Supplementary Appendix to: "The Impact of Child Work on Cognitive Development: Results from Four Low to Middle Income Countries"

by Michael Keane, Sonya Krutikova, and Timothy Neal

October 14, 2020

Appendix 1: This appendix presents a table of the summary statistics.

- **Appendix 2:** This appendix presents regressions of the lagged test score on the current test score across countries and ages.
- **Appendix 3:** This appendix presents details on the instruments that were used for the IV results.

Appendix 1: Summary Statistics

Table A1: Summary Statistics

	Ethiopia		India		Peru		Vietnam	
	YC (age 5)	OC (age 12)						
Male	0.528	0.509	0.534	0.495	0.503	0.533	0.512	0.496
	(0.499)	(0.500)	(0.499)	(0.500)	(0.500)	(0.499)	(0.500)	(0.500)
Mother's Education:								
None	0.496	0.442	0.512	0.582	0.0829	0.101	0.116	0.0954
	(0.500)	(0.497)	(0.500)	(0.493)	(0.276)	(0.301)	(0.320)	(0.294)
primary	0.397	0.485	0.279	0.289	0.454	0.506	0.484	0.493
	(0.489)	(0.500)	(0.449)	(0.454)	(0.498)	(0.500)	(0.500)	(0.500)
secondary	0.0869	0.0616	0.177	0.103	0.279	0.247	0.327	0.363
	(0.282)	(0.241)	(0.382)	(0.304)	(0.448)	(0.432)	(0.469)	(0.481)
post secondary	0.0197	0.0117	0.0321	0.0259	0.184	0.146	0.0730	0.0488
	(0.139)	(0.108)	(0.176)	(0.159)	(0.388)	(0.353)	(0.260)	(0.216)
Father's Education:								
none	0.209	0.171	0.333	0.397	0.0115	0.0136	0.0761	0.0679
	(0.407)	(0.377)	(0.471)	(0.490)	(0.107)	(0.116)	(0.265)	(0.252)
primary	0.628	0.715	0.320	0.333	0.346	0.563	0.460	0.446
	(0.483)	(0.452)	(0.467)	(0.472)	(0.476)	(0.496)	(0.499)	(0.497)
secondary	0.114	0.0850	0.255	0.201	0.452	0.274	0.378	0.406
	(0.318)	(0.279)	(0.436)	(0.401)	(0.498)	(0.446)	(0.485)	(0.491)
post secondary	0.0487	0.0287	0.0925	0.0685	0.190	0.149	0.0855	0.0795
	(0.215)	(0.167)	(0.290)	(0.253)	(0.393)	(0.356)	(0.280)	(0.271)
Household size	6.038	6.532	5.521	5.182	5.490	5.560	4.673	4.900
	(2.058)	(2.041)	(2.232)	(1.826)	(2.072)	(1.970)	(1.514)	(1.391)
No. brothers in house	1.454	1.885	0.653	0.968	0.953	1.286	0.536	0.887
	(1.380)	(1.448)	(0.725)	(0.842)	(1.118)	(1.161)	(0.707)	(0.846)
No. sisters in house	1.384	1.861	0.810	0.890	0.951	1.127	0.616	0.881
	(1.333)	(1.411)	(0.875)	(0.901)	(1.153)	(1.067)	(0.842)	(0.972)
No. grandparents in house	0.238	0.181	0.959	0.507	0.617	0.346	0.680	0.278
	(0.548)	(0.467)	(0.963)	(0.752)	(0.891)	(0.674)	(0.907)	(0.564)
No. other adults in house	0.419	0.407	0.764	0.341	0.709	0.475	0.452	0.148
	(1.067)	(1.012)	(1.424)	(1.054)	(1.271)	(1.044)	(1.003)	(0.542)
No. other elderly in house	0.00547	0.0106	0.00105	0	0.0132	0.0121	0.00573	0.00212
	(0.0738)	(0.103)	(0.0324)	(0)	(0.127)	(0.109)	(0.0755)	(0.0460)
Both parents in house	0.759	0.651	0.945	0.863	0.791	0.709	0.919	0.917
	(0.428)	(0.477)	(0.228)	(0.344)	(0.407)	(0.454)	(0.273)	(0.276)
Mother's age	31.45	38.26	27.63	34.70	31.27	38.40	31.18	38.43

	(6.404)	(6.946)	(4.315)	(5.634)	(6.566)	(6.555)	(5.782)	(5.713)
Father's age	40.72			40.89				
Father's age		47.95	33.44		35.31	42.47	34.09	40.77
	(8.650)	(8.390)	(5.203)	(6.160)	(6.889)	(6.942)	(5.948)	(6.025)
Child's age (in months)	62.37	145.2	64.74	148.5	63.98	148.2	63.67	147.6
	(3.796)	(3.718)	(3.713)	(4.080)	(4.681)	(5.227)	(3.619)	(3.782)
Child lives in urban area	0.401	0.407	0.252	0.245	0.695	0.747	0.205	0.196
	(0.490)	(0.492)	(0.434)	(0.431)	(0.461)	(0.435)	(0.404)	(0.397)
Height for age z-score	-1.441	-1.374	-1.638	-1.641	-1.532	-1.531	-1.338	-1.472
	(1.121)	(1.264)	(1.111)	(1.672)	(1.131)	(1.157)	(1.109)	(1.088)
Wealth index	0.287	0.303	0.459	0.469	0.470	0.504	0.490	0.512
	(0.178)	(0.169)	(0.195)	(0.200)	(0.230)	(0.222)	(0.181)	(0.171)
Child Religion (1)	0.717	0.727	0.875	0.874	0.809	0.837	0.857	0.839
	(0.451)	(0.446)	(0.330)	(0.331)	(0.393)	(0.369)	(0.350)	(0.368)
Child Religion (2)	0.114	0.112	0.0699	0.0674	0.133	0.133	0.143	0.161
	(0.318)	(0.315)	(0.255)	(0.251)	(0.340)	(0.339)	(0.350)	(0.368)
Child Religion (3)	0.160	0.154	0.0547	0.0581	0.0577	0.0301	NA	NA
	(0.366)	(0.361)	(0.227)	(0.234)	(0.233)	(0.171)		
Child Religion (4)	0.00984	0.00744	NA	NA	NA	NA	NA	NA
	(0.0987)	(0.0860)						
Child Ethnicity (1)	0.288	0.287	0.181	0.206	0.916	0.926	0.857	0.874
	(0.453)	(0.453)	(0.385)	(0.405)	(0.277)	(0.262)	(0.350)	(0.332)
Child Ethnicity (2)	0.214	0.207	0.148	0.109	0.0561	0.0422	0.143	0.126
	(0.410)	(0.406)	(0.355)	(0.312)	(0.230)	(0.201)	(0.350)	(0.332)
Child Ethnicity (3)	0.228	0.228	0.469	0.469	0.0278	0.0316	NA	NA
	(0.420)	(0.420)	(0.499)	(0.499)	(0.164)	(0.175)		
Child Ethnicity (4)	0.270	0.277	0.139	0.152	NA	NA	NA	NA
	(0.444)	(0.448)	(0.346)	(0.360)				
Child Ethnicity (5)	NA	NA	0.0636	0.0633	NA	NA	NA	NA
			(0.244)	(0.244)				
N	1829	941	1903	964	1906	664	1919	943

Notes: Standard deviations in brackets; Sample includes all those for whom at least one of the main models can be estimated. Height-for-age z-scores calculated using WHO 2006 reference tables. Religion codes by country: Ethiopia 1=Christian Orthodox, 2=Other Christian, 3=Muslim, 4=Other; India 1=Hindu, 2=Muslim, 3=Other (includes Christian, Buddhist); Peru 1=Catholic, 2=Evangelist, 3=Other (biggest group = none); Vietnam 1=none, 2=Other (biggest groups include Buddhist, ancestor worship). Ethnicity codes by country: Ethiopia 1=Amhara, 2=Oromo, 3=Tigrian, 4=Other (biggest groups include Gurage, Hadiva, Sidama, Wolavta); India 1= Scheduled Caste, 2=Scheduled Tribe, 3=Backward Caste, 4=Other Hindu, 5=Other non-Hindu; Peru 1=Mestizo, 2=White, 3=Other; Vietnam 1=Majority (Kinh), 2=Minority (biggest groups include H'mong, Dao, Tay, Nung). Wealth index, constructed and publicly archived by the Young Lives team is a simple average of three separate indexes that range from 0 to 1: housing quality, consumer durables, and access to services. The housing quality index is a mean of (1) rooms per person (number of rooms divided by number of household members), set to take a maximum value of 1; (2) floor quality (a dummy variable which takes the value of 1 if the floor is made of finished material); and (3) roof quality (a dummy variable that takes the value of 1 if the roof is made of iron, concrete tiles, or slate. The consumer durables index is the proportion of durables a household owns from a list of seven (radio, motorbike/scooter, bicycle, TV, motorized vehicle or truck, landline telephone, modern bed or table). The services index is the proportion of key services that a household has access to: electricity, piped water, own pit latrine/flush toilet, and modern cooking fuel (gas, kerosene, or electricity).

Appendix 2: Lagged Test Scores and Time Inputs

Two key features of our extended value added model are: (i) we control for lagged test scores and lagged time inputs, and (ii) we allow the effects of these lagged inputs (as well as the coefficients on current inputs and background variables) to differ by both age and country. Table A2 reports the coefficients on lagged math scores in the math ability equations. Notice that the lagged math score, which the value added model uses to control for unobserved ability, is highly significant in all instances. At ages 12, 15 and 19, the lagged score coefficients range from 0.36 to 0.74, with a median point estimate of 0.46.

However, the coefficients on the age 5 score are much smaller in magnitude. As we note in Section IV.B, age 5 math skills were assessed using the CDA, an instrument designed for very young children, while math skills at later ages were assessed using paper and pencil tests. The changing nature of the test explains the smaller coefficient on lagged score at age 8.

	Ethio	pia	India		Peru		Vietnam	
	coef	se	coef	Se	coef	se	coef	Se
Age 8: Lagged math score age 5	0.049***	0.013	0.151***	0.018	0.103***	0.016	0.091***	0.021
Age 12: Lagged math score age 8	0.560***	0.053	0.477***	0.033	0.541***	0.029	0.352***	0.045
Age 15: Lagged math score age 12	0.361***	0.037	0.395***	0.031	0.400***	0.034	0.393***	0.037
Age 19: Lagged math score age 15	0.562***	0.038	0.677***	0.042	0.721***	0.046	0.448***	0.044

Table A2: Coefficients on Lagged Test Score – MATH equation

Table A3 reports the coefficients on lagged verbal (PPVT test) scores in the verbal ability equations. Note that the lagged verbal score is highly significant in all instances. The lagged score coefficients range from 0.20 to 0.60, with a median point estimate of 0.31. Thus, PPVT is slightly less persistent over time than the math score.

	Ethiopia		India		Peru		Vietnam	
	coef	se	Coef	Se	coef	se	coef	se
Age 8: lagged PPVT score age 5	0.221***	0.051	0.267***	0.039	0.363***	0.038	0.218***	0.028
Age 12: lagged PPVT score, age 8	0.394***	0.077	0.339***	0.027	0.494***	0.032	0.278***	0.045
Age 15: lagged PPVT score, age 12	0.284***	0.073	0.431***	0.060	0.626***	0.040	0.278***	0.053

Table A3: Coefficients on Lagged Test Score – PPVT equation

As we discussed in Section V, if test scores measure ability with error the coefficient on the lagged score will tend to be biased downward, causing the coefficients on time inputs to be biased in an ambiguous direction. To address this issue, we re-estimated our models using the lagged PPVT score to instrument for the lagged math score, and vice versa. We find that these are strong instruments. Instrumenting causes the lagged test score coefficients to increase as expected if there is independent measurement error in the two tests (typically by 50% to 100% of the values reported in Tables A2 and A3). However, we find the coefficients on the time use variables are hardly affected (results available on request).

The lagged time input coefficients are too numerous for us to report in detail. But we can highlight some key results. First, the lagged 5 time inputs are jointly significant (based on the F-test) in 12 of the 16 math equations (8 times at the 1% level, 4 times at the 5% level). Thus, including lagged inputs (as suggested by Todd and Wolpin, 2007) clearly improves the fit of the math ability equations. The lagged time inputs are somewhat less important in the PPVT equations (i.e., they are jointly significant in only 4 of the 12 verbal ability equations).

Another potential concern with our results is measurement error in the parent reports of child time use. But at ages 12 and 15 children were also asked about their own time use. Thus, in this subset of cases, we can use the child's reports to instrument for the parent's reports. While the IV estimates are slightly less precise, they are very similar to the OLS estimates (results available on request).

Appendix 3: Details of the Construction of Instruments

This appendix will list each of the instruments that were included in the 2SLS models in Table 5 and 6 in the paper, the methodology used to construct them, and finally the instruments that were picked by LASSO for the constrained instrument set. There are two types of instruments that we considered: instruments that vary across communities and those that vary across households. Several of the household-level instruments and control variables were also interacted with the community-level instruments and rainfall to capture additional variation in the first stage regression.

Table A4 details the community-level instruments, which are principal components of community prices, wages and services. Due to changes in the survey questions between Round 3 and 4 in the Young Lives surveys, the components of each category might change between the IV regression for Age 8 (which is Round 3) and Age 12 (which is Round 4).

Instrument	Methodology
Community	Young Lives conducts community-level surveys with a representative from each community. The
Prices	two instruments used here are the first two principal components of prices for the following goods
	Grade 3 textbook, Notebook, Shoes, Boy's shirt, Girl's shirt, Boy's pants, Girl's skirt, Oral rehydration salts, Paracetamol, Amoxicillin, Mebendazol, Cigarettes, Detergent, Kerosene, Fertiliser – dap, Fertiliser – urea, Barley, Wheat, Corn, Sorghum, Coffee, Sugar, Salt, White teff, Black teff, Millet, Oats, Horse Beans, Cow peas, Chick peas, Field peas, Other pulse, Milk, Yoghurt, Butter, Eggs, Beef, Mutton, Chicken meat, Goat meat, Other meat, Chat, Araqi, Tej, Cooking oil, Karia, Berbere, Other spices, Bread, Enset, Potato, Gommen Selata, Jinjibel, Tikl gommen, Nech shinkuri, Fasolia, Fenugreek, Onions, and Vegetable.
Community	Young Lives conducts community-level surveys with a representative from each community. The
Wages	two instruments used here are the first two principal components of the following average daily wage information for males and females (when available):
	Age 8 (Round 3): Prepare the land for agricultural use, Plant/sow, Weed out agricultural land,
	Harvest, Perform post-harvest duties, To pasture / to put cattle to graze, shepherding, Helping at workshops, Construction worker, Factory worker, Taxi / minibus driver, Security guard, Maid /
	domestic worker, Shop assistant, Teacher, Police, Mechanic, Cook, Fisherman, Tailor, Military, Computer Operator, and other tasks.
	Age 12 (Round 4): Prepare the land for agricultural use, Plant/sow, Weed out agricultural land,
	Harvest, Perform post-harvest duties, To pasture / to put cattle to graze, shepherding, and other tasks.
Community	The 'service/infrastructure factors' are the first two principal components derived from a range of
Services	variables in the community-level surveys that seek to measure the range and quality of
	infrastructure or services in the community.

Table A4: Community-level instruments

The variables are defined as S_{ic} for type of infrastructure/service *i* and community *c*. $S_{ic} = 0$ if the type of infrastructure *i* is absent from the community, $S_{ic} = 1$ if *i* is present but considered 'bad' in quality, $S_{ic} = 2$ if *i* is present and considered 'so-so' in quality, and finally $S_{ic} = 3$ if present and considered 'good' in quality. The infrastructure or services included are as follows:

Age 8 and Age 12: space exclusively assigned for little children (i.e. children's playground), sporting fields, camping zones or family recreational areas, indoor/outdoor movie theatres, video games, public telephones, private telephones, public internet cabin, electricity, drinkable water, sewerage, police station, public bank, and private bank.

Age 8 only: fairgrounds for temporary recreational activities, religious institutions (church, mosque, etc.), and schools.

Age 12 only: private internet cabin, mobile phone service, primary agricultural cooperative society, and local government credit/savings schemes.

Table A5 details the household-level instruments. Several of them have been interacted with the community-level factors, as well as the rainfall instruments, to capture additional variation. These are time to school, livestock prices, agricultural prices, frost, and the month of interview dummies. Several control variables were also interacted, which are gender, the wealth index, the urban dummy, the number of brothers in the household, and the number of sisters in the household. Lastly, the community-level factors and rainfall were also interacted with the number of older siblings in the house. The number of older siblings was not included in the instrument set in levels out of concern for potential endogeneity. Table A6 lists the instruments that were picked by LASSO at each Age Group and whether the lagged test score was included in the IV regression as a control variable.

Table A5: Household-level instruments

Instrument	Methodology
Agricultural Prices (as levels and interacted with Gender)	$AP_{it} = \sum_{c=1}^{C} w_{it,c} R_{it,c} / \sum w_{it,c}$ where $w_{it,c}$ is the quantity of crop <i>c</i> that is grown by individual <i>i</i> 's family at time <i>t</i> (if the household does not grow crops then $w_{it,c} = 0 \forall c$ and $AP_{it} = 0$). $R_{it,c}$ is the relative price of crop <i>c</i> at time <i>t</i> , defined as the price of crop <i>c</i> at time <i>t</i> for household <i>i</i> (which varies by community) divided by the average price for the same price measured in the same community at all observed time periods. The crops included in the set <i>C</i> are: barley, black/mixed teff, white teff, cabbage, chat, chick peas,
	cow peas, wheat, sorghum, potatoes, onions, oats, maize, horse beans, vegetables, and coffee.

Livestock Prices (as levels and	$LP_{it} = \sum_{l=1}^{L} w_{it,l} R_{it,l} / \sum w_{it,l}$
interacted with	$\sum_{l=1}^{l} w_{ll,l} (l,l) \sum_{l=1}^{l} w_{ll,l}$
Gender)	where $w_{it,l}$ is the value of livestock l that is held by individual i 's family at time t (if the household does not hold livestock then $w_{it,l} = 0 \forall l$ and $LP_{it} = 0$). $R_{it,l}$ is the relative price of livestock l at time t , defined as the price of livestock l 's output at time t for household i (which varies by community) divided by the average price for the same price measured in the same community at al observed time periods.
	The livestock considered in this analysis, and the definition of its output, is as follows: The output price for cows is the average price of milk, beef, and yoghurt. The output price for chickens is the average price of chicken meat and eggs, the output price for sheep is the price of mutton, and the output price for goats is the price of goat meat.
Whether the father is employed in Agriculture	Constructed dummy variable from Young Lives survey data. It equals 1 if child father's 'most important activity' is reported as either: 'self-employed (food crops)', 'self-employed (aquaculture)', 'self-employed (livestock)', 'wage employment (agriculture)', 'annual farm servant', or 'other agricultural work', and 0 otherwise.
Whether the local school is public	The Young Lives survey asks respondent to classify the school that is being attended. This instrument = 1 if the school is publicly operated, and 0 otherwise.
Month of	Young Lives provides date of interview in DD/MM/YY format. The month of the interview is
Interview Dummies	extracted from the date and turned into a series of dummies. For ET age 8 the following month dummies are used: 1, 10, 11, 12. For ET Age 12 we use dummies for month: 1, 2, 10, 11, 12.
Environmental Shock - Frost	The Young Lives survey asks respondents if they had experienced frosts.
Time to school	Young Lives survey question which asks respondent to report 'Travel time to school (in minutes)'.
Rainfall in last 2	Step1: Obtain Rainfall Raster Images at the Relevant Dates
and 12 months	Satellite imagery of rainfall was obtained using the Tropical Rainfall Measuring Mission (TRMM) a joint mission by NASA and the Japanese Space Agency (JAXA). The rainfall maps are at 0.25 degrees resolution and collected for all months around the date of interviews in Round 3 and 4 of the Ethiopia Young Lives Surveys. Available online at: https://neo.sci.gsfc.nasa.gov/view.php?datasetId=TRMM_3B43M
	Step 2: Estimate the GPS location of Ethiopian communities in the Young Lives surveys
	The GPS locations of the communities surveyed in Young Lives were estimated from the sentinel site information provided by Table 1 of the Young Lives Survey Design and Sampling (Round 5) document for Ethiopia. Available online at: <u>https://www.younglives.org.uk/content/survey-design-and-sampling-round-5-ethiopia</u>
	Step 3: Construct the Instruments
	The two instruments measure the demeaned rainfall in the last 2 and 12 months. The monthly sum of precipitation at the child's community GPS location is aggregated for 2/12 months prior to the household-level date of interview (usually interviews in a single community is conducted over two months, so there is a little variation within communities). This sum of rainfall over the last 2/12 months is then demeaned from the summer sinfall over the sum 2000, 2010

month is then demeaned from the average rainfall over those same months over 2000-2019.

For example, if the household was interviewed was in October 2008, the rainfall in the last 2 months instrument is calculated from: Rainfall at GPS location in October 2008 + Rainfall at GPS location in September 2008 – Average rainfall at GPS location in October 2000-2019 – Average rainfall at GPS location in September 2000-2019.

Table A6: Instruments chosen by LASSO

Regression	Instruments chosen by LASSO in the first stage regressions
Age 8 (no lagged test score)	<i>Work:</i> Agricultural Prices * Gender, Father in Agriculture, Rain (last two months) * Agricultural Prices, Wage factor 1 * Time to School, Wage factor 1 * Livestock Prices, Price factor 1 * Gender, Price factor 1 * # of Older Siblings
	<i>Chores</i> : Rain (last two months) * Urban, Wage factor 1 * Agricultural Prices, Price factor 1 * Urban, Services factor 1 * # of Sisters, Services factor 1 * Agricultural Prices, Services factor 2 * Time to School
	<i>Leisure</i> : Rain (last 12 months), Rain (last 12 months) * Agricultural Prices, Wage factor 2 * Time to School, Wage factor 2 * # of Older Siblings, Price factor 2 * Gender, Price factor 2 * Urban
	<i>Sleep</i> : Services factor 2, Wage factor 2 * October dummy, Price factor 2 * Frost Shock, Services factor 2 * Agricultural Prices
Age 8 (with lagged test score)	<i>Work:</i> Agricultural Prices * Gender, Time to School, Father in Agriculture, Rain (last two months) * Agricultural Prices, Wage factor 1 * Time to School, Wage factor 1 * Livestock Prices, Price factor 1 * Gender
	<i>Chores</i> : Services factor 2, Price factor 1 * Urban, Price factor 1 * # of Older Siblings, Services factor 2 * Time to School
	<i>Leisure</i> : Rain (last 12 months) * Agricultural Prices, Wage factor 2 * Time to School, Wage factor 2 * # of Older Siblings, Price factor 2 * Gender, Price factor 2 * Urban
	<i>Sleep</i> : Rain (last 12 months) * November Dummy, Price factor 2, Price factor 2 *Frost Shock, Services factor 1 * # of Brothers
Age 12 (no lagged test score)	<i>Work:</i> Livestock Prices * Gender, Public School dummy, Rain (last 12 months) * Livestock Prices, Price factor 2 * Gender, Price factor 2 * Time to School, Services factor 2 * Gender, Services factor 2 * Frost Shock, Services factor 2 * # of Older Siblings
	<i>Chores</i> : December dummy, Wage factor 1, Wage factor 2 * Gender, Wage factor 2 * Urban, Wage factor 2 * # of Brothers, Price factor 2 * Wealth index, Price factor 2 * Urban, Price factor 2 * Livestock Prices, Services factor 1 * Wealth index, Services factor 1 * Urban, Services factor 1 * October Dummy, Services factor 2 * Wealth index, Services factor 2 * Urban
	Leisure: Rainfall (last two months) * Urban, Rainfall (last 12 months) * Wealth index
	<i>Sleep</i> : Wage factor 2, Wage factor 2 * Livestock Prices, Wage factor 2 * Frost Shock, Wage factor 2 * October dummy

Age 12 (with	Work: Livestock Prices * Gender, Public School dummy, Rain (last 12 months) * Livestock Prices
lagged test score)	Price factor 2 * Gender, Price factor 2 * Time to School, Services factor 2 * Gender, Services factor 2 * Frost Shock, Services factor 2 * # of Older Siblings
	<i>Chores</i> : Wage factor 1, Wage factor 2 * Gender, Wage factor 2 * # of Brothers, Price factor 2 * Wealth index, Price factor 2 * Urban, Services factor 1 * Wealth index, Services factor 2 * Urban
	<i>Leisure</i> : Rainfall (last two months) * Urban, Rainfall (last 12 months) * Gender, Rainfall (last 12 month) * Wealth index, Wage factor 1 * November Dummy, Wage factor 2 * Urban, Services factor 1 * November Dummy,
	Sleep: Wage factor 2, Wage factor 2 * Frost dummy, Wage factor 2 * October Dummy