WPS7973

Public Disclosure Authorized

Public Disclosure Authorized

Policy Research Working Paper 7973

Dynamics of Child Development

Analysis of a Longitudinal Cohort in a Very Low Income Country

Emanuela Galasso
Ann Weber
Lia C. H. Fernald

WORLD BANK GROUP
Development Research Group
Poverty and Inequality Team
February 2017

Policy Research Working Paper 7973

Abstract

Longitudinal patterns of child development and socioeconomic status are described for a cohort of children in Madagascar who were surveyed when they were 3–6 and 7–10 years old. Substantial wealth gradients were found across multiple domains: receptive vocabulary, cognition, sustained attention, and working memory. The results are robust to the inclusion of lagged outcomes, maternal endowments, measures of child health, and home stimulation. Wealth gradients are significant at ages 3–4, widen with age, and flatten out by ages 9–10. For vocabulary and sustained attention, the gradient grows steadily between ages three and six; for cognitive composite and memory of phrases, the gradient widens later (ages 7–8) before flattening out. These gaps in cognitive outcomes translate into equally sizeable gaps in learning outcomes. Between 12 and 18 percent of the predicted gap in early outcomes is accounted for by differences in home stimulation, even after controlling for maternal education and endowments.

This paper is a product of the Poverty and Inequality Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at egalasso@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

Dynamics of Child Development: Analysis of a Longitudinal Cohort in a Very Low Income Country

Emanuela Galasso[a], Ann Weber[b], and Lia C. H. Fernald[c]

JEL codes: I14, I25, J13, O15

Keywords: Sub-Saharan Africa, child development, nutrition, socioeconomic gradients

---

Contact information: (a) Emanuela Galasso (Corresponding author) (egalasso@worldbank.org). Development Research Group, The World Bank, 1818 H St, NW, MSN MC3-306, Washington DC 20433. Tel +1-202-458-1649, fax +1-202-522-1153; (b) annweber@stanford.edu, Center for Infant Studies, Stanford University; (c) fernald@berkeley.edu, School of Public Health, University of California, Berkeley.
Acknowledgement: We thank Valerie Ranaivo and the INSTAT teams in Madagascar for their hard work on the survey and Lisy Ratsifandrihamanana for her commitment and professionality in the adaptation of the child tests. We thank the National Nutrition Office, Voahirana Rajoela and Jumana Qamruddin for their support in the research effort. We are grateful to Harold Alderman and Italo Lopez for comments. Funding for this research was generously provided by the Bank-Netherlands Partnership Program (BNPP), The World Bank Knowledge for Change Program (KCPII), and the National Nutrition Office in Madagascar.

I. INTRODUCTION

Children living in poverty are exposed to a wide range of negative influences, such as poor housing, inadequate nutrition, parental stress, and lack of cognitive stimulation (Walker et al. 2011). Additionally, these children are often exposed to multiple stressors simultaneously and repeatedly over time (Najman et al. 2009). There is consistent evidence of socioeconomic gradients for child health outcomes in both high-income countries (Case et al. 2002; Currie 2009; Shonkoff et al. 2009) and low- to middle-income countries (Currie and Vogl 2013; Petrou and Kupek 2010).

Low socioeconomic status (SES) is also a risk factor for poor child development outcomes (e.g., language, cognition, and behavior), and children of parents with less income or education perform worse than children whose parents have more income or education (Cunha and Heckman 2008; Kelly et al. 2011), according to studies from high-income countries. There is evidence of socioeconomic gradients in child development outcomes from lower and middle-income countries as well, but existing data are primarily cross-sectional. For example, wealth gradients in child development scores are evident as early as 3–23 months in India, Indonesia, Peru, and Senegal (Fernald et al. 2012). Preschool children (3–6 years old) exhibit significant gaps in receptive language in Ecuador (Paxson and Schady 2007), several other Latin American countries (Schady et al. 2015), and the four countries of the Young Lives Study (Engle et al. 2011). Unfortunately, these studies are limited by the lack of panel data with repeated measures of child development outcomes across critical and sensitive periods. Cross-sectional studies cannot untangle issues of timing where age effects are confounded with cohort effects and where the role of early skills on later development cannot be modeled. With few exceptions (e.g., [Kobrosly et al. 2011]) and the Young Lives cohorts, longitudinal panels have primarily been focused in the US and Europe.

This paper is one of the few examples of a longitudinal cohort study set in a very low-income country in Sub-Saharan Africa that has assessed children with direct measures of cognitive development at two points in time: early childhood (3–6 years) and school age (7–10 years). The cohort was also evaluated with early math and literacy measures at school age. Previous work demonstrated that these same children showed significant differences across socioeconomic status in standardized scores for a wide range of cognitive and language tests; the largest gaps in the SES gradient were evident in receptive language and sustained attention scores (Fernald et al. 2011).

The primary objective of this paper is to take advantage of these longitudinal data and examine how the magnitude and age profile of SES gradients change over time, modeling skill formation as a cumulative and dynamic process. Differences in child development are assessed across wealth groups, and the evolution of the SES gradient across early childhood and into school age is documented. While other papers based on cross-sectional evidence have already documented that the SES gradient widens at early ages and then flattens out after age six during the school-age window (Schady 2011), this study relies on both the cross-sectional and the longitudinal aspect of the data to document the critical timing of the emergence and evolution of the SES differences in child development and learning outcomes.

We further add to the literature by assessing the magnitude of the SES gradients across *multiple* developmental domains. Much of the current focus on children living in resource-poor circumstances is on height, weight, or morbidity outcomes or on a single dimension of child development (such as receptive language). In this study, multiple domains of child development are examined, including measures of cognition, language, and executive function, to examine whether the magnitude and evolution of SES gradients differ across domains. There is evidence that certain regions of the brain may be differentially vulnerable to the effects of SES (Hackman and Farah 2009), particularly language (Noble et al. 2007) and executive function (Farah et al. 2006). Our results show that receptive vocabulary and sustained attention exhibit the largest gradients after the inclusion of lagged outcomes, maternal endowments (education and vocabulary), and home environment. These outcomes also exhibit a steeper widening of the gradient with age, in contrast to the other outcomes.

Next, we assess how much of the wealth gap in child development outcomes could be accounted for by household inputs such as home stimulation. Home stimulation as measured at the time of school age is shown to account for 12-18% of the wealth gap in z-scores across the various domains with the largest findings for receptive vocabulary and memory of phrases, which is a verbal measure of working memory. The estimate represents a lower bound on the role of the home environment as it does not account for the role of lagged home inputs on developmental outcomes.

Finally, the mapping of different domains of child outcomes onto learning outcomes in math and early literacy at school age is documented. The magnitudes of the gap for school learning outcomes are found to map closely with those obtained using our contemporaneous developmental outcomes (sustained attention, receptive vocabulary, and cognitive composite).

The remainder of the paper documents the data and outcomes being studied (section 1), outlines the empirical methods (section 2), describes the results (section 3), and conclusions (section 4).

II. DATA

This analysis draws from a nationally representative survey of nutrition and child development in Madagascar (*Enquete Anthropometrique et de Developpement de l'Enfant*) that was first administered in 2004. A random subsample of 150 of the original 446 communities representative of all rural and urban areas, excluding the largest urban cities, was drawn for this longitudinal study. About 1,500 children who were ages zero to three years in 2004 were identified for a prospective cohort from these communities (Galasso and Umapathi 2007). Children's development in multiple domains (including cognitive, language, and executive function) was assessed at two follow-up periods, 2007 and 2011, when the children were three to six years (preschool age) and seven to ten years of age (primary school age). If children changed residence between surveys, they were tracked to different households or villages.

Child Development Tests

The assessments used in this study were specifically selected to focus on the domains that have been shown to be neurologically correlated with poverty or sensitive to nutrition interventions. The choices of which specific tests to use for each domain were based on the adaptability of the tests to the social and cultural realities of Malagasy children and on the practical aspects of their use given the constraints of the investigation (e.g., length of test, sample size, cost and cultural acceptability of materials, and supervision required) (Fernald et al. 2010).

A battery of six cognitive and language tests was administered both in 2007 (Fernald et al. 2011) and 2011[1]; more detail is provided in appendix S1 (figure S.1 in the supplemental appendix).

1. Essential adaptations to the assessment tools were made during the initial development of the battery of tests, followed by adjustments and improvements based on pretest results from the field. No questions were eliminated from any of the child development measures. All instructions and words or phrases were translated into Malagasy and then

Three subscales of the Stanford Binet Intelligence Scales (visual-spatial processing, working memory, and fluid reasoning) were combined to estimate children's non-verbal cognitive ability. The Leiter test was used to evaluate the capacity to maintain attention, which has been identified as a critical metric of school readiness and is an important component of executive function (Duncan et al. 2007). The Woodcock Muñoz (WM) memory of phrases test requires children to repeat words and sentences with an increasing order of difficulty; it is also considered a measure of working memory, which is a component of executive function. The Peabody Picture Vocabulary Test (PPVT-III) is a test of receptive vocabulary knowledge and was administered to the children and their primary caregivers.

Z-scores were calculated for each of the above tests from two-month, age-specific means and standard deviations in each age category.

Learning Tests and Child Enrollment

In addition to the measures of cognitive and language development, two tests of learning outcomes were incorporated when children were school age in 2011: one for mathematics and the other for early literacy skills. The tests were adaptations of the Early Grade Reading Assessment (EGRA) (Gove and Wettenberg 2011) and the Early Grade Mathematics Assessment (EGMA) (Rubens 2009) instruments developed by USAID and widely used in other French-speaking African countries. The literacy test ranged from basic assessment of auditory comprehension, recognition of letters and key basic words, to more complex ability of reading and comprehension of a text. The mathematics test includes basic number identification, comparison of numbers, basic arithmetic (addition, subtraction, multiplication, and division) and word problems involving addition and subtraction. After translation, the Malagasy versions of these tests were systematically and rigorously pretested and were found to be informative even among children who had not attended school. For example, many children knew how to count and answer the simplest word problems even if they had not attended formal schooling. Short stories that were part of the original test were rewritten during the pretesting phase so that questions associated with auditory comprehension would be accessible to children who were illiterate.

SES and Sociodemographic Characteristics

The key proxy of SES used in the analyses is a household wealth index in 2004 when children were 0–3 years, divided into quintiles. The index was constructed using principal component analysis of key durables, assets and dwelling conditions[2] (Filmer and Pritchett 2001). This SES proxy is less likely to suffer from the concern about measurement error that generally affects income variables. Furthermore, an index for wealth is often used in the analysis of SES gradients in child outcomes in both developed and developing country contexts, and commonly adopted in descriptive analysis of Demographic and Health Surveys. As described in greater detail below, a full set of dummy variables for wealth quintiles were included in our regressions, and the results

---

translated back into English to ensure linguistic and functional equivalence. More information on each child test is available in appendix 1.

2. The wealth index is the first principal component of the following variables: walls, floors, roof, home ownership, whether the toilette is shared, whether household has access to safe water, ownership of chairs, phone, motor vehicle, cattle ownership, and poultry ownership. The sample is representative of rural and small urban villages (excluding large cities in all regions and the capital).

included in appendices. In the key summary tables, however, estimates are presented contrasting the wealthiest, or fifth, quintile with the poorest, or first, quintile. Household density (or crowding) is also included not as a direct proxy for SES, but as an additional regressor, documented to be an environmental stressor and important determinant of child wellbeing (Evans 2004).

Child Nutritional Status

Two measures of nutritional outcomes are included: birth weight (to reflect nutrition during gestation) and height-for-age z-score, HAZ (to reflect early childhood nutrition). For children with recorded birth weight information (48%), actual birth weight was highly correlated with maternal report of perceived birth weight ($r=0.60$, $p<0.001$), suggesting that perceived birth weight is a reasonable proxy for actual recorded birth weight.[3] Thus, perceived birth weight was included in the regressions (binary indicator variable for perceived "low" or "very low" birth weight relative to "normal," "high," or "very high"). Significant associations exist between low height-for-age (stunting) and delayed cognitive development, psychomotor development, behavior, and later success in school (Grantham-McGregor et al. 2007). In Madagascar, the prevalence of stunting among children was 58% at our first measurement when children were 0–3 years.[4]

Maternal Education and Endowments

Maternal education is expressed with indicators of completed primary or secondary education. In addition, maternal height and maternal receptive vocabulary are included as proxies for maternal ability. Maternal height is an important proxy for maternal cognitive development (Case and Paxson 2008) and represents a mechanism for the intergenerational transmission of economic status since it represents maternal genetic endowment in addition to her own prenatal and childhood exposure. Children born to mothers who are shorter are more likely to have lower birth weight/length and higher neonatal mortality than those born to taller mothers (Bhalotra and Rawlings 2013). To measure maternal receptive vocabulary, the same test used for children was used for mothers, the PPVT-III. Maternal vocabulary ability was estimated using item-response theory (IRT) to ensure a common metric for the variable across the sample (Masters and Wright 1984).

Parenting Inputs

To measure parenting inputs, an adaptation of the Home Observation for Measurement of the Environment Inventory (HOME) for Middle Childhood (MC) in 2011 was administered. The adapted HOME scale captures learning environments which represent key pathways through which poverty can affect school readiness; the HOME scale captures both risk factors that increase vulnerability (e.g., harsh parental punishment, lack of sensitive parenting, limited books or toys in the home) as well as buffering factors that can help children thrive in spite of poverty (e.g., exposure to responsive parenting, presence of books and toys in the home) (Blair and Raver 2012;

3. Children whose mothers assessed their weight as small, and children with low measured weight have been shown to be significantly less likely to survive than are children who weigh more at birth (Alderman et al. 2011).

4. The prevalence of stunting in our sample is slightly higher than the prevalence reported by UNICEF (50%, 2013) and the DHS (53% in children 6–59 months old, 2008–9), as our sample is representative of rural and small urban centers in the country, excluding the capital and larger cities.

Bradley and Corwyn 2002; Burchinal et al. 2000; Duncan and Brooks-Gunn 1997; Evans 2004; Hart and Risley 1995; Hoff 2003; Walker et al. 2011; Yeung et al. 2002). We focused on two subdomains from the HOME that were found to be correlated with math and literacy scores among children 7–10 years of age in the reference population: the domain relating to fostering growth and education and the domain relating to the provision of active stimulation (Bradley and Corwyn 2005).[5] The adapted HOME scale was included in the model as an ordinal scale from 0–10, with one point per item (see appendix S2 for complete list of questions).

III. Empirical Methods

The empirical methods described below for the analyses are parametric in nature, and, as such, are correlational. We are careful throughout the analysis in claiming neither causality nor inference about structural parameters in our results.

Magnitude and Evolution of the Socioeconomic Gradients

The primary objective of our analysis is to document the magnitude of the association between child development outcomes and disparities in SES and to document its evolution over time. We have a multidimensional set of domains (indexed by j) with data on cognition (visual spatial processing, working memory, and reasoning), language, and executive function spanning the preschool years when children were 3–6 years old to the school years when children are seven and older. For disparities in SES, household wealth quintiles from when children were 0–3 years are used[6].

In a parametric framework, SES gradients are estimated by OLS using the following specification (Todd and Wolpin 2007):

$$Z_{ijt} = \alpha + + \sum_{k=2}^{5} \delta_k^j(a) * Q_{ik} + X_{it}^{'}\gamma + \varepsilon_{ijt} \qquad j = 1, \dots, J \qquad t=1,2 \quad (1)$$

where Z_{ij} is the adjusted z-score (internally standardized) obtained for child i on the developmental domain j at time t. δ_k^j is the magnitude of the gradient for k-th quintile Q_{ik}, expressed as the size of the difference in mean z-scores obtained by wealth quintile Q_{ik} relative to the bottom quintile (k=1, the excluded category). The error term includes unobserved child and household characteristics clustered at the village level, which is our primary sample unit. The base specification is performed on the pooled sample and includes demographics (age and gender) and the full set of wealth quintile dummies (and urban and regional dummies). The evolution of the gap over time is modeled by adding a full set of interactions between quintile indicators and age dummies, thereby allowing δ_k^j to flexibly vary with age.

The robustness of the results is assessed by controlling for a large set of childhood and parental characteristics X_{it}. First, the longitudinal dimension of the panel is exploited and lagged child

5. Items were adapted to what was appropriate for a mostly poor and rural Malagasy context based on input from a local clinical psychologist.

6. Results are robust to using wealth rankings using the contemporaneous wealth indicators. These results are not presented but are available upon request.

development outcomes added in specification (i), which captures the full range of past observed and unobserved inputs as well as child-specific early conditions (perceived birth weight and stunting status, which are used as proxy for early investments in health and nutrition during the early years).[7] Subsequently, time-invariant parental endowments are added such as maternal education and height and receptive vocabulary, in specification (ii). Finally, variation in the results is assessed by the inclusion of key observed inputs such as the HOME, in specification (iii).

Closing the Gap

The estimates in (1) are used to assess the extent to which differences in inputs (such as the adapted HOME) can account for the magnitude of the estimated wealth gradients in child development. Specifically, the estimates of the full specification (iii) are used to assess how much of the gap in test scores can be closed by equalizing those determinants to the level of the highest wealth quintile, all other things being equal. Note that the adapted HOME environment scale is available only in 2011. Lagged inputs have been shown to be jointly significant determinants of child development in other settings (as in Todd and Wolpin 2007). In the full specification (iii), lagged HOME environment is subsumed into lagged child outcomes and therefore unaccounted for in the closing the gap exercise. As such, these estimates will likely be a lower bound estimate on the role of differences in inputs in accounting for differences in child outcomes.

Mapping School Readiness to School Learning Outcomes

Finally, we assess how the gaps in the child development scores map to learning outcomes at time t=2. As documented in the data section, the learning outcomes are censored: a fraction of the children in our sample (as in other developing countries such as in Paxson and Schady 2007) are not able to attain a basic score in either literacy or math. Censoring will bias OLS estimates of the gradient downwards (Chay and Powell 2001). Since estimation with MLE is consistent only under normality and homoscedasticity, a censored least absolute deviations (CLAD) estimator is used to estimate the following equation:

$$Y_{iL2} = \max(c, a + \sum_{k=2}^{5} \delta_{kL} * Q_{ik} \quad + \sum \beta^{j} Z_{ij2} + X_{i2}' \gamma^{L} + e_{iL2}) \qquad (2)$$

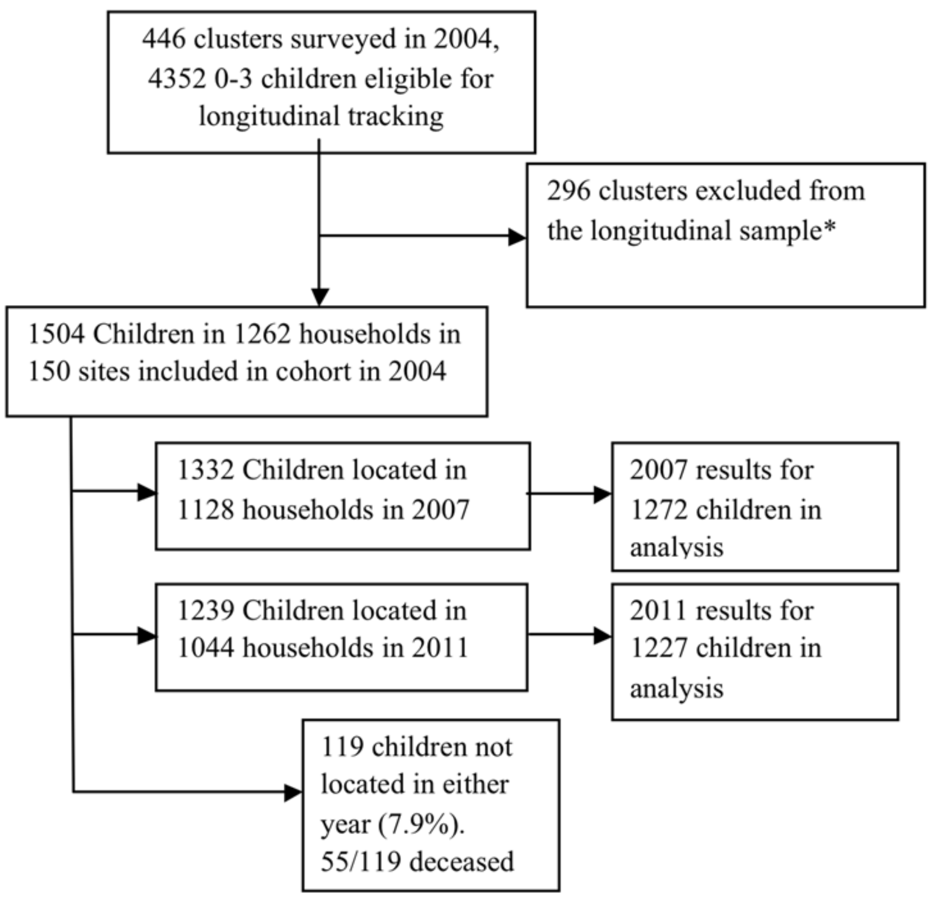
where Y_{iL2} are the observed, age standardized, learning (literacy and math) scores. For observed scores that are above the threshold c, learning outcomes are modeled as a function of SES, and an increasing set of controls as above (lagged outcomes Z_{1ij}, maternal education, and endowments, HOME). Finally, sensitivity to the inclusion of the concurrent/contemporaneous child development scores (Z_{2ij}) is checked. As in Cunha and Heckman (2007), this exercise has no causal interpretation, but it serves the purpose of mapping the magnitude of child development scores across different domains onto a metric of school achievement and learning.

7. While including a lagged dependent variable can lead to biased estimates if there is measurement error and latent heterogeneity in ability or endowments, in practice the extent of the bias in the dynamic value-added model appears to be limited in related education studies when the estimates are compared to experimental or quasi-experimental estimates (Angrist et al. 2013, Andrabi et al. 2011). It is also important to point out that we are not interested in estimating structural parameters in this setting but to describe the magnitude and evolution of the wealth gaps in child development.

IV. RESULTS

Of the 1504 age-eligible children in 2004, 92.1% (n=1385) were tracked in either or both years (figure 1). Only 119 children (7.9%) could not be found in either 2007 or 2011, with about half of those who could not be found reported to be deceased (n=55). The children lost to follow-up in both years were nearly three months younger on average than the children who were tracked; they were not significantly different by gender, birth weight, maternal education, or household crowding in 2004.

Figure 1. Sample attrition and selection



Note: Due to cost-considerations, a random subsample of 150 of the 446 communities (excluded 296), representative of all rural and urban areas excluding the largest urban cities, was selected for a longitudinal study

Table 1 presents characteristics of the sample in 2011, overall and by wealth quintiles. Wealth quintiles capture differences in housing conditions, assets, and livestock ownership. Wealth quintiles are evenly represented in urban and rural areas, with a slightly larger share of urban households in the top wealth quintile. Nearly a quarter of the mothers had received no formal education (up to half in the lowest wealth quintile) and about one-fifth of the children were perceived to be born with low birth weight by mothers (up to one-quarter in the lowest wealth quintile).

Table 1. Sample characteristics: overall and by wealth quintiles

	Wealth quintiles					*Whole sample*
	Q1	*Q2*	*Q3*	*Q4*	*Q5*	
<u>*Village characteristics*</u>						
Village classification ("urban"=1)	0.212	0.105	0.155	0.110	0.328	0.179
Household characteristics						
Rooms (number)	1.319	1.458	1.528	1.827	1.980	1.615
	(0.775)	(0.881)	(0.701)	(0.975)	(1.107)	(0.926)
House has cement floors (yes=1)	0	0.018	0.085	0.230	0.543	0.168
Access electricity (yes=1)	0	0	0	0	0.277	0.051
Thatched roof (yes=1)	0	0.014	0.057	0.049	0.113	0.045
Iron-concrete roof (yes =1)	0.962	0.921	0.833	0.739	0.539	0.804
Cooking fuel: charcoal (yes =1)	0.003	0.018	0.092	0.159	0.512	0.150
Cooking fuel: wood (yes=1)	0.997	0.975	0.904	0.834	0.465	0.842
Toilet open pit (yes=1)	0.014	0.152	0.327	0.551	0.766	0.354
Household has no toilet (yes=1)	0.986	0.819	0.659	0.410	0.148	0.614
House walls: bricks (yes=1)	0.017	0.058	0.131	0.201	0.270	0.133
Drinking water: public well (yes=1)	0.410	0.332	0.313	0.219	0.176	0.292
Drinking water: river (yes=1)	0.372	0.289	0.195	0.173	0.082	0.225
Own radio (yes=1)	0.188	0.473	0.567	0.731	0.844	0.554
Own tv (yes=1)	0.003	0.000	0.000	0.007	0.262	0.051
Own cart or bicycle (yes=1)	0.174	0.224	0.279	0.410	0.496	0.313
Own cattle or sheep (yes=1)	0.431	0.430	0.407	0.438	0.266	0.397
Own poultry=1	0.684	0.668	0.673	0.742	0.656	0.685
Home language local dialect (yes=1)	0.951	0.845	0.741	0.664	0.578	0.760
Maternal characteristics						
Unschooled =1	0.462	0.260	0.245	0.155	0.047	0.238
Primary education =1	0.469	0.628	0.620	0.625	0.430	0.556
Secondary/higher education =1	0.069	0.112	0.135	0.219	0.523	0.206
Height (cm)	153.14	152.48	152.63	152.03	153.27	152.70
	(5.83)	(6.17)	(6.33)	(5.92)	(6.44)	(6.15)
Receptive vocabulary* (log odds ratio)	-0.143	-0.054	0.015	0.179	0.587	0.112
	(0.484)	(0.537)	(0.567)	(0.602)	(0.713)	(0.636)
Child characteristics						
Male =1	0.451	0.455	0.480	0.516	0.559	0.491
Birth order =1	0.212	0.209	0.192	0.180	0.270	0.211
Birth order=2	0.160	0.199	0.224	0.230	0.234	0.209
Birth order=3 or higher	0.125	0.166	0.176	0.166	0.164	0.159

Perceived birth weight (=1 if low or very low)	0.271	0.213	0.202	0.141	0.164	0.199

*Maternal vocabulary represents estimated ability for the Peabody Picture Vocabulary Tests obtained from an item response model. Ability is reported in log odds units (logit) and ranges between -2 to 2.5 logits.

A comparison of descriptive statistics for the 2011 test scores by wealth quintile is presented in table 2. For all the child development test scores, there is a larger difference in test scores when comparing the middle wealth quintiles relative to the richest quintile (Q5 versus Q3) than from comparing the middle quintile to the poorest one (Q3 versus Q1). Across all age groups, 30% of our 2011 sample of school-age children did not attend school. The percentage of seven-year-olds who are not enrolled in school in our sample is about 50%, decreasing to 13% at age ten years. This pattern of delayed enrollment exhibits important differences across wealth quintiles, with 40% of children in the lowest wealth quintile not in school as opposed to 6% in the highest wealth quintile. Delayed enrollment has been shown to have important economic costs in terms of total grades completed (Alderman et al. 2001; Glewwe and Jacoby 1995). The learning test used in this study is designed to be administered to children both in and out of school; in spite of the fact that the test is supposed to be flexible, performance is highly censored, with three-quarters of the children not attending school reaching a high score of five (median two) in the literacy test as opposed to 40 (median 18) among those attending school (figure A1 in the appendix). As a consequence, wealth gaps in learning outcomes reflect a combination of delayed age of enrollment, as well as performance conditional on enrollment.

Table 2. Descriptive statistics (mean and SD): child development outcomes 2011, by wealth quintile

	Means (SD) by quintile					Test difference in means:		
	Q1	Q2	Q3	Q4	Q5	Q5-Q3	Q3-Q1	Q5-Q1.
Sustained attention z-score 2011	-0.579	-0.233	-0.037	0.189	0.688	-0.728***	-0.540***	-1.268***
	(0.814)	(0.975)	(0.917)	(0.865)	(0.927)	(0.084)	(0.078)	(0.080)
Memory of phrases z-score 2011	-0.569	-0.228	-0.052	0.174	0.687	-0.738***	-0.518***	-1.256***
	(0.688)	(0.858)	(0.922)	(0.991)	(1.010)	(0.089)	(0.074)	(0.079)
Cognition composite z-score 2011	-0.587	-0.247	-0.101	0.178	0.753	-0.855***	-0.486***	-1.341***
	(0.680)	(0.837)	(0.878)	(0.962)	(1.031)	(0.088)	(0.072)	(0.081)
Vocabulary z-score 2011	-0.686	-0.297	-0.106	0.226	0.910	-1.016***	-0.580***	-1.596***
	(0.641)	(0.819)	(0.855)	(0.899)	(0.953)	(0.083)	(0.068)	(0.074)
Adapted HOME scale 2011	3.657	4.645	5.390	6.081	7.170	-1.77***	-1.74***	-3.51****
	(2.247)	(2.240)	(2.259)	(2.218)	(1.786)	(0.19)	(0.20)	(0.19)
Child unschooled=1 2011 (proportion)	0.390	0.324	0.236	0.124	0.065	0.168***	0.157***	0.325***
Literacy raw score 2011 if not in school	2.22	4.40	3.68	2.73	10.20	-6.52**	-1.46*	-7.98***
	(3.64)	(6.45)	(5.05)	(2.85)	(11.26)	(2.08)	(0.81)	(1.64)
Literacy raw score 2011 if in school	11.50	16.08	20.67	25.47	36.69	-16.05**	-9.17****	-25.19***
	(13.72)	(15.89)	(18.07)	(20.42)	(19.39)	(1.87)	(1.82)	(1.89)
Literacy z-score 2011	-0.591	-0.350	-0.133	0.175	0.813	-0.944***	-0.460***	-1.404***
	(0.626)	(0.724)	(0.860)	(1.011)	(1.013)	(0.087)	(0.072)	(0.082)
Math raw score 2011 if not in school	13.78	17.20	19.46	22.21	28.05	-5.85**	-1.20	-7.05***
	(8.23)	(9.19)	(9.24)	(9.59)	(8.75)	(2.14)	(1.04)	(1.70)

Math raw score 2011 if in school	17.08	20.23	22.29	23.90	28.94	-5.21***	-6.65***	-11.86***
	(7.87)	(8.15)	(7.76)	(8.69)	(8.07)	(0.84)	(0.70)	(0.85)
Math z-score 2011	-0.674	-0.276	-0.072	0.231	0.826	-0.894***	-0.606***	-1.500***
	(0.816)	(0.876)	(0.871)	(0.906)	(0.818)	(0.078)	(0.076)	(0.075)

Note: The last three columns (Q5-Q3, Q3-Q1, Q5-Q1) provide the mean and the standard error of the difference in means across two quantile groups: *,**,*** denote significance of at 10%, 5% and 1% levels.

Although the mathematics test is well distributed, the literacy test is positively skewed. The bimodal distribution of scores for the literacy test is consistent with responding correctly to the comprehension questions administered after listening to a short story; any additional points for literacy required some knowledge of letters. In the analysis, the literacy test is defined to be censored at zero and the mathematics test to be censored at a raw score of five, representing about 12% of our non-missing responses in both cases.

Magnitude of the Socioeconomic Gradients

Table 3 presents the estimated wealth coefficients across four specifications. For parsimony, table 3 reports only coefficients for the top wealth quintile relative to the first quintile, but the full set of coefficients on all wealth quintiles and other regressors are presented in table A1 in the appendix. The four specifications are expressed in the different columns. The base model includes demographics and wealth (column base). The following group of regressors are added consecutively to test for robustness: (i) lagged test values and early child conditions; (ii) maternal education and endowments; and (iii) home environment.

Table 3. Magnitudes of SES gradient (Q5 v Q1), whole sample

	(base)	*(i)*	*(ii)*	*(iii)*
Receptive vocabulary	1.060***	0.892***	0.672***	0.532***
	(0.085)	(0.094)	(0.108)	(0.100)
Cognition composite	0.893***	0.601***	0.383***	0.288***
	(0.079)	(0.098)	(0.112)	(0.109)
Sustained attention	0.810***	0.646***	0.483***	0.400***
	(0.085)	(0.109)	(0.113)	(0.107)
Memory of phrases	0.715***	0.583***	0.423***	0.308***
	(0.079)	(0.109)	(0.117)	(0.111)
Demographics	Yes	Yes	Yes	Yes
Lagged outcomes & early conditions	No	Yes	Yes	Yes
Maternal education & endowments	No	No	Yes	Yes
Home environment	No	No	No	Yes

Note: The dependent variable is the standardized score of each test, normalized to have mean zero and standard deviation of one by 2-months age groups. Sample size is 1,234. Estimation is done using ordinary least squares, with robust standard errors clustered at the village level. Each cell presents the estimated coefficient (and robust standard error) on quintile 5 relative to quintile 1 for each specification. Coefficients on quintiles 2-4 are not reported for parsimony. Demographics include age (in years) dummies, gender, birth order dummies, urban and regional dummies. Lagged outcomes include all lagged z-scores as measured in 2007, as well as lagged stunting indicators and low birth weight indicators. Maternal endowments include the log of maternal height, maternal education defined as completed primary, secondary or higher (relative to no schooling), the maternal vocabulary score, and a crowding indicator. Home environment includes the total HOME score. *,**,*** denote significance at 10%, 5% and 1% levels.

The coefficients on wealth quintiles are significant for all outcome domains tested and are robust to the inclusion of a rich set of regressors across all four specifications. In the base model, accounting only for demographic variables, the estimated gap in z-scores between the fifth and first wealth quintiles ranges from approximately 1.1 SD for vocabulary to 0.7 SD for memory of phrases. After adding lagged child development outcomes and early conditions (i), the association between SES and child development outcomes for the school age period are reduced across all domains but still statistically significant.

The estimated wealth differences drop further when maternal education and endowments (ii) are added and again after adding home environment (iii); the magnitude of the estimated differences between the bottom and top wealth quintiles are reduced by up to one-half of the base model estimates while remaining statistically significant. The largest gradients, after controlling for all observed explanatory variables, are observed for receptive vocabulary and sustained attention.

In the full set of results presented in table A1, maternal endowments (education and vocabulary) are highly significant in all specifications, over and above wealth. Importantly, the home environment is significant in addition to controlling for maternal endowments, suggesting an important mediating role in family inputs explaining wealth differences in child development.

Evolution and Timing of Emergence of Wealth Gradients

In this section, the evolution of the wealth gradient is modeled across ages, with a fully flexible parametrization of wealth and age interactions. For the purpose of this analysis, ages 3–4 in 2007 and 6–7 in 2011 are merged to balance the size of each age category. Table 4 presents the coefficients of the top wealth quintile (relative to the bottom wealth quintile) at age 3–4 in the base model (first row). (The full set of results for all age and wealth quintile interactions are reported in table A2 in the appendix.) The bottom part of the table presents tests for the difference in the magnitude of the coefficients of the top wealth quintile interactions at different ages.

Table 4. Timing of wealth gradient widening

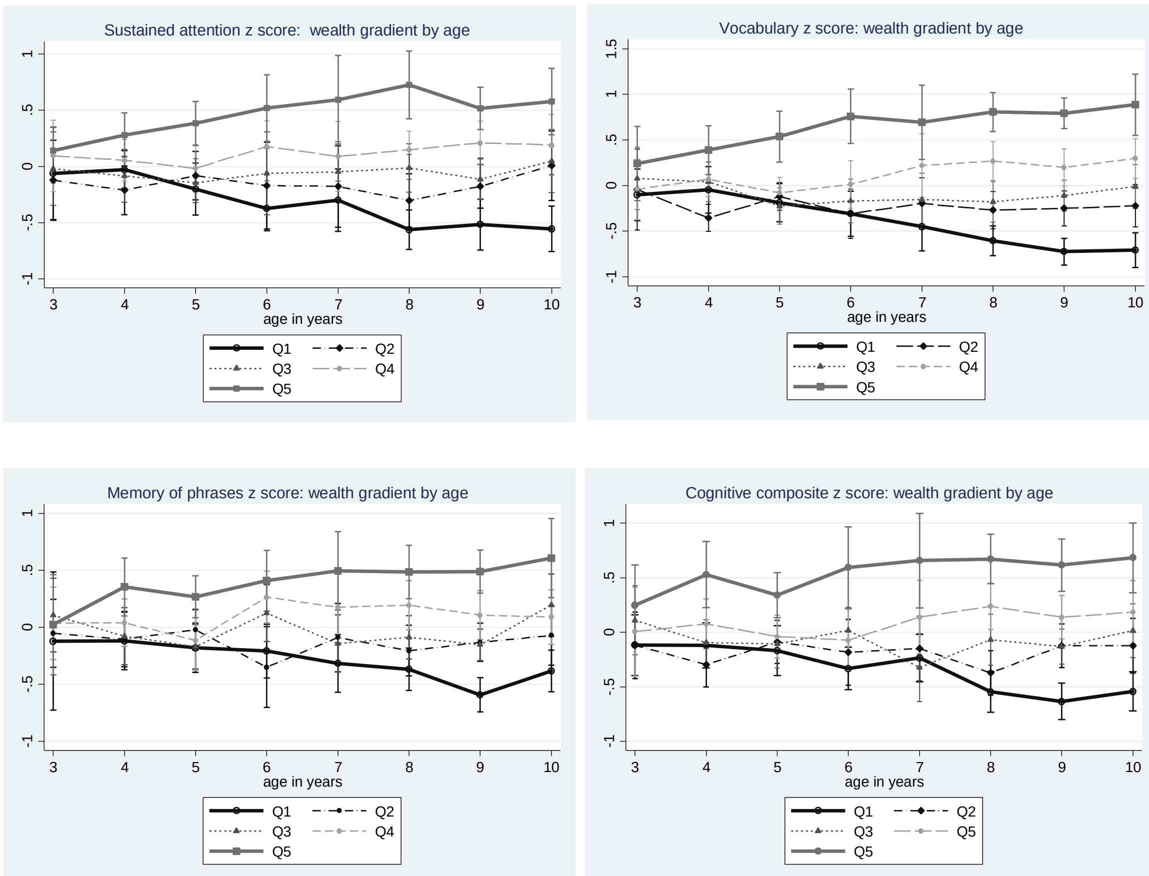
	Vocabulary	Cognitive composite	Sustained attention	Memory of phrases
Q5 (vs Q1) at age 3-4	0.414**	0.572***	0.280**	0.389**
	(0.168)	(0.152)	(0.116)	(0.177)
Difference Q5 (vs Q1) at different ages:				
Q5 (vs Q1) at age 5 vs age 3-4	0.312	-0.065	0.304*	0.058
	(0.234)	(0.210)	(0.162)	(0.207)
Q5 (vs Q1) at age 6 vs age 3-4	0.656**	0.351	0.611**	0.228
	(0.223)	(0.247)	(0.194)	(0.218)
Q5 (vs Q1) at age 7-8 vs age 3-4	0.915***	0.542**	0.886***	0.450**
	(0.177)	(0.192)	(0.180)	(0.211)
Q5 (vs Q1) at age 10 vs age 7-8	0.265	0.108	0.033	0.150
	(0.210)	(0.197)	(0.212)	(0.193)

Note: The dependent variable is the standardized score of each test, normalized to have mean zero and standard deviation of one by 2-months age groups in each wage of the full sample. Sample size is 1,234. Estimation is done using ordinary least squares, with robust standard errors are clustered at the village level. The first row present the estimated coefficient (and robust standard error) on quintile 5 relative to quintile 1 for the ages 3-4 (excluded category). Coefficients on quintiles 2-4 are not reported for parsimony. Demographics include age (in years) dummies, gender, birth order dummies, urban and regional dummies. The bottom part of the table presents the test of the difference between the coefficients on quintile 5*age interactions at different ages.
*,**,*** denote significance at 10%, 5% and 1% levels.

Gradients are large and significant by ages 3–4 for all outcome measures, in line with the evidence in the literature of the early emergence of SES gradients. Two patterns emerge from the comparison of the magnitude of the SES gradient at different ages. First, for vocabulary and sustained attention, the gradient grows steadily, as evidenced by the comparison of the coefficients of the interaction between wealth quintile and the corresponding ages (age five versus age 3–4, age six versus age 3–4, and age 7–8 versus age 3–4). The second contrasting pattern is for the cognitive composite and memory of phrases where the gradient widens later (age 7–8 versus age 3–4), which is the time of primary school enrollment. For all domains, the gradient flattens out at age 9–10, with no statistically significant difference between ages 7–8 and ten years (last row).

Figures 2 (a–d) mirror the results in table 4, with a visual representation of the evolution of the SES gradient: each figure displays the estimated coefficients of the full set of wealth quintiles age interactions plotted for a given outcome, with the corresponding 95% confidence intervals. The SES gradient for sustained attention originates early, widens up to six years of age but then flattens out after seven years old (figure 2a). For receptive vocabulary, the wealth gradient widens from age 4–6 and then continues to widen after age six (figure 2b). The widening across age is less marked for the memory of phrases and the cognitive composite (figures 2c and 2d).

Figure 2. SES age gradient estimates using cross-sectional data for four outcomes



Note: The graphs plot the coefficient estimates of the age and SES (wealth and maternal education) interaction, based on the cross sectional model (1) where z-score are regressed on a full set of interaction between age and SES, survey year, gender and birth order of the child, crowding, urban and regional dummies. Standard errors are clustered at the village level.

Closing the Gap

Next, we follow Todd and Wolpin (2007) and estimate how much of the predicted wealth gap in early outcomes is accounted for by differences in a key determinant of early child development: home stimulation. Table 5 presents how much of the wealth gap would be reduced if children from the lowest wealth group received the same inputs (e.g., home stimulation) on average as those received by children in the high wealth group, holding everything else constant.

Table 5. Closing the wealth gap in child outcomes

		Reduction Q5 vs Q1 obtained by equalizing home inputs	
	(1)	(2)	(3)
Outcome	Predicted wealth gap	Absolute gap reduction	Percent gap reduction
Receptive vocabulary	1.60 [1.43, 1.77]	0.253 [0.16, 0.34]	15.6% [9.97%, 21.3%]
Cognition	1.32 [1.14, 1.5]	0.173 [0.08, 0.26]	13.0% [6.10%, 19.9%]
Sustained attention	1.24 [1.06, 1.42]	0.151 [0.06, 0.24]	12.1% [4.61%, 19.5%]
Memory of phrases	1.22 [1.04, 1.4]	0.218 [0.13, 0.3]	17.8% [10.2%, 25.3%]

Note: Predicted gaps obtained from the full model (iii) with lagged test values and early child conditions, maternal education and endowments; and home environment. The cells in equalizing inputs use model estimates, where the home environment variable input for low SES (Q1) is set to be equal to the mean value of high SES (Q5), all else being equal. Confidence intervals were constructed from a non-parametric bootstrap with 10,000 replications.

The predicted mean difference in the outcomes between high and low wealth groups (i.e., the "gap") is approximately 1 to 1.6 SD in the age-adjusted z-scores (table 5). Home stimulation accounts for 12-18% of the wealth gap in z-scores across the various domains. The largest absolute reductions in the gap (column 2) are found for outcomes that appear to be more sensitive to verbal communication: receptive vocabulary and memory of phrases, which is a verbal measure of working memory.

Mapping School Readiness to School Learning Outcomes

We next turn to mapping the child development outcomes to learning outcomes. Table 6 presents the association between wealth and our two school learning outcomes, math and early literacy, estimated with a CLAD model that accounts for censoring (equation 2). Each cell represents the coefficient of the top wealth quintile (relative to the bottom quintile) in math and literacy scores across different specifications. The results from the base model show sizeable gaps in learning outcomes across SES groups: the differences in math and literacy outcomes between the top and bottom quintiles are about 1.3 SD (table 6, base specification). As with the school readiness outcomes (table 3), wealth differences remain robust to the inclusion of lagged developmental outcomes and early inputs as well as maternal education and endowments. The wealth gap remains significantly associated with math but borderline with literacy, with the inclusion of our proxy for the home environment (table 6, specifications [i]–[iii]). The magnitudes of the gap for school learning outcomes are comparable to those obtained for school readiness across each of the model specifications (i)– (iii).

In specification (iv), contemporaneous developmental outcomes are included in the model, and the wealth gradient ceases to be significant. These results suggest that the school-age developmental outcomes account for the remaining associations between learning outcomes and wealth, suggesting a close mapping of our school-based outcomes with our contemporaneous developmental outcomes (sustained attention, receptive vocabulary, and cognitive composite).

Table 6: Magnitudes of SES gradient for school learning outcomes (math and literacy)

		Base	(i)	(ii)	(iii)	(iv)
Math	Q5 vs Q1	1.315	0.904	0.684	0.485	0.213
		[0.979, 1.651]	[0.569, 1.239]	[0.319, 1.05]	[0.126, 0.845]	[-0.074, 0.5]
	R^2	0.11	0.148	0.163	0.173	0.26
	N. obs	1003	997	909	906	906
Literacy	Q5 vs Q1	1.317	0.74	0.519	0.395	-0.039
		[0.954, 1.68]	[0.372, 1.108]	[0.116, 0.921]	[-0.015, 0.806]	[-0.401, 0.323]
	R^2	0.125	0.165	0.194	0.199	0.273
	N.obs	960	955	878	876	876
Demographics		Yes	Yes	Yes	Yes	Yes
Lagged outcomes & early conditions		No	Yes	Yes	Yes	Yes
Maternal education & endowments		No	No	Yes	Yes	Yes
Home environment		No	No	No	Yes	Yes
Concurrent ability		No	No	No	No	Yes

Note: The dependent variable is the standardized score of each test, normalized to have mean zero and standard deviation of one by 2-months age groups in each wage of the full sample. Estimation is done using censored least absolute deviations estimator (CLAD) and bootstrap estimates of its sampling variance. We present the estimated coefficient (and 95% confidence intervals of bootstrap estimates using the normal approximation) on quintile 5 relative to quintile 1, and the pseudo R-squared for each model. Coefficients on quintiles 2-4 are not reported for parsimony. Demographics include age (in years) dummies, gender, birth order dummies, urban and regional dummies. Lagged ability and early conditions include all lagged z-scores for developmental measures as assessed in 2007, as well as lagged stunting indicators and low birth weight indicators. Maternal endowments include the log of maternal height, maternal education defined as completed primary, secondary or higher (relative to no schooling), the maternal vocabulary score, and the crowding indicator. Home environment includes the total HOME score. Concurrent ability include the concurrent developmental measures as assessed in 2011.

.

The full set of regressors is presented in appendix table A2. Lagged cognition is a significant predictor of math across the different specifications (i), (ii), and (iii). In addition, all four lagged readiness outcomes are significantly associated with literacy at school age in specification (i).

IV. Conclusions

There are wealth gradients in child development that start early, change over time, and accumulate with age. Data from a longitudinal cohort of children living in Madagascar are used to document that, even in the context of a very low income country where more than 2/3 of the population lives below the poverty line, wealth gradients are substantial. The gradients are robust to the inclusion of a large set of explanatory variables, from lagged outcomes and early conditions to maternal education and endowments to the inclusion of parenting inputs.

Our four main findings are as follows. First, the SES gradients observed at preschool age are shown to persist into school age across all developmental domains tested. Substantial gradients in school readiness outcomes by household wealth are evident when the children were 3–4 years old; these gradients widen with increasing child age and flatten out by ages 9–10 for all domains. Second, the magnitude and evolution of the wealth gradient are found to vary by domain, widening earlier for receptive vocabulary and sustained attention (between ages three and six) and later for the cognitive composite and memory of phrases (between ages 7–8). As reported previously, when the children were 3–6 years, vocabulary and sustained attention demonstrate the largest gradients when these same children are 7–10 years after controlling for lagged outcomes, maternal endowments, and home environment. Third, an estimated 12–18% of the predicted wealth gap in school readiness at ages 7–10 can be accounted for by differences in home stimulation at the same age, even when controlling for early skills and initial endowments. Finally, there are also significant SES gaps in learning outcomes (math and early literacy) when the children are 7–10 years old, mapping closely onto the gaps for our school readiness outcomes. As with the readiness measures, the significance of the wealth gradients in learning outcomes is robust to the inclusion of lagged outcomes and initial endowments.

Our results align with previous evidence that measures of executive function and language would be potentially more sensitive to the effects of household wealth or maternal education. The differential widening gradient may be due, in part, to the differential timing of development of various regions of the brain (Lenroot and Giedd 2006), although we do not have the data to understand these mechanisms.

The timing and widening of the SES gradients confirms the importance of investing during the early years, because of the brain's greater plasticity (i.e., capacity to benefit from environmental interventions) and the timing of physiologic development (Shonkoff 2011). These developmental benefits of early intervention are supported by a strong economic argument for intervening as early as possible (Doyle et al. 2009). During the first five years of life, children's language, early understanding of mathematics and reading, and self-control emerge. The extent to which children master these skills during this sensitive period has implications for success in school (Lerner 1998). There is also evidence that performance in early measures of executive function is associated with performance in later measures of academic skill, such as early math and reading ability (Blair and Razza 2007).

In addition to highlighting the importance of addressing poverty in early childhood, we find that SES continues to play a critical role into middle childhood. Our estimates from Madagascar of how much of the wealth gap in developmental outcomes can be accounted for by differences in the home environment (12–18%) are in line with studies in middle- and high-income countries. In the analysis of an SES gradient in Ecuador, a measure of parenting was found to be associated with a 21% decrease in the effect of maternal education and a 41% decrease in the effect of paternal education (Paxson and Schady 2007). Todd and Wolpin (2007) found that equalizing HOME inputs accounted for about 10–20% of the racial gap in test scores in the US. We acknowledge that our estimates represent a lower bound estimate of the effects of home environment as the effect of lagged home investments is subsumed by lagged outcomes in our analysis.

Our results related to "closing the gap" have significant policy implications. We have shown that bringing home stimulation in all groups up to the mean score of the higher SES group would result in a substantial reduction of the SES gradient for several key outcomes. While it is hard to influence SES disparities in the short term, our research and that of others suggest that programs that improve home environments and early child stimulation would have the potential to close 10–20% of the wealth gap. A full assessment of the specific policies would require more in-depth knowledge about the costs of such policies as well as knowledge of how mothers and families make decisions about the inputs they provide for their children, which is beyond the scope of our paper. We are currently investigating the possible benefits (and associated cost) of introducing an integrated nutrition and early stimulation intervention in Madagascar that encourages parents to talk, play, and provide cognitive stimulation to their infants and young children (aged 0–3 years) for improved development (Fernald et al. 2016). The magnitudes of the wealth gap for school learning outcomes is comparable to those obtained for school readiness even when controlling for lagged outcomes, maternal endowments, and home environment. Additionally, the wealth gradient for school learning outcomes ceases to be significant when contemporaneous school readiness outcomes are added to the model, indicating that the school-age developmental outcomes account for the remaining association between learning outcomes and wealth. This close mapping between school readiness and learning outcomes suggests that school readiness measures are a reasonable way to assess children's potential without censoring in settings of low school attendance, which is instrumental in projecting the estimated wealth differences in child outcomes to adult outcomes.

REFERENCES

Alderman, H., J. Behrman, V. Lavy, and R. Menon. 2001. “Child Health and School Enrollment: A longitudinal analysis.” *Journal of Human Resources* 36 (1): 185-205.

Andrabi, T., J.Das, A.I. Khwaja, and T. Zajonc. 2011. “Do Value-Added Estimates Add Value? Accounting for Learning Dynamics.” *American Economic Journal: Applied Economics* 3 (3): 29-54.

Angrist, J.D., P.A. Pathak, and C.R. Walters. 2013. “Explaining Charter School Effectiveness.” *American Economic Journal: Applied Economics* 5 (4): 1-27.

Bhalotra, S., and S. Rawlings. 2013. “Gradients of the Intergenerational Transmission of Health in Developing Countries.” *The Review of Economics and Statistics* 95 (2): 660-72.

Blair, C., and C.C. Raver. 2012. “Child development in the context of adversity: experiential canalization of brain and behavior.” *Am Psychol* 67 (4): 309-318.

Blair, C., R.P. Razza. 2007. “Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten.” *Child Dev.* 78 (2): 647-63.

Bradley, R.H., and R.F. Corwyn. 2002. “Socioeconomic status and child development.” *Annual Review of Psychology* 53: 371-99.

——. 2005. “Caring for children around the world: A view from HOME.” *International Journal of Behavioral Development* 29 (6): 468-78.

Burchinal, M.R., J.E. Roberts, S. Hooper, and S.A. Zeisel. 2000. “Social risk and developmental patterns during the first four years.” *Developmental Psychology*.

Case, A., D. Lubotsky, and C. Paxson. 2002. “Economic status and health in childhood: The origins of the gradient.” *American Economic Review* 92 (5): 1308-34.

Case, A., and C. Paxson. 2008. “Stature and Status: Height, Ability, and Labor Market Outcomes.” *Journal of Political Economy* 116 (3): 499-532.

Chay, K.Y., and J.L. Powell. 2001. “Semiparametric Censored Regression Models.” *Journal of Economic Perspectives* 15 (4): 29-42.

Cunha, F., and J. Heckman. 2008. “Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation.” *Journal of Human Resources* 43 (4): 738-82.

Currie, J. 2009. “Healthy, Wealthy, and Wise? Socioeconomic Status, Poor Health in Childhood, and Human Capital Devleopment.” *Journal of Economic Literature* 47 (1): 87-122.

Currie, J., and T. Vogl. 2013. “Early-Life Health and Adult Circumstances in Developing Countries.” *Annual Review of Economics* 5: 1-36.

Doyle, O., C.P. Harmon, J.J. Heckman, and R.E. Tremblay. 2009. “Investing in early human development: Timing and economic efficiency.” *Economics and Human Biology* 7: 1-6.

Duncan, G.J., and J. Brooks-Gunn. 1997. “Income effects across the life span: Integration and interpretation.” In G. J. Duncan and J. Brooks-Gunn eds. Consequences of growing up poor . New York: Russell Sage Foundation, 596-610

Duncan, G.J., et al. 2007. “School readiness and later achievement.” *Developmental Psychology* 43 (6): 1428-46.

Engle, P.L., et al. 2011. “Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries.” *Lancet,* 378 (9799): 1339-53.

Evans, G. W. 2004. “The environment of childhood poverty.” *Am Psychol* 59 (2), 77-92.

Farah, M. J., et al. 2006. “Childhood poverty: Specific associations with neurocognitive development. ” *Brain Research* 1110: 166-74.

Fernald, L.C., A. Weber, E. Galasso, and L. Ratsifandrihamanana. 2011. “Socioeconomic gradients and child development in a very low income population: evidence from Madagascar.” *Dev Sci* 14 (4): 832-47.

Fernald, L.C.H., et al. 2016. A cluster-randomized, controlled trial of nutritional supplementation and promotion of responsive parenting in Madagascar: The MAHAY study design and rationale *BMC Public Health*.

Fernald, L.C.H., P.K. Kariger, P. Engle, and A. Raikes. 2010. “Examining child development in low-income countries: A toolkit for the assessment of children in the first five years of life.” Available at: http://siteresources.worldbank.org/INTCY/Resources/395766-1187899515414/Examining_ECD_Toolkit_FULL.pdf. Accessed June 6, 2010. Washington DC: The World Bank.

Fernald, L.C.H., P.K. Kariger, M. Hidrobo, and P.J. Gertler. 2012. “Socio-economic gradients in child development in very young children: Evidence from India, Indonesia, Peru and Senegal.” *Proceedings of the National Academy of Science (PNAS),* Oct 16,109 Suppl 2, 17273-80.

Filmer, D., and L.H. Pritchett. 2001. “Estimating wealth effects without expenditure data--or tears: An application to educational enrollments in states of India.” *Demography* 38 (1): 115-32.

Galasso, E., and N. Umapathi. 2007. Improving Nutritional Status through Behavioral Change: Lessons from Madagascar. World Bank Policy Research Working Paper No. 4424.

Glewwe, P., and H. Jacoby. 1995. “An economic analysis of delayed primary school enrollment in a low income country: the role of early childhood nutrition.” *Review of Economics and Statistics* 77 (1): 156-69.

Gove, A., and A. Wettenberg. eds. 2011. “The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy.” Available at http://www.rti.org/pubs/bk-0007-1109-wetterberg.pdf. Accessed June 18, 2013. Research Triangle Park: Research Triangle Institute.

Grantham-McGregor, S., et al. 2007. “Developmental potential in the first 5 years for children in developing countries.” *Lancet* 369 (9555): 60-70.

Hackman, D.A., and M.J. Farah. 2009. “Socioeconomic status and the developing brain.” *Trends in Cognitive Science* 13 (2): 65-73.

Hart, B., and T.R. Risley. 1995. Meaningful differences in the everyday experience of young American children. Baltimore, MD: Paul Brookes.

Hoff, E. 2003. The specificity of environmental influence: “Socioeconomic status affects early vocabulary via maternal speech.” *Child Development* 74 (5): 1368-78.

Kelly, Y., A. Sacker, E. Del Bono, M. Francesconi, and M. Marmot. 2011. “What role for the home learning environment and parenting in reducing the socioeconomic gradient in child development? Findings from the Millennium Cohort Study.” *Arch Dis Child* 96 (9): 832-7.

Kobrosly, R.W., et al. 2011. “Socioeconomic position and cognitive function in the Seychelles: a life course analysis.” *Neuroepidemiology* 36 (3): 162-8.

Lenroot, R.K., and J.N. Giedd. 2006. “Brain development in children and adolescents: insights from anatomical magnetic resonance imaging.” *Neurosci Biobehav Rev* 30: 718-29.

Lerner, R.M. 1998. “Theories of human development: contemporary perspectives.” In W. Damon and R.M. Lerner eds. *Handbook of child psychology, 5th edition*. Volume 1: Theoretical models of human development. New York: John Wiley and Sons, Inc.

Masters, G.N., and B.D. Wright. 1984. “The essential process in a family of measurement models.” *Psychometrika* 49 (4): 529-44.
Najman, J.M., et al. 2009. “The impact of episodic and chronic poverty on child cognitive development.” *J Pediatr Health Care* 154 (2): 284-9.
Noble, K.G., B.D. McCandliss, and M.J. Farah. 2007. “Socioeconomic gradients predict individual differences in neurocognitive abilities.” *Developmental Science* 10 (4): 464-80.
Paxson, C., and N. Schady. 2007. “Cognitive development among young children in Ecuador: The roles of wealth, health, and parenting.” *Journal of Human Resources* 42 (1): 49-84.
Petrou, S., and E. Kupek. 2010. Poverty and childhood undernutrition in developing countries: a multi-national cohort study.” *Soc Sci Med* 71 (7): 1366-73.
Rubens, A. (Ed.) 2009. Early Grade Mathematics Assessment (EGMA): A Conceptual Framework Based on Mathematics Skills Development in Children: Research Triangle Institute.
Schady, N. 2011. “Parents’ Education, Mothers’ Vocabulary, and Cognitive Development in Early Childhood: Longitudinal Evidence From Ecuador.” *American Journal of Public Health* 101 (12): 2299-307.
Schady, N., J. Behrman, M.C. Araujo, R. Azuero, R. Bernal, D. Bravo, F. Lopez-Boo, K. Macours, D. Marshall, C. Paxson, and R. Vakis. 2015. “Wealth gradients in early childhood cognitive development in five Latin American countries.” *Journal of Human Resources* 50 (2): 446-63.
Shonkoff, J.P. 2011. “Protecting Brains, Not Simply Stimulating Minds.” *Science* 333.
Shonkoff, J.P., W.T. Boyce, and B.S. McEwen. 2009. “Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention.” *JAMA : the Journal of the American Medical Association* 301 (21): 2252-9.
Todd, P.E., and K.I. Wolpin, 2007. “The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps.” *Journal of Human Capital* 1 (1 (Winter 2007)), 91-136.
Walker, S.P., et al. 2011. “Inequality begins in early childhood: Risk and protective factors for early child development.” *Lancet* 378 (9799): 1325-38.
Yeung, W.J., M.R. Linver, and J. Brooks-Gunn. 2002. “How Money Matters for Young Children's Development: Parental Investment and Family Processes.” *Child Development* 73 (6): 1861-79.

Appendix 1. Tests of cognitive and language tests administered to children

Cognition: Three sub-scales of the Stanford Binet Intelligence Scales for Early Childhood, 5th edition (ESB5) (Roid, 2003) were combined to estimate children's non-verbal cognitive ability:

(i) Visual-spatial processing (ESB5): Consisted of inserting geometric shapes into a board, or arranging them to match a given model (internal consistency α = 0.71 and 0.79 in '07 and '11 respectively, assumes items after stopping rule have score zero).

(ii) Working memory (ESB5): In a series of items of increasing difficulty, the child is asked to find hidden objects or to reproduce a sequence of tapping blocks according to a given pattern (α = 0.83 and 0.71 in '07 and '11 respectively). This sub-scale evaluates the child's capacity for executive function, spatial perception and for retaining and manipulating information.

(iii) Fluid Reasoning (ESB5): Consists of identifying an element that completes a series, or a missing panel in a matrix (α = 0.73 and 0.69 in '07 and '11 respectively). It evaluates the capacity for comprehension of logical associations and the skill to solve a figurative problem.

No changes were made to the original version of the sub-scales, although two simplified matrices were created and added in 2011 as training items for the fluid reasoning task.

Sustained Attention: From the completely non-verbal Leiter International Performance Scales (Leiter-R) (Roid & Miller, 1997), this test includes 4 trials and consists of marking images that match a model image (e.g. a sun) as quickly as possible (α = 0.81 and 0.83 in '07 and '11 respectively). Some of the images in the commercially available test were replaced to reflect the local context (e.g. an image of a zebu, a type of ox, replaced that of a cow).

Memory of phrases: Adapted from a subscale from the Woodcock-Muñoz assessment (Woodcock & Munoz, 1996), this test requires children to repeat words and sentences spoken by the interviewer that are increasingly long and difficult (α = 0.87 and 0.83 in '07 and '11 respectively). A Malagasy version of this exercise was not a literal translation and was created to match the English version's principle of increasing order of difficulty of the items.

Receptive language: Receptive language was assessed using the Peabody Picture Vocabulary Test, 3rd edition, version B (PPVT-III) (Dunn, 1965), which consists of asking the subject to point to an image on a sheet of four images that corresponds to a word spoken by the interviewer (α = 0.90 and 0.93 in '07 and '11 respectively). When no direct translations into Malagasy existed for an English word, the words were either spelled using Malagasy orthography or translated to French. Images that were culturally inappropriate or ambiguous in the local language were replaced with a comparable image (e.g. a depiction of US dollars and cents was replaced with a picture of ariary and iraimbilanja, the currency used in Madagascar).

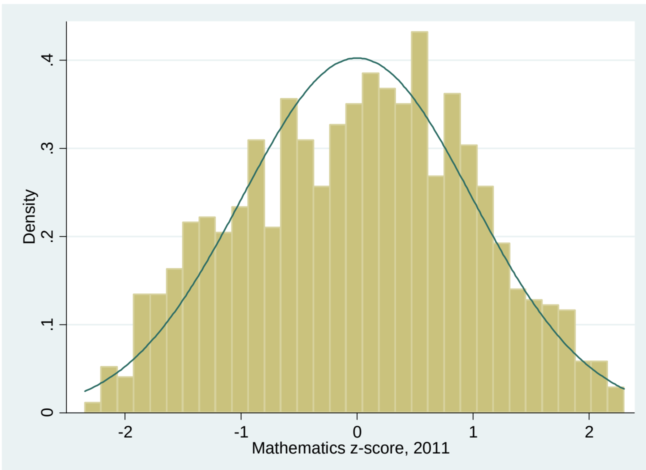
For all tests, we ensured overlapping items or tasks were incorporated in both time periods to take advantage of the longitudinal nature of the study. Many of the same interviewers were hired in both survey years, worked in pairs (one administering the test and one scoring), and received extensive classroom and field training prior to both surveys. Excellent inter-rater reliability (ρ > 0.95) was obtained during the training session for the 2007 survey.

Mathematics Skills: The mathematics test was adapted from USAID's Early Grade Mathematics Assessment (EGMA), building on the previous experience in adapting the test in a prospective school feeding evaluation carried out by J-PAL and the Ministry of Education. The math test includes sections on figures recognition, counting, math operations and logical skills. There is a large difference in the languages used by children for mathematics in Madagascar: many schools teach math in French, others teach in Malagasy, especially in the rural areas. The system of reading numbers differs between the two languages, and differs even among the dialects. Therefore, the children were allowed to read the numbers and reply to the problems in whichever language they preferred. All problems were also presented visually on a large card. An indicator was added to capture the language that the child used for the math test: French, Malagasy, or both.

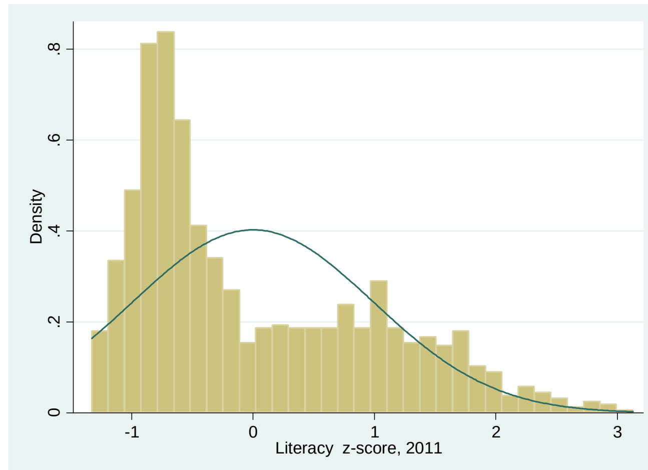
Early literacy Skills: The early literacy test was adapted from USAID's Early Grade Reading Assessment (EGRA), also building on the previous administration in country by the Ministry of Education and J-PAL. The test includes sections on sound recognition, recognition of letters and words, and two short stories, one for auditory comprehension and a second for reading out loud with subsequent comprehension questions. Several iterations of the stories were pre-tested during the training of the examiners before the stories and related questions were finalized. We wanted to assure that all children would be able to answer at least the first auditory comprehension question, even if they did not attend school, and assure that the comprehension questions were posed in such a way as to elicit a single correct response.

Figure A1: Learning z-scores: overall

Mathemathics

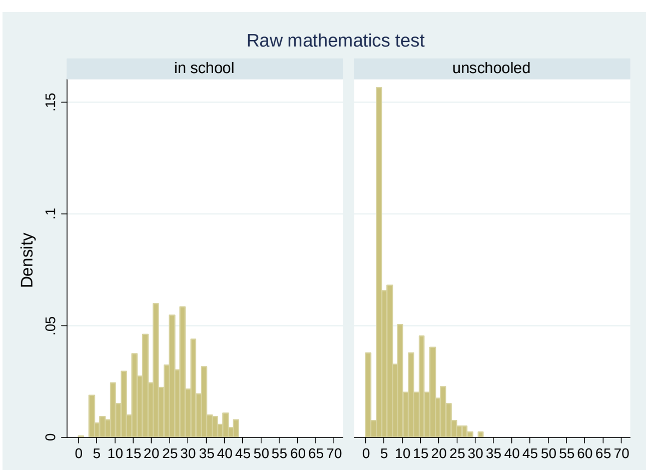


Literacy

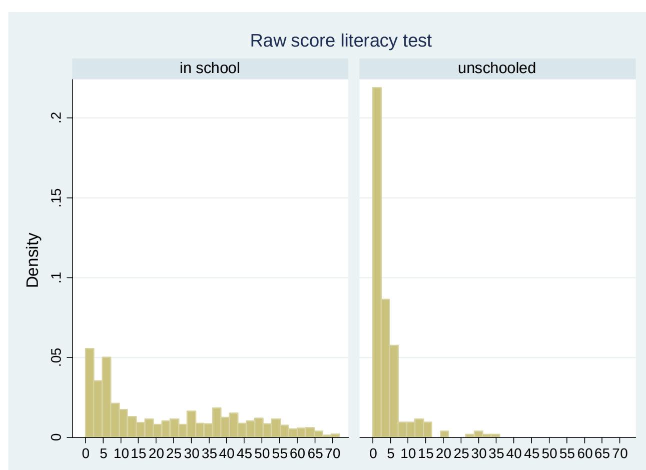


Learning z-score, by current enrollment

Mathemathics



Literacy



Appendix 2. Adaptation of questions from the HOME MC

1. Is there a music player or radio that child can listen to at home?

2. Is there something your child can use to make music at home (for example drum, harp, guitar, flute)?

3. About how many books are there in the household (including the bible, dictionary, school and picture books)? (*1 point given for any book*)

4. Do you have any other reading material at home, such as newspapers, magazines, pamphlets, or brochures?

5. Do you have any pictures, posters, calendars, or other type of art work on your walls at home?

6. Does child have paper and pencil, pen, or art supplies (crayons, paints) to write or draw with at home?

7. Does child make his/her own toys to play with, such as a football or dolls?

8. Does child play with any games of strategy such as checkers (Fanarona) or “Katro”?

9. Does child participate in any special organized activities (outside of school) such as boy or girl scouts, sports, speech (Hainteny), music, art, dance, or volunteering?

10. In the last year, how often has a family member taken or arranged to take child to another region or city? (*1 point given for any trip*)

Table A1. table 3, full set of regressors

	sustained attention				memory of phrases				Receptive vocabulary				cognitive composite			
	(base)	(i)	(ii)	(iii)	(base)	(i)	(ii)	(iii)	(base)	(i)	(ii)	(iii)	(base)	(i)	(ii)	(iii)
Wealth quintiles (Q1 excluded category)																
Q2	0.178***	0.308***	0.262***	0.246***	0.192***	0.258***	0.204**	0.185**	0.184***	0.338***	0.265***	0.241***	0.168***	0.276***	0.224***	0.205**
	(0.066)	(0.087)	(0.096)	(0.095)	(0.056)	(0.078)	(0.085)	(0.078)	(0.060)	(0.083)	(0.087)	(0.079)	(0.058)	(0.083)	(0.086)	(0.084)
Q3	0.269***	0.306***	0.251***	0.207**	0.254***	0.269***	0.207***	0.146*	0.309***	0.344***	0.261***	0.200**	0.274***	0.211***	0.151**	0.106
	(0.072)	(0.092)	(0.093)	(0.091)	(0.067)	(0.079)	(0.080)	(0.078)	(0.070)	(0.085)	(0.085)	(0.079)	(0.061)	(0.073)	(0.076)	(0.073)
Q4	0.451***	0.430***	0.325***	0.267***	0.393***	0.391***	0.325***	0.246**	0.529***	0.582***	0.434***	0.339***	0.439***	0.411***	0.277***	0.213**
	(0.072)	(0.084)	(0.091)	(0.088)	(0.081)	(0.094)	(0.099)	(0.101)	(0.080)	(0.097)	(0.095)	(0.093)	(0.071)	(0.091)	(0.096)	(0.093)
Q5	0.810***	0.646***	0.483***	0.400***	0.715***	0.583***	0.423***	0.308***	1.060***	0.892***	0.672***	0.532***	0.893***	0.601***	0.383***	0.288***
	(0.085)	(0.109)	(0.113)	(0.107)	(0.079)	(0.109)	(0.117)	(0.111)	(0.085)	(0.094)	(0.108)	(0.100)	(0.079)	(0.098)	(0.112)	(0.109)
Lagged outcomes & early conditions																
lagged sust.attention		0.154***	0.139***	0.139***		0.024	0.023	0.023		0.055*	0.042	0.040		0.066*	0.049	0.047
		(0.032)	(0.034)	(0.034)		(0.027)	(0.027)	(0.027)		(0.030)	(0.030)	(0.029)		(0.035)	(0.035)	(0.034)
lagged mem. phrases		0.067*	0.043	0.028		0.208***	0.181***	0.160***		0.125***	0.103***	0.079**		0.030	0.009	-0.007
		(0.038)	(0.041)	(0.041)		(0.035)	(0.035)	(0.036)		(0.033)	(0.034)	(0.033)		(0.034)	(0.033)	(0.033)
lagged receptive vocab.		0.041	0.023	0.019		0.106***	0.090***	0.084***		0.145***	0.127***	0.127***		0.087**	0.070*	0.067*
		(0.030)	(0.031)	(0.031)		(0.030)	(0.031)	(0.030)		(0.035)	(0.033)	(0.030)		(0.040)	(0.038)	(0.038)
lagged comp. cognition		0.139***	0.120***	0.122***		0.078**	0.066*	0.069**		0.057*	0.038	0.047		0.175***	0.149***	0.153***
		(0.032)	(0.034)	(0.034)		(0.032)	(0.034)	(0.033)		(0.032)	(0.031)	(0.031)		(0.030)	(0.033)	(0.032)
low birth weight		-0.058	-0.058	-0.070		0.127**	0.152**	0.136**		0.062	0.027	0.003		-0.014	-0.033	-0.045
		(0.065)	(0.070)	(0.069)		(0.061)	(0.060)	(0.059)		(0.068)	(0.067)	(0.064)		(0.074)	(0.077)	(0.077)
=1 if stunted		-0.224***	-0.210**	-0.206**		-0.057	-0.069	-0.063		-0.146**	-0.111	-0.100		-0.197***	-0.151*	-0.146**
		(0.076)	(0.082)	(0.081)		(0.076)	(0.080)	(0.078)		(0.073)	(0.069)	(0.066)		(0.071)	(0.078)	(0.074)
lagged stunting		-0.007	-0.055	-0.048		-0.041	-0.039	-0.030		-0.138**	-0.163**	-0.151**		-0.026	-0.041	-0.035
		(0.084)	(0.084)	(0.081)		(0.069)	(0.074)	(0.070)		(0.065)	(0.071)	(0.065)		(0.069)	(0.078)	(0.073)
Maternal education & endowments																
mother: primary school			0.139*	0.094			0.274***	0.211***			0.197***	0.116*			0.099	0.056
			(0.082)	(0.083)			(0.068)	(0.071)			(0.072)	(0.069)			(0.065)	(0.063)

mother: secondary school			0.264**	0.196*			0.316***	0.220***			0.349***	0.227***			0.324***	0.255***
			(0.109)	(0.111)			(0.082)	(0.082)			(0.082)	(0.080)			(0.088)	(0.087)
log maternal height			-0.746	-0.659			-0.303	-0.201			-0.550	-0.441			-0.135	-0.059
			(0.872)	(0.875)			(0.679)	(0.687)			(0.638)	(0.639)			(0.683)	(0.673)
maternal vocabulary			0.162***	0.140***			0.124**	0.091			0.219***	0.180***			0.204***	0.177***
			(0.052)	(0.052)			(0.059)	(0.059)			(0.052)	(0.053)			(0.060)	(0.060)
HOME environment				0.049***				0.071***				0.081***				0.056***
				(0.016)				(0.017)				(0.016)				(0.015)
R2	0.137	0.346	0.357	0.364	0.241	0.495	0.513	0.529	0.229	0.470	0.509	0.531	0.165	0.393	0.426	0.437
No. obs	1,235	1,030	936	933	1,235	1,018	927	924	xx	1,034	939	936	1,201	1,005	915	912

Note: All regressions subsume age (in years) dummies, birth order dummies, a rural/urban indicator and regional dummies. Robust standard errors (in parentheses) are clustered at the community level. *,**,*** denote significance at 10%, 5% and 1% levels.

Table A2. table 6, full set of regressors

	Math					Literacy				
	Base	(i) lagged	(ii) maternal	(iii) home	(iv) concurrent	base	(i) lagged	(ii) maternal	(iii) home	(iv) concurrent
Wealth quintiles (Q1 excluded category)										
Q2	0.61	0.535	0.41	0.352	0.172	0.217	0.197	0.151	0.09	-0.049
	[0.272, 0.948]	[0.2, 0.871]	[0.058, 0.763]	[0.016, 0.688]	[-0.088, 0.433]	[-0.009, 0.443]	[-0.116, 0.51]	[-0.188, 0.489]	[-0.25, 0.43]	[-0.313, 0.215]
Q3	0.593	0.535	0.402	0.321	0.231	0.232	0.179	0.175	0.074	0.037
	[0.263, 0.923]	[0.225, 0.846]	[0.074, 0.729]	[-0.02, 0.661]	[-0.018, 0.48]	[-0.041, 0.505]	[-0.131, 0.488]	[-0.165, 0.515]	[-0.253, 0.401]	[-0.233, 0.307]
Q4	0.926	0.737	0.569	0.443	0.249	0.534	0.349	0.326	0.182	0.008
	[0.576, 1.276]	[0.412, 1.063]	[0.227, 0.912]	[0.12, 0.765]	[-0.017, 0.515]	[0.179, 0.889]	[-0.013, 0.711]	[-0.055, 0.707]	[-0.18, 0.544]	[-0.282, 0.298]
Q5	1.315	0.904	0.684	0.485	0.213	1.317	0.74	0.519	0.395	-0.039
	[0.979, 1.651]	[0.569, 1.239]	[0.319, 1.05]	[0.126, 0.845]	[-0.074, 0.5]	[0.954, 1.68]	[0.372, 1.108]	[0.116, 0.921]	[-0.015, 0.806]	[-0.401, 0.323]
Lagged outcomes & early conditions										
Lagged cognition		0.163	0.14	0.134	0.041		0.173	0.125	0.111	0.018
		[0.06, 0.266]	[0.032, 0.248]	[0.026, 0.242]	[-0.042, 0.124]		[0.054, 0.293]	[-0.001, 0.251]	[-0.024, 0.246]	[-0.089, 0.125]
Lag.sust.attention		0.097	0.051	0.057	0.041		0.109	0.099	0.084	0.046
		[-0.024, 0.218]	[-0.076, 0.177]	[-0.064, 0.178]	[-0.042, 0.125]		[0, 0.217]	[-0.012, 0.21]	[-0.035, 0.203]	[-0.059, 0.151]
Lag.mem.phrases		0.101	0.077	0.089	0.062		0.165	0.123	0.139	0.065
		[-0.022, 0.224]	[-0.052, 0.205]	[-0.043, 0.221]	[-0.035, 0.159]		[0.026, 0.303]	[-0.024, 0.27]	[-0.014, 0.292]	[-0.069, 0.199]
Lag. vocabulary		0.084	0.097	0.072	0.037		0.158	0.107	0.114	0.066
		[-0.022, 0.189]	[-0.013, 0.208]	[-0.033, 0.177]	[-0.046, 0.12]		[0.043, 0.274]	[-0.012, 0.226]	[-0.005, 0.234]	[-0.035, 0.166]
Low birth weight		0.144	0.104	0.073	-0.018		0.103	0.15	0.097	0.041
		[-0.083, 0.372]	[-0.121, 0.329]	[-0.162, 0.308]	[-0.203, 0.167]		[-0.147, 0.353]	[-0.108, 0.409]	[-0.159, 0.352]	[-0.156, 0.239]
=1 if stunted		-0.332	-0.274	-0.223	-0.148		-0.257	-0.164	-0.171	-0.008
		[-0.576, -0.088]	[-0.521, -0.027]	[-0.481, 0.036]	[-0.355, 0.058]		[-0.546, 0.031]	[-0.484, 0.156]	[-0.468, 0.125]	[-0.251, 0.235]
Lag.stunted		0.084	0.078	0.078	0.128		0.025	-0.085	-0.05	-0.011
		[-0.144, 0.312]	[-0.175, 0.331]	[-0.164, 0.32]	[-0.068, 0.325]		[-0.248, 0.298]	[-0.379, 0.209]	[-0.316, 0.215]	[-0.253, 0.231]
Maternal education & endowments										
M:primary edu			0.261	0.195	0.047			0.267	0.184	0.123
			[-0.028, 0.55]	[-0.1, 0.49]	[-0.174, 0.269]			[-0.052, 0.586]	[-0.154, 0.521]	[-0.16, 0.406]
M.secondary ed			0.448	0.401	0.083			0.605	0.46	0.218
			[0.125, 0.77]	[0.08, 0.722]	[-0.196, 0.362]			[0.243, 0.967]	[0.077, 0.843]	[-0.097, 0.532]
Log m.height			-0.68	-0.058	0.039			0.05	-0.604	0.409
			[-2.997, 1.637]	[-2.548, 2.431]	[-1.873, 1.951]			[-2.559, 2.659]	[-3.257, 2.049]	[-1.932, 2.75]
M.vocabulary			0.162	0.103	0.004			0.215	0.195	0.035
			[-0.033, 0.356]	[-0.081, 0.287]	[-0.151, 0.159]			[0.001, 0.429]	[-0.016, 0.406]	[-0.137, 0.207]
HOME				0.087	0.03				0.079	0.046
				[0.028, 0.146]	[-0.012, 0.072]				[0.023, 0.136]	[-0.003, 0.095]
Concurrent test outcomes										
Cognition					0.23					0.118
					[0.13, 0.331]					[-0.007, 0.243]
Sust.attention					0.248					0.273
					[0.153, 0.343]					[0.145, 0.4]
Mem.phrases					0.058					0.165
					[-0.062, 0.178]					[0.016, 0.314]

Vocabulary	0.261 [0.135, 0.387]	0.32 [0.166, 0.473]