THE
ECONOMIC
JOURNAL

The Economic Journal, **125** (*November*), F347–F371. Doi: 10.1111/ecoj.12291 © 2015 Royal Economic Society. Published by John Wiley & Sons, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

EARLY HEALTH SHOCKS, INTRA-HOUSEHOLD RESOURCE ALLOCATION AND CHILD OUTCOMES*

Junjian Yi, James J. Heckman, Junsen Zhang and Gabriella Conti

In response to health shocks, parents make compensatory and reinforcing investments in different dimensions of human capital across children. Using household data on Chinese child twins whose average age is 11, we find that, compared with the twin sibling who did not suffer from negative early health shocks at age 0–3, the other twin sibling who did received 305 yuan more health investment, but received 182 yuan less educational investment. Overall, the family acts as a net equaliser in response to child early health shocks across children.

The importance of the family in fostering child human capital is well understood by economists. How parents invest in children with different endowments is not well-studied and there is no consensus in the literature. Becker and Tomes (1976) and Tomes (1981) suggest that parental investments reinforce initial endowments and that such behaviour increases inequality. Griliches (1979) conjectures that parental human capital investments compensate for gaps in children's endowments and that the family is an equalising agent. A number of empirical studies find evidence for reinforcing behaviour (Behrman *et al.*, 1994; Rosenzweig and Zhang, 2009). Other studies find empirical support for compensating behaviour (Behrman *et al.*, 1982; Pitt *et al.*, 1990).

This article studies how early health shocks to children affect intra-household resource allocation and the human capital formation of children. For two reasons, early health shocks in developing countries are likely to affect children's human capital accumulation and long-run outcomes negatively (Strauss and Thomas, 2007; Currie and Vogl, 2013). First, young children are especially vulnerable to health shocks in developing countries. For example, more than 10% of children suffer from diarrhoea in many developing countries such as Bangladesh (Strauss and Thomas, 1998). Second, in the absence of public health insurance and in the presence of pervasive poverty, a child affected by a negative health shock may not receive appropriate medical treatment. Consequently, early health shocks may have long-lasting consequences. In the absence of a well-functioning public education system, the consequences of an early health shock may be exacerbated and thus impair human capital formation (Glewwe and Miguel, 2007).

* Corresponding author: Junsen Zhang, Department of Economics, Chinese University of Hong Kong, Hong Kong. Email: jszhang@cuhk.edu.hk.

The research was supported in part by the American Bar Foundation, the Pritzker Children's Initiative, the Buffett Early Childhood Fund, NIH grants NICHD R37HD065072 and R01HD54702, an anonymous funder, a European Research Council grant (DEVHEALTH 269874), a grant for the FIS from the Chinese University of Hong Kong and a grant from the Institute for New Economic Thinking (INET) to the Human Capital and Economic Opportunity Global Working Group (HCEO)– an initiative of the Becker Friedman Institute for Research in Economics (BFI). The views expressed in this article are those of the authors and not necessarily those of the funders or persons named here.

This article formulates and estimates a model with two channels through which early health shocks affect child human capital formation. The first is a biological channel operating directly through the production function for human capital. The second is an intra-household resource allocation effect arising from parental responses to the shock.

Parental responses to early health shocks on children may be more important in developing countries with weaker health infrastructure and less well-established credit markets and social protection systems. The absence of an old-age pension system and the presence of tight credit constraints may drive parents to base their intra-household resource allocation decisions on efficiency rather than on equity concerns. In this case, parents are more likely to reinforce the harmful effects of an early health shock by devoting less resources to the less-endowed child. The role of the family must be considered when designing public policies to remedy the effects of inequality at birth or the early childhood stage.

Following recent developments in the economics of human capital (Cunha and Heckman, 2007; Heckman, 2007; Cunha *et al.*, 2010), our theoretical analysis extends the conventional literature on intra-household resource allocation in two ways. First, we allow for multidimensionality in human capital. The conventional literature assumes a single dimension of human capital on which parents can compensate or reinforce. Borghans *et al.* (2008), Cunha *et al.* (2010) and Almlund *et al.* (2011) extend the analysis of human capital from a single dimension to multiple dimensions and emphasise the cross-productivity of different types of human capital. These studies focus on the human capital production process of individuals and do not explore the implications of multidimensionality of human capital for the intra-household resource allocation across children.

Second, we examine the interaction of parental preferences and the human capital production function in the intra-household resource allocation process. The conventional literature focuses on parental preferences and has made special assumptions about the role of human capital production in the intra-household resource allocation process. For example, Becker and Tomes (1976) and Pitt *et al.* (1990) assume a linear production function with respect to child endowments, whereas Behrman *et al.* (1982, 1986) assume a Cobb–Douglas technology. Under both specifications, parental investment strategies are determined only by parental inequality aversion across children.

Our study combines the traditional literature on intra-household resource allocation with recent developments in the economics of human capital. We demonstrate that parents make reinforcing and compensatory investments in different dimensions of human capital in children in response to health shocks. Whether parents exhibit a reinforcing, compensating or neutral investment strategy is ultimately an empirical question.

We shed light on these mechanisms by estimating the effect of child early health shocks on family investments. We use data from the Chinese Child Twins Survey (CCTS), which contains detailed information on family health and educational investments separately for each child whose average age is 11. Our data are described in greater detail below. To the best of our knowledge, CCTS is the first census-type household survey on child twins around the world. We find evidence of compensating investments in child health and reinforcing investments in education in response to early health shocks for one twin child in the family. Our empirical results show that, compared with the twin sibling who did not suffer from negative early health shocks at ages 0–3, the other twin sibling who did received 305 yuan more health investments in the 12 months prior to the survey. This amount is substantial and is more than one third of the average family investment in child education per year or about one half of maternal monthly wage. Offsetting this in a different dimension, on average, the sick child received 182 yuan less on educational investments than its twin sibling.

We estimate the child human capital production function. Holding constant family investments, the estimated coefficient on early health shocks in the production function reflects only a biological effect. We separate the biological effect from the intra-household resource allocation effect.

Because family investments are chosen, and unobservables in choice and outcome equations are likely to be correlated, we correct for spurious correlation bias using 2SLS with price and non-labour income as instruments. Our estimates show that family investments have positive productivity effects and that early health insults negatively affect child outcomes, including health, education and socio emotional skills.

We also estimate a human capital production function excluding family investments. These reduced-form estimates correspond to the total effects of early health shocks on child human capital. From them we can infer the importance of the intra-household resource allocation effect by comparing the reduced-form estimates with the estimates from the structural production function. We find that reduced-form estimates understate the biological effect by one-half for anthropometric measures, such as body mass index (BMI), but overstate the biological effect by one-third for a number of educational outcome measures.

Our findings have important policy implications. When parents simultaneously compensate in health investments and reinforce in educational investments across their children, the effect of family investments on inequality in the society becomes more nuanced. A multidimensional perspective on inequality is warranted. The overall level of inequality is overestimated if one focuses solely on inequality in education because inequality in health is reduced by family investments. Our estimates suggest that the Chinese family appears to be a net equaliser in terms of the aggregate investment response in value terms child human capital investment when both education and health are combined. Parental responses should be considered when designing intervention policies to remedy disadvantages among children because parents can exacerbate or eliminate these effects by reallocating resources within the family.

Our results shed light on recent literature on the effect of early-life conditions on late-life outcomes (Case *et al.*, 2005; Almond and Currie, 2011). Although the literature has achieved a consensus on the negative effects of early-life health insults on both short-run (Currie *et al.*, 2010) and long-run outcomes (Smith, 2009), the role played by parental behaviour remains unclear. The reduced-form estimates of the effect of early-life shocks that disregard intra-household responses do not necessarily represent a biological effect. When parents make compensating and reinforcing investments along different dimensions of human capital, the reduced-form estimates cannot be unambiguously interpreted as upper or lower bounds of biological effects.

Our empirical analysis suggests that reduced-form estimates of early health shocks on child health understate the biological effect, whereas those for child education overstate the biological effect. These results imply that caution must be taken in interpreting reduced-form estimates as biological effects and confirms the importance of considering parental behavioural responses when studying the consequences of early-life health insults.

The remainder of the article is organised as follows. We derive our theoretical model in Section 1. In Section 2, we describe the CCTS data used to test our theoretical prediction. Our econometric specification is presented in Section 3. We present the estimation results in Section 4 and conclude in Section 5.

1. The Conceptual Framework

This Section analyses child early health shocks, intra-household resource allocation and child human capital formation. We show that an early health shock can affect child human capital through two channels: a direct channel (the biological effect through the production of human capital) and an indirect one (the intra-household resource allocation effect through parental responses). By introducing the multidimensionality of child human capital, we show that parents can compensate and reinforce along different dimensions of a child's human capital with respect to an early health shock to one child.

1.1. *The Model*

We assume that each family has two children ($\imath = i, j$).[1] Each child has two components of human capital: health (H) and a bundle of cognitive and socio-emotional skills (C). We treat the latter as an aggregate in this subsection. Child prenatal endowment such as birthweight is $\omega^k_{\imath,\tau}$; parental human capital investments is $I^k_{\imath,\tau}$, and child human capital is $\theta^k_{\imath,\tau}$, where $k = H, C$. We further denote a child's characteristics such as gender and ethnicity by $\xi_{\imath,\tau}$, and parental characteristics by h_τ. The human capital production function of type k for child i in family τ is specified as follows:

$$\theta^k_{i,\tau} = f^k(\omega^H_{i,\tau}, \omega^C_{i,\tau}, I^k_{i,\tau}, e^H_{i,\tau}; \xi_{i,\tau}, h_\tau), \tag{1}$$

where $e^H_{i,\tau}$ is defined as a postnatal negative health shock affecting child i at early stages, and $\partial\theta^k_{i,\tau}/\partial e^H_{i,\tau} < 0$. Child human capital is determined by the child's endowments, human capital investment and health shocks. Parental and individual characteristics also affect the formation of child human capital. The production technology of health differs from that of cognitive skills. The production function is the same for all children in family τ, but may differ from one family to another.[2]

In (1), we assume that the health shock of child j does not directly enter child i's human capital production function, although e^H_j can indirectly affect child i's human capital through parental investment. Thus, we assume away a contagious effect of early

[1] It is straightforward to extend the model to a general case with n children in the family.
[2] We suppress the family subscript τ on the technology to simplify notation.

health shocks in the model, but we provide evidence for this assumption in online Appendix A.[3] Second, we focus on health shocks; our data show that the most prevalent health shock is diarrhoea.

Parents are assumed to value child outcomes. They also care about their own consumption and leisure. Parental preferences are represented by the utility function:

$$U = U(c, l, q_i, q_j), \tag{2}$$

where c is parental consumption, l is parental leisure time and q_ι is the quality of child ι. Denoting T as the parental labour supply and normalising the parental time endowment to one, we have $l + T = 1$.[4] Child quality is a combination of health and cognitive skills such that $q_\iota = q(\theta_\iota^H, \theta_\iota^C)$. Both children have the same quality function. However, they may have different qualities because the endowments, early health shocks and human capital investments could be different. The budget constraint is specified as follows:

$$p_I\left(\sum_\iota \sum_k I_\iota^k\right) + c + wl = Y + w, \tag{3}$$

where p_I is the price of human capital investment and is assumed to be independent of the type of investment; w and Y are parents' wage rate multiplied by time available (assumed to equal 1) and non-labour income. The price of parental consumption is normalised to one. We further assume that parents provide all the resources for their children's human capital.[5]

1.2. *Child Early Health Shock and Parental Responses*

We now analyse how parents adjust intra-household resource allocation in response to an early health shock on their children. The parents' problem is to maximise the utility function (2) subject to the budget constraint (3) and the production technology (1). If the utility function and the production function are strictly concave and continuously twice-differentiable, then the existence, uniqueness and continuity of the solution to the intra-household resource allocation problem directly follow. Deleting family subscripts, we denote the optimal human capital investment of type k in child i as a function of the following form:

$$I_i^{k*} = \psi^k(\omega_i^H, \omega_i^C, \omega_j^H, \omega_j^C, e_i^H, e_j^H, \xi_i, \xi_j, h, p_I, w, Y). \tag{4}$$

A reinforcement strategy for investment of type k arises if $\partial I_i^{k*}/\partial e_i^H \leq 0$ and $\partial I_i^{k*}/\partial e_j^H \geq 0$, i.e., parents put less investment of type k in the child who has suffered from an early health shock and place more investment in his or her twin sibling who has not suffered. In this case, $(\partial I_i^{k*}/\partial e_i^H) - (\partial I_i^{k*}/\partial e_j^H) \leq 0$. In contrast, if

[3] In online Appendix A, we show that early health shocks measured in our empirical analysis are non-infectious diarrhoea.

[4] Economists have stressed the importance of parental tutoring time as a determinant of child human capital production. We ignore this argument in our analysis because the parental tutoring time in our sample is small, and we find little effect of early health shocks on parental tutoring time.

[5] This assumption is approximately valid in the context of developing countries where public education and medical insurance are absent, which is discussed in later Sections.

$\partial I_i^{k*}/\partial e_i^H \geq 0$ and $\partial I_i^{k*}/\partial e_j^H \leq 0$, we say that parents use a compensatory strategy in investment of type k in children. Thus, $(\partial I_i^{k*}/\partial e_i^H) - (\partial I_i^{k*}/\partial e_j^H) \geq 0$.

The definition of reinforcing or compensating strategies does not presuppose any specific model of intra-household resource allocation. The parameters determining which type of strategy parents adopt depends on the specific model employed.

For example, Behrman *et al.* (1982) assume a Cobb–Douglas production function and a constant elasticity of substitution (CES) parental utility function. They show that the optimal investment strategy is uniquely determined by parental preference parameters. Almond and Currie (2011) assume a CES production function and a Cobb–Douglas parental utility function. In their case, the optimal investment strategy is uniquely determined by production technology parameters.

In a more general model with multidimensional human capital, the child human capital investment strategy reflects not only parental preferences but also the production technology available to them. Parents can compensate and reinforce along different dimensions of human capital with respect to early health shocks.

Family investment can exacerbate inequality in one dimension but, at the same time, ameliorate inequality in another dimension. Our analysis does not impose particular functional forms. Online Appendix B discusses the consequences of specific choices of functional forms further.

1.3. *Early Health Shocks, Parental Responses and Child Human Capital*

Early health shocks affect child human capital through two different channels: a biological effect and a behavioural effect. From (1), the total effect of an early health shock on child i on the child's human capital k can be decomposed as follows:

$$\underbrace{\frac{\mathrm{d}\theta_i^k}{\mathrm{d}e_i^H}}_{A} = \underbrace{\frac{\partial\theta_i^k}{\partial e_i^H}}_{B} + \underbrace{\frac{\partial\theta_i^k}{\partial I_i^k}}_{C} \times \underbrace{\frac{\partial I_i^k}{\partial e_i^H}}_{D}. \tag{5}$$

The term on the left-hand side (A) is the total effect of an early health shock, which corresponds to the reduced-form estimate in the literature. The first term on the right-hand side (B) is a biological effect that directly operates through the production function and is assumed to be negative. The second term ($C \times D$) is a behavioural effect that operates through parental responses in adjusting family investment. The behavioural effect is the product of the productivity effect of the investment (C) and the intra-household resource allocation effect (D). In general, the effect of an early health shock is not the same as the biological effect. We assume that the productivity effect of family investment is positive, an assumption tested below. The sign of the behavioural effect is determined by the intra-household resource allocation effect, which cannot be determined *a priori*, and can vary across different dimensions of human capital.

The reduced-form estimate (A) is interpreted as an upper (lower) bound of the biological effect if we know that the parents adopt a reinforcing (compensatory) strategy (which is characterised by the sign of D). We can separate the biological effect (B) from the behavioural effect by estimating (1) controlling for family investment. By comparing the estimate of the biological effect (B) with the reduced-form estimate of the total effect (A), we can quantify the importance of the behavioural response in family investments to account for the total effect of early health shocks on child human capital.

1.4. *Early Health Shocks, Parental Labour Supply and Consumption*

Early health shocks to children could result in external effects on other family members. In this article, we study the effect of early health shocks on parental labour supply and consumption. The optimal parental labour supply (T^*) and parental consumption (c^*) is a vector function of the following form:

$$y = y(\omega_i^H, \omega_i^C, \omega_j^H, \omega_j^C, e_i^H, e_j^H, \xi_i, \xi_j, h, p_I, w, Y), \tag{6}$$

where $y = (T^*, c^*)$. Adopting specific functional forms in the example presented in online Appendix B, we show that a child's early health shock decreases parental consumption and increases parental labour supply.

2. Data

2.1. *The Chinese Child Twins Survey* (*CCTS*)

The CCTS was conducted by the Urban Survey Unit (USU) of the National Bureau of Statistics in late 2002 and early 2003 in Kunming, China. Kunming, which is the capital city of Yunnan Province, has a total population of approximately 5 million. Yunnan is a relatively underdeveloped province located in the far south-western corner of China.

To the best of our knowledge, CCTS is the first census-type household survey on twin children. The survey includes almost all households with twins aged between 6 and 18 years living in Kunming in 2002. The average age of the twin children is 11. The households have been initially identified by the USU on the basis of the 2000 population census according to whether the children have the same birth year and month and whether they have the same relationship with the household head. The addresses of these households are then obtained from the census office, and the presence of twins is verified with a visit to the household. Starting from 2,300 pairs of potential twins identified in the census, 1,694 households with twins are successfully interviewed. The survey covers an extensive range of information about family investment in each child separately and child outcomes, in addition to a wide range of demographic, social and economic information at the household level. See Rosenzweig and Zhang (2009) for a detailed description of the CCTS.

2.2. *Descriptive Statistics*

We now describe the variables in our data that are the empirical counterparts of those in the theoretical analysis. Table 1 tabulates the descriptive statistics. Early health shocks (e^H) are defined by a dummy variable that indicates whether a child suffered from a serious disease from ages 0 to 3. The complete list of diseases include serious

Table 1

Descriptive Statistics

Theoretical variables (1)	Empirical counterparts (2)	Mean (3)	SD (4)	Within-twin SD/overall SD (5)
e^H	Early health shocks (dummy)	0.09	0.28	0.35
I^H	Health investment (¥/year)	225.83	772.00	0.62
I^C	Educational investment (¥/year)	910.44	1,225.73	0.20
θ^H	Height (cm)	137.33	19.30	0.13
	Weight (kg)	33.60	11.82	0.16
	BMI	17.39	3.22	0.27
	General health status*	2.92	0.63	0.34
θ^C	Literature (score: 1–100)	81.93	13.79	0.42
	Literature (relative measure)†	3.53	0.84	0.48
	Mathematics (score: 1–100)	80.90	16.41	0.43
	Mathematics (relative measure)†	3.48	0.92	0.50
	Good student awards (dummy)	0.24	0.43	0.51
	Awards in contests (dummy)	0.07	0.25	0.49
	Grade repetition (dummy)	0.04	0.21	0.52
	Doing minor actions in class‡	1.73	0.76	0.49
	Always feel lonely§	1.20	0.50	0.34
	Easily distracted§	1.59	0.68	0.34
	Easily frightened§	1.37	0.60	0.28
	Emotionally unstable§	1.12	0.35	0.28
T	Paternal labour supply (days/month)	25.51	4.68	
	Maternal labour supply (days/month)	25.32	4.72	
c	Paternal consumption (¥/six months)	700.33	869.38	
	Maternal consumption (¥/six months)	283.81	736.59	
ω	Birthweight (kg)	2.46	0.47	0.36
ξ	Male	0.49	0.50	0.50
	Age	11.19	3.09	
	Born at the first parity	0.79	0.40	
h	Maternal age	36.85	4.84	
	Maternal ethnicity (Han = 1)	0.86	0.35	
	Maternal schooling years	8.65	3.28	
w, p_I	Maternal working sector (Public = 1)	0.08	0.27	
	Rural	0.53	0.50	
Y	Household asset (score)	−0.05	1.73	

Notes. ¥ stands for Chinese Yuan. The sample includes 1,456 pairs of twins. Column (5) presents the share of the within-twin standard deviation out of the overall sample standard deviation. *4-point Likert scale: 1 (worst) to 4 (best). †1 (top quintile in the class) to 5 (lowest quintile). ‡4-point Likert scale: 1 (never) to 4 (always). §3-point Likert scale: 1 (disagree), 2 (agree), 3 (strongly agree).

diarrhoea, calcium deficiency, asthma, fracture, attention deficit disorder, heart disease, serious hearing difficulties, whooping cough, stammer and serious eyesight problems. The first three constitute 92% of the serious diseases suffered by children, and such a number is consistent with the evidence for children in developing countries (Strauss and Thomas, 1998).[6] Table 1 shows that the prevalence rate of early health

[6] Unfortunately, we cannot distinguish between mental and physical diseases because the former has low prevalence in our sample.

shocks in our sample is 9%. We address potential concerns with various types of measurement errors in constructing the variable of child early health shocks in online Appendix C.

Our main dependent variables are measures of the family investments in children in the 12 months prior to the survey (I^k). The investments are separately recorded for each child. Health investments include money spent on medical treatments and on the purchase of medicine or health products.[7] Educational investments include school tuition and money spent on purchasing books and stationery, hiring home tutors and attending tutoring classes. We note that the medical and educational investments on children constitute a substantial fraction of the family income. Educational investments on one child alone amount to 912 yuan per year out of a *per capita* family income of 3,030 yuan per year.

The CCTS has rich information measuring child human capital. As measures of child health $\left(\theta^H\right)$, we use anthropometric indicators (i.e. height, weight and BMI) and general health status, which are all reported by both parents. We use child academic and schooling performance to measure child educational outcomes. As measures of academic performance (components of θ^C), we use both objective (examination transcripts) and subjective (self-reported evaluations in comparison with the class norm) measures in two different subjects: literature and mathematics. They are compulsory courses from primary school to high school (from age 6 to 18). We also analyse several outcomes related to school performance, which are recorded from transcripts. They include good student awards, awards in contests, grade repetition and whether the child often does naughty actions in class as reported by teachers.

Our data are also rich in terms of socio-emotional measures, which are categorical and reported by both parents. We have four measures: always feels lonely, easily distracted, easily frightened and emotionally unstable. These variables are derived from the Strengths and Difficulties Questionnaire (SDQ) used by CCTS. The SDQ is one of the most commonly used instruments for screening child psychiatric morbidities. See Du *et al.* (2008) for the discussion on the reliability and validity on Asian children of SDQ.

We also analyse the effect of early health shocks on parental labour supply (T) and consumption (c). Parental labour supply is measured as days worked per month. Parental consumption is measured as the total expenditure on cigarettes, alcohol, clothes and cosmetics in the past six months prior to the survey. Consumption is separately recorded for mothers and fathers.

Birthweight is used to proxy child pre-birth endowments (ω). This variable is used to measure both health and cognitive endowments. We have three variables for measuring child characteristics (ξ): age, gender and birth order.[8] The variables of maternal age, ethnicity and schooling years are used to proxy parental characteristics (h).

[7] Grossman (2000) also measures medical care by personal medical expenditures on doctors, dentists, hospital care, prescribed and non-prescribed drugs, non-medical practitioners and medical appliances.

[8] The one-child policy is strictly implemented in urban areas in Kunming. However, households in rural areas are encouraged to have one child, but are exempted from the strict one-child policy, although they are allowed to have two children at most (Family Planning Commission of Yunnan Province, 2003).

We use both the rural *hukou* and maternal working sector to proxy wages (w) and medical and educational prices (p_I). The rural indicator is commonly used to proxy for wages in studies in developing countries because of the large wage gap between rural and urban areas (Johnson, 1947; Rosenzweig, 1980). We also use *hukou* to proxy the price of health services. At the time of the survey, the medical insurance system was almost completely absent in rural areas, whereas medical expenditures on children could be partly reimbursed for urban residents. Although public education is not free in both urban and rural areas at the time of the survey, the quality of education in urban areas is higher than that in rural areas. Note that Chinese residents find it difficult to change their *hukou*.

The maternal working sector is a good variable proxying for prices on health and educational investments. Child medical treatment and education are usually subsidised by the government if mothers are working in the public sector. Moreover, job turnover from public to private sectors is rare for Chinese women. We use household assets as a proxy for non-labour income (Y). ‘Household asset’ is a score generated by using principal component analysis (Kolenikov and Angeles, 2009). The durable goods used to generate the household asset scores include washing machines, refrigerators, air-conditioners, shower heaters, water dispensers, sterilised cupboards, motorcycles, cars, radios/recorders, televisions, video recorders, video displayers, hi-fi’s, cameras, telephones and mobile phones.

Finally, column (5) of Table 1 shows substantial within-twin-pair variations in early health shocks, family investments and child outcomes. For example, the within-twin standard deviations accounts for about 35% of the total sample standard deviations of child early health shocks. This ratio is as high as 62% for child health investments.

3. Our Econometric Model and Its Identification

This Section presents our econometric model and a discussion of how we identify it. We first analyse how parents respond to early child health shocks. We linearise the optimal human capital investment (4) as a first-order approximation and add a disturbance term ($\epsilon^k_{i,\tau}$), which is assumed to reflect measurement errors in the investment of type k specific to child i in family τ; and assume that children are treated symmetrically (so that statistically they are exchangeable):

$$I^k_{i,\tau} = \alpha^k_1 e^H_{i,\tau} + \alpha^k_2 e^H_{j,\tau} + \alpha^k_3 \omega_{i,\tau} + \alpha^k_4 \omega_{j,\tau} + \alpha^k_5 \xi_{i,\tau} + \alpha^k_6 \xi_{j,\tau} + \alpha^k_7 \boldsymbol{\zeta}_\tau + (\mu_\tau) + \epsilon^k_{i,\tau}. \quad (7)$$

The assumption of symmetry implies that the coefficients are the same across children. We normalise the coefficient on parental preferences (μ_τ) to be 1. $\boldsymbol{\zeta}_\tau$ is a vector of variables including the price for human capital investment (p_I), the wage rate (w), non-labour income (Y) and other observable parental characteristics, which is denoted as κ. We denote μ as the unobserved parental characteristics or preferences. Thus, $h = (\kappa,\mu)$. Both ζ_τ and μ_τ are shared by twin siblings. Unlike (4), we do not make any distinction between health and cognitive endowments because we have only one variable – birthweight – to measure prenatal endowment in our data set. Conditioning on the unobservable μ_τ, we further assume $\epsilon^k_{i,\tau}$ to be *i.i.d.*

The major problem in identifying the own effect (α^k_1) and cross effect (α^k_2) of an early health shock on family investment is that the early health shock may reflect unobserved parental preferences μ_τ. For example, on the one hand, parents who prefer child human capital may be more likely to report that their children have suffered from early health shocks even if their children did not. On the other hand, these parents may spend more on human capital investments. In this case, the least squares estimates of (7), ignoring unobserved parental characteristics or preferences, would be biased. If the cross-sibling effects are symmetric and equal, we can eliminate this source of bias by the following within-twin fixed-effects (FE) estimator:

$$\Delta I_\tau^k = (\alpha_1^k - \alpha_2^k)\Delta e_\tau^H + (\alpha_3^k - \alpha_4^k)\Delta\omega_\tau + (\alpha_5^k - \alpha_6^k)\Delta\xi_\tau + \Delta\epsilon_\tau^k, \tag{8}$$

where Δ is an operator forming within-twin differences. We test the symmetric regression assumption of cross-sibling effects below and find that it is not rejected in our data.

We are interested in the estimates of $\alpha_1^k - \alpha_2^k$ in (8) which reflect parental investment strategy or the intra-household resource allocation effect of human capital investment of type k. If parents reinforce investment of type k in response to an early health shock on child i, then $\alpha_1^k - \alpha_2^k < 0$; compensation implies that $\alpha_1^k - \alpha_2^k > 0$. The self and cross effects reflect the interaction between parental preference and production technology. Our theory predicts that parents could make compensating and reinforcing investments along different dimensions of human capital in response to an early health shock. Therefore, the signs of $\alpha_1^H - \alpha_2^H$ and $\alpha_1^C - \alpha_2^C$ are not necessarily the same.

Our basic identifying assumption is that the within-twin variation in the error term of (8) is uncorrelated with the within-twin variation in early health shocks. This assumption may not necessarily hold. For example, the within-twin difference in individual specific prenatal endowment might not be fully controlled for by $\Delta\omega$, leading to a correlation between Δe and $\Delta\varepsilon$. If so, Δe would be endogenous in (8), and the OLS estimates of (8) would be biased.[9]

We measure early health shocks by incidence of serious diseases at ages 0–3 as discussed in the Section above. Diseases can either reflect a random shock which is consistent with the definition of early health shocks in the model, or be an indication of an individual-specific health endowment. In other words, though early health shocks may or may not be exogenous in (7), what we need for identification is that the within-twin variation in early health shocks is random and exogenous in (8).

We propose and implement a test to examine the identifying assumption. We assume that the unobserved in health shocks are uniform across children within a family and that elimination of a family fixed effect eliminates any spurious correlation. A test of this assumption compares the OLS estimates of the coefficients of (7) for the sick and healthy child. The difference in the OLS estimates should accord with the fixed-effect estimates assuming all other covariates are equally correlated with the health shocks.

Specifically, under the null hypothesis that within-twin variations in early health shocks are exogenous such that e_i^H and ε are uncorrelated in (7) after controlling for the unobservable μ_τ, we have:

[9] A similar issue has been extensively investigated in the literature on twin-based estimation of returns to schooling (Griliches, 1979; Bound and Solon, 1999; Neumark, 1999).

(*i*) $plim\ \widehat{\alpha}_1^k = \alpha_1^k + (\sigma_{\mu,e_i^H}/\sigma^2_{e_i^H})$ and $plim\ \widehat{\alpha}_2^k = \alpha_2^k + (\sigma_{\mu,e_j^H}/\sigma^2_{e_j^H})$ in (7), where σ_{μ,e_i^H} is the covariance of μ (family fixed effects) and $e_\imath^H$ (early health shock for child τ), and $\sigma^2_{e_\imath^H}$ is the variance of $e_\imath^H$;[10]

(*ii*) $\sigma_{\mu,e_i^H}/\sigma^2_{e_i^H} = \sigma_{\mu,e_j^H}/\sigma^2_{e_j^H}$ by symmetry; and

(*iii*) $plim\ (\widehat{\alpha_1^k - \alpha_2^k}) = \alpha_1^k - \alpha_2^k$ in (8).

Therefore, the null hypothesis of symmetry (really exchangeability) gives the cross-equation restriction:

$$plim(\widehat{\alpha}_1^k - \widehat{\alpha}_2^k) = plim(\widehat{\alpha_1^k - \alpha_2^k}), \tag{9}$$

where $\widehat{\alpha}_1^k$ and $\widehat{\alpha}_2^k$ are the OLS estimates of (7), and $(\widehat{\alpha_1^k - \alpha_2^k})$ are the fixed-effects estimates of (8). Thus, we test differences of estimates obtained from OLS *versus* estimates obtained from a fixed-effects procedure. If there are unobserved, individual-specific prenatal endowments and $e_\imath^H$ and ε are correlated in (7), the null hypothesis of symmetry and the equality in (9) would not hold.[11] We test and do not reject this restriction in our empirical analysis reported below.

We analyse the effect of early health shocks on child human capital. By linearising the production function (1) as a first-order approximation and by adding a disturbance term that reflects measurement errors $(v_{i,\tau})$, which are individual-specific, we specify the following regression equation:

$$\theta_{i,\tau}^k = \beta_1^k e_{i,\tau}^H + \beta_2^k \omega_{i,\tau} + \beta_3^k I_{i,\tau}^k + \beta_4^k \xi_{i,\tau} + \beta_5^k \kappa_\tau + \mu_\tau + v_{i,\tau}, \tag{10}$$

where κ denotes observable parental characteristics. The cross-sectional variation in early health shocks may reflect the unobservable parental characteristics. For example, better child-rearing practices may lead to fewer early health shocks and, at the same time, more child human capital. To remove the possible bias induced by unobserved parental characteristics, we use the following fixed effect (FE) specification:

$$\Delta\theta_\tau^k = \beta_1^k \Delta e_\tau^H + \beta_2^k \Delta I_\tau^k + \beta_3^k \Delta\omega_\tau + \beta_4^k \Delta\xi_\tau + \Delta v_\tau, \tag{11}$$

where β_1^k captures the biological effects corresponding to the first term on the right-hand side of (5), which is expected to be negative; β_2^k measures the productivity effect of family investments, which is expected to be positive.

Because family investment is a parental choice, we need to worry about simultaneous equations bias. We use 2SLS to estimate (11). Our theoretical analysis guides us to choose instrumental variables (IV). The price of investments (p_I), wage rate (w) and non-labour income (Y) in the family investment (4) are excluded from the production function (1). However, these variables drop out from the within-twin differences of (11). We thus use p_I, w and Y interacted with the within-twin variation in child prenatal endowment and characteristics ($\Delta\omega \times Y$, $\Delta\omega \times p_I$, $\Delta\omega \times h$, $\Delta\xi \times Y$, $\Delta\xi \times p_I$, and $\Delta\xi \times h$) as IVs for within-twin variation in investments (ΔI^k $(k = H,C)$) in (11). Specifically, we use two types of interaction terms as IVs for within-twin differences in

[10] We ignore other covariates in the equation to simplify our discussion. We further assume no contagious effect of early health shocks. The result below holds if the contagious effect is symmetric within twin sibling pairs.

[11] The cross-equation restriction also applies if e^H is measured with errors which are related to family investment behaviour. See the discussion in online Appendix C.

health or educational investments. The first type includes the interaction terms between the within-twin difference in birthweight and the household and parental level variables; the second type includes the interaction terms between the within-twin difference in gender and the household and parental-level variables. The household and parental-level variables include a rural indicator, maternal working sector, age, ethnicity and schooling years.

Administrative data usually do not have information on family investments in children's human capital. Thus, the recent literature in health economics does not estimate health production functions. Instead, the literature specifies a reduced-form regression equation that ignores the intra-household response such as:

$$\Delta\theta_\tau^k = \phi_1^k \Delta e_\tau^H + \phi_2^k \Delta\omega_\tau + \phi_3^k \Delta\xi_\tau + \Delta u_\tau. \tag{12}$$

The reduced-form estimate of ϕ_1^k captures the total effect of an early health shock on child human capital, and corresponds to the left hand side of (5). When $\beta_2^k \neq 0$ and $\alpha_1^k - \alpha_2^k \neq 0$, β_1^k differs from ϕ_1^k. The total effect deviates from the biological effect. To compare our results with those in the literature, we also perform reduced-form estimation. By comparing β_1^k with ϕ_1^k, we can infer the importance of the intra-household resource allocation effects by using (5).

4. Empirical Evidence

4.1. *Child Early Health Shock and Family Investments*

Before presenting our estimation results of the investment equation, we first examine the determinants of early health shocks. The results are reported in Table 2. Column (1) reports the OLS estimates. We find that low birthweight is positively correlated with the probability of suffering early health shocks. We also observe that males are more likely to suffer. Results also show a positive correlation between maternal schooling and the probability of reporting children suffering from early health shock. We interpret this positive correlation as a reporting bias that is commonly found in he literature in the health economics (Strauss and Thomas, 1998). If the reporting bias is determined by parental observed and unobserved characteristics, the issue of reporting bias can be addressed by the within-twin FE estimation because these parental characteristics are removed by within-twin differences. The FE estimates in column (2) show that the occurrence of an early health shock is unrelated to birthweight. In column (3), we report a regression of the within-twin difference in early health shocks on both within-twin differences in birthweight and gender and other level covariates. We find that the within-twin variations in early health shocks are uncorrelated with the level variables. Therefore, the results in Table 2 do not reject our identifying assumption of the randomness of within-twin variation in early health shocks.

We now turn to our main estimation results about child early health shocks and family investments in children. Column (1) in Table 3 reports the OLS estimates of (7) with respect to health investments. We find that the own effect of an early health shock on health investment is significantly positive, whereas the cross-sibling effect is almost zero. We also find that being a male, being born at first parity, maternal schooling years and household asset are positively correlated with child health investments. By

Table 2

Determinants of Child Early Health Shocks

	Dependent variable: early health shocks		
	(1)	(2)	(3)
Birthweight: < 2 kg	0.044**	0.020	0.021
	(0.020)	(0.022)	(0.022)
Birthweight: 2–2.5 kg	0.008	−0.000	0.001
	(0.016)	(0.018)	(0.018)
Birthweight: 2.5–3 kg	−0.012	−0.024	−0.022
	(0.016)	(0.015)	(0.015)
Male	0.040***	0.020*	0.020*
	(0.010)	(0.010)	(0.010)
Age	−0.001		0.001
	(0.002)		(0.002)
Born at the first parity	0.008		−0.007
	(0.014)		(0.014)
Maternal age	−0.002		−0.000
	(0.001)		(0.001)
Maternal ethnicity (Han = 1)	−0.014		0.013
	(0.015)		(0.015)
Maternal schooling years	0.007***		−0.000
	(0.002)		(0.002)
Maternal working sector (public = 1)	0.009		0.012
	(0.022)		(0.021)
Rural	−0.019		−0.014
	(0.013)		(0.013)
Household asset	−0.007*		0.004
	(0.004)		(0.004)
R^2	0.024	0.009	0.015
No. of pair of twins	1,456	1,456	1,456

Notes. Columns (1) and (2) report the OLS and within-twin fixed-effects estimates, respectively. In column (3), we regress the within-twin differences in early health shock on both within-twin differences in birthweight and gender and other variables at the household level. Standard errors are in parenthesis; *significant at 10%; **significant at 5%; ***significant at 1%.

contrast, the variable of 'born in rural areas' is negatively correlated with child health investment.

To test the symmetry assumption of cross-sibling effects, we use the same sample as in column (1) but employ the family health investment in child j as a dependent variable. The results are reported in column (2). We then conduct a Wald test. The χ^2 statistic is 18.87 and the p-value is 0.34. These findings indicate that we cannot reject the symmetry assumption. Therefore, we conduct an FE estimation of (8) with respect to health investments. The results are reported in column (3). On the basis of these estimates, we find that the gap in health expenditures on average increases by 305 yuan (1.39×225) in favour of the sick twin child, which amounts to more than one third of the average family investment in child education per year.

To test the null hypothesis that within-twin variations in early health shocks are exogenous such that Δe and $\Delta \varepsilon$ are uncorrelated in (8), we conduct a Wald test on (9). If any individual specific component is omitted from (8) such that Δe and $\Delta \varepsilon$ are correlated, within-twin variations in early health shocks would be endogenous. Thus,

Table 3

Child Early Health Shocks and Parental Investment in Health

	Dependent variable:		
	Health investment in child *i* (*log*)	Health investment in child *j* (*log*)	Difference in health investment
	(1)	(2)	(3)
Early health shocks (*i*)	1.192***	−0.288	1.349***
	(0.325)	(0.312)	(0.243)
Early health shocks (*j*)	0.001	1.194***	
	(0.322)	(0.309)	
Birthweight: < 2 kg (*i*)	0.113	−0.297	0.534***
	(0.286)	(0.275)	(0.204)
Birthweight: 2–2.5 kg (*i*)	0.237	−0.286	0.484***
	(0.232)	(0.223)	(0.163)
Birthweight: 2.5–3 kg (*i*)	0.248	−0.192	0.418***
	(0.213)	(0.204)	(0.139)
Birthweight: < 2 kg (*j*)	0.203	0.889***	
	(0.289)	(0.277)	
Birthweight: 2–2.5 kg (*j*)	0.054	0.463**	
	(0.227)	(0.218)	
Birthweight: 2.5–3 kg (*j*)	−0.073	0.317*	
	(0.200)	(0.192)	
Male (*i*)	0.291**	0.226*	0.071
	(0.137)	(0.132)	(0.095)
Male (*j*)	0.021	0.087	
	(0.137)	(0.132)	
Age	−0.041*	−0.046**	
	(0.024)	(0.023)	
Born at the first parity	0.387**	0.361**	
	(0.162)	(0.155)	
Maternal age	0.007	0.003	
	(0.015)	(0.015)	
Maternal ethnicity (Han = 1)	0.006	0.102	
	(0.173)	(0.166)	
Maternal schooling years	0.083***	0.114***	
	(0.024)	(0.023)	
Maternal working sector (public = 1)	−0.286	−0.378	
	(0.246)	(0.237)	
Rural	−0.298**	0.047	
	(0.145)	(0.139)	
Household asset	0.124***	0.067	
	(0.045)	(0.043)	
χ^2		18.87	0.14
p-values		0.34	0.71
No. of pair of twins	1,456	1,456	1,456

Notes. Columns (1) and (2) report the OLS estimates; column (3) reports the within-twin fixed-effect estimates. Standard errors are in parenthesis; *significant at 10%; **significant at 5%; ***significant at 1%.

the equality in (9) would break down. We report the test result at the base of column (3). The χ^2 statistic is 0.14, and the p-value is 0.71. Therefore, the null hypothesis is not rejected.

Table 4 presents the estimates of child early health shocks and family educational investment. Column (1) reports the ordinary least square (OLS) estimates of (7) with respect to educational investments. In contrast to the estimates with health investments

Table 4

Child Early Health Shocks and Parental Investment in Education

	Dependent variables:		
	Education investment in child i (*log*)	Education investment in child j (*log*)	Difference in education investment
	(1)	(2)	(3)
Early health shocks (i)	−0.142	0.065	−0.204***
	(0.131)	(0.124)	(0.047)
Early health shocks (j)	0.213	0.008	
	(0.130)	(0.123)	
Birthweight: < 2 kg (i)	0.012	0.059	−0.015
	(0.115)	(0.110)	(0.039)
Birthweight: 2–2.5 kg (i)	0.131	0.123	0.016
	(0.094)	(0.089)	(0.031)
Birthweight: 2.5–3 kg (i)	0.007	0.031	−0.013
	(0.086)	(0.081)	(0.027)
Birthweight: <2 kg (j)	−0.143	−0.127	
	(0.117)	(0.111)	
Birthweight: 2–2.5 kg (j)	−0.060	−0.045	
	(0.091)	(0.087)	
Birthweight: 2.5–3 kg (j)	−0.098	−0.103	
	(0.081)	(0.077)	
Male (i)	−0.042	−0.014	−0.024
	(0.055)	(0.053)	(0.018)
Male (j)	0.027	0.008	
	(0.055)	(0.053)	
Age	0.083***	0.088***	
	(0.010)	(0.009)	
Born at the first parity	0.042	0.041	
	(0.065)	(0.062)	
Maternal age	0.014**	0.014**	
	(0.006)	(0.006)	
Maternal ethnicity (Han = 1)	0.090	0.111*	
	(0.070)	(0.066)	
Maternal schooling years	0.038***	0.041***	
	(0.010)	(0.009)	
Maternal working sector (public = 1)	0.083	0.024	
	(0.099)	(0.094)	
Rural	−0.150**	−0.184***	
	(0.058)	(0.055)	
Household asset	0.130***	0.124***	
	(0.018)	(0.017)	
χ^2		19.39	0.70
p-values		0.31	0.41
No. of pair of twins	1,456	1,456	1,456

Notes. Columns (1) and (2) report the OLS estimates; column (3) reports the within-twin fixed-effect estimates. Standard errors are in parenthesis; *significant at 10%; **significant at 5%; ***significant at 1%.

(column 1 in Table 3), we find that the own effect of an early health shock on educational investment is not statistically significant, although the estimate is negative. In contrast, we find that the cross-sibling effect is positive. The other covariates, namely, child age, maternal age, maternal schooling years and household asset are significantly positively correlated with child educational investments, whereas the covariate 'born in rural areas' is negatively correlated with child educational investment.

The result of testing the symmetry assumption of cross-sibling effects is reported at the base of column (2). The χ^2 statistic is 19.39 with a p-value of 0.31. We cannot reject the symmetry assumption. Therefore, we conduct the FE estimation of (8) with respect to educational investments. The results are reported in column (3). We find that, in contrast to the estimates with health investment, the gap in educational expenditures increases by 182 yuan (0.2×910), on average, in favour of the healthy child. The χ^2 statistic of testing the cross-equation restriction (9) is 0.70, and the p-value is 0.41. We cannot reject the null hypothesis that no endogenous within-twin variation exists in early health shock.

Summarising Tables 3 and 4, our results indicate that parents adopt a compensating strategy with respect to health investment and a reinforcing strategy with respect to educational investment in response to an early health shock which affects one of the twin children.

We also estimate the family investment equation by different subsamples based on *hukou* status, maternal education, household wealth and gender composition of twin children. We find significant differences in the compensating and reinforcing patterns across subsamples. First, the increase in health expenditures in favour of the sick twin in rural areas is not accompanied by a corresponding decrease in educational expenditures. In urban areas, instead, the fall in the amount of educational resources on the sick child almost exactly offsets, in monetary terms, the amount redistributed to pay for the medical expense. Second, both the compensating health investment and the reinforcing education investment are more precisely determined when mothers have a higher education level. Third, there is no significant difference in the compensating health investment behaviour between poor and rich households. In contrast, the reinforcing education investment behaviour is more significant in rich households than that in poor households. Finally, we find significant differences by gender. The compensating health investment and the reinforcing education investment are more significant in the female twin than male twin samples. The results are reported and discussed in online Appendix D.

Our results have important implications. First, the results are consistent with our two major observations in the theoretical analysis on the intra-household resource allocation. The reinforcement or compensatory child human capital investment strategy reflects not only parental preferences but also production technology. Furthermore, parents can compensate and reinforce along different dimensions of human capital with respect to an early health shock on one child, which helps us reconcile the seemingly conflicting findings in the literature.

Second, the results deepen our understanding of the role of family investment in the overall level of inequality in a society. Our estimates indicate that intra-household resource reallocation in child investment is not trivial. The gap in educational investment between sick and healthy children accounts for up to one-fifth of the average educational investment per person. The gap in health investment is even larger. Therefore, the family plays an important role in accounting for the difference in investment in children. However, the role of the family in this case is complicated. On the one hand, parents increase the gap in educational investment by allocating more resources to the child suffering from an early health shock. In this regard, the family increases the overall level of inequality in the economy. On the other hand, parents decrease the gap in health investment by allocating more resource on the sick child. Thus, in terms of health investment, the family decreases inequality. If we focus on educational investment, the role of the family as an inequality mitigator will be understated.

Finally, our empirical results clarify the interpretation of the recent reduced-form estimates on early-life conditions and late-life outcomes. When parents make compensating and reinforcing investments along different dimensions of human capital, these reduced-form estimates cannot be unambiguously interpreted as upper or lower-bounds on the biological effects. Our estimates on intra-household resource allocation suggest that the reduced-form estimates underestimate the biological effect of an early health shock on health but overestimate the biological effect on education. To verify this prediction empirically, we estimate the child human capital production function below.

4.2. *Early Health Shocks and Child Human Capital*

4.2.1. *Child health*

We next examine the effects of early health shocks on child human capital by estimating production function (11). Health investment is an endogenous variable in the child human capital equation. As discussed above, we use a 2SLS estimator. Guided by our theoretical model, we use $\Delta\omega \times Y$, $\Delta\omega \times p_I$, $\Delta\omega \times h$, $\Delta\xi \times Y$, $\Delta\xi \times p_I$ and $\Delta\xi \times h$ as the IVs for ΔI^k ($k = H,C$) in (11). Specifically, the IVs include the interaction terms between the within-twin variations in birthweights and gender and the level variables, which include a rural indicator, maternal working sector, age, ethnicity and schooling. The results in Table 2 show that these level variables are uncorrelated with the within-twin variation in early health shocks.

Panel (*a*) in Table 5 reports the 2SLS estimates of the child health production function.[12] Conditional on health investments, the estimate of early health shocks captures the biological effect. We find a long-lasting negative biological effect of early health shocks on the child's later health status. The estimates are statistically significant at a high level of 1% for three out of four measures. We also find that the productivity effects of the health investment are consistently positive. The estimates are statistically significant at the 5% level for current weight and BMI.

To compare our results with those from the recent literature in health economics and to qualitatively gauge the importance of intra-household resource allocation effects, we also estimate the reduced-form of (12). Table 3 shows that parents adopt a compensatory strategy in health investments, and the intra-household resource allocation effect is positive. Panel (*a*) of Table 5 shows a significantly positive productivity effect of health investment. By using the formula of (5), we expect that the reduced-form estimates understate the biological effect of early health shock on later health status.

This prediction is confirmed by the reduced-form estimates, which are reported in panel (*b*) of Table 5. The reduced-form estimates are also consistently negative,

[12] We report only the estimated coefficients on early health shock and health investment in the article. The full results are reported in online Appendix E.

Table 5

Early Health Shocks and Child Health

	Dependent variables:			
	Height z-score (1)	Weight z-score (2)	BMI z-score (3)	Health status (4)
Panel (a): 2SLS estimates				
Early health shocks	−0.100	−0.422***	−0.395***	−0.513***
	(0.130)	(0.118)	(0.151)	(0.078)
Health investments†	0.070	0.118**	0.160**	0.047
	(0.064)	(0.059)	(0.077)	(0.039)
Panel (b): reduced-form estimates				
Early health shocks	−0.004	−0.263***	−0.201*	−0.449***
	(0.096)	(0.086)	(0.113)	(0.057)
No. of pair of twins	1,418	1,430	1,408	1,450

Notes. Standard errors are in parenthesis; *significant at 10%; **significant at 5%; ***significant at 1%. Birthweight and child gender are controlled for in each regression. †Endogenous variable. The instrumental variables include two types of interaction terms. The first type includes the interaction terms between the within-twin difference in birthweight and the household and parental level variables, and the second type includes the interaction terms between the within-twin difference in gender and the level variables. The household and parental level variables include a rural indicator, maternal working sector, age, ethnicity and schooling years.

indicating that the direct biological effects, which are negative, outweigh the positive intra-household resource allocation effects. We find that the reduced-form estimates substantially underestimate the biological effects. Taking the dependent variable of BMI as an example (column 3), the reduced-form estimate (−0.20) accounts for one half of the biological effect (−0.40). Almost the same proportion of the negative biological effect (−0.42) is accounted for by the reduced-form estimate (−0.26) with respect to weight (column 2). The results have important implications, suggesting that although the negative effect of an early health shock may persist throughout the life cycle of children, remediation is possible. In other words, the negative effects can be partly offset by compensating investments within the household.

4.2.2. *Child education and socio-emotional skills*

We next examine the effects of an early health shock on child education by estimating the child education production function (11). The same set of variables are used as IVs for educational investment as discussed above. Panel (*a*) in Table 6 reports the 2SLS estimates of early health shocks on child educational achievements, both perceived and actual. We find that the twin child affected by an early health sock has poorer academic achievement compared with the healthy child. The productivity effects of educational investments on academic achievements are consistently positive across the four measures but are generally imprecisely estimated. We also conduct a 2SLS estimation of the effect of early health shocks on child schooling performance. The results are reported in panel (*c*) of the Table, which shows that an early health shock negatively affects the child's schooling performance. We further observe that

Table 6

Early Health Shocks and Child Education

	Dependent variables:			
	Literature		Mathematics	
	Score (1)	Relative measure (2)	Score (3)	Relative measure (4)
Panel (a): 2SLS estimates				
Early health shocks	−3.990*	−0.168	−4.697*	−0.496***
	(2.045)	(0.145)	(2.432)	(0.157)
Educational investments†	6.124	0.904**	3.598	0.170
	(6.313)	(0.426)	(7.459)	(0.466)
Panel (b): reduced-form estimates				
Early health shocks	−5.142***	−0.352***	−5.372***	−0.531***
	(1.665)	(0.110)	(1.996)	(0.127)
No. of pair of twins	1,355	1,426	1,332	1,420

	Dependent variables:			
	Good student awards (5)	Awards in contests (6)	Grade repetition (7)	Doing minor actions in class (8)
Panel (c): 2SLS estimates				
Early health shocks	−0.199***	−0.067	0.025	0.296**
	(0.072)	(0.042)	(0.038)	(0.121)
Educational investments†	0.075	0.103	−0.235**	−0.639*
	(0.216)	(0.126)	(0.113)	(0.387)
Panel (d): reduced-form estimates				
Early health shocks	−0.215***	−0.088***	0.073**	0.396***
	(0.058)	(0.033)	(0.029)	(0.101)
No. of pair of twins	1,456	1,456	1,456	1,440

Notes. Standard errors are in parenthesis; *significant at 10%; **significant at 5%; ***significant at 1%. Birthweight and child gender are controlled for in each regression. †Endogenous variable. The instrumental variables include two types of interaction terms. The first type includes the interaction terms between the within-twin difference in birthweight and the household and parental-level variables; the second type includes the interaction terms between the within-twin difference in gender and the level variables. The household and parental level variables include a rural indicator, maternal working sector, age, ethnicity and schooling years.

educational investments exert a positive productivity effect on the child's schooling performance.

The reduced-form estimates are reported in panels (*b*) and (*d*). We find that the reduced-form estimates are consistently larger than the 2SLS estimates in terms of absolute values. The result suggests that the reduced-form estimates overstate the negative biological effect of early health shocks on child education. By comparing the 2SLS estimates with the reduced-form estimates, we find that the biological effects constitute a major part of the total effect for most measures of child educational outcomes. The result suggests a strong cross-productivity effect of health shocks on education, broadly consistent with Cunha *et al.* (2010). The intra-household resource allocation effect has also played an important role. Take the transcript recorded score of literature as an example (column 1), the intra-household resource allocation effect accounts for about one-third of the total negative effect of an early health shock.

Table 7

Early Health Shocks and Child Socio-economic Skills

	Dependent variables:			
	Feel lonely (1)	Easily distracted (2)	Easily frightened (3)	Emotionally unstable (4)
Panel (a): 2SLS estimates				
Early health shocks	0.132***	0.121**	0.066	0.085***
	(0.041)	(0.056)	(0.046)	(0.024)
Educational investments†	−0.158	−0.150	−0.383***	−0.113
	(0.123)	(0.166)	(0.139)	(0.073)
Panel (b): Reduced-form estimates				
Early health shocks	0.165***	0.151***	0.144***	0.108***
	(0.032)	(0.044)	(0.033)	(0.019)
No. of pair of twins	1,456	1,456	1,456	1,456

Notes. Standard errors are in parenthesis; *significant at 10%; **significant at 5%; ***significant at 1%. Birthweight and child gender are controlled for in each regression. †Endogenous variable. The instrumental variables include two types of interaction terms. The first type includes the interaction terms between the within-twin difference in birthweight and the household and parental-level variables; the second type includes the interaction terms between the within-twin difference in gender and the level variables. The household and parental level variables include a rural indicator, maternal working sector, age, ethnicity and schooling years.

With regard to the effect of early health shocks on child socio-emotional skills, panel (*a*) in Table 7 reports the 2SLS estimates. The importance of these skills in determining an individual's income and other measures of well-being has been increasingly recognised by economists (Borghans *et al.*, 2008). We find that the twin child who suffered from an early health shock at age 0–3 has inferior personality traits. When we use the SDQ to measure the child's socio-emotional skills, the estimates show that children who experienced an early health shock are more likely to feel lonely, be easily distracted, easily frightened and be emotionally unstable. Although the estimates of the productivity of educational investment are consistently negative, the estimates are only statistically significant in column (3) with the measure of 'being easily frightened'. The results suggest that educational investment may not be an important determinant of the development of socio-emotional skills.[13] Panel (*b*) reports the reduced-form estimates. By comparing the reduced-form estimates with the 2SLS estimates, we find that biological effects constitute a major part of the total effects except for the measure of being easily frightened (column 3).

4.3. *Child Early Health Shock and Parental Labour Supply and Consumption*

Finally, we investigate how an early health shock on one child affects parental labour supply and consumption. We specify the following regression equation:

$$y_\tau = \varphi_1 d + \varphi_2 \overline{\varpi}_\tau + \varphi_3 \xi_\tau + \zeta_\tau \varphi_4 + \upsilon_\tau, \qquad (13)$$

[13] Cunha *et al.* (2010) also use educational investments as one argument in the production function of socio-emotional skills.

Table 8

Child Early Health Shocks and Parental Labour Supply and Consumption

	Dependent variables:			
	Labour supply		Consumption	
	Father (1)	Mother (2)	Father (3)	Mother (4)
Early health shocks	−0.048	0.054**	−128.808*	27.751
	(0.078)	(0.024)	(77.978)	(73.685)
Child birthweight (mean)	0.015	0.002	77.850*	27.692
	(0.020)	(0.014)	(44.919)	(30.173)
Child age	−0.000	0.005	−8.672	−9.728
	(0.004)	(0.005)	(7.799)	(7.860)
Male twins	0.005	−0.017	−60.571	−22.393
	(0.023)	(0.021)	(54.827)	(28.318)
Female twins	−0.023	−0.034	−28.210	42.148
	(0.022)	(0.021)	(57.334)	(38.050)
Born at the first parity	−0.003	−0.049**	−15.817	4.098
	(0.026)	(0.023)	(43.038)	(30.684)
Age	−0.003	−0.005	−13.630***	−5.438**
	(0.002)	(0.004)	(3.359)	(2.629)
Ethnicity (Han = 1)	−0.028*	−0.030*	−47.481	−127.626
	(0.016)	(0.016)	(59.532)	(97.041)
Schooling years	−0.007	0.001	0.218	29.564*
	(0.005)	(0.004)	(11.718)	(16.167)
Working in public sector	−0.012	−0.054***	21.079	294.633**
	(0.028)	(0.021)	(101.420)	(141.145)
Rural	0.017	0.017	40.632	82.299
	(0.018)	(0.016)	(55.064)	(58.220)
Household asset	−0.004	−0.010	230.205***	125.220***
	(0.006)	(0.007)	(29.027)	(13.827)
No. of households	1,158	1,044	1,416	1,437

Notes. Columns (1)–(4) reports the OLS estimates. Standard errors are in parenthesis; *significant at 10%; **significant at 5%; ***significant at 1%.

where y_τ measures parental labour supply or consumption in household τ; d is a dummy variable indicating whether the household has only one twin child who suffered from early health shock; $\overline{\varpi}$ is the average birthweight of the twin children. Equation (13) is a linear approximation of the optimal parental labour supply or consumption function (6) with slight modifications. First, we estimate (13) for mothers and fathers separately. We use the mean of birthweights to measure child prenatal endowments. Child characteristics are the same between twin siblings except for gender. We use two dummy variables to measure the gender composition. The first variable indicates that both children are boys, and the other variable indicates that both are girls. Second, given that within-twin variation in early health shocks is uncorrelated with unobservable characteristics, which is tested above (Table 2), we add d in the regression equation. Thus, φ_1 gives a causal interpretation even if we ignore the unobservable parental characteristics when estimating (13). Finally, the added disturbance term υ_τ reflects measurement errors.

The estimation results are reported in Table 8. We find that in households with only one twin child who suffered from an early health shock, the father is significantly less likely to spend money on goods for himself, and the mother is significantly more likely to work. Therefore, if we take other family members except the twin children into account, our results imply that the within-twin FE estimates of child outcomes understate the overall negative effect of an early health shock in a family.

5. Conclusions

This article studies how child early health shocks affect intra-household resource allocation and child human capital formation. We formulate a theoretical model that early health shocks can affect human capital through two channels: a direct channel – the biological effect through the production of human capital – and an indirect one – the intra-household resource allocation effect through parental responses. By introducing multidimensionality of human capital, parents could compensate and reinforce along different dimensions with respect to early health shocks on children. By using the CCTS data, we find evidence of compensating investment in child health but of reinforcing investment in education. We further estimate the child human capital production function. The results confirm that early health shocks negatively affect child human capital, ranging from health and education to socio-emotional skills. Reduced-form estimates, which ignore the intra-household allocation process, understate the biological effect for health but overstate the biological effect for some educational outcomes.

Our results have important implications for evaluating the role of the family in creating overall inequality in the economy. When parents can simultaneously make reinforcing and compensating family investments on children along different dimensions of human capital, the effects of intra-household resource allocation on inequality becomes complicated. On the basis of our estimates, we conclude that the family appears to be a net equaliser in terms of family investment in financial terms. Our findings also emphasise the importance of accounting for parental behavioural responses to early health shocks. Intra-household responses should be considered when designing public interventions to remediate negative health shocks on children, because parents can partly offset the effect of public interventions by reallocating resources within the family.

Further research on the intra-household resource allocation and human capital formation is desirable. First, in line with recent progress on human capital production technology, estimating a structural model separating production technology from parental preference would be valuable. Second, consistent with the empirical literature based on sibling or twin data, our regression analyses use linear specifications. Relaxing the linearity specification in the estimation would be fruitful. However, this approach relies on the development of the econometrics of non-linear fixed-effects estimators. Finally, fertility decisions should be explicitly introduced in the next generation of models of intra-household resource allocation.

National University of Singapore
University of Chicago
Chinese University of Hong Kong
University College London

Additional Supporting Information may be found in the online version of this article:

Appendix A. The Measure of Child Early Health Shocks.
Appendix B. An Example of how the Sensitivity of Parental Responses to Child Early Health Shocks Depends on Specific Functional Forms.
Appendix C. Discussion on Potential Measurement Errors in Constructing the Variable of Child Early Health Shocks.
Appendix D. Child Early Health Shocks and Family Investments by Subsamples.
Appendix E. Early Health Shocks and Child Human Capital (full results).
Data S1.

References

Almlund, M., Duckworth, A.L., Heckman, J.J. and Kautz, T. (2011). 'Personality psychology and economics', in (E. Hanushek, S. Machin and L. Woessman, eds.), *Handbook of the Economics of Education*, pp. 1–181, vol. 4, Amsterdam: Elsevier.

Almond, D. and Currie, J. (2011). 'Human capital development before age five', in (D. Card and O. Ashenfelter, eds.), *Handbook of Labor Economics*, vol. 4B, pp. 1315–486, Amsterdam: Elsevier.

Becker, G.S. and Tomes, N. (1976). 'Child endowments and the quantity and quality of children', *Journal of Political Economy*, vol. 84(4), pp. S143–62.

Behrman, J.R., Pollak, R.A. and Taubman, P. (1982). 'Parental preferences and provision for progeny', *Journal of Political Economy*, vol. 90(1), pp. 52–73.

Behrman, J.R., Pollak, R.A. and Taubman, P. (1986). 'Do parents favor boys?', *International Economic Review*, vol. 27(1), pp. 33–54.

Behrman, J.R., Rosenzweig, M.R. and Taubman, P. (1994). 'Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment', *Journal of Political Economy*, vol. 102(6), pp. 1131–74.

Borghans, L., Duckworth, A.L., Heckman, J.J. and ter Weel, B. (2008). 'The economics and psychology of personality traits', *Journal of Human Resources*, vol. 43(4), pp. 972–1059.

Bound, J. and Solon, G. (1999). 'Double trouble: on the value of twins-based estimation of the return to schooling', *Economics of Education Review*, vol. 18(2), pp. 169–82.

Case, A., Fertig, A. and Paxson, C. (2005). 'The lasting impact of childhood health and circumstance', *Journal of Health Economics*, vol. 24(2), pp. 365–89.

Cunha, F. and Heckman, J.J. (2007). 'The technology of skill formation', *American Economic Review*, vol. 97(2), pp. 31–47.

Cunha, F., Heckman, J.J. and Schennach, S.M. (2010). 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica*, vol. 78(3), pp. 883–931.

Currie, J. and Vogl, T. (2013). 'Early-life health and adult circumstance in developing countries', Annual Review of Economics, vol. 5(1), pp. 1–36.

Currie, J., Stabile, M., Manivong, P. and Roos, L.L. (2010). 'Child health and young adult outcomes', *Journal of Human Resources*, vol. 45(3), pp. 517–48.

Du, Y., Kou, J. and Coghill, D. (2008). 'The validity, reliability and normative scores of the parent, teacher and self report versions of the Strengths and Difficulties Questionnaire in China', *Child and Adolescent Psychiatry and Mental Health*, vol. 2(1), pp. 1–15.

Family Planning Commission of Yunnan Province (2003). 'Regulations on Population and Family Planning in Yunnan Province', Beijing: Family Planning Commission of Yunnan Province China Population Press.

Glewwe, P. and Miguel, E. (2007). 'The impact of child health and nutrition on education in less developed countries', in (T.P. Schultz and J. A. Strauss, eds.), *Handbook of Development Economics*, vol. 10, pp. 3561–606, Amsterdam: Elsevier.

Griliches, Z. (1979). 'Sibling models and data in economics: beginnings of a survey', *Journal of Political Economy*, vol. 87(5), pp. S37–64.

Grossman, M. (2000). 'The human capital model', in (A.J. Culyer and J.P. Newhouse, eds.), *Handbook of Health Economics*, vol. 1, pp. 347–408, Amsterdam: Elsevier.

Heckman, J.J. (2007). 'The economics, technology and neuroscience of human capability formation', *Proceedings of the National Academy of Sciences*, vol. 104(3), pp. 13250–5.

Johnson, D.G. (1947). *Forward Prices for Agriculture*, Chicago, IL: University of Chicago Press.

Kolenikov, S. and Angeles, G. (2009). 'Socioeconomic status measurement with discrete proxy variables: is principal component analysis a reliable answer?', *Review of Income and Wealth*, vol. 55(1), pp. 128–65.

Neumark, D. (1999). 'Biases in twin estimates of the return to schooling', *Economics of Education Review*, vol. 18 (2), pp. 143–8.

Pitt, M.M., Rosenzweig, M.R. and Hassan, M.N. (1990). 'Productivity, health, and inequality in the intrahousehold distribution of food in low-income countries', *American Economic Review*, vol. 80(5), pp. 1139–56.

Rosenzweig, M.R. (1980). 'Neoclassical theory and the optimizing peasant: an econometric analysis of market family labor supply in a developing country', *Quarterly Journal of Economics*, vol. 94(1), pp. 31–55.

Rosenzweig, M.R. and Zhang, J. (2009). 'Do population control policies induce more human capital investment? Twins, birth weight and China's one-child policy', *Review of Economic Studies*, vol. 76(3), pp. 1149–74.

Smith, J.P. (2009). 'The impact of childhood health on adult labor market outcomes', *Review of Economics and Statistics*, vol. 91(3), pp. 478–89.

Strauss, J. and Thomas, D. (1998). 'Health, nutrition, and economic development', *Journal of Economic Literature*, vol. 36(2), pp. 766–817.

Strauss, J. and Thomas, D. (2007). 'Health over the life course', in (T.P. Schultz and J.A. Strauss, eds.), *Handbook of Development Economics*, vol. 4, pp. 3375–474, Amsterdam: Elsevier.

Tomes, N. (1981). 'The family, inheritance, and the intergenerational transmission of inequality', *Journal of Political Economy*, vol. 89(5), pp. 928–58.

The Economic Journal, Doi: 10.1111/ecoj.12291 © 2015 Royal Economic Society. Published by John Wiley & Sons, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

Technical Appendix to

EARLY HEALTH SHOCKS, INTRA-HOUSEHOLD RESOURCE ALLOCATION AND CHILD OUTCOMES

Junjian Yi, James J. Heckman, Junsen Zhang and Gabriella Conti

Economic Journal, doi: 10.1111/ecoj.12291

Appendix A. The Measure of Child Early Health Shocks

We discuss the measure of child early health shocks in this Section. The CCTS questionnaire contains questions about the history of serious diseases for each child. The complete list of diseases includes serious diarrhoea, calcium deficiency, asthma, fracture, attention-deficit disorder, heart disease, serious hearing difficulties, whooping cough, stammer and serious eyesight problem. To minimise recall errors, this part of the questionnaire was answered by parents and children together. Furthermore, the parents and children were required to specify the timing and duration of each disease.[1] The most common disease reported is serious diarrhoea. This fact is consistent with the case for children in developing countries (Strauss and Thomas, 1998). The average age when children suffered from serious diarrhoea is 1.42. The duration of this disease is 1.54 months. For the case of both twin children suffered from the disease, 61.23% of the twins were suffered at the same time.

One assumption with the specification of our production function ((1) in the article) is that no contagious or spillover effect of child early health shocks exists in the production function. Does the serious diarrhoea reported in our survey violate this assumption? Our answer is probably no.

There are two types of diarrhoea. One is infectious and the other is non-infectious (Chen and Scrimshaw, 1983).[2] Viral gastroenteritis with rotavirus are the most common cause for infectious diarrhoea. Fang *et al.* (2005) show that rotavirus diarrhoea accounts for about 50% of the cases in children under five in China. In general, infections that cause diarrhoea are highly contagious. However, the symptoms of diarrhoea caused by viral gastroenteritis usually last just a few days (Chen and Scrimshaw, 1983). In the case of viral gastroenteritis, replacing lost fluid and salts is the only treatment needed.

The non-infectious diarrhoea is usually caused by poisoning food, such as poisoning mushrooms and polluted water. The non-infectious diarrhoea is usually very acute and needs medical treatment. If children suffering from non-infectious diarrhoea do not get timely and appropriate medical treatment, the diarrhoea becomes chronic, leading to ulcerative colitis, celiac disease or irritable bowel syndrome. Chronic diarrhoea can last much longer than the infectious one.

[1] See Appendix C for the discussion on potential measurement error problems.

[2] We thank several discussions with Dr. Zhixiang Zuo in the Medical School at the University of Chicago and with Dr. Linjun Wang in the Medical School at the University of Michigan.

We believe that the type of diarrhoea varying between twins is the non-infectious type caused by eating poisoning food or drinking polluted water.[3] On one hand, the duration of diarrhoea reported in the survey is much longer than that in the case of infectious diarrhoea. On the other hand, the interviewers asked the parents about serious diseases during the survey. So the parents were unlikely to report the infectious diarrhoea which lasted only for a short period and did not have medical treatment for their children.

We suggest a simple test to check potential contagious effects of early health shocks in the production function. If the early health shock is contagious, both twin siblings who shared a bedroom during childhood would more likely experience early health shocks than those who did not share a bedroom. The reason is that the rotavirus is more likely to spread to both if they sleep in the same bedroom. We test this hypothesis by conducting a multinominal logit estimation. Specifically, we regress the number of twin siblings who suffered from early health shocks on a dummy variable that indicates whether the twin children shared a bedroom. The results are

Table A1

Multinominal Logit Estimates of Sharing a Room and the Number of Children Suffered from Early Health Shocks

	Dependent variables			
	No. of children suffered from early health shocks:			
	Only one (1)	Both (2)	Only one (3)	Both (4)
Sharing a room	−0.006	0.008	−0.010	0.004
	(0.011)	(0.016)	(0.012)	(0.017)
Child birthweight (mean)			−0.013	−0.023
			(0.011)	(0.015)
Child age			0.001	−0.001
			(0.002)	(0.003)
Male twins			0.008	0.018
			(0.012)	(0.017)
Female twins			−0.006	−0.018
			(0.013)	(0.019)
Twin born at the first parity			0.025	−0.002
			(0.016)	(0.018)
Maternal age			−0.002	−0.001
			(0.001)	(0.002)
Maternal ethnicity (Han = 1)			−0.002	−0.013
			(0.014)	(0.018)
Maternal school years			0.002	0.005**
			(0.002)	(0.003)
Maternal working sector (public = 1)			−0.025	0.013
			(0.021)	(0.022)
Rural			−0.014	−0.010
			(0.011)	(0.015)
Household asset			−0.003	−0.005
			(0.003)	(0.005)
No. of households	1,456	1,456	1,456	1,456

Notes. The omitted category is that no one in a pair of twins suffered early health shocks. Columns (1)–(4) report the marginal effects of the multinominal logit estimates. Standard errors are in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

[3] Chen *et al.* (2013) and Ebenstein (2012) show water pollution is an increasingly important threat to health in China.

reported in Table A1. We find that sharing a bedroom does not significantly affect the probability that both twin children suffer from an early health shock (columns (2) and (4)). Our evidence is inconsistent with the hypothesis of contagious effects. In addition, Table A1 shows that no covariate significantly predicts that only one child suffers from early health shocks (columns (1) and (3)). This result confirms our assumption of the randomness of the within-twin variation in early health shock.

Appendix B. An Example of How the Sensitivity of Parental Responses to Child Early Health Shocks Depends on Specific Functional Forms

B.1. *An Example*

We characterise the dependence of the results in the previous literature on choices of functional forms. To derive analytic results, we first specify the parental utility function as

$$U = c^{a_1} l^{a_2} \left[(q_i^\rho + q_j^\rho)^{\frac{1}{\rho}} \right]^{1-a_1-a_2}, \tag{B.1}$$

where $0 < a_1, a_2 < 1$ and $\rho \leq 1$. The parameter ρ measures the degree of parental inequality aversion across twin siblings. For example, if $\rho = 1$, parents exhibit no inequality aversion. When $\rho = -\infty$, parents exhibit infinite inequality aversion. The latter is called a Rawlsian case. The child quality is a combination of health and cognitive skills such that

$$q_\iota = (\theta_\iota^H)^{\alpha_H} (\theta_\iota^C)^{\alpha_C},$$

where $0 < \alpha_H, \alpha_C < 1$. We specify the production functions as

$$\theta_\iota^H = (\eta_\omega \omega_\iota^C)^\gamma (\beta_\omega \omega_\iota^H + \beta_e e_\iota^H + \beta_I I_\iota^H)^{1-\gamma}, \tag{B.2}$$

$$\theta_\iota^C = (\eta_\omega \omega_\iota^H + \eta_e e_\iota^H)^\gamma (\beta_\omega \omega_\iota^C + \beta_I I_\iota^C)^{1-\gamma}, \tag{B.3}$$

where $0 < \gamma, \eta_\omega, \beta_\omega, \beta_I < 1$, and $\beta_e, \eta_e < 0$.[4] Comparing with the general production function (1) in the article, we omit parental and child characteristics. The specific production function form is used only for analytic convenience. By nesting a linear function into a Cobb-Douglas function, we can easily derive the analytical comparative static result. At the same time, the result is general enough to demonstrate the two major observations.

Solving the optimisation problem, we derive the difference between the self and cross-sibling effects of early health shock on family investment for child i as follows:[5]

$$\frac{\partial I_i^{H*}}{\partial e_i^H} - \frac{\partial I_i^{H*}}{\partial e_j^H} = \frac{\alpha_H}{\beta_I} \left(\frac{\partial \pi_i}{\partial e_i^H} - \frac{\partial \pi_i}{\partial e_j^H} \right) W - \frac{\beta_e}{\beta_I}, \tag{B.4}$$

$$\frac{\partial I_i^{C*}}{\partial e_i^H} - \frac{\partial I_i^{C*}}{\partial e_j^H} = \frac{\alpha_C}{\beta_I} \left(\frac{\partial \pi_i}{\partial e_i^H} - \frac{\partial \pi_i}{\partial e_j^H} \right) W, \tag{B.5}$$

where

$$W = \beta_\omega \sum_k \sum_\iota \omega_\iota^k + \beta_I \sum_k \sum_\iota I_\iota^k + \beta_e \sum_\iota e_\iota^H, \tag{B.6}$$

[4] This is because that e is defined as a negative shock.
[5] The mathematical derivation and detailed discussion are presented in the Section below.

and

$$\pi_i = \frac{q_i^\rho}{q_i^\rho + q_j^\rho}. \tag{B.7}$$

We note that the sign of $\partial\pi_i/\partial e_i^H - \partial\pi_i/\partial e_j^H$ is unambiguously determined by parental preferences. Specifically, $\mathrm{sign}(\partial\pi_i/\partial e_i^H) = \mathrm{sign}(-\rho)$ and $\mathrm{sign}(\partial\pi_i/\partial e_j^H) = \mathrm{sign}(\rho)$. If parents are sufficiently inequality averse ($\rho < 0$), then $\partial\pi_i/\partial e_i^H - \partial\pi_i/\partial e_j^H$ is positive. Otherwise, it is negative.

We now investigate the parental investment strategy or the intra-household resource allocation effect. Equation (B.4) shows that the reinforcing or compensatory health investment in response to early health shocks is determined not only by parental aversion to inequality but also the production technology. The sign of the first term on the right-hand side of the equation is determined by parental preferences, whereas the second term is determined by the production technology. This result is consistent with our first observation listed above. By combining (B.4) and (B.5), we note that parents could reinforce the human capital investment in one dimension and compensate in another dimension. Assume $\rho > 0$, parents reinforce investment in cognitive skills, but parental investment strategy with respect to health is undetermined. This result is consistent with our second observation listed above. Therefore, we conclude that intra-household resource allocation effects could have different signs depending on different types of human capital investment. The role of the intra-household resource allocation in the overall level of inequality in the economy is undetermined.

The effects of child early health shocks on parental consumption and labour supply are as follows:

$$\frac{\partial T^*}{\partial e_i^H} = -\frac{a_2\beta_e}{w[1-\gamma(1-a_1-a_2)]},$$

and

$$\frac{\partial c^*}{\partial e_i^H} = \frac{a_1\beta_e}{[1-\gamma(1-a_1-a_2)]}.$$

As $\beta_e < 0$, an early health shock increases parental labour supply but decreases parental consumption.

B.2. *Mathematical Derivation of the Comparative Static Results*

Parents maximise

$$\begin{aligned}\underset{c,l,I_{i,1}^H,I_{i,1}^C,I_{j,1}^H,I_{j,1}^C}{Max} \quad U &= U(c,l,Q)\\ &= c^{a_1} l^{a_2} [(q_i^\rho + q_j^\rho)^{\frac{1}{\rho}}]^{1-a_1-a_2},\end{aligned}$$

subject to the time constraint

$$l + T = 1,$$

and the budget constraint

$$\sum_k \sum_\imath p_I I_m^k + c = Y + wT \quad (k = H, C;\ \imath = i, j).$$

The child quality function is

$$q_\imath = (\theta_\imath^H)^{\alpha_H} (\theta_\imath^C)^{\alpha_C}.$$

The human capital production functions are

$$\theta_\iota^H = (\eta_\omega \omega_\iota^C)^\gamma (\beta_\omega \omega_\iota^H + \beta_e e_\iota^H + \beta_I I_\iota^H)^{1-\gamma},$$
$$\theta_\iota^C = (\eta_\omega \omega_\iota^H + \eta_e e_\iota^H)^\gamma (\beta_\omega \omega_\iota^C + \beta_I I_\iota^C)^{1-\gamma}.$$

The Lagrangian is

$$\mathcal{L} = U(c, l, Q) + \lambda\left(Y + w - p_I \sum_k \sum_\iota I_\iota^k - c - wl\right).$$

The first-order conditions are

$$\frac{a_1 U}{c} = \lambda \tag{B.8}$$

$$\frac{a_2 U}{l} = \lambda w \tag{B.9}$$

$$(1 - a_1 - a_2) U \pi_i \alpha_H (1-\gamma) \beta_I \frac{1}{\beta_\omega \omega_i^H + \beta_e e_i^H + \beta_I I_i^H} = \lambda p_I \tag{B.10}$$

$$(1 - a_1 - a_2) U \pi_i \alpha_C (1-\gamma) \beta_I \frac{1}{\beta_\omega \omega_i^C + \beta_I I_i^C} = \lambda p_I \tag{B.11}$$

$$(1 - a_1 - a_2) U \pi_j \alpha_H (1-\gamma) \beta_I \frac{1}{\beta_\omega \omega_j^H + \beta_e e_j^H + \beta_I I_j^H} = \lambda p_I \tag{B.12}$$

$$(1 - a_1 - a_2) U \pi_j \alpha_C (1-\gamma) \beta_I \frac{1}{\beta_\omega \omega_j^C + \beta_I I_j^C} = \lambda p_I, \tag{B.13}$$

where $\pi_i = q_i^\rho/(q_i^\rho + q_j^\rho)$ and $\pi_j = q_j^\rho/(q_i^\rho + q_j^\rho)$.

We first solve the optimal level of human capital investment. Specifically,

$\frac{(B.10)}{(B.12)}$:

$$\frac{\pi_i}{\pi_j}(\beta_\omega \omega_j^H + \beta_e e_j^H + \beta_I I_j^H) = \beta_\omega \omega_i^H + \beta_e e_i^H + \beta_I I_i^H; \tag{B.14}$$

$\frac{(B.10)}{(B.13)}$:

$$\frac{\alpha_H}{\alpha_C}\frac{\pi_i}{\pi_j}(\beta_\omega \omega_j^C + \beta_I I_j^C) = \beta_\omega \omega_i^H + \beta_e e_i^H + \beta_I I_i^H; \tag{B.15}$$

$\frac{(B.11)}{(B.12)}$:

$$\frac{\alpha_C}{\alpha_H}\frac{\pi_i}{\pi_j}(\beta_\omega \omega_j^H + \beta_e e_j^H + \beta_I I_j^H) = \beta_\omega \omega_i^C + \beta_I I_i^C; \tag{B.16}$$

$\frac{(B.11)}{(B.13)}$:

$$\frac{\pi_i}{\pi_j}(\beta_\omega \omega_j^C + \beta_I I_j^C) = \beta_\omega \omega_i^C + \beta_I I_i^C; \tag{B.17}$$

(B.15) + (B.17):

$$\frac{1}{\alpha_C}\frac{\pi_i}{\pi_j}(\beta_\omega \omega_j^C + \beta_I I_j^C) = \beta_\omega \sum_k \omega_i^k + \beta_I \sum_k I_i^k + \beta_e e_i^H; \tag{B.18}$$

$\frac{(B.14)}{(B.15)}$:

$$\frac{1}{\alpha_C}(\beta_\omega \omega_j^C + \beta_I I_j^C) = \beta_\omega \sum_k \omega_j^k + \beta_I \sum_k I_j^k + \beta_e e_j^H \tag{B.19}$$

(B.18) + (B.19):

$$I_j^{C*} = \frac{\alpha_C}{\beta_I}\pi_j W - \frac{\beta_\omega}{\beta_I}\omega_j^C \tag{B.20}$$

where $W = \beta_\omega \sum_k \sum_\imath \omega_\imath^k + \beta_I \sum_k \sum_\imath I_\imath^k + \beta_e \sum_\imath e_\imath^H$;

(B.14) + (B.16):

$$\frac{1}{\alpha_H}\frac{\pi_i}{\pi_j}(\beta_\omega \omega_j^H + \beta_I I_j^H + \beta_e e_j^H) = \beta_\omega \sum_k \omega_j^k + \beta_I \sum_k I_j^k + \beta_e e_j^H \tag{B.21}$$

$\frac{(B.17)}{(B.16)}$:

$$\frac{1}{\alpha_H}(\beta_\omega \omega_j^H + \beta_e e_j^H + \beta_I I_j^H) = \beta_\omega \sum_k \omega_j^k + \beta_I \sum_k I_j^k + \beta_e e_j^H \tag{B.22}$$

(B.21) + (B.22):

$$I_j^{H*} = \frac{\alpha_H}{\beta_I}\pi_j W - \frac{\beta_e}{\beta_I}e_j^H - \frac{\beta_\omega}{\beta_I}\omega_j^H. \tag{B.23}$$

By symmetry,

$$I_i^{H*} = \frac{\alpha_H}{\beta_I}\pi_i W - \frac{\beta_e}{\beta_I}e_i^H - \frac{\beta_\omega}{\beta_I}\omega_i^H$$

$$I_i^{C*} = \frac{\alpha_C}{\beta_I}\pi_i W - \frac{\beta_\omega}{\beta_I}\omega_i^C.$$

We first note that

$$W = \beta_\omega \sum_k \sum_\imath \omega_\imath^k + \beta_e \sum_\imath e_\imath^H + \beta_I[(Y + w - c - wl)/p_I].$$

W measures the total inputs used to produce the children's human capital, which include health and cognitive endowments of both children, health shocks and total investment allocated to children. The inputs are weighted by their relative importance in the production function.

We then consider $\pi_\imath$, which measures the relative importance of child $\imath$ in the utility that parents derive from their children. First, $\pi_i + \pi_j = 1$. Second, the sign of $\partial\pi_i/\partial e_i^H$ is unambiguously determined by the degree of parental inequality aversion:[6]

$$\mathrm{sign}\left(\frac{\partial \pi_i}{\partial e_i^H}\right) = \mathrm{sign}(-\rho).$$

Following Becker and Tomes (1976), we interpret $\mathrm{d}\pi_i/\mathrm{d}e_i^H$ as a price effect. As indicated in the first-order conditions of (B.11) and (B.12), an early health shock on child i changes the shadow prices of the human capital investments on the child. If parents are sufficiently inequality averse ($\rho < 0$), then the price effect is positive. Otherwise, the price effect is negative. The sign of the cross-sibling price effect, which is defined as $\partial\pi_i/\partial e_j^H$, is also unambiguously determined by ρ. However, the sign of the cross-sibling price effect is opposite to that of $\partial\pi_i/\partial e_i^H$:[7]

[6] We derive that

$$\frac{\partial \pi_i}{\partial e_i^H} = \rho \frac{(q_i q_j)^\rho}{e_i^H}\frac{(\varepsilon_{q_i,e_i^H} - \varepsilon_{q_j,e_i^H})}{(q_i^\rho + q_j^\rho)^2}.$$

where ε_{q_i,e_i^H} is the elasticity of child i's quality with respect to an early health shock on himself, and ε_{q_j,e_i^H} is the elasticity of child i's quality with respect to an early health shock on his sibling. We assume that $\varepsilon_{q_i,e_i^H} - \varepsilon_{q_j,e_i^H} < 0$. This assumption derives from the parental utility function. If parents have symmetric preference, that is $U(q_i, q_j) = U(q_j, q_i)$, then this assumption is automatically satisfied. The assumption of symmetric preference is also invoked in Behrman *et al.* (1982). We test this assumption in our empirical analysis, which is not rejected.

[7] The assumption of symmetry is a sufficient condition for this conclusion.

$$\operatorname{sign}\left(\frac{\partial \pi_i}{\partial e_j^H}\right) = \operatorname{sign}(\rho).$$

Therefore,

$$\operatorname{sign}\left(\frac{\partial \pi_i}{\partial e_i^H} - \frac{\partial \pi_i}{\partial e_j^H}\right) = \operatorname{sign}(-\rho).$$

The difference between self and cross-sibling effects of early health shock on investments in the health and cognitive skills of child i are derived as follows:

$$\frac{\partial I_i^{H*}}{\partial e_i^H} - \frac{\partial I_i^{H*}}{\partial e_j^H} = \frac{\alpha_H}{\beta_I}\left(\frac{\partial \pi_i}{\partial e_i^H} - \frac{\partial \pi_i}{\partial e_j^H}\right) W - \frac{\beta_e}{\beta_I},$$
$$\frac{\partial I_i^{C*}}{\partial e_i^H} - \frac{\partial I_i^{C*}}{\partial e_j^H} = \frac{\alpha_C}{\beta_I}\left(\frac{\partial \pi_i}{\partial e_i^H} - \frac{\partial \pi_i}{\partial e_j^H}\right) W.$$

We derive the optimal consumption and labour supply and conduct comparative static analysis. By summing up (B.11)–(B.14) and by using $\pi_i + \pi_j = 1$ and $\alpha_H + \alpha_C = 1$, we have

$$\frac{(1 - a_1 - a_2)U}{W}(1-\gamma)\beta_I = \lambda. \tag{B.24}$$

By substituting (B.25) into (B.9), we derive

$$c = \frac{a_1 W}{(1 - a_1 - a_2)(1-\gamma)\beta_I}. \tag{B.25}$$

and $\frac{(\mathrm{B.9})}{(\mathrm{B.10})}$:

$$c = \frac{a_{1w}}{a_2} l. \tag{B.26}$$

By substituting (B.27) into (B.26), we derive

$$l^* = \frac{a_2}{w}\frac{W}{\beta_I[1 - \gamma(1 - a_1 - a_2)]} = \frac{a_2}{w}\widetilde{W}, \tag{B.27}$$

where $W = \beta_\omega \sum_k \sum_\iota \omega_\iota^k + \beta_e \sum_\iota e_\iota^H + \beta_I[(Y + w)/p_I]$ is the weighted family total resources adding the child endowments and early health shocks, and $\widetilde{W} = \overline{W}/\beta_I[1 - \gamma(1 - a_1 - a_2)]$ is the weighted effective resources adjusted by the production technology. By substituting (B.27) into (B.26), we have

$$c^* = a_1 \widetilde{W}. \tag{B.28}$$

Equations (B.27) and (B.28) show that the demand for leisure and consumption has the standard form in the textbook when the utility function is a CD form. The optimal labour supply is

$$T^* = 1 - \frac{a_2}{w}\widetilde{W}.$$

The effect of early health shock on parental consumption is

$$\begin{aligned}\frac{\partial c^*}{\partial e_i^H} &= a_1\left(\frac{\partial \widetilde{W}}{\partial e_i^H}\right)\\ &= \frac{a_1\beta_e}{[1 - \gamma(1 - a_1 - a_2)]}\\ &< 0\end{aligned}$$

and the effect of early health shock on parental labour supply

$$\begin{aligned}\frac{\partial T^*}{\partial e_i^H} &= -\frac{a_2}{w}\left(\frac{\partial \widetilde{W}}{\partial e_i^H}\right)\\ &= -\frac{a_2\beta_e}{w[1-\gamma(1-a_1-a_2)]}\\ &> 0,\end{aligned}$$

as $\beta_e < 0$.

Appendix C. Discussion on Potential Measurement Errors in Constructing the Variable of Child Early Health Shocks

This Section addresses potential concerns with measurement errors in constructing the variable of child early health shocks. The issue of measurement errors with twin-based estimator has been extensively discussed in the literature on returns to schooling. Griliches (1979) demonstrates that the within-twin estimator may exacerbate the problem of measurement errors with reported schooling years in estimating returns to education. A special issue of *Economics of Education Review* contains a recent exchange about twin-based estimators with measurement errors (Behrman and Rosenzweig, 1999; Bound and Solon, 1999; Neumark, 1999; Rouse, 1999). Strauss and Thomas (1998) discuss the measurement error problem with self-reported health status in developing countries.

We believe that the measurement error problem is not a major concern in our study. For illustration, we assume that child early health shocks, $e_{\imath,\tau}$, are not observed; instead, we observe

$$e^m_{\imath,\tau} = e_{\imath,\tau} + v_{\imath,\tau}. \tag{C.1}$$

We discuss three types of measurement errors. The first one is the classic measurement error such that $\mathrm{cor}(v_{\imath,\tau}, e_{\imath,\tau}) = 0$, $\mathrm{cor}(v_{\imath,\tau}, e^m_{\imath,\tau}) \neq 0$ and $\mathrm{cor}(v_{\imath,\tau}, \epsilon^k_{i,\tau}) = 0$, where $\epsilon^k_{i,\tau}$ is the error term in the family investment equation (See (7) in the article).[8] This type of measurement error can be addressed by factor methods. Specifically, we have asked different types of serious diseases in the questionnaire for each child. Therefore, we have multiple measures on child early health shocks. We estimate the measurement system and generate a factor score for early health shocks. Denote $m^r_{\imath,\tau}$ as the rth measurement of early health shock on child $\imath$ ($\imath = i,j$) in family τ. We then write the measurement system as

$$m^r_{\imath,\tau} = \Psi^r_0 + \Psi^r_1 e_{\imath,\tau} + \varepsilon^r_{\imath,\tau}. \tag{C.2}$$

In this case, $e_{\imath,\tau}$ is a continuous variable and can be interpreted as the intensity of early health shocks. The classic measurement errors are no longer contained in $e_{\imath,\tau}$. The estimation results obtained by using the factor score are reported at the end of this Section, which are qualitatively the same as those reported in the article. Therefore, the results reported in the article are robust to the presence of classic measurement errors.

The second type of measurement errors is that respondents may use different thresholds. Our measure of early health shocks is based on reported health histories. Some of the differences in the reported illness across households may simply reflect differences in the standard. This type of measurement error is common with retrospectively constructed measures of health conditions (Strauss and Thomas, 1998; Smith, 2009). In this case,

$$v_{\imath,\tau} = \mu_\tau + \upsilon_{\imath,\tau}, \tag{C.3}$$

where μ_τ is the cross-household heterogeneity in reporting thresholds and $\upsilon_{\imath,\tau}$ is the classic measurement errors such that $\mathrm{cor}(\upsilon_{\imath,\tau}, \epsilon^k_{i,\tau}) = 0$.[9] If this is the case, $v_{\imath,\tau}$ may correlate with the

[8] The estimates are unbiased when $\mathrm{cor}(v_{\imath,\tau}, e_{\imath,\tau}) \neq 0$, $\mathrm{cor}(v_{\imath,\tau}, e^m_{\imath,\tau}) = 0$ and $\mathrm{cor}(v_{\imath,\tau}, \epsilon^k_{i,\tau}) = 0$.

[9] We assume that the cross-household heterogeneity in reporting thresholds, μ_τ, is linearly additive in the measurement equation above.

family investment decision because it reflects the cross-household heterogeneity, which cannot be solved by factor methods. The twin design overcomes this problem. Parents have the same standard of reporting child illness between twin siblings. When we construct a dummy variable of early health shocks, the child who is reported to suffer from a serious disease is more likely to experience early health shocks than his or her twin sibling. Therefore, $\mathrm{cor}(v_{\imath,\tau}, \epsilon^k_{i,\tau}) = 0$ after controlling for μ_τ in (8) in the article.

The third type includes the measurement error such that $\mathrm{cor}(v_{\imath,\tau}, e^m_{\imath,\tau}) \neq 0$ and $\mathrm{cor}(v_{\imath,\tau}, \epsilon^k_{i,\tau}) \neq 0$, which cannot be resolved by either the factor method or twin design. One source of this type of measurement errors comes from recall bias. Parents whose child is currently sick may report that the child had also been sick in the past. In this case, the measurement error correlates with family investments and the measures of child early health shocks simultaneously. We believe that the recall bias is less of concern in our study because of three reasons:

(*i*) health history questions for each child were answered together by the father, mother and children in our sample;
(*ii*) the average age of children in our sample is 11. Given the young age of the twin, the recall period is not long; and
(*iii*) parents and children are also asked to specify the timing and duration of each disease.

This contextualisation has the potential to increase recollection effort and further minimise recall error.

The third type of measurement error may come from other sources besides recall bias. For example, individual-specific heterogeneity which is unobservable may exist in the investment equation. If the measurement error were not common for each twin pair but correlated with the unobservable heterogeneity, it would be correlated with the family investment decision. In this case, $cor(v_{\imath,\tau}, \epsilon^k_{i,\tau}) \neq 0$, and there exists a combined impact of endogenous variation and measurement errors, invalidating our twin design. If this is the case, the null hypothesis of the cross-equation restriction presented in Section 4 in the article should be rejected. The test results reported in Tables 3 and 4 in the article fail to reject the null hypothesis. Therefore, we conclude that the combined impact of endogenous variation and measurement errors is not a major concern in our article.[10]

Table C1

Early Health Shocks and Family Investments (FE estimates)

	Dependent variable	
	Health investments (1)	Educational investments (2)
Early health shocks	0.306***	−0.056***
	(0.090)	(0.017)
Birthweight (kg): < 2	0.564***	−0.019
	(0.206)	(0.039)
Birthweight (kg): 2–2.5	0.481***	0.016
	(0.164)	(0.031)
Birthweight (kg): 2.5–3	0.390***	−0.009
	(0.140)	(0.027)
Male	0.088	−0.027
	(0.096)	(0.018)
No. of pairs of twins	1,456	1,456

[10] We thank an anonymous referee for pointing out the application of the cross-equation restriction to test the existence of endogenous measurement errors.

Appendix D. Child Early Health Shocks and Family Investments by Subsamples

This Section discusses the effect of child early health shocks on family investments by subsamples. We divide the sample by *hukou* status (rural *versus* urban), maternal education, household wealth and the gender composition of twins. Before presenting the estimation results, Table D1 summarises the family health and education investments by subsamples. We first note substantial differences between rural and urban groups. Both the health and education investments in the urban sample are about twice as much as those in the rural sample. These differences echo the big income gap between rural and urban areas in China. As far as maternal education is concerned, children of more educated mothers get many more investments in their health and education. Similarly, children from rich families get more investments in both health and education. In contrast, we do not find significant gender differences in family investments. The female twins get marginally more family investments than male twins do. But the family investment in the mixed sample with one male and one female twins is lower than the other two groups.[11]

Table D1

Summary Statistics on Family Health and Education Investments by Subsamples

	Health investments (Yuan) (1)	Education investments (Yuan) (2)
Whole sample		
	225.83	910.44
No. of pairs of twins	1,457	1,456
Rural sample		
	149.80	632.48
No. of pairs of twins	772	772
Urban sample		
	311.65	1,224.17
No. of pairs of twins	684	684
Less educated sample (maternal schooling years ≤ 9)		
	184.11	748.09
No. of pairs of twins	1,109	1,109
More educated sample (maternal schooling years > 9)		
	359.18	1,428.29
No. of pairs of twins	347	347
Poor families (Family asset (score) ≤ sample mean)		
	167.15	568.49
No. of pairs of twins	731	731
Rich families (Family asset (score) > sample mean)		
	289.88	1,254.07
No. of pairs of twins	725	725
Male sample		
	238.65	909.13
No. of pairs of twins	539	539
Female sample		
	254.54	977.63
No. of pairs of twins	558	558
Mixed sample (one male and one female twins)		
	161.98	807.95
No. of pairs of twins	359	359

[11] In the mixed sample, there is also no significant gender difference. The health investment on boys is 167.79 yuan and the education investment is 797.48 yuan. For girls, they are 156.18 yuan and 821.43 yuan respectively.

Table D2

Child Early Health Shocks and Family Investments by Subsamples (FE estimates)

	Dependent variables	
	Health investments (1)	Education investments (2)
Whole sample		
Early health shocks	1.349***	−0.204***
	(0.243)	(0.047)
No. of pairs of twins	1,456	1,456
Rural sample		
Early health shocks	1.548***	−0.054
	(0.351)	(0.047)
No. of pairs of twins	772	772
Urban sample		
Early health shocks	1.148***	−0.320***
	(0.343)	(0.079)
No. of pairs of twins	684	684
Lowly educated sample (maternal schooling years ≤ 9)		
Early health shocks	1.182***	−0.089*
	(0.279)	(0.048)
No. of pairs of twins	1,109	1,109
Highly educated sample (maternal schooling years > 9)		
Early health shocks	1.672***	−0.472***
	(0.502)	(0.115)
No. of pairs of twins	347	347
Poor families (Family asset (score) ≤ sample mean)		
Early health shocks	1.433***	−0.120***
	(0.371)	(0.043)
No. of pairs of twins	731	731
Poor families (Family asset (score) > sample mean)		
Early health shocks	1.260***	−0.280***
	(0.323)	(0.081)
No. of pairs of twins	725	725
Male sample		
Early health shocks	1.084***	−0.171***
	(0.380)	(0.052)
No. of pairs of twins	539	539
Female sample		
Early health shocks	2.078***	−0.410***
	(0.421)	(0.089)
No. of pairs of twins	558	558
Mixed sample (one male and one female twins)		
Early health shocks	0.845*	−0.006
	(0.494)	(0.114)
No. of pairs of twins	359	359

Notes. Each entry comes from a separate regression. Robust standard errors are in parentheses; *significant at 10%; **significant at 5%; ***significant at 1%. Birthweight is controlled for in each regression; gender has been controlled for in all estimations except the ones on the male and female samples; rural has been controlled for in all estimation except the ones on the rural and urban samples. *Source*. CCTS.

Table D2 presents the estimates of the effects of child early health shocks on family investments by subsamples. We find significant differences in the compensating and reinforcing pattern across subsamples. First, the increase in health expenditures in favour of the sick twin in rural areas is not accompanied by a corresponding decrease in educational expenditure. This finding may be because families live at a subsistence level in rural areas, so that no further reductions in educational expenditures are possible.[12] In urban areas, instead, the amount of educational resources subtracted from the sick child almost exactly offsets, in monetary terms, the amount redistributed to pay for the medical expense. Second, both the compensating health investment and the reinforcing education investment are more significant when mothers have a higher education level. Third, there is no significant difference in the compensating health investment behaviour between poor and rich households. In contrast, the reinforcing education investment behaviour is more pronounced in rich households than that in poor households. Finally, we find significant differences by gender. The compensating health investment and the reinforcing education investment are more pronounced in the female twin than male twin samples.[13]

Appendix E. Early Health Shocks and Child Human Capital (Full Results)

Table E1

Early Health Shocks and Child Health

	Dependent variables			
	Height *z*-score (1)	Weight *z*-score (2)	BMI *z*-score (3)	Health Status (4)
2SLS estimates				
Early health shocks	−0.100	−0.422***	−0.395***	−0.513***
	(0.130)	(0.118)	(0.151)	(0.078)
Health investments	0.070	0.118**	0.160**	0.047
	(0.064)	(0.059)	(0.077)	(0.039)
Birthweight (kg): < 2	−0.356***	−0.516***	−0.360***	−0.043
	(0.088)	(0.080)	(0.106)	(0.053)
Birthweight (kg): 2–2.5	−0.269***	−0.341***	−0.201**	−0.014
	(0.071)	(0.065)	(0.086)	(0.043)
Birthweight (kg): 2.5–3	−0.199***	−0.206***	−0.080	−0.001
	(0.061)	(0.056)	(0.074)	(0.037)
Male	0.030	0.024	−0.009	−0.021
	(0.038)	(0.035)	(0.046)	(0.023)
Reduced-form estimates				
Early health shocks	−0.004	−0.263***	−0.201*	−0.449***
	(0.096)	(0.086)	(0.113)	(0.057)
Birthweight (kg): < 2	−0.317***	−0.454***	−0.275***	−0.018
	(0.081)	(0.072)	(0.094)	(0.048)
Birthweight (kg): 2–2.5	−0.235***	−0.284***	−0.125*	0.009
	(0.064)	(0.057)	(0.075)	(0.038)
Birthweight (kg): 2.5–3	−0.170***	−0.156***	−0.012	0.019
	(0.055)	(0.049)	(0.064)	(0.033)
Male	0.034	0.032	0.003	−0.018
	(0.038)	(0.034)	(0.044)	(0.022)
No. of pairs of twins	1,418	1,430	1,408	1,450

[12] The family investment in children's education is 632.48 yuan in rural areas in the past 12 months prior to the survey, which is equivalent to USD 75.

[13] This may be due to the fact that the return to schooling is significantly higher for females than that for males (Zhang *et al.*, 2005).

Table E2

(a) Early Health Shocks and Child Education (1); (b) Early Health Shocks and Child Education(2)

	Dependent variables			
	Literature		Mathematics	
	Score (1)	Relative measure (2)	Score (3)	Relative measure (4)
(a) 2SLS estimates				
Early health shocks	−3.990*	−0.168	−4.697*	−0.496***
	(2.045)	(0.145)	(2.432)	(0.157)
Education investments	6.124	0.904**	3.598	0.170
	(6.313)	(0.426)	(7.459)	(0.466)
Birthweight (kg): < 2	−2.845**	−0.136	−2.336	−0.095
	(1.395)	(0.096)	(1.695)	(0.104)
Birthweight (kg): 2–2.5	−1.157	−0.091	−0.424	−0.023
	(1.093)	(0.077)	(1.316)	(0.083)
Birthweight (kg): 2.5–3	−1.255	−0.055	−0.128	0.020
	(0.929)	(0.066)	(1.121)	(0.071)
Male	−2.755***	−0.150***	−0.609	−0.019
	(0.663)	(0.046)	(0.791)	(0.050)
Reduced-form estimates				
Early health shocks	−5.142***	−0.352***	−5.372***	−0.531***
	(1.665)	(0.110)	(1.996)	(0.127)
Birthweight (kg): < 2	−3.041**	−0.149	−2.482	−0.098
	(1.381)	(0.091)	(1.674)	(0.105)
Birthweight (kg): 2–2.5	−1.070	−0.076	−0.408	−0.021
	(1.090)	(0.073)	(1.320)	(0.084)
Birthweight (kg): 2.5–3	−1.318	−0.066	−0.166	0.018
	(0.927)	(0.062)	(1.122)	(0.071)
Male	−2.903***	−0.173***	−0.680	−0.023
	(0.645)	(0.043)	(0.780)	(0.049)
No. of pairs of twins	1,355	1,426	1,332	1,420

	Dependent variables			
	Good student awards (1)	Awards in contests (2)	Grade repetition (3)	Doing minor actions in class (4)
(b) 2SLS estimates				
Early health shocks	−0.199***	−0.067	0.025	0.296**
	(0.072)	(0.042)	(0.038)	(0.121)
Health investments	0.075	0.103	−0.235**	−0.639*
	(0.216)	(0.126)	(0.113)	(0.387)
Birthweight (kg): < 2	−0.092*	0.007	−0.008	−0.067
	(0.048)	(0.028)	(0.025)	(0.087)
Birthweight (kg): 2–2.5	−0.048	0.005	0.004	−0.098
	(0.038)	(0.022)	(0.020)	(0.069)
Birthweight (kg): 2.5–3	−0.019	0.001	0.000	0.010
	(0.033)	(0.019)	(0.017)	(0.059)
Male	−0.129***	−0.029**	−0.002	0.256***
	(0.023)	(0.013)	(0.012)	(0.041)

Table E2

(*Continued*)

	Dependent variables			
	Good student awards (1)	Awards in contests (2)	Grade repetition (3)	Doing minor actions in class (4)
Reduced-form estimates				
Early health shocks	−0.215***	−0.088***	0.073**	0.396***
	(0.058)	(0.033)	(0.029)	(0.101)
Birthweight (kg): < 2	−0.093*	0.006	−0.005	−0.056
	(0.049)	(0.028)	(0.024)	(0.084)
Birthweight (kg): 2–2.5	−0.047	0.007	0.001	−0.108
	(0.039)	(0.022)	(0.019)	(0.067)
Birthweight (kg): 2.5–3	−0.02	−0.001	0.003	0.018
	(0.033)	(0.019)	(0.017)	(0.057)
Male	−0.131***	−0.032**	0.004	0.272***
	(0.023)	(0.013)	(0.011)	(0.039)
No. of pairs of twins	1,456	1,456	1,456	1,440

Table E3

Early Health Shocks and Child Socio-emotional Skills

	Dependent variables			
	Feel lonely (1)	Easily distracted (2)	Easily frightened (3)	Emotionally unstable (4)
2SLS estimates				
Early health shocks	0.132***	0.121**	0.066	0.085***
	(0.041)	(0.056)	(0.046)	(0.024)
Education investments	−0.158	−0.150	−0.383***	−0.113
	(0.123)	(0.166)	(0.139)	(0.073)
Birthweight(kg): < 2	0.000	−0.009	−0.016	0.018
	(0.027)	(0.037)	(0.031)	(0.016)
Birthweight(kg): 2–2.5	0.016	−0.02	0.013	0.010
	(0.022)	(0.030)	(0.025)	(0.013)
Birthweight (kg): 2.5–3	0.007	−0.027	−0.015	0.009
	(0.019)	(0.025)	(0.021)	(0.011)
Male	−0.004	0.071***	−0.016	0.007
	(0.013)	(0.018)	(0.015)	(0.008)
Reduced-form estimates				
Early health shocks	0.165***	0.151***	0.144***	0.108***
	(0.032)	(0.044)	(0.033)	(0.019)
Birthweight (kg): < 2	0.002	−0.007	−0.011	0.02
	(0.027)	(0.037)	(0.027)	(0.016)
Birthweight (kg): 2–2.5	0.013	−0.023	0.007	0.009
	(0.022)	(0.029)	(0.022)	(0.013)
Birthweight (kg): 2.5–3	0.009	−0.025	−0.011	0.011
	(0.018)	(0.025)	(0.019)	(0.011)
Male	0.000	0.075***	−0.006	0.01
	(0.013)	(0.017)	(0.013)	(0.007)
No. of pairs of twins	1,456	1,456	1,456	1,450

References

Becker, G.S. and Tomes, N. (1976). 'Child endowments and the quantity and quality of children', *Journal of Political Economy*, vol. 84(4), pp. S143–62.

Behrman, J.R. and Rosenzweig, M.R. (1999). 'Ability biases in schooling returns and twins: a test and new estimates', *Economics of Education Review*, vol. 18(2), pp. 159–67.

Behrman, J.R., Pollak, R.A. and Taubman, P. (1982). 'Parental preferences and provision for progeny', *Journal of Political Economy*, vol. 90(1), pp. 52–73.

Bound, J. and Solon, G. (1999). 'Double trouble: on the value of twins-based estimation of the return to schooling', *Economics of Education Review*, vol. 18(2), pp. 169–82.

Chen L.C. and Scrimshaw, N.S. (1983). *Diarrhea and Malnutrition: Interactions, Mechanisms, and Interventions*, New York: Plenum Press.

Chen, Y., Ebenstein, A., Greenstone, M. and Li, H. (2013). 'Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy', *Proceedings of the National Academy of Sciences*, vol. 110(32), pp. 12936–41.

Ebenstein, A. (2012). 'The consequences of industrialization: evidence from water pollution and digestive cancers in China', *Review of Economics and Statistics*, vol. 94(1), pp. 186–201.

Fang, Z.Y., Wang, B., Kilgore, P.E., Bresee, J.S., Zhang, L.J., Sun, L.W., Du, Z.Q., Tang, J.Y., Hou, A.C., Shen, H., Song, X.B., Nyambat, B., Hummelman, E., Xu, Z.Y. and Glass, R.I. (2005). 'Sentinel hospital surveillance for rotavirus diarrhea in the People's Republic of China, August 2001–July 2003', *Journal of Infectious Diseases*, vol. 192(Supplement 1), pp. S94–9.

Griliches, Z. (1979). 'Sibling models and data in economics: beginnings of a survey', *Journal of Political Economy*, vol. 87(5), pp. S37–64.

Neumark, D. (1999). 'Biases in twin estimates of the return to schooling', *Economics of Education Review*, vol. 18 (2), pp. 143–8.

Rouse, C.E. (1999). 'Further estimates of the economic return to schooling from a new sample of twins', *Economics of Education Review*, vol. 18(2), pp. 149–57.

Smith, J.P. (2009). 'The impact of childhood health on adult labor market outcomes', *Review of Economics and Statistics*, vol. 91(3), pp. 478–89.

Strauss, J. and Thomas, D. (1998). 'Health, nutrition, and economic development', *Journal of Economic Literature*, vol. 36(2), pp. 766–817.

Zhang, J., Zhao, Y., Park, A. and Song, X. (2005). 'Economic returns to schooling in urban China, 1988 to 2001', *Journal of Comparative Economics*, vol. 33(4), pp. 730–52.