CENTER FOR GLOBAL DEVELOPMENT

Preprimary Education and Early Childhood Development: Evidence from Government Schools in Rural Kenya

Pamela Jakiela, Owen Ozier, Lia C. H. Fernald, and Heather A. Knauer

Abstract

We estimate the impact of preprimary education on early childhood development in a sample of Kenyan three-year-olds. Our identification strategy exploits the fact that children in our sample are more likely to start school at age three rather than at age four if they live within a few hundred meters of the nearest primary school, though other household characteristics do not vary across such small distances. Instrumental variables estimation suggests that enrolling in preschool at age three has large positive impacts on vocabulary in children's mother tongue, which is the primary language of instruction in preprimary. However, we do not find evidence that these short-term gains translate into persistent advantages in vocabulary or other measures of child development one to three years later.

KEYWORDS

preschool, early childhood, preprimary education, human capital, school readiness, early literacy, mother tongue instruction, instrumental variables

JEL CODES

O12, I25, J24, H52

WORKING PAPER 661 · OCTOBER 2023

Preprimary Education and Early Childhood Development: Evidence from Government Schools in Rural Kenya

Pamela Jakiela
Williams College, BREAD, CGD, IZA, and J-PAL, email: pj5@williams.edu

Owen Ozier
Williams College, BREAD, IZA, and J-PAL, email: oo3@williams.edu

Lia C. H. Fernald
University of California at Berkeley, email: fernald@berkeley.edu

Heather A. Knauer
University of Michigan, email: heather.knauer@gmail.com

Orazio Attanasio, Blake Bathum, Dave Evans, Seema Jayachandran, Abbosbek Juraev, Shangbo Jing, Eva Jackson Mbeyela, Daniel Page, Riana Ramonjamanana, Justin Sandefur, Zia Saylor, and Faiza Zia provided helpful comments. We are grateful to William Blackmon, Emily Cook-Lundgren, Julian Duggan, Laura Kincaide, and Elyse Thulin for excellent research assistance. IRB approval for the larger RCT was provided by the University of California at Berkeley, by the Maseno University Ethics Review Committee, and by Innovations for Poverty Action. This research was funded by Echidna Giving and by the Strategic Impact Evaluation Fund and the Early Learning Partnership at the World Bank. All errors are our own.

Pamela Jakiela, Owen Ozier, Lia C. H. Fernald, and Heather A. Knauer. 2023. "Preprimary Education and Early Childhood Development: Evidence from Government Schools in Rural Kenya." CGD Working Paper 661. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/preprimary-education-and-early-childhood-development-evidence-government-schools-rural

CENTER FOR GLOBAL DEVELOPMENT
2055 L Street, NW Fifth Floor
Washington, DC 20036
202.416.4000

1 Abbey Gardens
Great College Street
London
SW1P 3SE

www.cgdev.org

1 Introduction

Investments in health and human capital during early childhood influence later life outcomes such as educational attainment and adult income (Behrman, Engle and Fernald 2013, Tanner et al. 2015, Britto et al. 2017). In low- and middle-income countries (LMICs), more than a third of all children under five do not receive adequate nutrition and cognitive stimulation (Black et al. 2017). Inadequate parental stimulation at home contributes to a lack of school-readiness, particularly among children from low-income households (Bradley and Caldwell 1976, Engle et al. 2011, Cabrera et al. 2020). High-quality early childhood education can help to address these disparities in early life investment in human capital (Burger 2010, Schady et al. 2015).

In recent years, many LMIC governments have expanded preprimary education programs to promote school readiness and improve young children's human capital (Berlinski and Schady, eds 2015). Between 2000 and 2020, gross preprimary enrollment in LMICs increased from 28 percent to 58 percent of preschool-aged children (World Bank 2023). Yet, the evidence on the developmental impacts of early childhood education programs is mixed. While many studies find that preschool improves child development (cf. Engle et al. 2011, Martinez, Naudeau and Pereira 2017, Speir et al. 2020), low-quality early childhood education does not always benefit children (Bouguen, Filmer, Macours and Naudeau 2018, Blimpo, Carneiro, Jervis and Pugatch 2022, Wong, Luo, Zhang and Rozelle 2013) and enrolling in preschool does not always benefit low-income children (Berkes, Bouguen, Filmer and Fukao 2019). When institutional quality is low relative to the counterfactual, access to daycare and preschool can have negative impacts on child development (Behrman, Fernald and Engle 2013). At present, there is very little credible causal evidence on the impacts of government preprimary education in Sub-Saharan Africa on child development outcomes. We are among the first to fill this gap.[1]

[1] A difficulty in estimating the impact of national preschool programs is that there is rarely any systematic collection of learning outcomes at ages where schooling is not compulsory, making it challenging to compare those who attend preschool to those who do not. As such, while there have been several randomized trials of specific private or NGO-assisted preschool programs (cf. Martinez, Naudeau, and Pereira 2017, Spier et

We estimate the impact of attending a public preschool on cognitive development in a sample of three year old children in rural Kenya. We make use of a rich cross-sectional data set containing multiple measures of early childhood development as well as detailed household survey data including information on the precise location of each household. In our setting, children are significantly more likely to start school at age three when they live very close to the school (less than 500 meters away). This allows us to use distance to the school as an instrument for enrollment among three year olds. Since the distances involved are quite small – all households in our sample live within 750 meters of a government primary school – other characteristics are unlikely to vary across households. Consistent with this, we see that the distance-based gradient in school enrollment disappears by age four.

Our two-stage least squares (2SLS) estimates suggest that preschool enrollment has large positive impacts on early childhood development. Enrollment increases an aggregate index of child development by between 0.78 (without covariates) and 0.83 (covariate-adjusted) SD. Impacts are driven by substantial improvements in receptive vocabulary in children's mother tongue (Luo), which improves by more than a full standard deviation. This impact represents more than a fifty percent increase in mother tongue receptive vocabulary among three year olds. Coefficient estimates also suggest developmentally meaningful improvements in expressive vocabulary and fine motor skills, though these are not statistically significant. Preschool does not improve receptive vocabulary in English – which makes sense, since Kenya's national language policy stipulates that preprimary instruction should be in children's mother tongue whenever possible.[2] The absence of effects on English vocab-

al. 2020, Bjortvan et al. 2022, Blimpo et al. 2023) and there is also a growing body of research on the impacts of programs intended to improve the quality of existing preprimary education facilities (cf. Özler et al. 2018, Wolf, Aber, Behrman and Tsinigo 2019, Blimpo et al. 2023), there is very little quasi-experimental evidence on the impact of national-scale government-provided preschool in Africa. Our work is most similar to Bietenbeck, Ericsson and Wamalwa (2019), which uses a household fixed effects approach to estimate the causal impact of attending preschool on subsequent educational outcomes among children in Kenya and Tanzania.

[2]See discussion in USAID (2021) and Begi (2014). Recent national policy documents make explicit that "In pre-primary education, the medium of instruction is the language of the catchment area" and that "The Constitution of Kenya 2010, accords English the status of one of the official languages while according to the language policy of 1976, it is the language of instruction from grade four onwards" (see pp. 32 and

ulary provides evidence that gains in other domains represent genuine increases in human capital, and not simply greater comfort with the modality of direct child assessment among children who have been in center-based childcare.

While we find evidence of meaningful impacts on three year olds induced to attend preschool by their proximity to the school, there is not a similar reduced form relationship between distance to the school and early childhood development among children aged four to six. This suggests that the developmental impacts of preschool may fade over time, once all children are enrolled in school. We do not find any evidence that children aged four to six who live nearer to the school are more likely to be at or above the appropriate grade-for-age, suggesting that those who start school later are placed in classes together with similarly-aged children who started school earlier. Hence, three year olds who enter school early may not be in a position to build on their early human capital gains in subsequent years.

Our research contributes to a large and growing literature on early childhood development interventions in LMICs (Engle et al. 2011, Black et al. 2017, Britto et al. 2017, Devercelli and Beaton-Day 2020).[3] Experimental and quasi-experimental impact evaluations of preprimary education in LMIC contexts have typically found that access to preschool has at least modest positive impacts on child development and school readiness, on average, in Latin America (Berlinski, Galiani and Manacorda 2008, Berlinski, Galiani and Gertler 2009, Bastos, Bottan and Cristia 2017), South Asia (Dean and Jayachandran 2020), Southeast Asia (Brinkman et al. 2017, Bloem and Wydick 2023), and Sub-Saharan Africa (cf. Martinez, Naudeau and Pereira 2017, Bjortvan et al. 2022). However, whether attending preschool leads to meaningful gains in human capital depends on both the quality of the preschool and the counterfactual educational experience of the child (Bougen et al. 2018, Dean and Jayachandran 2020, Blimpo et al. 2022). Our work contributes to the small but

49 of Kenya Institute for Curriculum Development 2019). See Piper and Miksic (2011) and Kerwin and Thornton (2021) for discussion of language-of-instruction policy in East Africa, and the implications for child development and literacy.

[3] There is also an extensive literature documenting the impacts of preschool in high-income country settings. For a review, see Almond and Currie (2010).

growing evidence base documenting the impacts of preschool in Sub-Saharan Africa. While previous studies using experimental and quasi-experimental methods have either focused on newly constructed schools (cf. Martinez, Naudeau and Pereira 2017, Blimpo et al. 2022) or private providers (Bjortvan et al. 2022; Gray-Lobe et al. 2022), our identification approach allows us to estimate the impact of existing government schools on those induced to enroll by their close proximity.[4] Our work complements Bietenbeck et al. (2019), who show that Kenyan and Tanzanian children who attended preschool have higher test scores at ages 13 to 16 than their siblings who did not attend preschool.

The rest of this paper is organized as follows. In Section 2, we describe our research design and the context of our study. We present our empirical results in Section 3. Section 4 concludes.

2 Research Design

Data for our study comes from a cross-sectional survey of households with children aged three to six in 73 ethnolinguistically-homogeneous rural communities in Kisumu District, Kenya.[5] Household and child surveys were conducted as part of the baseline for an impact evaluation of a pre-literacy intervention (Jakiela, Ozier, Fernald and Knauer 2020). In each of the 73 communities in our sample, we interviewed all the families living within 750 meters of the government primary school who had at least one child between the ages of three and six, generating a sample of 2,503 preschool-aged children from 2,013 households.

[4] Our work is also related to research on access to daycare (for children under three years old) in LMICs contexts (cf. Clark et al. 2019, Attanasio et al. 2022) and on other types of ECD interventions such as parenting education and home visits from child development specialists (cf. Knauer et al. 2019, Garcia et al. 2023).

[5] Communities were chosen from a sample frame that included all the government primary schools in rural areas of Kisumu District that were not within 1.5 km of another public primary school. We excluded 15 communities because they either had too few preschool-aged children (six communities), were selected for piloting (four communities), were hostile to survey enumerators (four communities), or were not predominantly Luo-speaking (one community).

2.1 Local Context

The rural villages in our sample are located in a predominantly Luo-speaking district in the western part of Kenya. Luo is a Nilotic language spoken by approximately five million Kenyans – just over ten percent of the country's population. Kenya's official policy is that children's mother tongue should be the language of instruction in preprimary and the first three years of primary school, while English is the language of instruction for the later years of primary as well as secondary school (USAID 2021, Kenya Institute of Curriculum Development 2019). In practice, mother tongue instruction is implemented unevenly (Piper, Zuilkowski and Ong'ele 2016), but Luo is widely used in primary schools throughout our study area, as it is the mother tongue of almost all local children (only four of the 634 three-year-olds in our sample do not have at least one parent whose mother tongue is Luo).

Data was collected in the second half of 2017. At that time, all the schools in our sample followed the 8-4-4 curriculum that was first implemented under President Daniel Arap Moi in 1985 (including eight years of primary school and four years of secondary school). The schools in our sample also offered three levels of preprimary education: two years of preschool ("Baby Class" followed by "Nursery") as well as a kindergarten year ("Pre-Unit").[6] Figure A1 demonstrates that school enrollment increases rapidly with age in our study area: almost no children are enrolled in school at 24 months of age, but by 48 months more than 80 percent of children are enrolled.[7] Our identification strategy exploits the fact that, in our setting, children who live very close to the school tend to begin their schooling slightly earlier.

2.2 Data

Our data set includes information on child, parent, and household characteristics. For children aged three to six, developmental outcomes were measured through direct child

[6]This system was replaced in 2018, when Kenya restructured preprimary education in an attempt to guarantee that all children complete two years of preprimary (PP1 and PP2) before beginning primary school.

[7]Data on enrollment is available for all children in the household, since it was collected as part of the household roster. Data on child development is available for children aged three to six.

assessment. Trained survey enumerators from Innovations for Poverty Action administered locally-adapted assessments of vocabulary and fine motor skills, and also measured children's heights. We assessed receptive vocabulary in English and Luo separately (the primary languages of instruction in upper and lower primary, respectively), and we also measured expressive vocabulary, which is the ability to produce an appropriate word in any relevant local language (English, Luo, or Swahili).[8,9] We measured fine motor skills using items from the Malawi Developmental Assessment Tool (Gladstone et al. 2010), a holistic instrument for assessing early childhood development in African contexts. All child development outcomes including height-for-age are converted into age-adjusted z-scores, following standard practice.[10]

In addition to child outcomes, we also collected data on parent and household characteristics including parental background and educational attainment, household composition, and information on household assets. Our study takes place in a rural, predominantly agricultural area, and all the households in our sample grow some of their own food. Our survey does not contain detailed information on parents' income-generating activities, household consumption, or other adult-focused outcomes. However, detailed information on durable assets provides a reasonable proxy for household wealth and socioeconomic status (Filmer and Pritchett 2001). Critically, our data set also includes the precise GPS coordinates of each family's dwelling, which allows us to construct a household-specific measure of distance to the local primary school.

[8]Receptive vocabulary is a measure of children's comprehension of words while expressive vocabulary is a measure of the ability to produce words. Receptive vocabulary begins developing first: children begin to understand what is being said to them before they develop the ability to articulate their own thoughts and ideas (Fernald, Prado, Kariger and Raikes 2017).

[9]Our measures of English and Luo receptive vocabulary are locally-adapted versions of the British Picture Vocabulary Scale, which is itself an adaptation of the Peabody Picture Vocabulary Test appropriate for people speaking British or Commonwealth English (Dunn and Dunn 1997; Dunn, Dunn and Styles 2009; Knauer et al. 2019b). We developed and validated a locally-appropriate tool for measuring receptive vocabulary as part of an earlier evaluation of an early literacy intervention (Knauer, Kariger, Jakiela, Ozier and Fernald 2019b).

[10]Height-for-age z-scores are calculated relative to external norms available from the World Health Organization.

2.3 Identification Strategy

We use an instrumental variables approach to identify the causal impact of preprimary education on young children's human capital. Building on a large literature in labor economics (cf. Card 1993, Kane and Rouse 1993, Dee 2004), we instrument for school enrollment using distance to the local school (measured in kilometers).[11] In a variety of contexts and at many levels of schooling, access to school predicts school participation (eg. Duflo 2001, Muralidharan and Prakash 2017), but in the present context of very young children, the distances involved are very small.[12] All households in our sample are within 750 meters of a government primary school, so distance is unlikely to impact enrollment for most school-aged children. Our identification exploits the fact that, empirically, children in our sample are more likely to start school at age three, as opposed to age four or five, when they live extremely close to the school – perhaps because parents are more willing to let a young child walk a short distance, or because less time is required to walk a young child to school.[13]

Figures 1 and A2 illustrate our instrumental variables strategy. Figure A2 plots the locations of all the households in our sample that include either a three-year-old (left panel) or four-year-old (right panel) child. The figure shows that four-year-olds are more likely to be enrolled in school than three-year-olds, irrespective of where they live (within our narrowly-defined catchment areas). Among three year olds, there are fewer out-of-school children near the school: within the first tercile of distance to the school (within approximately 360 meters of the school building) approximately 72 percent of three-year-olds are enrolled, but the proportion drops to 58 and 51 percent enrolled for, respectively, the second and third terciles of distance to the school. Figure 1 presents local polynomial regressions

[11] McKenzie and Sakho (2010) use a similar distance-based instrument to estimate the causal effect of formality (registering with the tax authority) on firm profitability in Bolivia. They use distance to the tax office as an instrument for legal registration.

[12] Attanasio, Maro and Vera-Hernández (2013) also use a similar distance-based instrument to estimate the effect of nursery participation on child nutrition in Colombia.

[13] This is analogous to the pattern that Muralidharan and Prakash (2017) show in their Figure 1: in India, those further from a secondary school are less likely to enroll; however, the distances involved in this study are much smaller. In India, there is a roughly 20 percentage point drop in the probability that 16 and 17 year olds are enrolled in secondary school as the distance from home to school rises from 0 to 15 kilometers. In the present setting, there is a roughly 20 percentage point drop in the probability a three year old attends preschool, but this occurs as the distance from home to school rises from 0.1 to 0.7 kilometers.

of enrollment on distance to the school. Among children aged four to six, the proportion enrolled is close to one and does not vary with distance from the school. However, among three-year-olds, the proportion enrolled declines substantially between about 300 meters and 500 meters from the school.

For our instrument to be valid, it must satisfy an exclusion restriction: distance from the primary school should not impact outcomes of interest except through enrollment in preprimary. We argue that this assumption is reasonable in our specific context because the distances involved are quite small: all households in our sample live within 750 meters of their local school. Households that are relatively further from the school are not more rural or more remote in any meaningful sense because all sample households are quite close together. We argue that these differences in distance from the school are only likely to matter for very young children, who may not yet be used to walking around their villages unaccompanied.

Though it is impossible to test the exclusion restriction, Table A3 presents evidence that households that are further from the school do not look different from those closer to the school in terms of their observable characteristics. The table reports the results of regressing each of twelve baseline characteristics on our measure of a household's distance from the primary school. None of the estimated coefficients on distance is statistically significant at the 5 percent level, and coefficient magnitudes suggest that households closer to the school are similar to those further away in terms of observable characteristics. If we regress distance from the school on all of these baseline covariates in a single OLS regression, we cannot reject the joint hypothesis that the coefficients are all equal to zero (p-value 0.3125). Though this does not prove that exclusion restriction is satisfied, the absence of meaningful differences in observable characteristics between households that are closer to and further from the local primary school is consistent with our identifying assumptions.

2.4 Estimation Approach

We are interested in estimating the impact of preprimary education on child development. In principle, we might wish to estimate an OLS regression of the form

$$Y_{ih} = \alpha + \beta P_{ih} + \lambda X_{ih} + \epsilon_{ih} \tag{1}$$

where P_{ih} is a dummy equal to one if child i in household h is enrolled in preprimary education, Y_{ih} is a child development outcome of interest, X_{ih} is a vector of controls, and ϵ_{ih} is a conditionally-mean-zero error term. However, such an estimation approach will yield a biased estimate of the causal impact of P on Y when P is correlated with ϵ. More informally, if parents who send their young children to school differ from those who do not, and those differences are also associated with child development outcomes, $\hat{\beta}_{OLS}$ will not provide an unbiased estimate of the impact of preschool on child development. To proceed, we need a source of plausibly exogenous variation in preprimary enrollment.

We estimate the impacts of preprimary education via two-stage least squares (2SLS), using the distance from household h to the local primary school as our instrument. In our first stage regression, we estimate the relationship between distance from the primary school and enrollment in preprimary via OLS. Our first-stage regression equation is:

$$P_{ih} = \alpha_1 + \delta D_h + \lambda_1 X_{ih} + \varepsilon_{ih} \tag{2}$$

where D_h is the distance from household h to the local primary school and ε_{ih} is a conditionally mean-zero error term, clustered at the household level. In our main specifications, we control for child age in months, height-for-age z-score, whether the child is male, whether the child's primary caregiver is their mother, mother's education, whether the mother is Luo, household size, the number of older siblings in the household, and an index of household durable assets. Our second stage equation is:

$$Y_{ih} = \alpha_2 + \beta \widehat{P}_{ih} + \lambda_2 X_{ih} + \xi_{ih} \tag{3}$$

where Y_{ih} is a child outcome of interest, and $\widehat{P}_{ih}$ is predicted enrollment (from the first stage), and ξ_{ih} is a conditionally mean-zero error term, clustered at the household level.

3 Results

3.1 The First Stage

We report our first-stage results in Table 1. Panel A includes no covariates, while Panel B reports results from regressions including controls (for child age, gender, and height-for-age z-score; mother's education and ethnicity; whether the child's mother is their primary caregiver; household size, the number of older siblings present in the household, and household wealth). Among three year olds (Column 1), distance from the school is a robust predictor of enrollment, consistent with the visual evidence in Figures A2 and 1. Coefficient estimates suggest that a child aged three who lived half a kilometer further from school would be almost 25 percentage points less likely to be enrolled. In contrast, distance is not a robust predictor of enrollment for children ages four and up (within the narrow range of distances observed in our sample). Among five year olds (Column 3), the estimated coefficients (on distance from the school) are approximately one tenth the size of those in Column 1, though they remain (at least marginally) statistically significant. Among four and six year olds (Columns 2 and 4, respectively), the relationship between distance from the school and enrollment is not statistically significant – and coefficient magnitudes are, again, much smaller than those in Column 1. Thus, within our sample, distance is a robust predictor of enrollment among three year olds, but not for older children.

3.2 The Impact of Preprimary Education on Child Development

Impacts on child development outcomes are reported in Table 2. We consider five outcomes: Luo (i.e. mother tongue) receptive vocabulary, English receptive vocabulary, expressive vocabulary, fine motor skills, and an aggregate index of child development that combines these four sub-indices.[14] Preprimary enrollment has large and statistically significant positive impacts on Luo receptive vocabulary. 2SLS estimates indicate that preschool increases Luo vocabulary by more than a standard deviation: the coefficient estimate is 1.16 SD when no controls are included (p-value 0.026) and 1.401 SD with controls (p-value 0.021).[15] Among three year olds, a standard deviation represents approximately four correct responses on the direct child assessment. The median number of correct responses is five, and the mean is six – so this is a developmentally meaningful increase in vocabulary.[16] 2SLS estimates also suggest quantitatively large improvements in expressive vocabulary and fine motor skills (coefficient estimates are between 0.5 SD and 0.8 SD), though these are not statistically significant (p-values range from 0.121 to 0.251). In contrast, IV estimates of the impact of preprimary on English receptive vocabulary are negative, smaller in magnitude, and not statistically significant: the coefficient estimate is -0.131 SD without controls and -0.238 SD with controls (p-values 0.647 and 0.786, respectively).[17] The absence of impacts on English language skills is to be expected, since Kenya's official policy is that children's mother tongue be used as the language of instruction in preprimary (through third grade) in linguistically homogeneous areas (see footnote 2). However, it also provides evidence that the positive impacts on other domains represent genuine increases in human capital, and not simply greater familiarity and/or comfort with adults outside the household or the procedures used for direct child assessment.

Column 5 of Table 2 reports IV estimates of the impact of preprimary on an overall index of early childhood development which aggregates our four individual outcome measures. Coefficients estimates suggest that preschool improves child development by between 0.78 SD (without controls, p-value 0.098) and 0.832 SD (with controls, p-value 0.090). Again,

[14]As discussed above and in footnote 2, Luo is the official language of instruction in preprimary, though English is also widely used in schools and is the official language of instruction in grades four and up.

[15]The 2SLS confidence intervals are relatively wide, so the 95-percent confidence interval associated with each statistically significant positive effect is of course consistent with a range of effect sizes besides our point estimates, including both more modest effects and more dramatic ones.

[16]For comparison, three year olds from the highest quintile of household wealth in our sample score approximately 0.25 SD higher on the Luo receptive vocabulary assessment than those from the lowest quintile of household wealth.

[17]Our results differ from those of Piper, Zuilkowski, Kwayumba and Oyanga (2018), who find that a mother tongue literacy program improved literacy in English.

these are developmentally meaningful effects, suggesting that preprimary education is an effective child development intervention.

3.3 Characteristics of the Compliers

Instrumental variables estimates provide estimates of the local average treatment effect on compliers, in this case children who are induced to enroll at age three by their proximity to the school. In Table A4, we test the extent to which our instrument has a differential impact on the likelihood of enrollment at age three for different sub-groups. To do this, we regress enrollment on distance to the school a specific characteristic (e.g. whether a child is male), and an interaction between that characteristic and distance to the school. The interaction term provides a test of the hypothesis that observations with the given characteristic are over-represented (or under-represented) among compliers.

We consider four characteristics: whether a child is male, whether a child has a below-median height-for-age z-score, whether a household has at or below the median number of durable assets, and whether a child's mother has below median education. We find no evidence of heterogeneity in the strength of the first stage as a function of child gender, height-for-age z-score, or mother's education. However, our results suggest that households with low levels of durable assets – in other words, low-SES households – are disproportionately represented among the compliers.

3.4 Comparing IV and OLS

For completeness, we also report the cross-sectional OLS relationship between preprimary enrollment and child development outcomes in Appendix Table A5. These regressions are potentially biased in relation to the causal effect estimates reported in Table 2: if children's home environments or levels of development determine when they are enrolled in school, coefficient estimates in Appendix Table A5 will be biased. Comparing Panel 1 to Panel 2 in Appendix Table A5, we find evidence consistent with this bias (what is often referred to as "ability bias" at higher grades), as the inclusion of controls substantially diminishes the magnitudes of these estimates, so it is more advanced pupils who are endogenously enrolled.[18] Our instrumental variables estimates provide evidence of large effects of preprimary enrollment for Luo language but none for English, while in contrast, the cross-sectional OLS relationships between preprimary enrollment and language skills are very similar in the two languages. (The reduced form coefficients are statistically different for Luo and English, p=0.0046, while the potentially biased OLS coefficients are statistically indistinguishable for the two languages, p=0.3043.) This suggests that endogenous enrollment decisions may substantially misstate effects of preprimary, and that the impacts on the asset-poor compliers are substantially larger than cross-sectional OLS estimates would suggest. We must emphasize, of course, that the 2SLS confidence intervals are relatively wide, so comparisons between those coefficients and the OLS coefficients would generally not yield any statistically significant differences.

3.5 Using Reduced Form Results to Explore Fade Out

Early childhood interventions often have large impacts on developmental outcomes in the short-term which disappear over time. We do not have detailed information about each child's school enrollment history, but if we are willing to assume that the distance instrument had similar impacts on three year olds in the past, we can assess the extent of fade out using reduced form regressions. Specifically, we estimate OLS regressions of the form

$$Y_{ih} = \gamma + \delta D_h + \theta X_{ih} + \nu_{ih} \tag{4}$$

where Y_{ih} is an outcome of interest for child i in household h and D_h is the distance from household h to the school. Reduced form regression results are reported in Table 3. Among three year olds, we see (as expected) a negative relationship between distance from the school and Luo vocabulary, expressive vocabulary, fine motor skills, and our aggregate index of early childhood development, though the relationship is imprecisely estimated.

[18]Here we use the term "ability bias" in the sense described by Griliches (1977).

We do not find similar relationships at older ages: coefficients are smaller in magnitude, in some cases positive, and never statistically significant. In general, results suggest that the developmental impacts of enrollment at age three become attenuated over time.

Our results are consistent with many other studies which document large impacts of early childhood interventions that eventually fade (Tanner et al. 2015, Andrew et al. 2018, Özler et al. 2018). In some cases, ECD interventions can have long-term impacts, even when their short-term effects initially fade (cf. Gertler et al. 2014). It is important to view our results in context: children who enroll in preprimary at age three receive an additional year of early childhood education, but this does not put them on a more advanced curricular trajectory than children who start later. Instead, our data suggests that children who start school at age three are no more likely to be at or above the appropriate grade-for-age level at age four or five (Table 4). Most three and four year olds in our sample are enrolled in "Baby Class," the initial year of preprimary education. Hence, one potential explanation for the rapid diminution of the impacts of early enrollment is that schools don't appear to push children forward to the next grade level after a year in the first level of preprimary. This is analogous to the situation described by Andrew et al. 2018, in which an intervention's effects on early childhood outcomes were no longer detectable two years after the intervention concluded.[19]

It is also possible that there is a cognitive or non-cognitive impact of preprimary that is not captured by the measurements conducted in this study along which there might be persistent gains. For early childhood programs in the United States such as Head Start and the Perry Preschool Program, a fade-out of an initial effect followed by a reappearance of effects for longer-term outcomes might be explained this way (cf. Bailey et al. 2020; Deming 2009; Gray-Lobe, Pathak and Walters 2023). We cannot rule out the possibility that the apparent fade-out we observe might eventually be followed by a reappearance of effects on other dimensions later in life.

[19] Özler et al. (2018) and Bernal, Giannola and Nores (2023) also estimate positive short-run impacts of center-based child-development programs that subsequently become indistinguishable from zero, in Malawi and Colombia, respectively.

4 Conclusion

High-quality preprimary education is a core component of early childhood development policy, though until recently few LMIC governments in Sub-Saharan Africa offered free and universal preprimary (Behrman, Engle, and Fernald 2013, Richter et al. 2017, Devercelli and Beaton-Day 2020). Many countries in Sub-Saharan Africa have expanded access to preprimary in recent years, but to date there is little evidence on the developmental impacts of public preschool in Sub-Saharan Africa. Many recent studies measure the impacts of NGO-sponsored or community preschools (cf. Martinez, Naudeau and Pereira 2017, Blimpo et al. 2022) or private preschools (Bjortvan et al. 2022, Gray-Lobe et al. 2022). We compliment these studies by estimating the impact of existing national-scale government preschools on a set of three-year-olds who are induced to enroll at a young age by their proximity to the local school.

Using an instrumental variables approach, we estimate the impacts of preprimary education on early childhood development. Among three year olds in rural Kenya, preschool enrollment leads to large improvements in child development, particularly mother tongue receptive vocabulary. However, we we do not find evidence of persistence over time. Whether the early gains in vocabulary and motor skills that we document lead to meaningful improvements in later life outcomes remains an open question.

References

Almond, Douglas and Janet Currie, "Human Capital Development Before Age Five," National Bureau of Economic Research Working Paper No. 15827, 2010.

Andrew, Alison, Orazio Attanasio, Emla Fitzsimons, Sally Grantham-McGregor, Costas Meghir, and Marta Rubio-Codina, "Impacts 2 years after a scalable early childhood development intervention to increase psychosocial stimulation in the home: A follow-up of a cluster randomised controlled trial in Colombia," *PLoS Medicine*, 2018, *15* (4), e1002556.

Attanasio, Orazio P, Vincenzo Di Maro, and Marcos Vera-Hernández, "Community nurseries and the nutritional status of poor children. Evidence from Colombia," *The Economic Journal*, 2013, *123* (571), 1025–1058.

Attanasio, Orazio, Ricardo Paes de Barros, Pedro Carneiro, David K Evans, Lycia Lima, Pedro Olinto, and Norbert Schady, "Public childcare, labor market outcomes of caregivers, and child development: Experimental evidence from Brazil," National Bureau of Economic Research Working Paper No. 30653, 2022.

Bailey, Drew H, Greg J Duncan, Flávio Cunha, Barbara R Foorman, and David S Yeager, "Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions," *Psychological Science in the Public Interest*, 2020, *21* (2), 55–97.

Bastos, Paulo, Nicolas L Bottan, and Julian Cristia, "Access to preprimary education and progression in primary school: Evidence from Rural Guatemala," *Economic Development and Cultural Change*, 2017, *65* (3), 521–547.

Begi, Nyakwara, "Use of mother tongue as a language of instruction in early years of school to preserve the Kenyan culture," *Journal of Education and Practice*, 2014, *5* (3), 37–49.

Behrman, Jere R., Lia C.H. Fernald, and Patrice Engle, "Preschool programs in developing countries," in Paul Glewwe, ed., *Education policy in developing countries*, University of Chicago Press, 2013.

Berkes, Jan Lukas, Adrien Bouguen, Deon Filmer, and Tsuyoshi Fukao, "Improving preschool provision and encouraging demand: Heterogeneous impacts of a large-scale program," *World Bank Policy Research Working Paper*, 2019, (9070).

Berlinski, Samuel and Norbert Schady, eds, *The early years: child well-being and the role of public policy*, New York: Palgrave Macmillan New York, 2015.

___, Sebastian Galiani, and Marco Manacorda, "Giving children a better start: Preschool attendance and school-age profiles," *Journal of Public Economics*, 2008, *92* (5-6), 1416–1440.

___, ___, and Paul Gertler, "The effect of pre-primary education on primary school performance," *Journal of Public Economics*, 2009, *93* (1-2), 219–234.

Bernal, Raquel, Michele Giannola, and Milagros Nores, "The Effect of Center-Based Early Education on Disadvantaged Children's Developmental Trajectories: Experimental Evidence from Colombia," 2023.

Bietenbeck, Jan, Sanna Ericsson, and Fredrick M Wamalwa, "Preschool attendance, schooling, and cognitive skills in East Africa," *Economics of Education Review*, 2019, *73*, 101909.

Bjorvatn, Kjetil, Denise Ferris, Selim Gulesci, Arne Nasgowitz, Vincent Somville, and Lore Vandewalle, "Childcare, labor supply, and business development: Experimental evidence from Uganda," 2022.

Black, Maureen M., Susan P. Walker, Lia C.H. Fernald, Christopher T. Andersen, Ann M. DiGirolamo, Chunling Lu, Dana C. McCoy, Günther Fink, Yusra R. Shawar, Jeremy Shiffman et al., "Early Childhood Development Coming of Age: Science through the Life Course," *Lancet*, 2017, *389* (10064), 77–90.

Blimpo, Moussa P., Pedro Carneiro, Pamela Jervis, and Todd Pugatch, "Improving Access and Quality in Early Childhood Development Programs: Experimental Evidencefrom the Gambia," *Economic Development and Cultural Change*, 2022, *70* (4), 1479–1529.

Bloem, Jeffrey R. and Bruce Wydick, "All I Really Need to Know I Learned in Kindergarten? Evidence from the Philippines," *Economic Development and Cultural Change*, 2023, *71* (2), 753–791.

Bouguen, Adrien, Deon Filmer, Karen Macours, and Sophie Naudeau, "Preschool and Parental Response in a Second Best World: Evidence from a School Construction Experiment," *Journal of Human Resources*, 2018, *53* (2), 474–512.

Bradley, Robert H and Bettye M Caldwell, "The relation of infants' home environments to mental test performance at fifty-four months: A follow-up study," *Child Development*, 1976, *47* (4), 1172–1174.

Brinkman, Sally Anne, Amer Hasan, Haeil Jung, Angela Kinnell, and Menno Pradhan, "The impact of expanding access to early childhood education services in rural Indonesia," *Journal of Labor Economics*, 2017, *35* (S1), S305–S335.

Britto, Pia R., Stephen J. Lye, Kerrie Proulx, Aisha K. Yousafzi, Stephen G. Matthews, Tyler Vaivada et al., "Nurturing care: promoting early childhood development," *Lancet*, 2017, *389* (10064), P91–P102.

Burger, Kaspar, "How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds," *Early Childhood Research Quarterly*, 2010, *25*, 140–165.

Cabrera, Natasha J, Ui Jeong Moon, Jay Fagan, Jerry West, and Daniela Aldoney, "Cognitive stimulation at home and in child care and children's preacademic skills in two-parent families," *Child Development*, 2020, *91* (5), 1709–1717.

Card, David, "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," National Bureau of Economic Research Working Paper No. 4483, 1993.

Clark, Shelley, Caroline W Kabiru, Sonia Laszlo, and Stella Muthuri, "The impact of childcare on poor urban women's economic empowerment in Africa," *Demography*, 2019, *56* (4), 1247–1272.

Dean, Joshua T. and Seema Jayachandran, "Attending kindergarten improves cognitive development in India, but all kindergartens are not equal," 2020.

Dee, Thomas S., "Are There Civic Returns to Education?," *Journal of Public Economics*, 2004, *88* (9-10), 1697–1720.

Deming, David, "Early childhood intervention and life-cycle skill development: Evidence from Head Start," *American Economic Journal: Applied Economics*, 2009, *1* (3), 111–134.

Devercelli, A. and F. Beaton-Day, *Better Jobs and Brighter Futures: Investing in Childcare to Build Human Capital*, Washington, DC: World Bank, 2020.

Duflo, Esther, "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment," *American Economic Review*, 2001, *91* (4), 795–813.

Dunn, L. M. and D. M. Dunn, "PPVT-III: Peabody Picture Vocabulary Test," 1997.

___ , ___ , and B. Styles, "British Picture Vocabulary Scale (3rd ed.)," 2009.

Engle, Patrice L., Lia C.H. Fernald, Harold Alderman, Jere Behrman, Chloe O'Gara, Aisha Yousafzi et al., "Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries," *Lancet*, 2011, *378* (9799), P1339–1353.

Fernald, Lia C. H., Elizabeth Prado, Patricia Kariger, and Abbie Raikes, *A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries*, Washington, DC: The World Bank, 2017.

Filmer, Deon and Lant H. Pritchett, "Estimating Wealth Effects Without Expenditure Data – Or Tears: An Application to Educational Enrollments in States of India," *Demography*, 2001, *38* (1), 115–132.

Garcia, Italo Lopez, J Luoto, F Aboud, and L Fernald, "Group Meetings and Boosters to Sustain Early Impacts on Child Development: Experimental Evidence from Kenya," IZA Discussion Paper No. 16392, 2023.

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M Chang, and Sally Grantham-McGregor, "Labor market returns to an early childhood stimulation intervention in Jamaica," *Science*, 2014, *344* (6187), 998–1001.

Gladstone, M., G. A. Lancaster, E. Umar, M. Nyirenda, E. Kayira, N. R. van den Broek, and R. L. Smyth, "The Malawi Developmental Assessment Tool (MDAT): The Creation, Validation, and Reliability of a Tool to Assess Child Development in Rural African Settings," *PLoS Medicine*, 2010, *7* (5), e1000273.

Gray-Lobe, Guthrie, Anthony Keats, Michael Kremer, Isaac Mbiti, and Owen Ozier, "Can education be standardized? Evidence from Kenya," University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2022-68, 2022.

——, **Parag A Pathak, and Christopher R Walters**, "The long-term effects of universal preschool in Boston," *Quarterly Journal of Economics*, 2023, *138* (1), 363–411.

Griliches, Zvi, "Estimating the returns to schooling: Some econometric problems," *Econometrica*, 1977, *45* (1), 1–22.

Jakiela, Pamela, Owen Ozier, Lia C.H. Fernald, and Heather Knauer, "Evaluating the Effects of an Early Literacy Intervention," *Journal of Development Economics*, 2020, *accepted based on pre-results review.*

Kane, Thomas J. and Cecilia Rouse, "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?," National Bureau of Economic Research Working Paper No. 4268, 1993.

Kenya Institute of Curriculum Development, "Basic education curriculum framework," 2019. Available online at `https://kicd.ac.ke/wp-content/uploads/2019/08/BASIC-EDUCATION-CURRICULUM-FRAMEWORK-2019.pdf`, accessed June 9, 2023.

Kerwin, Jason T and Rebecca L Thornton, "Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures," *Review of Economics and Statistics*, 2021, *103* (2), 251–264.

Knauer, Heather A., Pamela Jakiela, Owen Ozier, Frances Aboud, and Lia C. H. Fernald, "Enhancing young children's language acquisition through parent-child book-sharing: a randomized trial in rural Kenya," *Early Childhood Research Quarterly*, 2019, *50*, 179–190.

——, **Patricia A. Kariger, Pamela Jakiela, Owen Ozier, and Lia C. H. Fernald**, "Multilingual Assessment of Early Child Development: Analysis from Repeated Observations of Children in Kenya," *Developmental Science*, 2019, *22* (5), e12875.

Martinez, Sebastian, Sophie Naudeau, and Vitor Azevedo Pereira, "Preschool and Child Development Under Extreme Poverty: Evidence from a Randomized Experiment in Rural Mozambique," World Bank Policy Research Working Paper 8290 2017.

McKenzie, David and Yaye Seynabou Sakho, "Does it pay firms to register for taxes? The impact of formality on firm profitability," *Journal of Development Economics*, 2010, *91* (1), 15–24.

Muralidharan, Karthik and Nishith Prakash, "Cycling to school: Increasing secondary school enrollment for girls in India," *American Economic Journal: Applied Economics*, 2017, *9* (3), 321–350.

Özler, Berk, Lia CH Fernald, Patricia Kariger, Christin McConnell, Michelle Neuman, and Eduardo Fraga, "Combining pre-school teacher training with parenting education: A cluster-randomized controlled trial," *Journal of Development Economics*, 2018, *133*, 448–467.

Piper, Benjamin and Emily Miksic, "Mother tongue and reading: Using early grade reading assessments to investigate language-of-instruction policy in East Africa," in "The Early Grade Reading Assessment: Application and Intervention to Improve Basic Literacy," Research Triangle Park, NC: RTI Press, 2011, pp. 139–182.

____, Stephanie Simmons Zuilkowski, Dunston Kwayumba, and Arbogast Oyanga, "Examining the secondary effects of mother-tongue literacy instruction in Kenya: Impacts on student learning in English, Kiswahili, and mathematics," *International Journal of Educational Development*, 2018, *59*, 110–127.

____, Stephanie Z. Zuilkowski, and Salome Ong'ele, "Implementing Mother Tongue Instruction in the Real World: Results from a Medium-Scale Randomized Controlled Trial in Kenya," *Comparative Education Review*, 2016, *60* (4), 776–807.

Richter, Linda M., Bernadette Daelmans, Joan Lombardi, Jody Heymann, Florencia Lopez Boo, Jere R. Behrman et al., "Investing in the foundation of sustainable development: pathways to scale up for early childhood development," *Lancet*, 2017, *389*, 103–118.

Schady, Norbert, Jere Behrman, Maria Caridad Araujo, Rodrigo Azuero, Raquel Bernal, David Bravo, Florencia Lopez-Boo, Karen Macours, Daniela Marshall, Christina Paxson et al., "Wealth gradients in early childhood cognitive development in five Latin American countries," *Journal of Human Resources*, 2015, *50* (2), 446–463.

Spier, Elizabeth, Kevin Kamto, Adria Molotsky, Azizur Rahman, Najmul Hossain, Zannatun Nahar, and Hosneara Khondker, "Bangladesh Early Years Preschool Program Impact Evaluation," American Institutes for Research Endline Report, 2020.

Tanner, Jeffrey C., Tara Candland, and Whitney S. Odden, "Later Impacts of Early Childhood Interventions: A Systematic Review," World Bank Independent Evaluation Group Working Paper 2015/3, 2015.

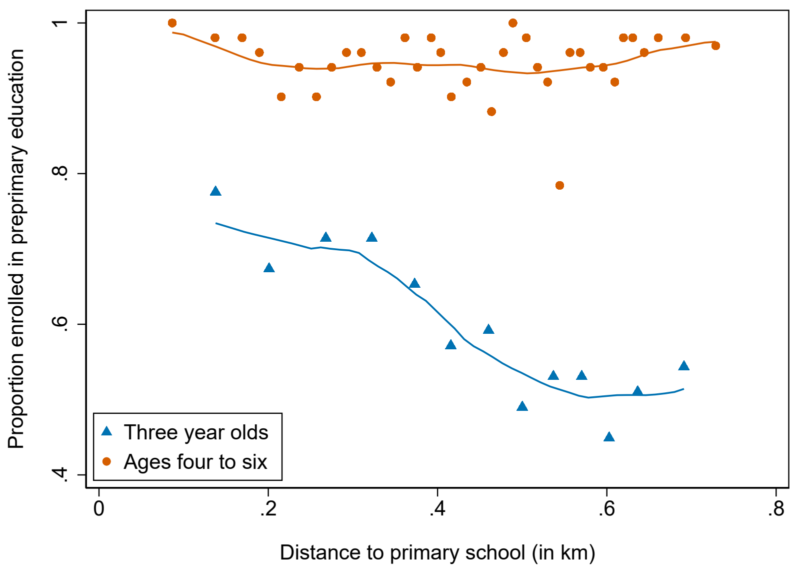
USAID, "Language of Instruction Country Profile: Kenya," Technical Report, Bureau for Africa, Office of Sustainable Development, Education and Youth Division 2021.

Wolf, Sharon, J. Lawrence Aber, Jere R. Behrman, and Edward Tsinigo, "Experimental impacts of the "Quality Preschool for Ghana" interventions on teacher professional well-being, classroom quality, and children's school readiness," *Journal of Research on Educational Effectiveness*, 2019, *12* (1), 10–37.

Wong, Ho Lun, Renfu Luo, Linxiu Zhang, and Scott Rozelle, "The impact of vouchers on preschool attendance and elementary school readiness: A randomized controlled trial in rural China," *Economics of Education Review*, 2013, *35*, 53–65.

World Bank, "World Development Indicators," 2023.

Figure 1: Probability of Enrollment by Distance from the Primary School



Data from 2,503 households with a child aged three to six. Figures plot local polynomial regressions of school enrollment on distance to the local school (measured in kilometers). Scatter plots depict averages in bins containing approximately fifty households.

Table 1: First Stage Estimates of the Impact of Distance on Preprimary Enrollment

Age in Years:	3 Years	4 Years	5 Years	6 Years
	(1)	(2)	(3)	(4)
Panel A: without covariates				
Distance to school (km)	-0.496	0.003	-0.041	0.009
	(0.116)	(0.075)	(0.025)	(0.012)
	[p<0.001]	[0.969]	[0.099]	[0.469]
Panel B: covariate-adjusted				
Distance to school (km)	-0.443	-0.034	-0.055	0.002
	(0.110)	(0.072)	(0.028)	(0.010)
	[p<0.001]	[0.642]	[0.049]	[0.832]
Obs.	634	610	669	590

All regressions estimated via OLS. Standard errors (clustered at the household level) in parentheses, p-values in brackets. Dependent variable is indicator for being enrolled in school. Covariates included in Panel B: child age in months (fixed effects), child gender (indicator for male), child height-for-age z-score, a dummy for having an imputed value of the height-for-age z-score, an indicator equal to one if a child's mother is their primary caregiver, mother's education, an indicator for having a Luo mother, household size, the number of older siblings in the household, and a household wealth index.

Table 2: 2SLS Estimates of the Impact of Preprimary Enrollment on Child Outcomes

	Vocabulary				
	Luo	English	Expressive	Fine Motor	ECD Index
	(1)	(2)	(3)	(4)	(5)
Panel A: without covariates					
Enrolled in preprimary	1.160	-0.131	0.701	0.581	0.780
	(0.520)	(0.483)	(0.475)	(0.444)	(0.471)
	[0.026]	[0.786]	[0.140]	[0.191]	[0.098]
Panel B: covariate-adjusted					
Enrolled in preprimary	1.401	-0.238	0.774	0.526	0.832
	(0.608)	(0.520)	(0.504)	(0.458)	(0.491)
	[0.021]	[0.647]	[0.124]	[0.251]	[0.090]
Obs.	634	634	634	634	634

All specifications estimated via 2-stage least squares (2SLS). First-stage F-statistics: 18.28 (Panel A) and 17.25 (Panel B). Standard errors (clustered at the household level) in parentheses, p-values in brackets. Covariates included in Panel B: child age in months (fixed effects), child gender (indicator for male), child height-for-age z-score, a dummy for having an imputed value of the height-for-age z-score, an indicator equal to one if a child's mother is their primary caregiver, mother's education, an indicator for having a Luo mother, household size, the number of older siblings in the household, and a household wealth index.

Table 3: Reduced Form Impacts of Distance to School on Child Development

Age in Years:	3 Years	4 Years	5 Years	6 Years
	(1)	(2)	(3)	(4)
Panel A: dependent variable is early childhood development index				
Distance to school (km)	-0.369	-0.217	0.019	-0.005
	(0.223)	(0.231)	(0.209)	(0.202)
	[0.099]	[0.346]	[0.927]	[0.981]
Panel B: dependent variable is Luo receptive vocabulary index				
Distance to school (km)	-0.621	-0.323	-0.064	-0.297
	(0.242)	(0.240)	(0.230)	(0.222)
	[0.011]	[0.179]	[0.780]	[0.183]
Panel C: dependent variable is expressive vocabulary index				
Distance to school (km)	-0.343	-0.029	0.165	0.336
	(0.230)	(0.224)	(0.234)	(0.206)
	[0.136]	[0.899]	[0.481]	[0.102]
Panel D: dependent variable is fine motor skills index				
Distance to school (km)	-0.233	0.163	-0.211	0.044
	(0.207)	(0.245)	(0.210)	(0.249)
	[0.261]	[0.505]	[0.314]	[0.861]
Obs.	634	610	669	590

All regressions estimated via OLS. Standard errors (clustered at the household level) in parentheses, p-values in brackets. All specifications include controls for child age in months (fixed effects), child gender (indicator for male), child height-for-age z-score, an indicator equal to one if a child's mother is their primary caregiver, mother's education, an indicator for having a Luo mother, household size, the number of older siblings in the household, and a household wealth index.

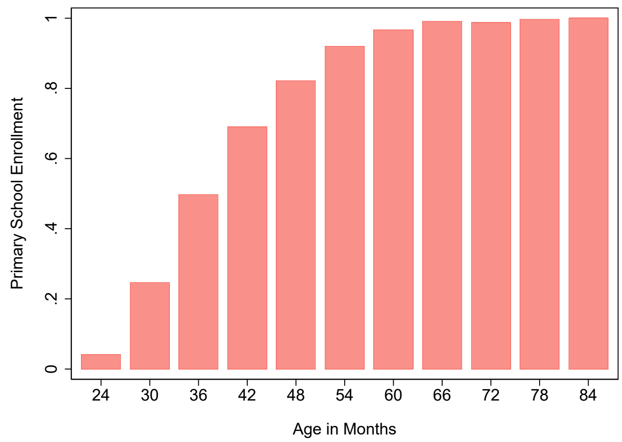
Table 4: The Impact of Distance on Likelihood of Being at Appropriate Grade-for-Age

Age in Years:	3 Years	4 Years	5 Years	6 Years
	(1)	(2)	(3)	(4)
Panel A: without covariates				
At or above appropriate grade-for-age level	-0.496	-0.151	0.041	0.005
	(0.116)	(0.114)	(0.114)	(0.106)
	[p<0.001]	[0.183]	[0.716]	[0.964]
Panel B: covariate-adjusted				
At or above appropriate grade-for-age level	-0.443	-0.258	-0.012	-0.039
	(0.110)	(0.100)	(0.100)	(0.095)
	[p<0.001]	[0.010]	[0.902]	[0.678]
Obs.	634	610	669	590

All regressions estimated via OLS. Standard errors (clustered at the household level) in parentheses, p-values in brackets. Dependent variable is indicator for being at or above appropriate grade-for age (i.e. in first year of preprimary at age three, in second year of preprimary at age four, in third year of preprimary at age five, and in standard one at age six). Covariates included in Panel B: child age in months (fixed effects), child gender (indicator for male), child height-for-age z-score, a dummy for having an imputed value of the height-for-age z-score, an indicator equal to one if a child's mother is their primary caregiver, mother's education, an indicator for having a Luo mother, household size, the number of older siblings in the household, and a household wealth index.

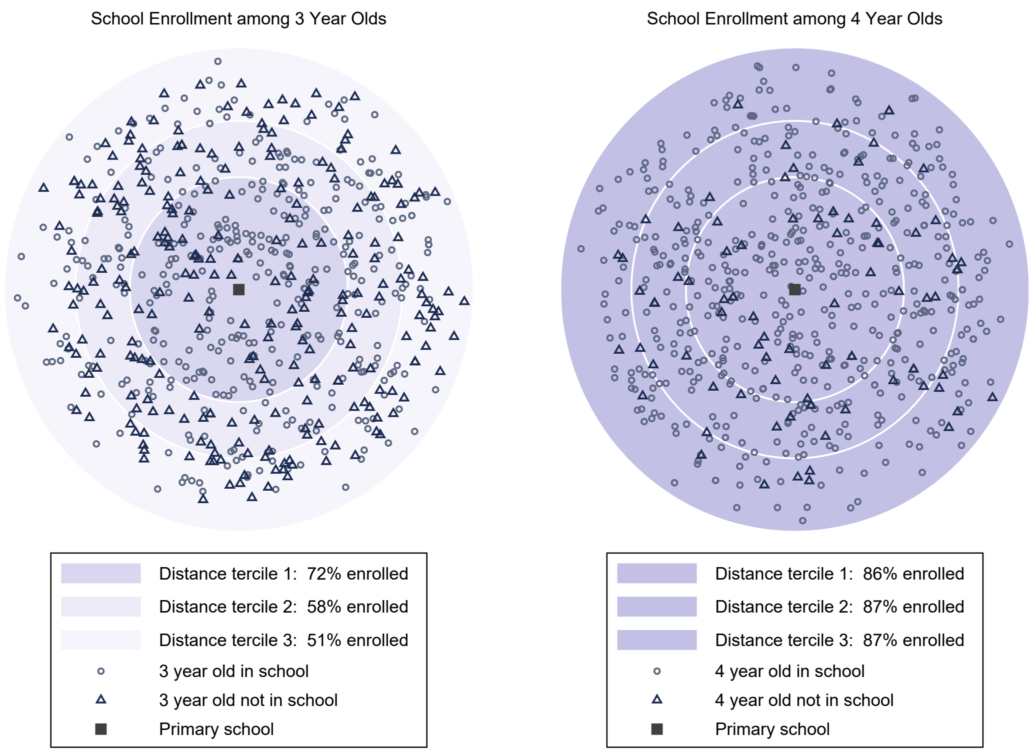
A Online Appendix: not for print publication

Figure A1: School Enrollment by Age



Data from 2,869 children aged 24 to 83 months in 2,013 households in 75 rural villages in western Kenya. Bars indicate the proportion of children in each 6 month age bin who were enrolled in school, based on parent reports.

Figure A2: Probability of Enrollment in Local School Among Three and Four Year Olds



Data from 1,188 households with a child aged three (left panel) or four (right panel). Figure plots the GPS location of each household, relative to the GPS location of the primary school. Households in the first tercile of distance are less than 359.2 meters from their local school; those in the second tercile are between 359.3 and 532.5 meters from their local school, and those in the furthest tercile are between 532.7 and 750 meters from their local school.

Table A1: Summary Statistics on Three Year Olds

	Mean	S.D.	Median	Min.	Max.	N
Child age (in months)	41.70	3.37	42	36	47	634
Height-for-age z-score	-0.57	1.44	-0.66	-4	4	634
Child is male	0.49	0.50	0	0	1	634
Mother is child's primary caregiver	0.88	0.33	1	0	1	634
Mother's education in years	7.88	2.48	8	0	13	634
Mother is Luo	0.95	0.22	1	0	1	634
Father absent from household	0.14	0.35	0	0	1	625
Father's education in years	8.49	2.67	8	0	13	568
Father is Luo	0.98	0.12	1	0	1	568
Household size	5.80	1.98	6	2	14	634
Older siblings in household	1.40	1.29	1	0	6	634
Asset index (out of 10)	3.43	1.47	3	0	9	634
Distance to school (in km)	0.44	0.17	0.46	0.05	0.75	634
Child is enrolled in school	0.60	0.49	1	0	1	634

Data on 634 children aged 36 to 47 months. Children are from 622 unique households (12 households include two three-year-old children). Asset index is the sum of indicators for having a cement floor, iron roof, latrine, or connection to the electricity grid, and indicators for owning a motorized vehicle, a bicycle, a television, a mobile phone, a computer, or a radio.

Table A2: Summary Statistics on All Children Age Three to Six

	Mean	S.D.	Median	Min.	Max.	N
Child age (in months)	59.23	13.65	60	36	83	2503
Height-for-age z-score	-0.41	1.36	-0.46	-4	4	2503
Child is male	0.50	0.50	1	0	1	2503
Mother is child's primary caregiver	0.86	0.35	1	0	1	2503
Mother's education in years	7.72	2.39	8	0	13	2503
Mother is Luo	0.95	0.22	1	0	1	2503
Father absent from household	0.17	0.37	0	0	1	2486
Father's education in years	8.55	2.63	8	0	13	2270
Father is Luo	0.98	0.12	1	0	1	2270
Household size	5.88	1.92	6	2	17	2503
Older siblings in household	1.53	1.25	1	0	8	2503
Asset index (out of 10)	3.55	1.50	3	0	9	2503
Distance to school (in km)	0.43	0.16	0.45	0.02	0.75	2503
Child is enrolled in school	0.86	0.35	1	0	1	2503

Data on 2,503 children aged 36 to 83 months. Children are from 1,994 unique households. Asset index is the sum of indicators for having a cement floor, iron roof, latrine, or connection to the electricity grid, and indicators for owning a motorized vehicle, a bicycle, a television, a mobile phone, a computer, or a radio.

Table A3: Does Distance Predict Observable Characteristics of Children and Households?

	Coefficient	S.E.	p-value
Child age (in months)	-0.511	0.793	0.519
Height-for-age z-score	-0.483	0.330	0.143
Child is male	0.025	0.121	0.836
Mother is child's primary caregiver	0.000	0.080	0.995
Mother's education in years	0.347	0.619	0.575
Mother is Luo	-0.044	0.054	0.420
Father absent from household	-0.000	0.084	0.996
Father's education in years	1.137	0.666	0.088
Father is Luo	0.074	0.040	0.062
Household size	0.419	0.510	0.412
Older siblings in household	0.528	0.301	0.080
Asset index (out of 10)	0.541	0.355	0.128

Coefficients from OLS regressions of outcome variables on distance from the school (in km). Data on 634 children aged 36 to 47 months. Children are from 622 unique households (12 households include two three-year-old children). Standard errors clustered at the household level. Asset index is the sum of indicators for having a cement floor, iron roof, latrine, or connection to the electricity grid, and indicators for owning a motorized vehicle, a bicycle, a television, a mobile phone, a computer, or a radio.

Table A4: Complier Characteristics: OLS Regressions of School Enrollment on Distance

		Indicator for Below Median:		
Characteristic:	Male	HAZ	Assets	Mother's Ed.
	(1)	(2)	(3)	(4)
Distance to school (km)	-0.604	-0.470	-0.135	-0.381
	(0.149)	(0.144)	(0.155)	(0.142)
	[0.000]	[0.001]	[0.384]	[0.007]
Characteristic	-0.176	0.013	0.268	0.062
	(0.102)	(0.114)	(0.109)	(0.111)
	[0.086]	[0.907]	[0.014]	[0.575]
Characteristic × distance	0.341	0.052	-0.606	-0.154
	(0.221)	(0.228)	(0.215)	(0.235)
	[0.124]	[0.820]	[0.005]	[0.513]

All specifications estimated via OLS. Standard errors (clustered at the household level) in parentheses, p-values in brackets. The outcome variable in all specifications is an indicator for school enrollment. The sample is restricted to three year olds. All specifications include the following covariates child age in months (fixed effects), child gender (indicator for male), child height-for-age z-score, a dummy for having an imputed value of the height-for-age z-score, an indicator equal to one if a child's mother is their primary caregiver, mother's education, an indicator for having a Luo mother, household size, the number of older siblings in the household, and a household wealth index.

Table A5: Cross-Sectional Relationship Between Preprimary Enrollment and Outcomes

	Vocabulary				
	Luo	English	Expressive	Fine Motor	ECD Index
	(1)	(2)	(3)	(4)	(5)
Panel A: without covariates					
Enrolled in preprimary	0.451	0.492	0.654	0.720	0.782
	(0.078)	(0.075)	(0.077)	(0.076)	(0.075)
	[p<0.001]	[p<0.001]	[p<0.001]	[p<0.001]	[p<0.001]
R^2	0.048	0.060	0.101	0.125	0.140
Panel B: covariate-adjusted					
Enrolled in preprimary	0.299	0.403	0.478	0.530	0.578
	(0.087)	(0.081)	(0.082)	(0.082)	(0.078)
	[p<0.001]	[p<0.001]	[p<0.001]	[p<0.001]	[p<0.001]
R^2	0.133	0.130	0.221	0.253	0.293
Obs.	634	634	634	634	634

All cross-sectional specifications estimated via OLS. A separate regression for each outcome is presented in each column of each panel. Child outcomes are listed at the top of each column. Standard errors (clustered at the household level) in parentheses, p-values in brackets. As in other estimates, the sample includes 634 children aged 36 to 47 months. Covariates included in Panel B: child age in months (fixed effects), child gender (indicator for male), child height-for-age z-score, a dummy for having an imputed value of the height-for-age z-score, an indicator equal to one if a child's mother is their primary caregiver, mother's education, an indicator for having a Luo mother, household size, the number of older siblings in the household, and a household wealth index.