Health at Birth, Parental Investments and Academic Outcomes

Prashant Bharadwaj, Juan Eberhard & Christopher Neilson[†]

August 2014

Abstract

This paper explores the relationship between health at birth and academic outcomes using administrative panel data from Chile. Twins fixed effects models estimate a persistent effect of birth weight on academic achievement while OLS and siblings fixed effects models find this relationship to decline over time. We make sense of these findings in the context of a model of human capital accumulation where parental investments respond to initial endowments. Using detailed data on parental investments, we find that investments are compensatory with regard to initial health, but that within twins, parents do not invest differentially. These findings suggest that initial health shocks significantly affect academic outcomes and that parental investments are a potential channel via which the negative impacts of health shocks can be mitigated over the long run.

JEL Codes: I10, I20, J20

[†] University of California, San Diego; University of Southern California and Universidad Adolfo Ibañez; NYU Stern and Princeton University. The authors wish to thank Danuta Rajs from the Departamento de Estadisticas e Informacion de Salud del Ministerio de Salud (MINSAL) and Francisco Lagos of the Ministry of Education (MINEDUC) of the government of Chile for facilitating joint work between government agencies that produced the data used in this study. Finally, the authors also wish to thank Joseph Altonji, Michael Boozer, Ryan Cooper, Julie Cullen, Adrian de la Garza, Adam Kapor, Bhash Mazumder, Karthik Muralidharan, Seth Zimmerman, and seminar participants at Yale, ITAM, Oxford, Berkeley and the Chicago Federal Reserve for their comments and suggestions. The authors have no relevant or financial interests related to this project to disclose. Previous versions of this paper were circulated beginning in July 2010 under the title: *Do initial endowments matter only initially? Birth Weight, Parental Investments and Academic Achievement in School*

1 Introduction

Recent empirical work has shown evidence that initial health endowments are important determinants of later life labor market and cognitive outcomes (Almond and Currie 2011b). However, there is much less evidence on the relationship between initial health endowments and school outcomes, the evolution of this relationship during early childhood, and how investments in human capital adjust in response to these endowments. We contribute to this literature by examining the relationship between health at birth (as measured by birth weight[1]), subsequent parental investments, and academic outcomes from childhood to early adolescence using administrative data covering the entire student population of Chile. This empirical evidence is important as it sheds light on the mechanisms through which initial health affects later life labor market outcomes (Black, Devereux, and Salvanes 2007).

We use administrative data from Chile to link birth records of children born between 1992 and 2002 to their academic records between 2002 and 2012. This panel data set follows cohorts of students from first grade through high school and college entrance exams. In addition to this unique linkage of records, the data allows for the estimation of models with rich heterogeneity as well as models with siblings and twins estimators which have been used in the literature to account for unobserved characteristics affecting both birth weight and the outcome of interest. We supplement this large dataset of birth records and school achievement with data on parental investments recorded at the individual child level, from both parent and child reports. We use this data to examine whether parental investments systematically vary by birth weight and, in particular, whether parents differentially invest

[1]Birth weight is the measure of health at birth used in this paper. The data lacks other health measures such as APGAR scores. Importantly, the data also does not contain information on birth order, which recent work has shown to be important (Choi 2013). Given all the other measures of health at infancy that are correlated with birth weight, we view our study using birth weight as a proxy for health at birth.

within twin pairs.

We find that birth weight significantly affects academic outcomes throughout the schooling years. Our estimates that include twins fixed effects (the standard in this literature for estimating causal impacts) suggest that in first grade a 10% increase in birth weight increases outcomes in math and language scores by 0.04-0.06 standard deviations. We find this result to be stable from first grade through to middle and high school and even for college entrance exams. This implies a persistent effect of birth weight among twins that is seemingly not undone (or exacerbated) by the behavioral responses of parents and teachers. The effect of being born low birth weight (less than 2500 grams) or very low birth weight (less than 1500 grams) is greater, a decrease of around 0.1-0.2 standard deviations, suggesting non-linearities in the birth weight-academic outcomes relationship. To put the magnitude of our results in perspective, consider that recent examples of large-scale interventions in education in developing countries show increases in test scores between 0.17 SD to 0.47 SD (Duflo and Hanna 2005, Muralidharan and Sundararaman 2009, Banerjee, Cole, Duflo, and Linden 2007).

These results contrast with siblings fixed effects and OLS estimators which show a steady decline in the effect of birth weight on test scores. However, the decline is less among siblings who are closer together in age than among siblings who are further apart. Using detailed data on parental investments, we find that education-related investments are negatively correlated with birth weight; i.e. parents invest more via time spent reading, time spent helping out with home work, etc in children with lower birth weight. We find that within twins, however, parental investments are *not* systematically correlated with birth weight, which is the assumption typically made when using twins fixed effects estimators.

We present a model of human capital accumulation and parental investments to rationalize the empirical results described above. This model suggests that over time, depending on parental preferences (whether parents compensate or reinforce initial conditions), test score differences within sibling or twin pairs will converge or diverge over time. To this fairly standard model of academic achievement, we add a dimension of public goods in parental investments within the household to explain the differences we observe when using twins and sibling fixed effects. The main intuitive insight of the model is that if there are public goods within the household with regards to parental investments, then test score differences will converge or diverge *less* over time compared to a case with no public goods in investments. We argue that in the case of twins the role of public goods in investments could be large (if a parent reads to one twin, it is difficult to actively prohibit the other twin from listening in), implying that even if parents wish to invest differentially they are unable to do so. Hence, the model would predict that over time, twins fixed effects estimates diverge or converge less than OLS, and in this way, the twins estimates bring us closer to the causal effect of birth weight over time. We emphasize that the time component is critical to our model and results, as twins fixed effects and OLS differences at any given point in time (in cross sectional data) can be explained by things such as measurement error.

This paper bridges a gap in the literature investigating the lasting role of initial endowments, in particular initial health endowments. By examining repeated educational performance outcomes for children between the ages of 6-18, we are able to provide a more complete picture of how initial health affects human capital accumulation, which in turn is a potential mechanism for explaining adult labor market outcomes. Papers by Black, Devereux, and Salvanes (2007), Torche and Echevarría (2011) and Oreopoulos, Stabile, and Walld (2008) look at long term cognitive outcomes in their analysis of the impact of birth weight using twins and sibling estimators. However, these papers do not have repeated observations on cognitive achievement to study how the health endowment effect evolves over time.

This paper also adds to the literature on parental investments and initial endowments (Aizer and Cunha 2010, Rosenzweig and Zhang 2009, Ashenfelter and Rouse 1998, Adhvaryu and Nyshadham 2012). Like Loughran, Datar, and Kilburn (2004) and others, we use birth weight as a summary measure of initial endowments. We find that parental investments are negatively correlated with birth weight, which viewed through the lens of our model would explain the difference between the sibling fixed effects, OLS, and twins fixed effects estimates. Additionally, this paper partially addresses an important assumption used in many twins based studies. Most twins papers that examine the role of birth weight on long term outcomes have to *assume* that parental investments are not related to individual birth weight. We find that while parents in general invest more in lower birth weight children, they do not differentiate based on birth weight within twins.

A recent related paper by Figlio, Guryan, Karbownik, and Roth (2013) finds similar persistent effects of birth weight on test scores using data on twins from Florida in elementary through middle school years. We view these two papers as jointly providing a more complete picture of the role of early childhood endowments in determining school outcomes. Their paper focuses on understanding whether the birth weight effect varies by socio-economic background and by school quality. Using twins fixed effects, their findings suggest that the effect of initial differences in birth weight is not undone in the long run. While our twins estimates would suggest a similar conclusion, we build on their paper by pushing on the role of parental investments. Our theoretical model and direct data on parental investments in conjunction with a close comparison of OLS, siblings fixed effects, and twins fixed effects estimates suggest that parental investments might have the ability to reduce initial health inequalities among the general population. This finding, in the context of a middle income country like Chile where reinforcement of initial endowments might be the *a priori* hypothesis, is surprising and novel. Arguably, our results can be explained by other unobserved characteristics that drive the differences between OLS, twins and sibling fixed effects models (Section 6.6 of this paper is devoted to thinking about alternative explanations).

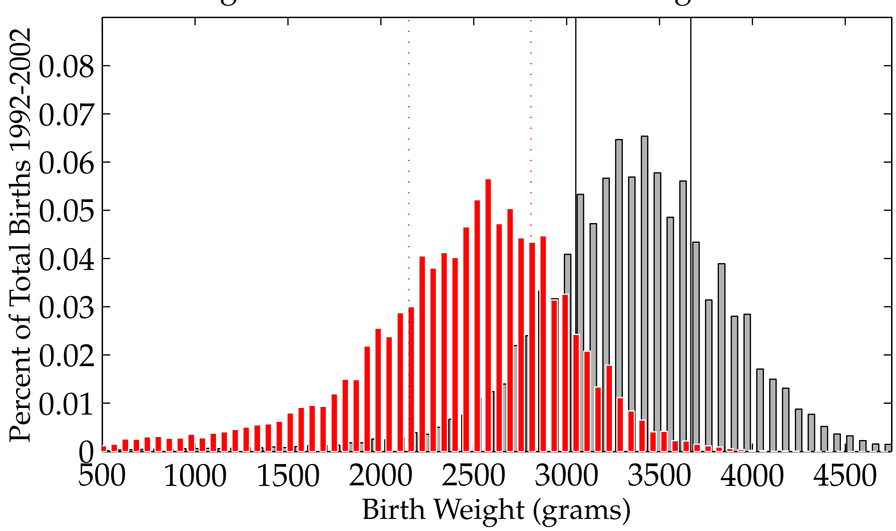
We, therefore, like to think of our results on parental investments as a starting point for thinking about the dynamics of early childhood health and its interaction with investments and intra household resource allocation. This approach is important as it highlights that some of the inequalities at birth can potentially be undone through the efforts made by parents and possibly public policies aimed at investing in the health and human capital of children. Providing a framework and empirical evidence for understanding the differences between OLS and twin/sibling fixed effects estimates is a key contribution of this paper. The rest of this paper is organized as follows: Section 2 provides a brief medical background, Section 3 describes the data used, Sections 4 and 5 present the theoretical and empirical framework respectively, Section 6 discusses the results, and Section 7 concludes.

2 Medical Background

2.1 Birth Weight and Cognitive Development

Medical research suggests a few pathways by which birth weight and the incidence of low birth weight affect cognitive development. Hack, Klein, and Taylor (1995) suggest an association between brain damage and low birth weight leading to poorer performance by low birth weight children on tests. The extent of brain damage and lesions associated with low birth weight can be as severe as resulting in extreme forms of cerebral palsy. Another pathway that is highlighted in Lewis and Bendersky (1989) is that of intraventricular hemorrhage (IVH, or bleeding into the brain's ventricular system). However, IVH is often thought to be due to shorter gestational periods and, therefore, less likely to be the mechanism in the case of twins (Annibale and Hill 2008). Using detailed MRI data from very low birth weight and normal birth weight babies, Abernethy, Palaniappan, and Cooke (2002) suggest that learning disabilities might be related to the growth of certain key brain structures like the caudate nuclei (pertaining to learning and memory) and the hippocampus. Hence, it appears from our reading of a sampling of the medical literature that low birth weight is correlated with developmental problems of the brain, which might lead to lower cognitive ability later in life. Figure 1 shows the distribution of birth weight for the population and for twins.

Figure 1: Distribution of Birth Weight



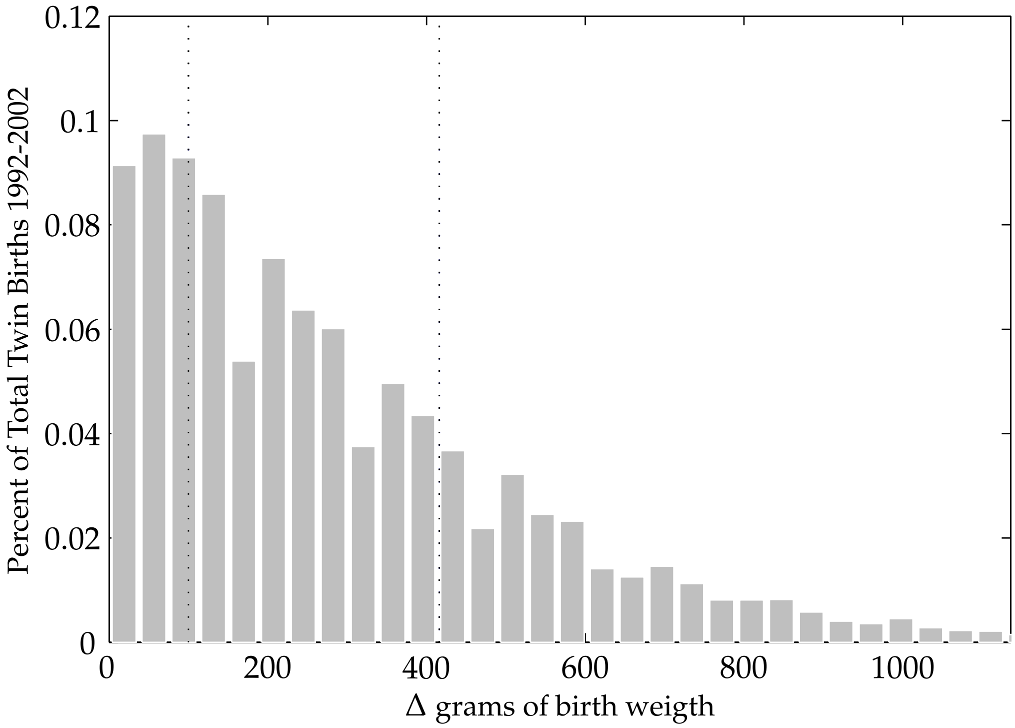
This histogram shows all live births in Chile between 1992 and 2002 (in grey) and also only twin births (in red). The two vertical lines indicate the 25st and 75th percentile of each distribution respectively.

2.2 Why do twins differ in birth weight?

Empirical estimation strategies that use twins fixed effects identify the relationship between birth weight and outcomes using the variation of birth weight between twins. This makes it important to understand why these differences arise. In this section we capitalize on the excellent reviews of the medical literature regarding why differences in birth weight arise within twin pairs provided in Almond, Chay, and Lee (2005) and Black, Devereux, and Salvanes (2007),and we summarize their arguments. Figure 2 shows the density of birth weight differentials within twin pairs in our sample of twins. The average birth weight differential is around 290 grams in our sample, while the median is around 230 grams, and the 25th and 75th percentile are at 100 and 416 grams. The main reason why birth weight differentials arise within twins is due to IUGR (intrauterine growth retardation).[2] The leading reason for differential fetal growth is nutritional intake, and in the case where two placentae are present, nutritional differences can arise due to position in the womb. Among monozygotic twins (which most often share a placenta), the placement of the umbilical chord affects nutritional intake. For details and references on the subject, we refer the reader to footnote 13 in Almond, Chay, and Lee (2005). Figure 2 shows the distribution of birth weight differences within twins for our sample.

[2]The other common reason for low birth weight is gestational age. However, gestational age is identical for twins; hence, the birth weight differentials must arise from fetal growth factors.

Figure 2: Histogram of Birth Weight Differentials among Twins



Note: This histogram shows the distribution of birth weight differentials among twins born in Chile between 1992 and 2002. The mean difference is 294 grams. The dotted lines indicate the 25th and 75th percentile difference.

3 Data

The data used in this paper is largely similar to the data used for the Chile-specific analysis in Bharadwaj, Loken and Neilson (2013). While what follows is a brief summary, we refer the interested reader to the Online Appendix in Bharadwaj, Loken and Neilson (2013) for details on merge rates and attrition across the various data sets used.

3.1 Birth Data

The data on birth weight and background information on parents come from a dataset provided by the Health Ministry of the government of Chile. This dataset includes information on all children born between 1992-2002. It provides data on the sex, birth weight, length, and weeks of gestation as well as demographic information on parents such as the age, education, and occupational status. In addition, the dataset provides a variable describing the type of birth (single or multiple). Twins and siblings are identified by using a mother-specific ID made available for our purposes. Unfortunately, the data does not provide information on zygosity of the twins.

3.2 Education Data

The data on school achievement comes from the SIMCE and RECH database that consists of administrative data on the grades and test scores of every student in the country between 2002 and 2008. This database was provided by the Ministry of Education of Chile (MINEDUC).

3.2.1 RECH

The RECH is the Registro de Estudiantes de Chile (the student registry). This database consists of the grades by subject of each student in a given year and is a census of the entire student population. This database provides the information on the educational results of twins broken up by subjects and allows the construction of the ranking and level measures of academic success at the school/class/grade level. For our purposes, we standardize the grades at the classroom level for each student. While these are classroom grades, we note that performance in the classroom as captured by these grades is highly correlated with performance on national exams such as the SIMCE and PSU.

3.2.2 SIMCE

Chile began to use SIMCE tests and surveys in 1988 as a way of providing information to parents on the quality of schools. This is important in the Chilean context as the education system is compromised of a large private and voucher school system. The tests are administered to all children in a given grade. Between 1988 and 2005, the test alternated between 4th, 8th, and 10th grades. Since 2006, the test is administered to 4th grade every year and alternates between 8th and 10th grade every other year. The total number of children varies between 250,000 and 280,000 across approximately 8000 schools. The response rate to the test is generally over 95%. The SIMCE test covers three main subjects: Mathematics, Science, and Language Arts. The education data sets were subsequently matched to the birth data using individual level identifiers. Since we observe grades for all students who take the test in a given year, we standardize the SIMCE scores at the national level.

3.2.3 PSU

The PSU or *Prueba de Selecion Universitaria* test is the college entrance exam and is the main criteria used in determining admission to the higher education system in Chile. The data included in this study covers both Mathematics and Language. The test is voluntary but required for most forms of financial aid and for the current years includes the majority of graduating seniors. The test is standardized each year. For more information on the PSU and college admissions in Chile see Hastings, Neilson, and Zimmerman (2013).

3.3 Parental Investments Data

The SIMCE test is also accompanied by two surveys, one to parents and one to teachers. The survey to parents include questions about household income and other demographics. The parent survey has a response rate above 80% and is a large endeavor that requires visiting even the most remote schools in the northern and southern regions of the country, and substantial efforts are made to evaluate all schools, both private and public.

The parent survey covers questions about the demographics of the household as well as the parents' opinion of the school and the teacher. In some years the survey covered specific questions regarding parental investments. These years are 2002 and 2007. In 2009, the latest year available, SIMCE surveyed not only the parents but also the students. This allowed students to give their opinions regarding their perceptions of school in many dimensions. One component of the survey asked about the help they received from their parents and how they perceived their parents' role in their education. We use this data in conjunction with the data on parental investments.

The investments (measured in grade 4) are on a scale of 1-5 where 5 denotes "very often" and 1 denotes "never". We aggregate these responses into a dummy variable that takes on the value of 1 if parents report "often" or "very often" and 0 if parents report "never", "not often" or "sometimes". Since there are a wide range of investment questions, we aggregate these into a single index and also perform factor analysis to get summary measures of investments. These factors appear to be easily interpretable (in the parent responses for example) into educational and non-educational inputs. Educational inputs for example include questions like: "How often do you read to your child?"; "Do you help your child with homework?"; etc. On the other hand, non-educational inputs include questions like: "How often do you talk to your child?"; "How often do you write messages for your child?"; "How often do you run errands with your child?" In the case of child responses about parental investments, the factors lump into what we can term as more straightforward educational inputs and "educational encouragement". "Educational encouragement" contains statements such as: "Parent congratulates me on good grades in school"; "Parent challenges me to get better grades"; etc. A detailed list of the investment questions and its correlation with birth weight appears in Table 5.

4 Economic Framework

We build on prior work (Heckman 2007, Almond and Currie 2011a, Conti, Heckman, Yi, and Zhang 2010, Todd and Wolpin 2007) to construct a simple model of human capital formation, taking as its inputs health at birth and parental investments. The key aspect of this model is that it allows for parental responses to health at birth based on an inequality aversion parameter in the parental utility function. The model provides a concise framework for thinking about how health at birth can affect the trajectory of test scores in school, while taking into account how parents might invest differentially across siblings.

An important caveat to the model is that we suppress forces other than parental investments in charting out the evolution of school achievement. For example, inputs by teachers and the history of teacher inputs could be just as important as parental inputs. However, not only do we lack such data on other sources of investments in children, but our model also quickly loses tractability if we were to include say the behaviors of teachers with respect to individual children within the classroom. Hence, while we think our model provides an interesting way to interpret the results, we wish to emphasize that this interpretation is *not* unique; i.e. other ways of rationalizing the data using different sources of investments might be possible. Given the long history of understanding intra household resource allocation and the importance of parental investments in the development of children, we consider our framework a relevant starting point.

4.1 Model of Human Capital Accumulation

Our simple model shows how, in a two child ($i = \{1, 2\}$) household, test scores of twins evolve when a) parents derive utility from how well children do in school (test scores T_{ig}, where g denotes current grade), b) test scores in grade g are a function of parental investments X_{ig} and cognitive endowment θ_{ig}, c) cognitive endowment *evolves* based on prior endowments (θ_{ig-1}) and past parental inputs (X_{ig-1}), and d) parental inputs may have a public good component based on the age difference between siblings (in the case of twins, this difference is zero).

With these preliminaries in mind, we construct the model in two steps. In the first step, we only consider a single time period to illustrate how the main preference parameter in the utility function governs resource allocation across siblings, and then we introduce a dynamic component of how endowments evolve over time. Doing so gives us traction on how the preference parameter in the utility function determines the evolution of the test score *gap* within twins. In the second step, we introduce a public good component to parental investments and show how different levels of public goods affect parental allocations as well as the evolution of test score gaps.

We begin with the single period maximization problem that parents face in the case of twins.[3] There are two main features to note. First, parents derive utility (CES) from the test scores of their children and face an overall constraint on how much they can invest in their

[3]The main simplifying feature here is that we consider both children in the family to be at the same grade at the same time. The implications of the model are the same when children are in different grades but makes the notations needlessly more complicated.

children (T_E). We assume a CES utility function for child test scores because we associate two different behaviors with regards to the elasticity of substitution parameter ρ. ρ in this case governs what Behrman, Pollak, and Taubman (1982) call "inequality aversion". This implies that depending on ρ, parents either behave in ways that allocate more investments to the child with the higher returns, or they are "inequality averse" and invest in the child with lower returns in a bid to lower test score gaps. We study the implications of the model for a broad range of ρ, and thus, we can test which ρ or parental behavior better fits the empirical evidence. Second, test scores are produced (Cobb-Douglas) using current endowments and parental investments. In Appendix A.5, we consider a situation when test score production is also CES and the results are very similar; however, we do not start with CES in test score production so as to obtain closed form solutions on optimal parental investments.[4]

Since this formulation considers only 2 children per household, we denote each child by numbers 1 and 2. Parents maximize:

$$\max_{X_{1g},X_{2g}} \quad U(T_{1g},T_{2g}) = \Big((T_{1g})^{\rho} + (T_{2g})^{\rho}\Big)^{\frac{1}{\rho}} \qquad (1)$$
$$X_{1g} + X_{2g} \leq T_E$$

Next, school achievement is produced using a Cobb Douglas production function. Recall that *current* cognitive endowment (θ_{ig}) and *current* parental investment (X_{ig}) are the

[4] Note that Almond and Currie (2012) use a CES for test score production and a Cobb-Douglas for parental preferences. These choices, however, in our setting are made for simplicity as the main implications of the model go through with more general production functions. See Appendix A.5.

inputs for the current test score. Thus, school achievement is expressed as

$$T_{ig} = \theta_{ig}^{\gamma} X_{ig}^{1-\gamma} \qquad i = \{1, 2\} \tag{2}$$

It is important to note upfront that one of the main simplifying assumptions we make is that parents solve each period's problem in the given period - i.e. parents do not solve this problem dynamically in period zero. We make this assumption to keep the model tractable and to have closed form solutions for the optimal investment patters; however, a fully dynamic model is simulated in Appendix A.5 and yields similar results. Hence, in each period, parents solve the following maximization problem:

$$\begin{aligned} \max_{X_{1g}, X_{2g}} \quad & \left(\theta_{1g}^{\gamma\rho}(X_{1g}^{1-\gamma})^{\rho} + \theta_{2g}^{\gamma\rho}(X_{2g}^{1-\gamma})^{\rho}\right)^{\frac{1}{\rho}} \\ \text{s.t.} \quad & X_{1g} + X_{2g} \leq T_E \end{aligned} \tag{3}$$

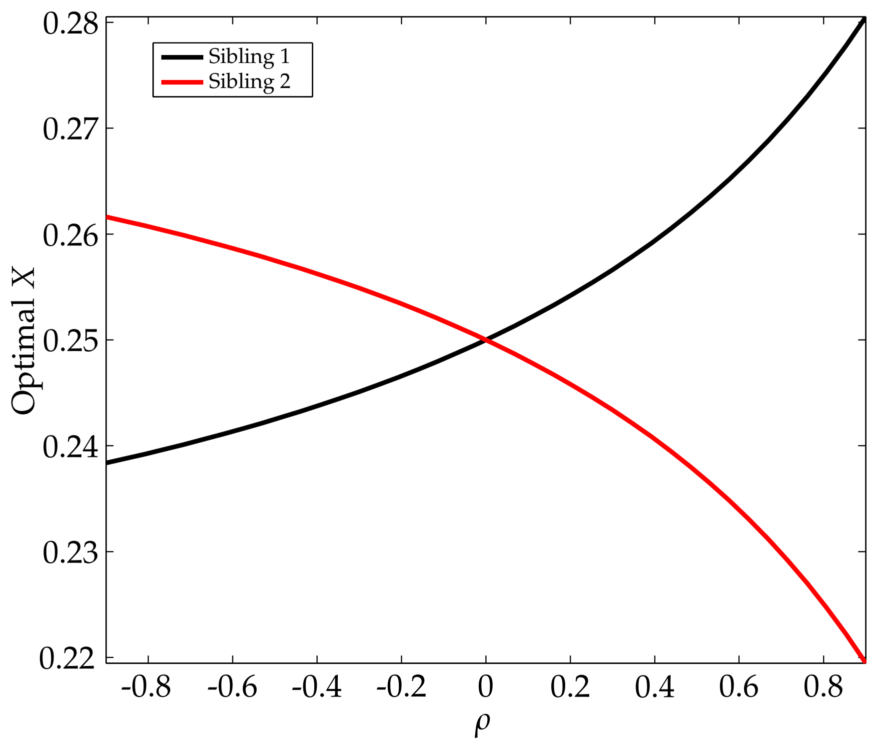
Equation (3) shows that each child's cognitive endowments act as loading factors in the CES utility function. Large positive ρ would suggest that the parents should invest more in the child with better endowments to raise their utility. However, parents may have aversion for inequality captured by a small or negative ρ. When $\rho \rightarrow -\infty$, parents invest in order to equalize test scores across siblings. Hence, ρ is the fundamental parameter governing whether parents invest more in the child with lower endowments or whether they invest more in the child with better endowments.

For any ρ, the optimal allocations are

$$X_{1g} = \frac{T_E}{1+\left(\frac{\theta_{2g}}{\theta_{1g}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}}\left(\frac{\theta_{2g}}{\theta_{1g}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}} = \frac{T_E \cdot \theta_{2g}^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}}{\theta_{1g}^{\frac{\gamma\rho}{(1-\gamma)\rho-1}} + \theta_{2g}^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}} \quad (4)$$

$$X_{2g} = \frac{T_E}{1+\left(\frac{\theta_{2g}}{\theta_{1g}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}} = \frac{T_E \cdot \theta_{1g}^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}}{\theta_{1g}^{\frac{\gamma\rho}{(1-\gamma)\rho-1}} + \theta_{2g}^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}} \quad (5)$$

Figure 3: Optimal Investment Time X



Note: This figure displays optimal allocations, for different values of ρ at a specific point in time t. In this case, we assume that sibling 1 has a higher cognitive endowment at time t.

The optimal parental allocations for each child take into account the endowments of the other child. This is crucial since it shows that there are two ways in which a shock to child i's cognitive endowment affects test scores: through a direct effect of the endowment on test scores and through an indirect "parental resource" allocation effect (this is easily seen by substituting the optimal parental allocation into the production function in equation (2) and taking a derivative with respect to endowment of child *i*). Moreover, it shows that a shock to child 1 affects the test scores of child 2 through just the resource allocation effect.

Thus far, we have not introduced the concept of how endowments at *birth* affect test scores or parental resources across twins. Since our goal is to understand how shocks to the endowment at birth affect test scores in school over time, we introduce the idea that current period cognitive endowments are a function of past endowments and past parental investments. In other words:

$$\theta_{ig} = \beta \theta_{i(g-1)}^{\eta} X_{i(g-1)}^{\zeta} \tag{6}$$

We assume that current cognitive endowment is an increasing function of past parental investment and past cognitive endowment. It is important to note that the choice of β, η, and ζ are not relevant for our results as long as θ_{ig} increases over time when parental investments are positive. Equation (6) is crucial since, along with the optimal allocation equations, it introduces the idea that birth weight (endowment at time 0) can have an impact on test scores in each grade and that its impact could differ based on how parents invest or disinvest.

A fully dynamic version of this model would utilize the fact that parents can solve each stage's problem using backward induction starting at time 0, i.e. at the time of birth. We show simulations for the fully dynamic model in Appendix A.5 as there is no closed form solution to the dynamic problem. Hence, we choose to discuss our main theoretical predictions using this simplification. This is particularly useful when we introduce the public good component.

4.2 Public Good Dimension of Parental Investment

Thus far, we have solved the parents optimal allocation problem assuming that they can completely differentiate the educational input dedicated to each child. In other words, they can potentially invest $X_{1g} \neq X_{2g}$. However, parental investment may have a public good dimension, or spillover effects across siblings.

For instance, parents may read books to both children, or they may simultaneously help them with their homework. The fundamental assumption for our model with public goods is that when siblings are close in age, we expect the degree of spillover to be greater. Therefore, under certain conditions, it can be difficult for parents to invest differentially across children. Twins are an extreme example of this issue, in the sense that they are of the same age, and, if they attend the same school and classroom (85% of twins in our sample are observed in the exact same classroom for example), their homework and other educational needs are probably very similar. For these reasons, we conjecture that it might be difficult for parents to differentially invest across siblings when they are very close together in age.

To formalize public goods in parental investments, we assume that the effective parental investment $\hat{X}_{ig}$ received by child i in grade g corresponds to

$$\hat{X}_{1g} = X_{1g} + \delta(1,2)X_{2g} \tag{7}$$

where X_{1g} is the optimal parental investment, coming from the problem without the public good dimension in parental investments, as expressed in equation (4) and X_{2g} corresponds to the optimal decision in equation (5).

The loading function $\delta(1,2)$ captures the degree of public good dimension of parental investment. If $\delta(1,2)$ is zero, parental investments have no public good dimension, and we return to the original problem. The bigger $\delta(1,2)$ is, the more important the public good dimension is in the provision of parental investment. We assume that the degree of public good dimension depends on the age difference. Thus, $\delta(1,2)$ is larger when the sibling age difference is smaller. For example, $\delta(1,2) = C^{(\text{Siblings' Age Difference})}$, where C is some constant between 0 and 1, would be a candidate loading function.[5]

Note that as far as parents are concerned, a public good dimension in X increases the effective time endowment available for educational activities.

$$\hat{T}_E = \hat{X}_{1g} + \hat{X}_{2g} = (1 + \delta(1,2))(X_{1g} + X_{2g}) = (1 + \delta(1,2))T_E \tag{8}$$

In order to compare the results between an environment with and without public good dimension, we can derive time endowments from a "first stage" where parents decide between educational and non-educational inputs.[6] Under certain conditions specified in the Appendix, we show that the total time allocation for educational inputs *reduces* as the public good dimension in educational investment increases. This is important when comparing across models because we want to isolate the direct effect of public good dimension from the additional effect on the time constraint. In the case of twins, however, the total time allocation component does not matter for our overall results as the public good dimension simply results in equal investments across both twins.

We assume that parents are aware of the public good dimension and solve the follow-

[5] C = 0.71 in the simulation presented here.
[6] The derivation is in the Appendix.

ing within-sibling allocation problem:[7]

$$\max_{X_{1g},X_{2g}} \quad \left(\theta_{1g}^{\gamma\rho}(\hat{X}_{1g}^{1-\gamma})^{\rho} + \theta_{2g}^{\gamma\rho}(\hat{X}_{2g}^{1-\gamma})^{\rho}\right)^{\frac{1}{\rho}} \tag{9}$$
$$\text{s.t.} \quad \hat{X}_{1g} = X_{1g} + \delta(1,2)X_{2g}$$
$$\hat{X}_{2g} = X_{2g} + \delta(1,2)X_{1g}$$
$$\hat{X}_{1g} + \hat{X}_{2g} \leq \hat{T}_E$$

Defining $T_E^{**} = \frac{\hat{T}_E}{1+\delta(1,2)}$, the new optimal allocations are

$$X_{1g} = \frac{T_E^{**}}{(1-\delta(1,2))\left[1+\left(\frac{\theta_{2g}}{\theta_{1g}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right]}\left[\left(\frac{\theta_{2g}}{\theta_{1g}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}} - \delta(1,2)\right] \tag{10}$$

$$X_{2g} = \frac{T_E^{**}}{(1-\delta(1,2))\left[1+\left(\frac{\theta_{2g}}{\theta_{1jg}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right]}\left[1-\delta(1,2)\left(\frac{\theta_{2g}}{\theta_{1g}}\right)^{\frac{\gamma\rho}{(1-\gamma)\rho-1}}\right] \tag{11}$$

For the specific case of twins, where $\delta(1,2) = 1$, optimal effective parental investment is equal across twins. This is because optimal allocations are not defined for $\delta(1,2) = 1$ (i.e. a case with no age gap between siblings, as in the twins case). The problem has infinite solutions for X_{1g} and X_{2g}. However, parents know that any feasible solution in this case implies equal *effective* parental investment among twins. Hence, for simplicity we assume that the solution for twins[8] $X_{1g} = X_{2g}$. In this case, parents may try to differentiate across twins, but the public good dimension of their investment counters any strategic behavior (either mitigating or reinforcing). Consequently, parents simply invest the same amount for each twin.

[7]Parents solving for the effective parental investment or just the parental investment, but knowing the nature of the public good feature leads to the same solution.

[8]This solution also corresponds to the limit of the optimal allocation when δ approaches one.

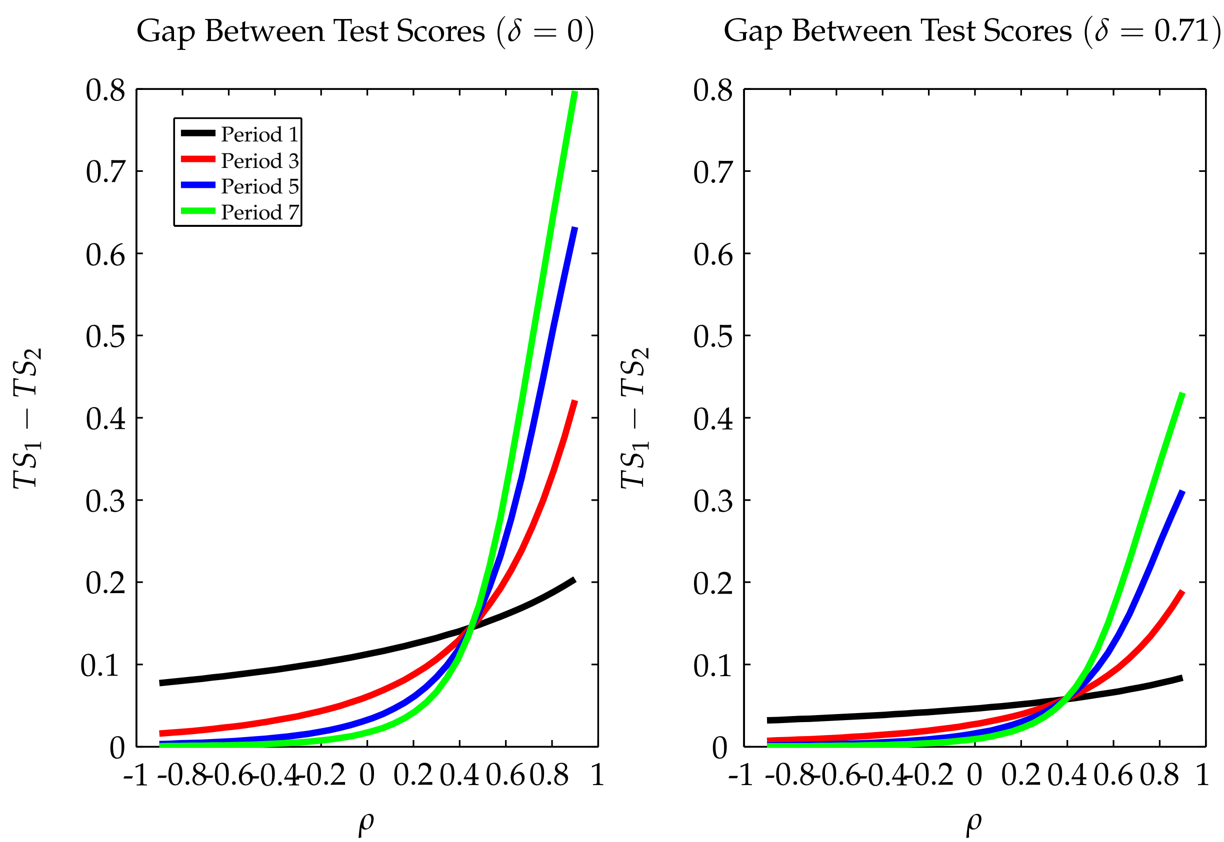
Test scores are a function of initial conditions and the history of educational inputs; hence, in the twins case, initial conditions differ, but educational inputs are the same for both children at any time. The history dependent feature of test scores implies that the relative importance of initial endowment diminishes over time.

Siblings offer additional insight about the underlying strategic behavior of parents. In this case, we are able to deduce the evolution of the test score gap between siblings, because $\delta(1, 2) < 1$. Moreover, using variation across families in age differences, we can test whether the public good dimension decreases with increasing age difference. When doing this, we ignore any strategic decisions by the parents regarding the spacing of children. Hence, we assume that spacing is exogenous in this case (see Aizer and Cunha (200X) for a model where spacing is one of the decisions parents make along with differential investments across children).

Figure 4 displays the evolution of test scores using the structure on optimal parental investments, test score production, and endowment evolution for different parameters values of ρ. We show these dynamics for a model with and without public good dimension.[9] The figure displays the evolution of the gap in test scores (Test score child 1 - Test score child 2), over time, for different values of ρ. The simulations assume that child 1 has higher birth weight compared to child 2.

[9]Other parameters and details of how we create these graphs are presented in the Appendix.

Figure 4: Evolution of Test Scores with Public Good in Parental Investment



Note: This figure displays how differences in achievement change over time under different assumptions about parental preferences ρ. The right panel assumes parental investment is a public good among twins and on the left panel there is no public good aspect to parental investments. It can be seen in this simulation that differences are muted and change less over time in the presence of public goods.

In the left panel we can see that for high values of ρ, the solid black line representing period 1 is below the red line representing period 3, which in turns is below the blue line representing period 5, and so on. This sequence means that the gap is increasing over time. The original gap is positive because child 1 has higher initial cognitive endowment that child 2. This is the graphical representation of the effect of a reinforcing parental behavior on the dynamics of the gap in test scores. We observe the opposite evolution when $\rho < 0$. In this case, the gap diminishes over time as a reaction of the parents' compensating efforts. Note from this graph that the switch from divergence over time to convergence over time in test score gaps does not occur precisely around ρ=0. This means that just observing whether test scores diverge or converge over time is not enough to discern whether parents want to compensate or reinforce. However, combined with knowledge of the correlation between investments and endowments, we can make an informed guess of whether parents compensate or reinforce initial endowments.

The right panel of Figure 4 shows the test score evolution in the presence of some public goods in parental investments ($\delta = 0.71$),[10] Y axis scales are purposely kept the same as in Figure 4 to show how the evolution in differences is muted with a higher δ. Hence, the public good dimension diminishes the effectiveness of parental investments in either compensating or reinforcing initial differences.

The implications of our model for the evolution of test scores over time can be summarized by a fairly intuitive proposition and two corollaries:

Proposition 1 *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 1

Please see Appendix A.

Corollary 1 *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 2

[10] In our calibrations, $\delta = 0.71$ corresponds to an age difference of 1.5 years.

Please see Appendix A.

Corollary 2 *For twins, in the presence of public goods in parental investments, the test score gap will be quite stable over time.*[11]

Proof 3

In this case the actual (effective) parental investment is equal across twins. Over time, the only change in test score gap comes from the evolution of the cognitive endowment. In particular, when $\eta_1 = 1$ and $\beta \cdot (\frac{\hat{T}_E}{2})^{\eta_2} < 1$, the evolution of cognitive endowment will imply a convergence of test scores over time.

5 Empirical estimation

One of the implications of the model is that current test scores are a function of past endowments and the history of parental inputs. This is because the optimal allocations of parental investment are function of current cognitive endowment (equations 4, 5, 10, and 11) . Moreover, given the dynamics of cognitive endowment presented in equation (6), the difference in test scores is a function of initial cognitive endowment and the history of optimal parental investment. This is easily seen by taking logs of equation (2) and iteratively replacing the endowment term (θ_{ig}) with the terms from equation (6) until we derive an expression for test scores in grade g as a function of initial endowments and the history of parental investments. Doing so, we obtain:

[11] If $\eta_1 = 1$, and $\beta \cdot (\frac{\hat{T}_E}{2})^{\eta_2} = 1$ the test score gap will be exactly constant.

$$ln(T_{ig}) = A + \lambda_g ln(\theta_{i0}) + (1-\gamma)ln(X_{ig}) + \gamma\zeta \sum_{k=1}^{g} \eta^{k-1} ln(X_{g-k}) \qquad i = \{1,2\} \tag{12}$$

Where

$$\begin{aligned} A &= \gamma \sum_{k=o}^{g-1} \eta^k ln(\beta) \\ \lambda_g &= \gamma \eta^g \end{aligned}$$

Since θ_0 is captured by birth weight, it is easy to see how not capturing parental investments in each period can create endogeneity problems while estimating the effects of birth weight on test scores. Note that while the only source of endogeneity in this specification is the one due to unobserved parental investments, a second, major source is due to parental/maternal characteristics that could affect both birth weight and test scores. Hence, the strategy of using twins or sibling fixed effects overcomes both sources of bias. An important thing to note is that while we stay true to estimating versions of equation (12), we use *standardized* test scores rather than log of the raw test score as our dependent variable. This is done largely for ease of interpretation and comparability with prior work in education.

Given this, our estimating equation using OLS (adding an error term and rewriting coefficients on parental investments for simplicity) takes the form:

$$T_{ig} = \lambda_g \theta_{i0} + \beta_1 X_{ig} + \beta_2 X_{ig-1} + \ldots + \beta_g X_{i0} + \epsilon_{ig} \tag{13}$$

Where T is the standardized test outcome measured with error ϵ. We estimate OLS for the entire sample graphically, but we also focus on the sample that shares a common support with twins between 700-3000 grams (3000 grams represents the 90th percentile of the twin birth weight distribution). Since twins are significantly smaller than the rest of the population, valid comparisons across twins and singletons for our purposes are only derived by focusing on singletons on the same birth weight support as twins.

5.1 Twins Fixed Effects

Before we write down the twins fixed effects estimator, it is useful to rewrite equation (13) with a new error term that captures all the unobservables:

$$T_{ig} = \lambda_g \theta_{i0} + \underbrace{\beta_1 X_{ig} + \beta_2 X_{ig-1} + \ldots + \beta_g X_{i0} + \epsilon_{ig}}_{u_{ig}} \tag{14}$$

A twins estimator is particularly useful in estimating λ_g from equation (14). As a twins fixed effects estimator essentially differences equation (14) within twins, it would difference out observable and unobservable time invariant family level components (while we have not modeled these variables like parental education explicitly, we believe that they would play a role in the bias that exists in OLS) since these are shared within twin pairs. Returning to the notation where we define siblings as 1 and 2, a twins estimate of equation (14) results in:

$$T_{1g} - T_{2g} = \lambda_g (BW_1 - BW_2) \tag{15}$$

$$+ \underbrace{\beta_1 (X_{1g} - X_{2g}) + \ldots + \beta_g (X_{10} - X_{20}) + \epsilon_{1g} - \epsilon_{2g}}_{u_{1g} - u_{2g}} \tag{16}$$

The model in the previous section would suggest that rather than *assuming* that parental investments are the same within twins, one way to think of why they might *effectively* be the same even when parents wish to invest differentially based on birth weight is due to public goods in the parental investment component. Under the conditions of our model in the previous section, *if* there are perfect spillovers within twins, then the effective parental investment is the same within twins, and equation (16) will result in consistent estimation of λ_g. In what follows, we estimate equation (16) for first through eighth grade for math and language classroom grades; fourth, eighth, and tenth SIMCE test scores in math and language; and for the college entrance exam also for math and language.

We wish to note an important caveat at this point. Twins fixed effects are useful in estimating λ_g only if there are no heterogenous returns to birth weight by parental investment. Empirically, this implies that we cannot have interaction terms between investments and birth weight in equation (13). While the model we presented in section 2 was quite general, the specific empirical application uses stricter functional form assumptions on the production of test scores and the evolution of the endowment. This is, however, essential to keep the empirical component tractable and meaningful, but we are aware that this is indeed a (perhaps drastic) simplification of reality.

5.2 Siblings Fixed Effects

A siblings fixed effects estimator is similar in spirit to the twins fixed effects estimator, the difference being that we expect a "greater" bias if we believe the lesser degree of public goods in parental investment within siblings as per the model in section 2 and proposition 1. For siblings (i and i') who are observed in grade g, we estimate a siblings fixed estimator of the form:

$$T_{ijg} - T_{i'jg} = \lambda_g (BW_{ij} - BW_{i'j}) \tag{17}$$

$$+ \underbrace{\beta_1 (X_{ijg} - X_{i'jg}) + \ldots + \beta_t (X_{ij1} - X_{i'j1}) + \epsilon_{ijt} - \epsilon_{i'jg}}_{u_{ijt} - u_{i'jg}} \tag{18}$$

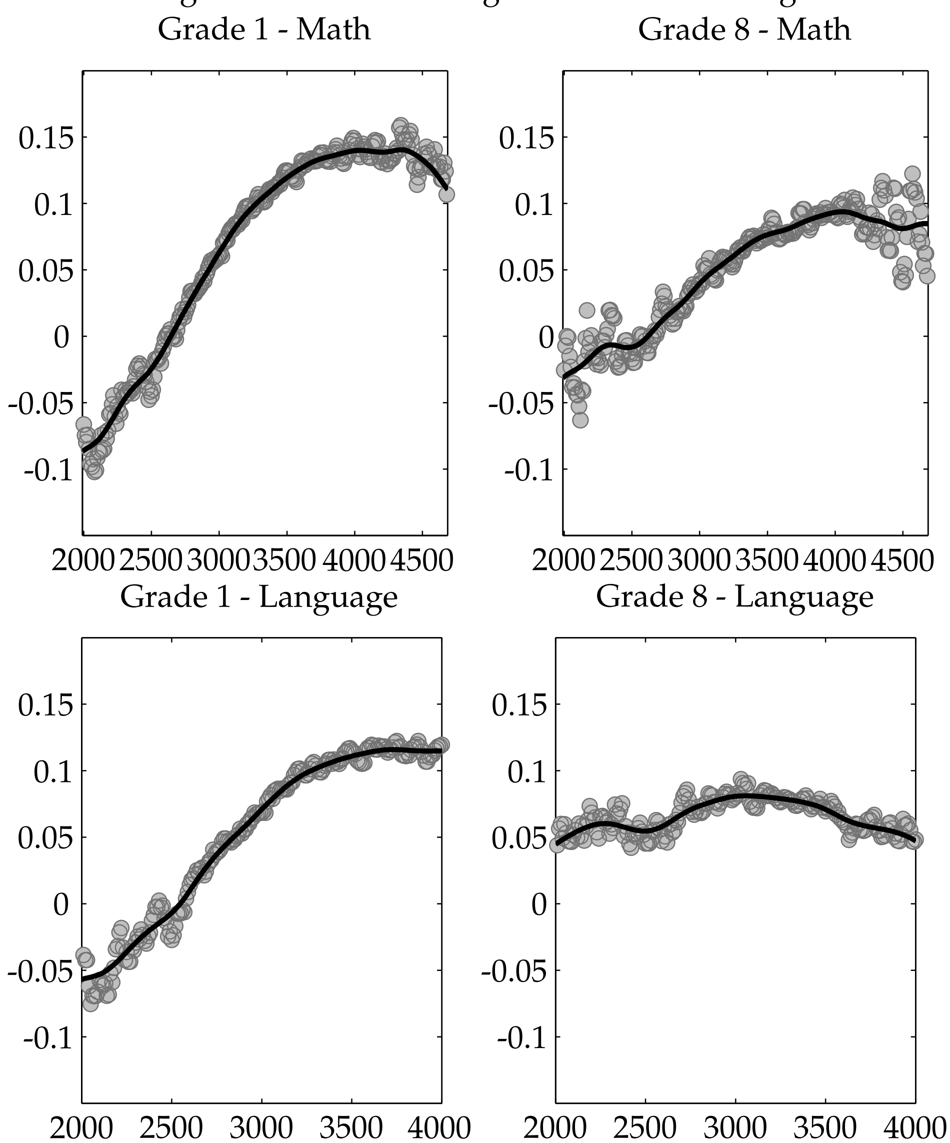
We estimate equation (18) for siblings varying in age difference from 1-5 years. Data limitations do not allow us to estimate this equation for very large age differences.

6 Results

6.1 Nonparametric and OLS Results

Figure 5 shows the relationship between academic achievement in Math and Language and birth weight in first and eighth grade. The relationship between birth weight and both math and language achievement is remarkably linear and upward sloping up until approximately 3300 grs (which is the approximate average birth weight), with higher birth weight babies doing better in both measures.

Figure 5: Standardized grades and Birth Weight



Note: This graph shows the relationship between birth weight and achievement in math (top panel) and language arts (bottom panel) for students born from 1992-2002 in Chile. The grades have been standardized at the classroom level. The black solid line represents a local second order polynomial regression. The dots represent a moving average with a centered window width of 30 grams.

Further exploration of this relationship via regressions confirms that this correlation is robust to the addition of various controls. The regressions estimated in Table 1 show the OLS coefficient for the birth weight effect at each grade on standardized math grades for various samples of the data using a specification similar to that in equation (13), with the exception that we do not have controls for the history of parental inputs. Moreover, since twins are quite different from the rest of the population, we wanted to focus our attention to siblings and singletons with the same birth weight support which is between 0-3000 grams. As is evident from Figure (5), most of the effects of birth weight on the outcome of interest is observed within this support. Row 1 shows λ_g estimated for the sample that shares the same birth weight support as the twins sample. In all OLS specifications, we control for gestational age, mother's education, mother's age at birth, and sex of the child. The second row shows the same specifications but restricting the sample to just the twins sample.

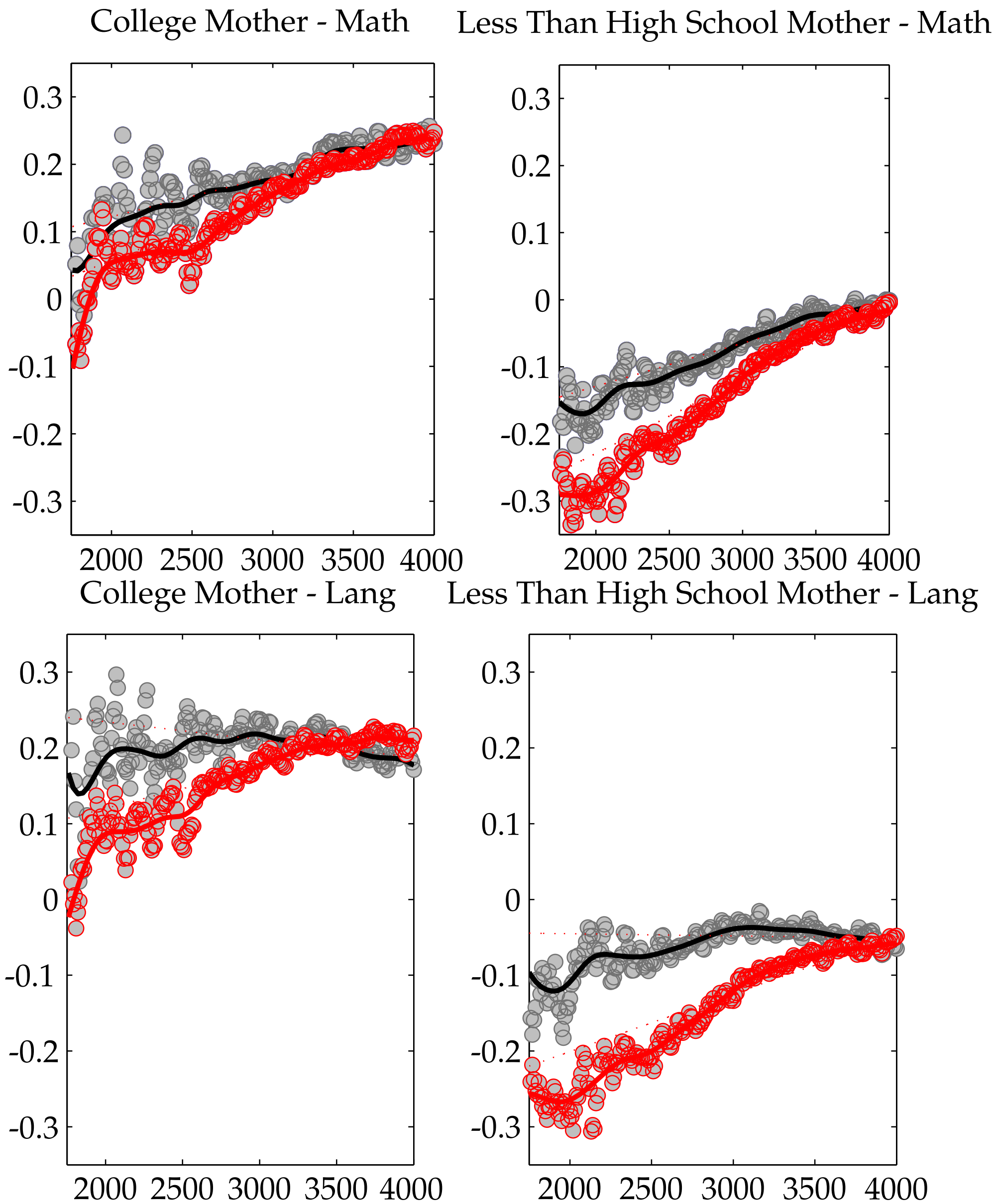
Across all rows, the results appear fairly similar and the main pattern among the coefficients is the decline in the birth weight effect in later grades. In first grade, the effect of birth weight appears to be around 0.35-0.4 SD, and by eighth grade the birth weight effect declines to 0.2 SD. Examining test scores in fourth, eighth, and tenth grades, we find similar results. The OLS regression coefficient declines over time in each case.

6.1.1 Heterogeneity

We also examine whether the OLS relationship between birth weight and standardized math grades varies by observable characteristics of the mother. The following graphs show that students with mothers with college education preform better than those of mothers with lower education levels but that the positive relationship between birth weight and academic achievement is similar in both groups in 1st grade. It can also be seen that, over time, this relationship diminishes in strength for both groups with lower birth weight children raising their relative performance.[12] The results from this section show that the simple correlation between initial health endowment and academic outcomes is quite significant but that this relationship seems to weaken over time.

[12]One reason for this pattern could be that maternal education influences parenting skills in ways that particularly benefit lower birth weight children.

Figure 6: Standardized Math and Language grades and Birth Weight by Mothers Education



Note: This graph shows the relationship between birth weight and achievement in math (top panel) and language arts (bottom panel) for students born from 1992-2002 in Chile to mothers with college education and with less than high school education. The red circles and lines indicate first grade results and the darker colors represent eighth grade achievement. The grades have been standardized at the classroom level. The solid line represents a local second order polynomial regression. The dots represent a moving average with a centered window width of 30 grams.

6.2 Twins Fixed Effects Estimates

To tackle the problem of unobserved characteristics and inputs, we modify equation (13) by including a dummy for the mother - i.e. a twins fixed effect. As suggested earlier, under certain assumptions, a twins estimate does a good job of recovering the true λ_g. Table 2 estimates equation (16) using log birth weight and a dummy variable for low birth weight in separate regressions as the independent variables of interest. In Table 2, statistical tests reveal that λ_8 and λ_1, as obtained under the fixed effects estimation, are not different, suggesting that the twins estimates of the impact of birth weight on test scores do not appear to diminish over time.

Table 2 suggests that a 10% increase in birth weight (corresponding to a 250 gram increase) raises test scores in math by 0.046 SD in 1st grade and that this effect is largely persistent.[13] This is in sharp contrast to the OLS estimates discussed earlier. Table 2 also shows that the impact of being born with low birth weight and very low birth weight is fairly severe on math grades - on average, being low birth weight reduces math scores by 0.1 SD.

6.2.1 Heterogeneity

We can also examine whether twins fixed effects results vary by observable characteristics of the mother. In Table 3 we show that examining twins fixed effects for mothers with high school and above is very similar to the effects obtained for mothers without a high school degree. To interpret this result in the context for our model, we require some assumptions about whether more educated and less educated mothers have different preferences with

[13]The sample size changes across columns since the overlap between birth data and test score data is not perfect for grades 1 through 8. Constraining the sample to cohorts for whom we have 8 years or 7 years of test score data does not change the pattern of results observed in this table. These results are available on request.

regards to inequality aversion across their children. To the extent we think that inequality aversion does not vary across high and low educated mothers, this result is not all together surprising.

The next two rows in Table 3 examine the results by type of school and the socioeconomic background of the children at the school. The SIMCE survey categorizes schools into five SES brackets using household data on the parents of the students that attend each school. We take the two lowest levels and designate them as "Low SES". Twins fixed effects results restricted to this school type shows largely similar results, although the birth weight effect seems to increase slightly over time. The next panel shows the results by private schools in Chile, and while the pattern over time is similar in that the effect remains the same, the levels are quite a bit larger. We interpret these results as evidence that there does appear to be some heterogeneity in the birth weight effect by school type and socioeconomic background.

6.3 Differences in Twins and OLS Estimates: The Role of Parental Investments

Twins fixed effects and OLS estimates contrast in patterns that are worth exploring further. In particular, while all estimation methods show a similar effect in grade 1, twins estimates stay persistent, while OLS estimates steadily decline over time (i.e. the effect of birth weight appears to lessen in later grades). Our model in Section 2 suggests that part of the reason for the differences in twins and OLS estimates is the role of parental investments.

Recall that under OLS, we estimate λ_t with bias:

$$\lambda_g^{OLS} = \lambda_g + Cov(BW_{ij}, \sum_{s=0}^{g-1} \beta_{s+1} X_{ijg-s} + \epsilon_{ijg}) \tag{19}$$

where ϵ_{ijg} is the current shock to T and $\sum_{s=0}^{g-1} \beta_{s+1} X_{ijg-s}$ contains the complete history of unobserved parental inputs (the same X_{ijg}'s from equation (13)). Given that OLS is smaller than twins fixed effects, we can conclude, if twins fixed effects are unbiased, that the direction of bias is negative. The results and the model would imply that parental investments and birth weight are negatively correlated. We can test this correlation in the data. We acknowledge that while we view these correlations as a partial explanation for why the differences in twins and OLS estimations arise, these results are by no means causal; neither do we attempt to get at a causal relationship between the role of parental investments and test scores. We also recognize that OLS and twins fixed effects can vary for a host of reasons, but within the context of our model and the data, the role of parental inputs appears to be the most tractable.

An important aspect of parental investments might be the choice of school that parents send their children to. However, in the case of twins in Chile, nearly 95% of the time both twins attend the same school and grade. Hence, there is simply very little variation in terms of school choice *within* twin pairs. Statistical tests also reveal that within the context of a twin fixed effects regression, birth weight does not matter for choice of school (in this case, the dependent variable examined was whether or not a child attends a private school - these results are available upon request). Since we cannot study school choice within families as a credible source of varying parental investments, we turn to other data that more directly measure parental time investments at the individual child level.

Table 4 estimates the relationship between parental investments (as reported by parents and children in separate columns) and birth weight for a subset of the data (see the data section on why we only have this data for a subset of our overall sample). The investments (measured in grade 4) are on a scale of 1-5 where 5 denotes "very often" and 1 denotes "never". We aggregate these responses into a dummy variable that takes on the value of 1 if parents report "often" or "very often" and 0 if parents report "never", "not often", or "sometimes". Since there are a wide range of investment questions, we aggregate these into a single index and also perform factor analysis to get summary measures of investments. These factors appear to be easily interpretable (in the parent responses, for example) into educational and non-educational inputs. Educational inputs, for example, include questions like: "How often do you read to your child?"; "Do you help your child with homework?"; etc. On the other hand, non-educational inputs include questions like: "How often do you talk to your child?"; "How often do you write messages for your child?"; and "How often do you run errands with your child?". In the case of child responses about parental investments, the factors lump into what we can term as more straightforward educational inputs and "educational encouragement". "Educational encouragement" contains statements such as: "Parent congratulates me on good grades in school"; "Parent challenges me to get better grades"; etc. A detailed list of the investment questions and its correlation with birth weight appears in Table 5.

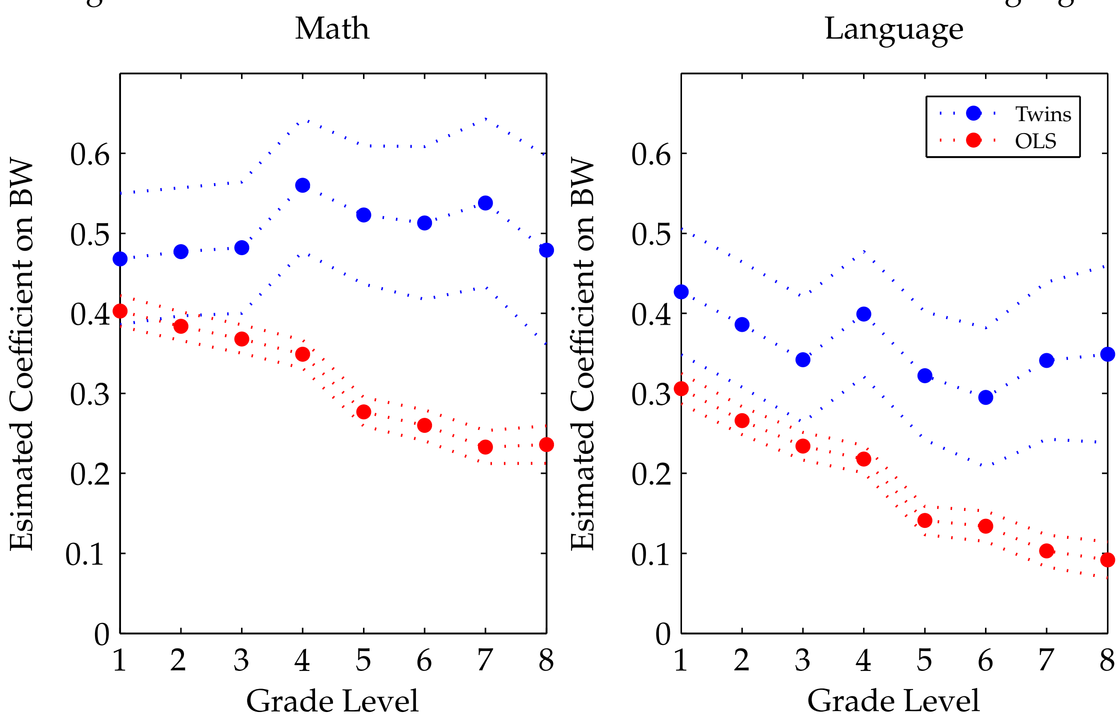
The broad results from Table 4 and 5 are quite obvious: OLS estimates reveal a negative relationship between investments and birth weight. In particular, this appears to be true in the case of educational inputs. What is interesting is that both parent and child responses to the questions reveal similar correlations. This is important as parents might be more likely to misreport how much they invest in their own children. However, in the twins only sample most of the correlations are not statistically significant (except the coefficient on educational investments in the parent reports).

A crucial assumption for interpreting twins fixed effects as revealing the unbiased effect of birth weight on test scores is that parental investments are the same within twins. The model in Section 2 suggested why this might be the case for twins due to public goods and spillovers in investments in households with twins. Given the data on parental investments, we can test within a twins fixed effects framework whether investments vary by birth weight. Table 6 shows that with a twins fixed effect there appears to be no significant correlation between birth weight and parental investments. While this makes it easier to interpret the twins fixed effects results in Table 2, it should be noted that the magnitudes of these correlations, even with the twins fixed effect are comparable to those in Tables 4 and 5, even if the signs are reversed in some cases. In summary, what is clear from Tables 4-6 is that in the overall sample, parental investments appear to be negatively correlated with birth weight, but that, within twins, these relationships are less precisely estimated.

It is important to realize that these parental investments are positively correlated with test scores.[14] While the model might suggest that controlling for parental inputs will make the OLS estimates closer to twins estimates, we do not find this to be true. We believe this is due to the fact that, ultimately, we only observe a small subset of various investments that parents engage in. Moreover, we certainly do not believe that the *entire* difference between OLS and twins is due to parental investments. There could be other biases at play, such as the role of schools or teachers that could mitigate or exacerbate the role of initial endowments. Finally, if parental investments are indeed responding to initial endowments, controlling for endogenous variables complicates the interpretation of the independent variable of interest (birth weight).

[14] Correlations between school performance and parental investments (using the parental responses) suggest that moving from ”Never” to ”Often” in terms of studying with the child, is correlated with an increase in test scores of 0.04 SD. This is perhaps a rather small increase given the importance of parental investments to long term outcomes (Gertler et al. 2013).

Figure 7: OLS and Fixed Effects Estimates for Twins: Math and Language

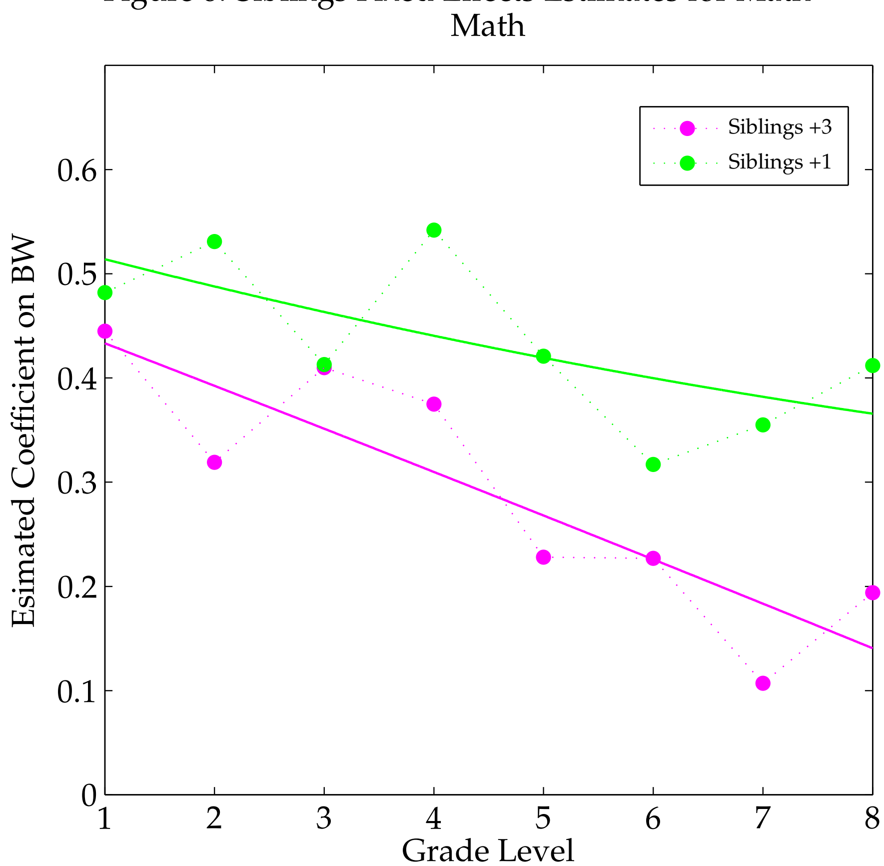


Note: This graph shows how the coefficient on log birth weight changes as children become older using different estimation strategies. These coefficients are from Tables 1, 2, and 8. In the legend *OLS* estimates are carried out using the sample of births under the twins birth weight support (0-3000 grams). In the legend *Twins* indicates estimated carried out using twins fixed effects. 95% confidence intervals are shown for each set of coefficents.

6.4 Siblings Fixed Effects Estimates

Siblings fixed effects in our case are useful to validate the "degree" of public goods argument in Section 2. Proposition 1 suggests that over time, in the presence of public goods, test scores should converge less than without public goods or spillovers. Siblings can provide a validation check on this idea by tracking test scores differences within siblings who are close together in age and siblings far apart in age. Table 7 estimates equation (18) for two types of sibling groups - those who are 1 year apart and those between 3 and 4 years apart.[15] The results across grades suggest that siblings 1 year apart show patterns quite similar to twins, whereas siblings 3-4 years apart show patterns similar to OLS in that the test score differences over time show declines. Siblings fixed effects, while validating our idea of public goods within the household for parental investments, in a more general setting also show the importance of health at birth in determining school performance. Since twins form a small portion of the overall population, it is useful to show that birth weight matters for school achievement in a setting with sibling fixed effects.

Figure 8: Siblings Fixed Effects Estimates for Math



Note: This graph shows how the coefficient on log birth weight changes as children become older using different estimation strategies. These coefficients are from Table 7. We have added a fitted quadratic curve to each set of coefficients.

[15] Note that our sibling fixed effects estimates only use families that have exactly two children. This choice was made to avoid complications that might arise due to birth order or aspects of being a middle child, etc.

6.5 Other School Achievement Variables

While mathematics grades in school is the main subject we have focused on, the data allows us to examine the effects of birth weight on language grades as well as nationalized tests such as SIMCE and the PSU. Table 8 shows our main estimates using OLS, twins and sibling fixed effects strategies for language scores between grades 1 and 8. The patterns for language mirror the patterns seen in math. While twins fixed effects estimates show a stable coefficient across each grade, OLS and larger sibling differences show a steady decline. Estimates for siblings 1 year apart are quite close to the twins estimates.

Table 9 uses the SIMCE and PSU as the main dependent variable. In each case we have examined both math and language scores. The main difference here is that we are able to examine the birth weight effect up to grade 10 and even up to grade 12 (PSU). Hence, we find that the birth weight effect, in the case of twins, appears to last throughout the schooling period. The OLS counterpart in these tables show some decline in the effect, but the decline is less than what is seen using classroom standardized grades. Moreover, we are unable to estimate sibling fixed effects models in the case of SIMCE and PSU given the timing of the tests and the data availability. We view these results as supportive of our overall findings, but ultimately, given that the tests are only administered in 4th, 8th, and 10th grade, we do not view these results as the core of the paper which is focused on understanding the dynamics of the birth weight effect over time.

6.6 Caveats & Alternative Explanations

While we believe this paper pushes the literature focusing on health at infancy and its impact on later life achievement, a few important caveats are worth mentioning. These are not only important for proper interpretation of our paper but can perhaps be considered important avenues for future research to explore. For example, birth weight differentials among twins and siblings are certainly not the result of the same biological process ; as a result, it is possible that what birth weight differentials mean in the case of twins and siblings are quite different. One way to assess this is to examine outcomes soon after birth that are presumably not complicated by the behavioral responses of parents. In analysis conducted by the authors (results available on request), we examine the impact of birth weight using twins and siblings fixed effects framework on 24hr, neonatal, and infant mortality. While sibling and twins fixed effects reveal different effect sizes, the pattern over time of these effects appear the same. Hence, while this paper acknowledges that twins and sibling fixed effects could result in different effect sizes, it is the evolution of these effects over time that is crucial to our model and empirical findings on the role of parental investments.

However, we acknowledge that even the patterns of how birth weight affects school achievement in this paper are not uniquely explained by the behavior of parents and the public good dimension of these investments in the case of twins. As suggested earlier in the paper, teachers and schools might also react to initial endowments resulting in different patterns of how birth weight affects outcomes over time. The idea of interventions in education showing a fade-out effect is also important to consider (Cascio and Staiger 2012). In our context, while the birth weight effect shows a fade out over time in the OLS and sibling fixed effects specifications, we want to highlight that using direct data on parental investments and showing that the patterns vary by sibling age-gap helps support our preferred hypothesis.

Yet another explanation could just be that the birth weight differentials between siblings and twins results in different degrees of complementarities with respect to parental investments. If sibling differences in birth weight are more meaningful in terms of reflecting true endowment differences, then the declining pattern could just be a feature of different returns to parental investment in the case of siblings as opposed to twins. Ultimately, there are many differences between siblings that are not likely different in the case of twins. Spacing for example, is an important dimension that affects siblings but not twins, and importantly, spacing can directly affect parental time investments. Price (2008) notes that this is indeed the case, showing that first born children get more reading related investments compared to their siblings. Given the nature of our investment data, we are unable to test this. However, to the extent that our sibling based results do not change whether or not we control for an indicator variable for the older sibling, we consider our results less susceptible to this concern. In a sense, our results in this paper assume that such aspects have a constant effect over time in that, while considerations like spacing might affect the level difference between twins and sibling fixed effects, they do not affect the evolution of the birth weight effect over time. We view the objective of the paper as showing evidence of an important behavioral input (parental investments) that can result in the pattern of birth weight effects over time ; the paper does not set out to prove that this is the only possible explanation.

7 Conclusion

This paper examined the relationship between health at birth, subsequent parental investments, and academic outcomes from childhood to adolescence using administrative data from Chile, a middle income OECD member country. Using data on all births in the country from 1992 to 2002 merged with schooling records for the entire education system, we construct a panel following children from birth to high school graduation. We find a declining correlation between initial health measured by birth weight and academic outcomes as children progress through school. In contrast, siblings and twins fixed effects estimators show a more persistent relationship between initial health and academic outcomes throughout schooling years. In particular, twins fixed effects models show strikingly persistent effects throughout 1st to 8th grade - a 10% increase in birth weight is associated with nearly 0.05 standard deviations higher performance in math. Similar results are found for national tests taken in fourth, eighth, and tenth grade as well as for the national college entrance exam after high school graduation. In addition, we find evidence that parental investments are larger for children of lower birth weight across families with similar observable characteristics suggesting a compensatory relationship between initial health and investments. We find suggestive evidence that this differential parental investment is decreasing in the age difference among siblings and is virtually absent among twins.

We present a simple model of human capital accumulation and extend existing models of intra household allocations to include a dimension of parental investment spillovers. This model is able to rationalize three empirical features found in the data : 1) the observed behavior of parental investments, 2) the declining correlation between birth weight and academic achievement in the population and 3) the persistent twins estimates. This framework interprets the different empirical results through the lens of a simple human capital accumulation model that implies varying degrees of bias in estimates of the relationship between initial health and later academic outcomes depending on the relationship between parental investments and endowments and how these accumulate over time. Thus, this model rationalizes both the observed behavior of parental investments and the different OLS, siblings and twins estimates of the relationship between initial health and academic achievement in school as well as its evolution over time.

We conclude, within the context of our model, that because parents do not differentially invest among twins, these fixed effects models effectively identify the structural relationship between initial conditions at birth measured by birth weight and later academic outcomes described in the model presented. However, given that the evidence presented shows parental investments are compensatory in this context, twins estimates overestimate the empirical relationship in the general population and suggest that differential parental investments seem to mitigate to some extent initial differences in endowments, and this becomes more relevant over time as parents have more time to adjust. This result helps put prior empirical work using twins estimators into context with regard to the general population. It also highlights that some of the inequalities at birth can potentially be undone through the efforts made by parents and possibly public policies aimed at investing in the health and human capital of children. A deeper understanding of how parents invest and precisely what types of investments matter more would be a fruitful topic for future research in this area.

References

Abernethy, L., M. Palaniappan, and R. Cooke (2002): "Quantitative magnetic resonance imaging of the brain in survivors of very low birth weight," *Archives of disease in childhood*, 87(4), 279.

Adhvaryu, A. R., and A. Nyshadham (2012): "Endowments at Birth and Parents Investments in Children," *Unpublished manuscript. Yale University*.

Aizer, A., and F. Cunha (2010): "Child Endowments, Parental Investments and the Development of Human Capital," *Brown University Working Paper*.

Almond, D., K. Chay, and D. Lee (2005): "The Costs of Low Birth Weight," *The Quarterly Journal of Economics*, 120(3), 1031–1083.

Almond, D., and J. Currie (2011a): *Human Capital Development before Age Five*vol. 4 of *Handbook of Labor Economics*, chap. 15, pp. 1315–1486. Elsevier.

——— (2011b): "Killing me softly: The fetal origins hypothesis," *The Journal of Economic Perspectives*, 25(3), 153–172.

Annibale, D., and J. Hill (2008): "Periventricular Hemorrhage–Intraventricular Hemorrhage," *URL: www. emedicine. com/ped/topic*, 2595.

Ashenfelter, O., and C. Rouse (1998): "Income, Schooling, and Ability: Evidence from A New Sample of Identical Twins," *Quarterly Journal of Economics*, 113(1), 253–284.

Banerjee, A., S. Cole, E. Duflo, and L. Linden (2007): "Remedying Education: Evidence from Two Randomized Experiments in India*," *The Quarterly Journal of Economics*, 122(3), 1235–1264.

BEHRMAN, J. R., R. A. POLLAK, AND P. TAUBMAN (1982): "Parental preferences and provision for progeny," *The Journal of Political Economy*, pp. 52–73.

BLACK, S., P. DEVEREUX, AND K. SALVANES (2007): "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes*," *The Quarterly Journal of Economics*, 122(1), 409–439.

CASCIO, E. U., AND D. O. STAIGER (2012): "Knowledge, tests, and fadeout in educational interventions," Discussion paper, National Bureau of Economic Research.

CONTI, G., J. J. HECKMAN, J. YI, AND J. ZHANG (2010): "Early Health Shocks, Parental Responses, and Child Outcomes," *Unpublished manuscript, Department of Economics, University of Hong Kong*.

DUFLO, E., AND R. HANNA (2005): "Monitoring works: getting teachers to come to school," *NBER Working Paper*.

FIGLIO, D. N., J. GURYAN, K. KARBOWNIK, AND J. ROTH (2013): "The Effects of Poor Neonatal Health on Children's Cognitive Development," Discussion paper, National Bureau of Economic Research.

GERTLER, P., J. HECKMAN, R. PINTO, A. ZANOLINI, C. VERMEERSCH, S. WALKER, S. M. CHANG, AND S. GRANTHAM-MCGREGOR (2013): "Labor market returns to early childhood stimulation: A 20-year followup to an experimental intervention in Jamaica," Discussion paper, National Bureau of Economic Research.

HACK, M., N. KLEIN, AND H. TAYLOR (1995): "Long-term developmental outcomes of low birth weight infants," *The future of children*, 5(1), 176–196.

HASTINGS, J. S., C. A. NEILSON, AND S. D. ZIMMERMAN (2013): "Are Some Degrees

Worth More than Others? Evidence from college admission cutoffs in Chile," Working Paper 19241, National Bureau of Economic Research.

HECKMAN, J. (2007): "The economics, technology, and neuroscience of human capability formation," *Proceedings of the National Academy of Sciences*, 104(33), 13250.

LEWIS, M., AND M. BENDERSKY (1989): "Cognitive and motor differences among low birth weight infants: Impact of intraventricular hemorrhage, medical risk, and social class," *Pediatrics*, 83(2), 187.

LOUGHRAN, D., A. DATAR, AND M. KILBURN (2004): "The Interactive Effect of Birth Weight and Parental Investment on Child Test Scores," *RAND Labor and Population Working Paper WR-168*.

MURALIDHARAN, K., AND V. SUNDARARAMAN (2009): "Teacher performance pay: Experimental evidence from India," *Working Paper*.

OREOPOULOS, P., M. STABILE, AND R. WALLD (2008): "Short, Medium, and Long Term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins," *Journal of Human Resources*, 43, 88–138.

PRICE, J. (2008): "Parent-Child Quality Time Does Birth Order Matter?," *Journal of Human Resources*, 43(1), 240–265.

ROSENZWEIG, M., AND J. ZHANG (2009): "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.

TODD, P., AND K. WOLPIN (2007): "The production of cognitive achievement in children: Home, school, and racial test score gaps," *Journal of Human capital*, 1(1), 91–136.

TORCHE, F., AND G. ECHEVARRÍA (2011): "The effect of birthweight on childhood cognitive development in a middle-income country," *International journal of epidemiology*, 40(4), 1008–1018.

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

$$\max_{T_E, T_{NE}} \quad V(T_E, T_{NE}) \qquad (20)$$
$$\text{s.t.} \quad T_E + T_{NE} \leq T$$

Where V is the utility coming from educational and non educational activities[16]. T is

[16] 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.

the total time allocated to raise the children in the household. Note that, if there are more than one child in the household, parents use the aggregate educational and non educational times and utilities to make the allocation decision.

We denote T_E^* and T_{NE}^* the optimal allocation of time, coming from the solution of the maximization problem in equation (20). 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

$$\begin{aligned} \max_{T_E, T_{NE}} \quad & V(T_E, T_{NE}) \\ \text{s.t.} \quad & T_E + T_{NE} \leq T \end{aligned} \tag{21}$$

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

$$\begin{aligned} \frac{\partial V(T_E^*, T_{NE}^*)}{\partial T_E} &= \lambda \\ \frac{\partial V(T_E^*, T_{NE}^*)}{\partial T_{NE}} &= \lambda \\ \text{combined:} \quad \frac{\partial V(T_E^*, T_{NE}^*)}{\partial T_E} &= \frac{\partial V(T_E^*, T_{NE}^*)}{\partial T_{NE}} \end{aligned} \tag{22}$$

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

$$\begin{aligned} \max_{\hat{T}_E, T_{NE}} \quad & V(\hat{T}_E, T_{NE}) \\ \text{s.t.} \quad & \frac{\hat{T}_E}{1+\delta(i,i')} + T_{NE} \leq T \end{aligned} \tag{23}$$

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

$$\frac{\partial V(T_E^{**}, T_{NE}^{**})}{\partial T_E}(1+\delta(i,i')) = \frac{\partial V(T_E^{**}, T_{NE}^{**})}{\partial T_{NE}} \tag{24}$$

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

$$T_E^* = \frac{a_2}{a_1 + a_2} T \quad \wedge \quad T_{NE}^* = \frac{a_1}{a_1 + a_2} T \tag{25}$$

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

$$T_E^{**} = \frac{a_2}{a_1(1 + \delta(i, i')) + a_2} T \quad \wedge \quad T_{NE}^* = \frac{a_1(1 + \delta(i, i'))}{a_1 + a_2} T \tag{26}$$

Comparing the allocation assigned to educational activities in (25) with the one displayed in (26), 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, equations (4) and (5) 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) 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[17]. 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 investment received by the children is a weighted average of the optimal parental investment expressed in equations (4) y (5). 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.

[17]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.

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 equations (4) and (5). We simulate the solutions with the following parameters:

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

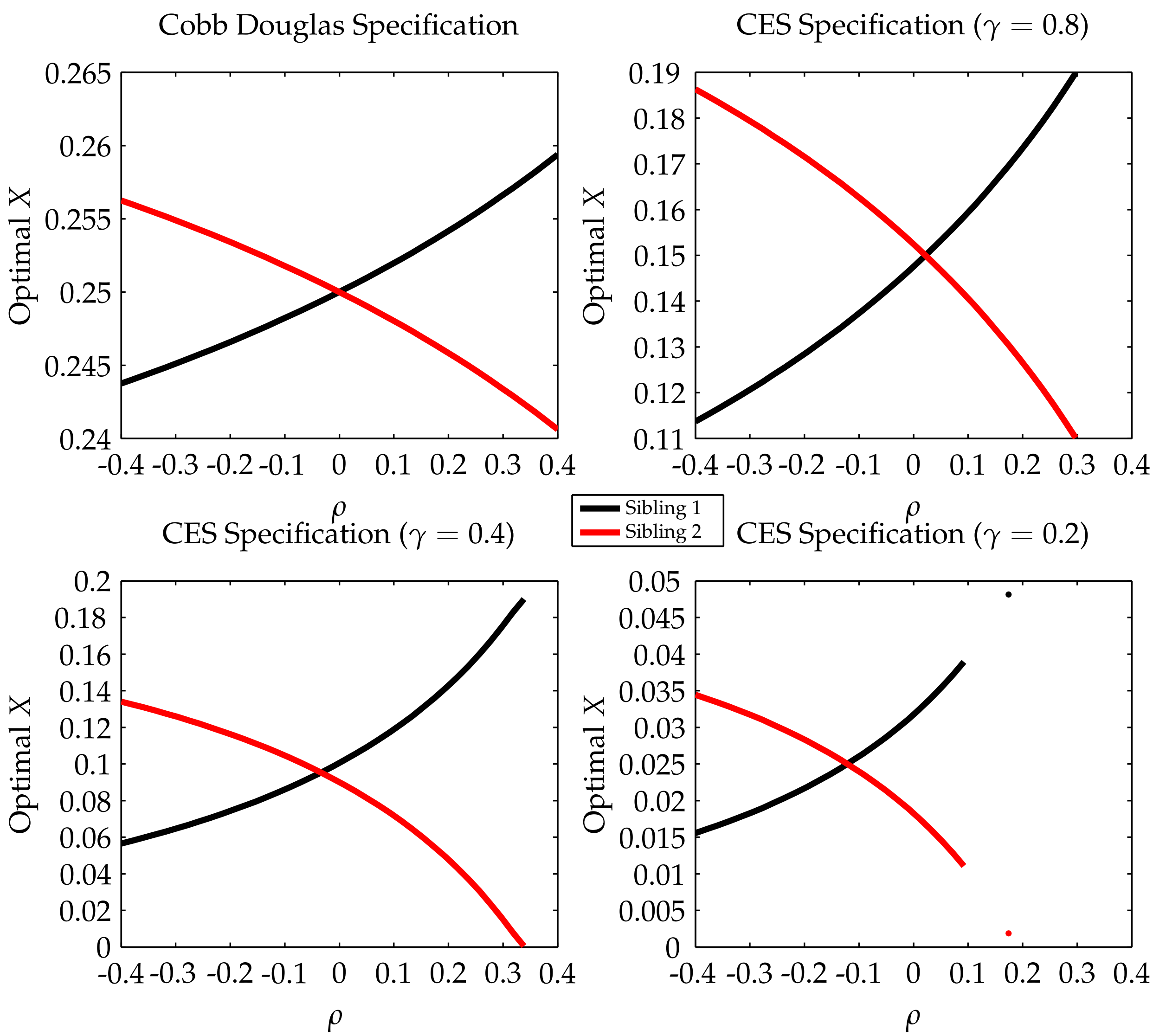
We assume that test scores take as input current cognitive endowment and current parental investment. In order to obtain a closed form solution, we use a Cobb-Douglas production function. However, a more general production function can be used.

Figure 10 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, that 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 9: CES Case



Note:

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

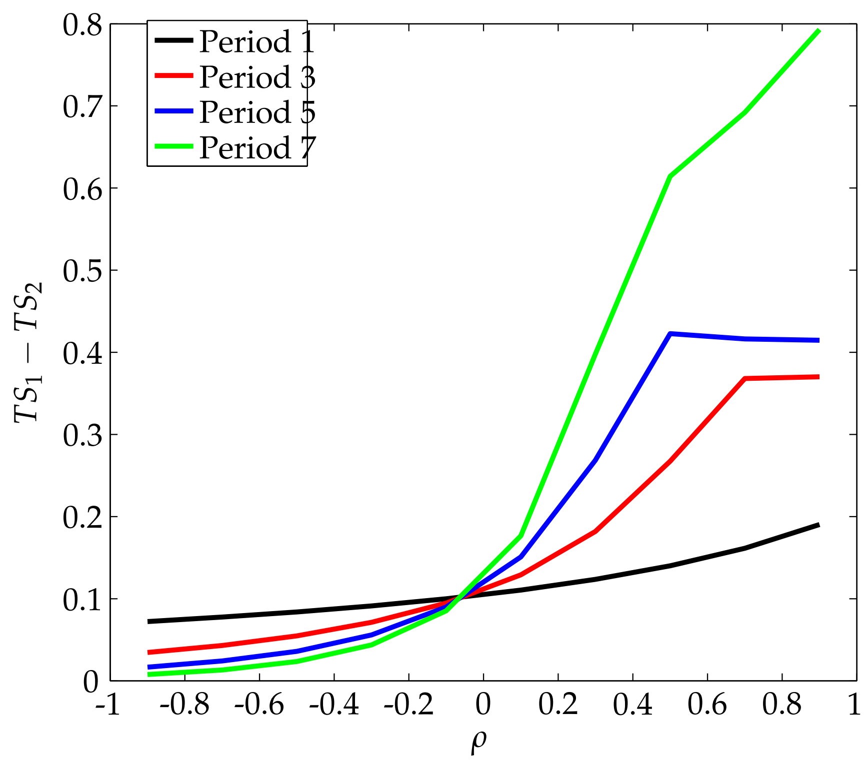
$$
\begin{aligned}
V_g(\theta_{1g}, \theta_{2g}) = \max_{X_{1g}, X_{2g}} \quad & \left(\theta_{1g}^{\gamma\rho}(X_{1g}^{1-\gamma})^{\rho} + \theta_{2g}^{\gamma\rho}(X_{2g}^{1-\gamma})^{\rho}\right)^{\frac{1}{\rho}} + V_{g+1}((\theta_{1(g+1)}, \theta_{2(g+1)}) \quad (27) \\
\text{s.t.} \quad & X_{1g} + X_{2g} \leq T_E \\
& \theta_{i(g+1)} = \beta\theta_{ig}^{\eta} X_{ig}^{\zeta} \qquad i \in \{1, 2\}
\end{aligned}
$$

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 10 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.

Figure 10: Gap Between Test Scores ($\delta = 0$)



Note:

TABLE 1: Birth Weight and Performance in Math - OLS Estimates

	Grade							
Standardized Math Scores	1	2	3	4	5	6	7	8
OLS: Sample uses same birth weight support as twins (0-3000 grams)								
Log Birth Weight	0.398	0.379	0.355	0.337	0.269	0.256	0.229	0.226
	(0.0106)***	(0.00990)***	(0.00986)***	(0.00967)***	(0.00994)***	(0.0106)***	(0.0113)***	(0.0128)***
Observations	371,259	421,804	443,531	451,065	425,070	374,552	329,077	272,753
OLS: Twins sample								
Log Birth Weight	0.357	0.298	0.329	0.333	0.277	0.282	0.244	0.202
	(0.0322)***	(0.0308)***	(0.0335)***	(0.0321)***	(0.0322)***	(0.0352)***	(0.0380)***	(0.0465)***
Observations	30,353	31,586	31,212	30,849	28,478	24,919	21,755	17,874

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: All estimates control for gestational age, mother's age and education and sex of the child. The dependent variable is standarized classroom grades in math.

TABLE 2: Birth Weight and Performance in Math - Twins Fixed Effect Estimates

	Grade							
Standardized Math Scores	1	2	3	4	5	6	7	8
Log Birth Weight	0.468	0.477	0.482	0.560	0.523	0.513	0.538	0.479
	(0.0410)***	(0.0408)***	(0.0410)***	(0.0415)***	(0.0432)***	(0.0477)***	(0.0524)***	(0.0590)***
Low Birth Weight (<2500 grams)	-0.0777	-0.0815	-0.0861	-0.104	-0.109	-0.0902	-0.108	-0.103
	(0.0134)***	(0.0133)***	(0.0134)***	(0.0136)***	(0.0140)***	(0.0154)***	(0.0169)***	(0.0189)***
Very Low Birth Weight (<1500 grams)	-0.162	-0.190	-0.230	-0.182	-0.190	-0.238	-0.297	-0.276
	(0.0397)***	(0.0407)***	(0.0432)***	(0.0440)***	(0.0461)***	(0.0512)***	(0.0591)***	(0.0668)***
Observations	30,353	31,586	31,212	30,849	28,478	24,919	21,755	17,874
Number of Twin Pairs	15,740	16,496	16,350	16,187	14,961	13,160	11,572	9,564

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: All estimates control for sex of the child.

TABLE 3: Birth Weight and Performance in Math - Heterogeneity: Twins Estimates

	Grade							
Standardized Math Scores	1	2	3	4	5	6	7	8
All coefficients reported are on log birth weight using twin fixed effects								
Mother with high school and above	0.476	0.520	0.536	0.613	0.541	0.514	0.563	0.465
	(0.0477)***	(0.0473)***	(0.0476)***	(0.0484)***	(0.0503)***	(0.0555)***	(0.0616)***	(0.0697)***
Mothers with less than high school	0.436	0.339	0.302	0.397	0.456	0.497	0.478	0.517
	(0.0809)***	(0.0812)***	(0.0820)***	(0.0815)***	(0.0854)***	(0.0935)***	(0.101)***	(0.112)***
Mother Employed	0.482	0.604	0.572	0.531	0.472	0.454	0.555	0.343
	(0.0784)***	(0.0802)***	(0.0805)***	(0.0837)***	(0.0899)***	(0.0976)***	(0.108)***	(0.123)***
Mother Unemployed	0.459	0.421	0.445	0.569	0.539	0.532	0.533	0.523
	(0.0477)***	(0.0470)***	(0.0475)***	(0.0476)***	(0.0491)***	(0.0546)***	(0.0599)***	(0.0672)***
Santiago	0.486	0.513	0.443	0.544	0.505	0.549	0.497	0.404
	(0.0643)***	(0.0639)***	(0.0630)***	(0.0644)***	(0.0671)***	(0.0740)***	(0.0835)***	(0.0941)***
Non-Santiago	0.454	0.450	0.514	0.572	0.535	0.484	0.565	0.531
	(0.0531)***	(0.0529)***	(0.0540)***	(0.0541)***	(0.0566)***	(0.0624)***	(0.0672)***	(0.0756)***
Private schools	0.319	0.804	0.813	0.748	0.751	0.743	0.790	0.713
	(0.191)*	(0.194)***	(0.182)***	(0.179)***	(0.187)***	(0.195)***	(0.205)***	(0.254)***
Poor Schools	0.432	0.329	0.339	0.504	0.465	0.483	0.515	0.506
	(0.0562)***	(0.0573)***	(0.0590)***	(0.0599)***	(0.0634)***	(0.0721)***	(0.0790)***	(0.0878)***

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: All estimates control for sex of the child. School categories are based on a 2010 categorization of schools in Chile. Hence, a school's classification as of 2010 is assumed to be the same between 2002-2008.

Table 4: Parental Investments and Birth Weight - OLS Estimates

	Parent report of Investments				Child's report of parental investments		
	Standardized Investment 2002	Standardized Investment 2007	PCA: Non-Educational Investments	PCA: Educational Investments	Standardized Investment 2009	PCA: Educational Investments	PCA: Educational Encouragement
OLS: Full Sample							
Log Birth Weight	-0.0128	-0.0588	0.0240	-0.100	-0.0460	-0.0367	0.00766
	(0.0146)	(0.0165)***	(0.0147)	(0.0145)***	(0.0105)***	(0.0119)***	(0.0119)
Observations	192,833	169,234	193,017	193,017	377,853	295,137	295,137
OLS: Sample uses same birth weight support as twins							
Log Birth Weight	-0.00989	-0.0813	0.0703	-0.121	-0.0736	-0.0507	0.0146
	(0.0276)	(0.0347)**	(0.0277)**	(0.0275)***	(0.0210)***	(0.0298)*	(0.0313)
Observations	58,806	48,010	60,027	60,027	105,893	48,635	48,635
OLS: Twins Sample							
Log Birth Weight	-0.180	-0.0936	-0.0799	-0.309	0.0338	0.00455	-0.0784
	(0.117)	(0.101)	(0.119)	(0.116)***	(0.0848)	(0.132)	(0.131)
Observations	2,833	2,617	2,900	2,900	2,583	2,583	2,583

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: All regressions control for gestational age, mother's age and education and sex of the child. "Standardized" investments use all investment related questions to create a single composite measure. "PCA" denotes measures obtained from Pricipal Components Analysis. Details of this procedure are available upon request. "PCA" components for parental responses are computed over their responses to the 2002 survey, and child responses are obly available from 2009. All investment meausres are asked of children in grade 4.

Table 5: Parental Investments and Birth Weight - OLS Estimates Details

Common support sample							
Details on Investments (Parent Responses)	Review Homework	Help with Homework	Study with Child	Read to Child	Give math problems	Talk to Child	Run errands with child
Log Birth Weight	-0.0348	-0.0520	-0.0450	0.00463	-0.00913	0.00674	-0.0151
	(0.0129)***	(0.0144)***	(0.0148)***	(0.0110)	(0.0149)	(0.00999)	(0.0141)
Observations	45,106	45,106	45,106	45,106	45,106	45,106	45,106
Mean of dependent variable	0.777	0.679	0.634	0.322	0.643	0.882	0.708
Details on Investments (Child Responses)	Parent explains things	Parent helps study	Parent helps with chores	Parent knows grades in school	Parent congratulates me on good performance	Parent challenges me to get good grades	Parent willing to help
Log Birth Weight	-0.0376	-0.0447	-0.0344	0.00767	-0.00572	-0.0180	0.00518
	(0.0112)***	(0.0117)***	(0.0119)***	(0.0106)	(0.00915)	(0.0120)	(0.0116)
Observations	79,839	79,762	78,676	78,759	68,489	73,551	78,486
Mean of dependent variable	0.555	0.500	0.484	0.752	0.835	0.408	0.618
Twins sample							
Details on Investments (Parent Responses)	Review Homework	Help with Homework	Study with Child	Read to Child	Give math problems	Talk to Child	Run errands with child
Log Birth Weight	-0.104	-0.162	-0.110	-0.0834	0.0242	0.0112	-0.0125
	(0.0518)**	(0.0561)***	(0.0573)*	(0.0305)***	(0.0575)	(0.0380)	(0.0538)
Observations	2,900	2,900	2,900	2,900	2,900	2,900	2,900
Mean of dependent variable	0.744	0.665	0.633	0.367	0.637	0.885	0.719
Details on Investments (Child Responses)	Parent explains things	Parent helps study	Parent helps with chores	Parent knows grades in school	Parent congratulates me on good performance	Parent challenges me to get good grades	Parent willing to help
Log Birth Weight	0.0173	0.00570	-0.0222	-0.0386	0.0243	0.00229	0.0356
	(0.0420)	(0.0433)	(0.0437)	(0.0376)	(0.0360)	(0.0438)	(0.0426)
Observations	5,652	5,641	5,548	5,583	4,857	5,206	5,540
Mean of dependent variable	0.543	0.486	0.467	0.737	0.824	0.405	0.615

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: All regressions control for gestational age, mother's age and education and sex of the child.

Table 6: Parental Investments and Birth Weight - Fixed Effects Estimates

	Parent report of Investments				Child's report of parental investments		
Overall measures	Investment 2002	Investment 2007	PCA: Non-Educational Investments	PCA: Educational Investments	Standardized Investment	PCA: Educational Investments	PCA: Educational Encouragement
Log Birth Weight	0.109	0.120	0.105	-0.0186	-0.0397	0.0998	0.299
	(0.0835)	(0.0907)	(0.0882)	(0.101)	(0.146)	(0.238)	(0.263)
Observations	2,833	2,617	2,900	2,900	5,701	2,583	2,583
Details on Investments (Parent Responses)	Review Homework	Help with Homework	Study with Child	Read to Child	Give math problems	Talk to Child	Run errands with child
Log Birth Weight	0.0502	-0.0699	0.0382	0.0249	0.0499	-0.00482	0.0449
	(0.0466)	(0.0490)	(0.0495)	(0.0270)	(0.0488)	(0.0355)	(0.0430)
Observations	2,900	2,900	2,900	8,541	2,900	2,900	2,900
Details on Investments (Child Responses)	Parent explains things	Parent helps study	Parent helps with chores	Parent knows grades in school	Parent congratulates me on good performance	Parent challenges me to get good grades	Parent willing to help
Log Birth Weight	0.0622	-0.0795	-0.0144	-0.107	0.0285	-0.0604	0.0893
	(0.0764)	(0.0785)	(0.0850)	(0.0727)	(0.0713)	(0.0847)	(0.0812)
Observations	5,652	5,641	5,548	5,583	4,857	5,206	5,540
Mean of dependent var	0.543	0.486	0.467	0.737	0.824	0.405	0.615

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: all regressions control for sex of the child.

TABLE 7: Birth Weight and Performance in Math - Sibling Fixed Effect Estimates

	Grade							
Standardized Math scores	1	2	3	4	5	6	7	8
Siblings 1 year apart								
Log Birth Weight	0.487	0.546	0.428	0.567	0.439	0.346	0.373	0.444
	(0.122)***	(0.102)***	(0.101)***	(0.0967)***	(0.102)***	(0.111)***	(0.125)***	(0.138)***
Observations	2383	2659	2796	3052	2967	2607	2265	1775
Siblings 3-4 years apart								
Log Birth Weight	0.447	0.322	0.418	0.380	0.234	0.235	0.118	0.203
	(0.0748)***	(0.0720)***	(0.0720)***	(0.0707)***	(0.0728)***	(0.0856)***	(0.0984)	(0.140)
Observations	6434	7062	7215	7388	6647	5293	3989	2494

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Sample uses siblings on common birth weight support as twins (0-3000 grams). All regressions control for gestational age, mother's age and education, an indicator for the older sibling and sex of the child.

TABLE 8: Birth Weight and Performance in Language

	Grade							
Standardized Language Scores	1	2	3	4	5	6	7	8
Twins FE	0.427	0.386	0.342	0.399	0.322	0.295	0.341	0.349
	(0.0394)***	(0.0392)***	(0.0391)***	(0.0391)***	(0.0400)***	(0.0435)***	(0.0491)***	(0.0553)***
OLS (Twins Sample)	0.278	0.204	0.229	0.186	0.141	0.0918	0.0906	0.0569
	(0.0316)***	(0.0315)***	(0.0306)***	(0.0326)***	(0.0321)***	(0.0353)***	(0.0378)**	(0.0478)
OLS (Birth weight support of twins)	0.296	0.258	0.220	0.204	0.131	0.129	0.0954	0.0825
	(0.0105)***	(0.00962)***	(0.00954)***	(0.00947)***	(0.00961)***	(0.0105)***	(0.0110)***	(0.0123)***
Siblings 1 year apart (FE)	0.232	0.501	0.220	0.369	0.254	0.0644	0.176	0.316
	(0.120)*	(0.0991)***	(0.0969)**	(0.0946)***	(0.0970)***	(0.106)	(0.119)	(0.135)**
Siblings 3-4 years apart (FE)	0.368	0.335	0.291	0.205	0.180	0.165	0.0732	-0.105
	(0.0730)***	(0.0690)***	(0.0697)***	(0.0691)***	(0.0713)**	(0.0839)**	(0.0954)	(0.139)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: All estimates control for sex of the child. OLS and Sibling estimates contain other controls, see notes under Table 1 & 7.

Table 9: SIMCE and PSU test scores

	SIMCE			College Entrance (PSU)
	4th grade	8th Grade	10th Grade	
All estimates are the coefficient on log birth weight				
Math				
Twins FE	0.601	0.578	0.432	0.465
	(0.0503)***	(0.0975)***	(0.102)***	(0.109)***
OLS (Twins Sample)	0.308	0.306	0.178	0.329
	(0.0291)***	(0.0598)***	(0.0634)***	(0.0770)***
Observations	22790	6180	5416	5052
Language				
Twins FE	0.397	0.338	0.327	0.281
	(0.0531)***	(0.101)***	(0.112)***	(0.109)***
OLS (Twins Sample)	0.115	0.102	0.121	0.142
	(0.0292)***	(0.0607)*	(0.0662)*	(0.0763)*
Observations	22,790	6,180	5,416	5,052

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: All estimates control for sex of the child. OLS estimates contain other controls, see notes under Table 1.