
Constant	-3.618	-4.242	-3.952	-4.932	-4.274	-4.675	-4.169	-4.167
	[0.507]***	[0.468]***	[0.473]***	[0.469]***	[0.498]***	[0.554]***	[0.634]***	[0.762]***
Number of Same Sex Pairs	5,908	6,726	6,434	6,279	5,505	4,480	3,599	2,617
Panel B - Boy-Girl Twin Pairs								
Log Birth Weight	0.595	0.547	0.308	0.608	0.620	0.434	0.657	0.729
	[0.145]***	[0.134]***	[0.138]**	[0.139]***	[0.151]***	[0.178]**	[0.201]***	[0.255]***
Constant	-4.770	-4.326	-2.374	-4.721	-4.752	-3.241	-5.022	-5.598
	[1.136]***	[1.047]***	[1.084]**	[1.091]***	[1.183]***	[1.394]**	[1.576]***	[2.002]***
Number of Twin Pairs	1,945	2,138	2,021	1,937	1,725	1,335	1,055	764
Panel C - Boy-Boy Twin Pairs								
Log Birth Weight	0.362	0.407	0.431	0.558	0.467	0.641	0.590	0.527
	[0.104]***	[0.094]***	[0.096]***	[0.096]***	[0.102]***	[0.115]***	[0.136]***	[0.161]***
Constant	-2.965	-3.217	-3.361	-4.328	-3.560	-4.895	-4.485	-4.002
	[0.809]***	[0.732]***	[0.753]***	[0.754]***	[0.795]***	[0.904]***	[1.065]***	[1.260]***
Number of Twin Pairs	2,568	2,932	2,774	2,699	2,339	1,871	1,496	1,094
Panel D - Girl-Girl Twin Pairs								
Log Birth Weight	0.509	0.625	0.553	0.686	0.629	0.597	0.523	0.559
	[0.083]***	[0.078]***	[0.078]***	[0.076]***	[0.081]***	[0.089]***	[0.100]***	[0.122]***
Constant	-4.101	-4.993	-4.376	-5.374	-4.814	-4.521	-3.955	-4.281
	[0.647]***	[0.608]***	[0.605]***	[0.594]***	[0.635]***	[0.697]***	[0.780]***	[0.949]***
Number of Twin Pairs	3,340	3,794	3,660	3,580	3,166	2,609	2,103	1,523

Notes: Twins fixed effects employed in all regressions. Regressions are based on Equation 15 in the main text. Robust standard errors in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.