Child work and cognitive development: Results from four low to middle income countries

Michael Keane
School of Economics, University of New South Wales

Sonya Krutikova
Institute for Fiscal Studies

Timothy Neal
School of Economics, University of New South Wales

We study the impact of child work on cognitive development in four Low- and Middle-Income Countries. We advance the literature by using cognitive test scores collected regardless of school attendance. We also address a key gap in the literature by controlling for children's complete time allocation budget. This allows us to estimate effects of different types of work, like chores and market/farm work, relative to specific alternative time-uses, like school or study or play/leisure. Our results show child work is more detrimental to child development to the extent that it crowds out school/study time rather than leisure. We also show the adverse effect of time spent on domestic chores is similar to time spent on market and farm work, provided they both crowd out school/study time. Thus, policies to enhance child development should target a shift from all forms of work toward educational activities.

Keywords. Child labor, child development, education, time use, item response theory, value added models.

JEL classification. I25, J13, J24, O15.

1. Introduction

Is child work harmful for child development? Existing evidence suggests the answer is "probably yes." But the results are inconclusive, because the prior literature suffers from several important data limitations: First, and most fundamentally, it relies on cognitive test scores collected only for children in school, creating an important selection problem. Second, it typically lacks test scores that are comparable across ages, limiting the

usefulness of panel data methods to control for unobserved ability and unmeasured inputs. Third, it lacks data on the complete time budget of children, so it fails to control for how children spend their non-work time. Hence, prior literature is not informative about effects of child work *relative* to other uses of time, such as school/study and leisure. Fourth, existing studies have failed to find reasonably strong instruments for child work, or other uses of time, limiting the usefulness of instrumental variables methods.

We make substantial progress in overcoming these limitations by using Young Lives, a multicountry panel data set that follows children in four low to middle income countries—Ethiopia, India (Andhra Pradesh and Telangana), Vietnam and Peru—from birth until age 22. The survey uses state-of-the-art methods to measure children's cognitive skills using scores on math and verbal tests. For the first time in the child labor literature, Young Lives collected test scores that are *both* comparable over time and obtained regardless of school attendance. The data also contain detailed information on children's time use over a "typical" 24-hour period, including time spent on market/farm work, household chores, school, home study, sleeping, and playing. And the data either contain, or can be linked with, detailed information on local weather, wages, and prices that allow us to form better instruments for child work.

We use Young Lives data to estimate child cognitive ability production functions that incorporate children's complete time budget, as well as controls for other inputs (e.g., parent education, wealth, family structure). Of course, a fundamental problem confronting the entire child development literature—not just our study—is that one cannot measure all inputs to child development, or control perfectly for latent child ability. We compare three ways to deal with this problem: Our preferred method is to exploit the panel aspect of the data to estimate value added (VA) production functions where omitted inputs and latent ability are proxied by a lagged test score (as in Todd and Wolpin (2007), Fiorini and Keane (2014)).[1] We also consider fixed effects (FE) models as an alternative method to deal with omitted inputs and latent ability. And we implement an instrumental variables approach using agricultural prices, wages, and weather to instrument for child time-use. This is only feasible for Ethiopia, where such instruments are strong because child labor is prevalent and heavily agricultural.

Our results are broadly consistent across the three approaches. We find time spent on all forms of child work—either market/farm work or chores—has a comparable negative effect on child cognitive development if it crowds out school/study time. This result is novel, as prior work typically finds time spent on chores is much less harmful for child development than time spent on market/farm work. We also find that market/farm work and chores do *not* have adverse effects on cognitive development if they only substitute for leisure.[2] Hence, policies that only shift work time to leisure will do little to enhance

---

[1]Several recent studies find VA models are a reliable way to control for latent ability. These studies rely on simulation (Guarino, Reccase, and Wooldridge (2014)) or comparison of experimental and VA estimates (Angrist, Pathak, and Walters (2013), Deming, Hastings, Kane, and Staiger (2014), Muralidharan and Venkatesh (2013)).

[2]Our IV results are similar to VA and FE in that work and chores both have similar negative effects relative to school/study. But our IV results imply larger negative effects of both work and chores relative to leisure. This may be because work/chores are particularly detrimental for "marginal" children pushed in and out of child labor by the variation of the instruments.

child development. It is important to design policies that shift children's work time toward educational activities.

We also examine heterogeneity in these patterns. We find young boys are a notable exception, as market/farm work is more detrimental for them than either chores or leisure. Market/farm work is particularly detrimental for young boys (ages 8 and 12).

We go on to show how failure to control for children's complete time-use budget—as is typical in prior work due to data limitations—leads to misleading conclusions. In fact, it would cause us to conclude that market/farm work is much more harmful for cognitive development than chores, simply because market/farm work tends to crowd out more school/study time.

We proceed as follows: Section 2 reviews the literature, Section 3 presents a very simple theoretical model to guide the analysis. Section 4 describes the Young Lives data and Section 5 describes our econometric methods. Section 6 to 9 present results obtained using different methods, and Section 10 concludes.

2. Background and literature review

A large proportion of the world's children are engaged in some form of work. According to the International Labour Organisation (ILO) one in ten children in the world today (or 152 million) participate in what they define as child labor.[3] Prevalence is highest in Africa where 20% of children are engaged in child labor, most of which is agricultural. The extent of child work is higher if household chores are included. For instance, 54 million 5 to 14-year olds spend at least 21 hours a week on household chores (ILO 2017a).

There is broad consensus in policy circles that child labor is detrimental for child development, making its elimination a high-profile issue in Low to Middle Income Countries (LMIC). The UN Sustainable Development Goals call for the eradication of all forms of child labor by 2025. The great majority of the world's children live in countries that have ratified the ILO's two main child labor conventions (ILO 2017b),[4] and, according to the ILO, from 2004 to 2014, 57 LMICs implemented 279 policies, plans, and programs aimed at reducing child labor. There is also growing international pressure on companies to avoid using child labor. An example is the US Trade Facilitation and Trade Enforcement Act of 2015, which banned import of goods made by forced or indentured child labor.

Given this broad policy consensus, it is perhaps surprising that the academic literature on the effect of child labor on child development has failed to reach definitive conclusions. There are several sources of ambiguity:

---

[3]We define "child work" to include not only paid market work, but also domestic chores, care responsibilities, and work on the family farm/business. "Child labor" is typically defined as a narrower set of activities, which excludes domestic chores (see ILO 2017b). The ILO defines child labor as "work that is harmful to children's physical and mental development [including] work that... interferes with their schooling by: depriving them of the opportunity to attend school; obliging them to leave school prematurely; or requiring them... to combine school attendance with excessively long and heavy work." (http://www.ilo.org/ipec/facts/lang–en/index.htm)

[4]Minimum Age Convention of 1973 (No. 138) and Worst Forms of Child Labour Convention, 1999 (No. 182).

Much of the evidence on child labor concerns its effects on schooling. Most studies find negative effects on schooling (see Assad, Levison, and Dang (2010), Beegle, Dehejia, and Gatti (2009), Boozer and Suri (2001), Buonomo Zabaleta (2011), Sedlacek, Duryea, Ilahi, and Sasaki (2009)), although some find complementarity (Patrinos and Psacharopoulos (1997), Ravallion and Wodon (2000), Ray and Lancaster (2005)).

But the weaknesses of schooling as a measure of human capital are well understood.[5] This suggests we should focus instead on *direct* measures of human capital, like ability test scores, which Glewwe (2002) finds predict wages better than schooling.

However, evidence on the impact of child labor on direct human capital measures is limited. Early studies found negative correlations between child labor and ability test scores (Akabayashi and Psacharopoulos (1999), Heady (2003)). Recent studies use panel data or IV methods to try to establish causality. IV studies by Gunnarsson, Orazem, and Sánchez (2006) and Bezerra, Kassouf, and Arends-Kuenning (2009) find negative effects, but they are very imprecisely estimated. Emerson, Ponczek, and Portela Souza (2017) may be the clearest evidence to date of a negative effect: Using fixed effects models to control for latent child ability, they find that working while in school lowers annual test score gains for children in Sáo Paulo municipal schools by roughly one quarter to three-fifths. But Dumas (2012), who controls for lagged test scores, finds that, in Senegal, child work actually has (small) positive effects on math scores.

There are important limitations to existing studies based on test scores: The studies cited above only have data on test scores for the subsample of children who are *at school.* This may exclude children most adversely affected by child labor (i.e., those who drop out). A key advantage of Young Lives is that it gathers test scores for *all* children *regardless of school attendance.* Another advantage of Young Lives is that it provides test scores that are comparable over time, something rarely achieved in existing data. This allows use of panel methods like VA and FE models to study the evolution of test scores over time (Das and Zajonc (2010)).

An additional limitation of the IV studies cited above is that they all suffer from very weak instruments. But Young Lives contains, or can be linked to, detailed data on local agricultural prices/wages and weather that let us form relatively strong instruments in the Ethiopia case.

A final important limitation of existing studies is lack of data on the complete time budget of children.[6, 7] This precludes estimating effects of child work *relative* to specific alterative activities like leisure or school/study. It also precludes testing whether different types of child work have different effects on child development. Most prior studies focus on market work (Beegle, Dehejia, and Gatti (2009), Gunnarsson, Orazem, and

---

[5]Child labor may hinder learning even if it does not reduce school enrolment (Gunnarsson, Orazem, and Sánchez (2006), Dumas (2012)). Conversely, nonschool activities, including child labor, may generate human capital.

[6]A notable exception is Akabayashi and Psacharopoulos (1999), who look at effects of school attendance and time spent working and studying on reading and math ability. But a key limitation of their study is that the math and reading measures are simply parental reports of whether the child could read or do math.

[7]Some theoretical models consider more complete time-use vectors (Cigno and Rosati (2005), Edmonds (2007)).

Sánchez (2006)). But Edmonds (2007) descriptive analysis of UNICEF's MICS data suggests the total time a child spends on all types of work is likely to matter for human capital, not just market activities. The Young Lives data enable us to fill this gap in the literature because it collects the complete time budget of children. It separately classifies market work, farm work, and household chores, allowing us to test the impact of different types of work on child development.

3. A SIMPLE THEORETICAL FRAMEWORK

In developed countries, school attendance is mandatory for young children. As a result, the child development literature *in that context* focusses on how parental investments—parental time and goods inputs, hours/quality of childcare, school quality—affect child development. Child quality is viewed as an *output* produced by families (Becker (1960)).

But in developing countries many children do not go to school, and the child plays an important dual role as both an output produced by families and an *input* into household production of goods/income. Allocation of a child's time between market/farm work, chores, and school is a first-order problem. Hence our simple theoretical framework (and subsequent empirical work) emphasizes the allocation of a child's time between school and work, while abstracting from the parental time inputs emphasized in the developed county literature. But the key econometric problem is similar in both contexts: Time inputs are chosen by parents, so they are likely to be correlated with latent child ability, creating an endogeneity problem.

To elucidate the endogeneity problem, as well as potential solutions, consider a very simple unitary household model, where a household makes decisions about adult work hours (h), child work hours (h_c), and child school hours (s). The household faces a budget constraint:

$$C = wh + w_c h_c - ps + N,$$

where C is consumption, w is the adult wage, w_c is the child wage, p is the price of school, and N is nonlabor income. The household also faces the time constraints:

$$l = H - h; \qquad l_c = H - h_c - s,$$

where H is available time, and l and l_c are adult and child leisure, respectively. The final constraint is the production function for child cognitive ability (Y), which takes the form:

$$Y = \mu + s\beta_s - h_c\beta_h. \tag{1}$$

Here, μ is the child skill endowment, β_s is the effect of s on Y, and β_h is the effect of h_c on Y, both measured relative to leisure (the residual category). The household maximizes a utility function $u(C, Y, l, l_c, s)$. Without loss of generality, we let utility depend on both child leisure and school time, with child work hours the residual category. In

this model, the three decision variables (h, h_c, s) are driven by four exogenous factors—namely w, w_c, p, and N—as well as the child skill endowment μ. An interior solution satisfies the first-order conditions:

$$u_c w = u_l, \qquad u_c w_c - u_Y \beta_h = u_{l_c}, \qquad u_Y \beta_s + u_s(\mu) - u_c p = u_{l_c}.$$

We write $u_s(\mu)$ to emphasize that the child skill endowment μ plausibly shifts the marginal utility of school. For example, youth with higher μ may have a greater marginal utility of school hours $u_{s\mu} > 0$ because they are better at school and find it easier (see Keane and Wolpin (1997)).

This simple model illustrates how the five factors w, w_c, p, N, and μ influence the solution (h^*, h_c^*, s^*). For clarity, consider a special case with $p = 0$, $u_{YY} = 0$ and parent work predetermined, and set all cross partials of $u(\cdot)$ to zero. Then total differentiation of the FOCs (see Appendix A in the Online Supplementary Material (Keane, Krutikova, and Neal (2022)) reveals that:

$$\frac{ds^*}{d\mu} = -u_{s\mu}/[u_{ss} + u_{l_c l_c} R_c] > 0, \qquad \frac{dh_c^*}{d\mu} = u_{s\mu}/[u_{ss} + u_{cc} w_c^2 R_l] < 0,$$

where $R_c = (u_{cc} w_c^2/[u_{cc} w_c^2 + u_{l_c l_c}])$ and $R_l = ([u_{ss} + u_{l_c l_c}]/u_{l_c l_c})$. That is, school hours are increasing in child ability, while child work hours are decreasing, in the plausible case that $u_{s\mu} > 0$, provided that $u_{ss} < 0$, $u_{l_c l_c} < 0$ and $u_{cc} < 0$. We also have

$$\frac{ds^*}{dN} = u_{cc} w_c/[u_{ss} + u_{cc} w_c^2 R_l] > 0, \qquad \frac{dh_c^*}{dN} = -u_{cc} w_c/[u_{cc} w_c^2 + u_{l_c l_c}/R_l] < 0$$

so wealthier households choose a higher level of child schooling and less child work.[8]

The main point is that the child time allocation (h_c^*, s^*) is driven by *both* exogenous factors (w, w_c, p, N) and the child skill endowment μ. Thus, in estimating the production function for child cognitive ability $Y = \mu + s\beta_s - h_c\beta_h$ we face the problem that s and h_c are likely to be correlated with μ. But on the plus side, there is also exogenous variation in s and h_c induced by (w, w_c, p, N). We seek to identify the production function from the latter.

There are three main approaches to this problem. *First*, we could estimate a structural model, estimating the cognitive ability production function jointly with households' decision rules for child schooling and child work. *Second*, we could control for μ when estimating the production function. Given panel data, we could use a FE model where we difference out μ, or a VA model, where we use a lagged ability measure to proxy for μ (see Todd and Wolpin (2003, 2007)). In these panel-data approaches, we identify β_s and β_h from variation in s and h_c induced by the variation in exogenous factors (w, w_c, p, N) holding μ fixed. *Third*, we could take an IV approach, using exogenous factors

---

[8]If we allow for $u_{YY} < 0$ the sign of $ds^*/d\mu$ becomes ambiguous, as it depends on the relative magnitude of u_{YY} and $u_{s\mu}$. But ds^*/dN remains unambiguously positive, while the sign of ds^*/dw_c is ambiguous, as it depends on income and substitution effects. Allowing for $p > 0$ complicates the analysis but does not change the key points.

(w, w_c, p, N) to instrument for s and h_c. We then identify β_s and β_h from exogenous variation in s and h_c induced by these instruments.

A structural approach would require us to expand the simple expository model above into a more realistic model that we could credibly use to fit data on household decisions. This would be incredibly challenging for several reasons: We would need to model how the time inputs of parents, siblings and relatives are jointly determined—a daunting task. Even if this approach were feasible, we lack the necessary data to implement it. We would also need to model farm households in Ethiopia, India, Peru, and Vietnam in different ways, as crops, farm technology, and the role of children in farming differ greatly by country. Furthermore, many households are urban so they face a whole different set of constraints, again differing by country.

The VA, FE, and IV approaches are more practical, as they enable us to abstract from the detailed structure of how households choose child time inputs. Instead, we can focus on estimating the production function for child cognitive ability in (1) separately from the rest of the model. These approaches deliver consistent estimates of the production function under different assumptions, and identify effects of time inputs using different sources of variation, so each has advantages/disadvantages. If the VA model assumptions hold then, conditional on the controls for latent ability (and other inputs), all remaining variation in the time inputs is used to identify the production function. The FE approach relies only on the within-child variation in the time inputs. The IV approach relies only on variation in the time inputs generated by the instruments. If the instruments are valid this variation is independent of latent ability, but it is only a subset of the variation used to identify the VA model. Thus, both IV and FE may entail a large efficiency loss if the assumptions that underlie the VA approach are valid.

As a concrete example, when we implement IV in Section 8, our instruments include a local wage index, assumed to capture local demand conditions. In our simple model, local demand shocks would shift parent and child wages (w, w_c), which in turn generates exogenous (independent of μ) variation in child time inputs.[9] The VA approach implicitly assumes such exogenous variation occurs "behind the scenes," generating some variation in child time inputs that is independent of latent ability. Thus, if we can adequately control for μ in the production function, the VA approach identifies the effects of child time inputs from *all* their observed variation that is independent of μ, including variation induced by factors besides the instrument. In contrast, the IV approach relies only on the independent variation in time use generated by the instruments, which is only a subset of the variation used to identify the VA estimates. This is why IV is less efficient (than VA) if the VA assumptions hold.

Prior to our work, only two papers used panel data to control for latent ability via lagged scores (Dumas (2012)), or fixed effects (Emerson, Ponczek, and Portela Souza (2017)). Young Lives is ideally suited to estimate VA and FE models as it collected, for

---

[9]It seems plausible that the dimension of latent skill (μ) that affects cognitive ability (Y) is not relevant for productivity in the sorts of menial work that children typically do. But even if it were, this is not a problem for the IV approach if instruments are uncorrelated with μ. Nor is it a problem for the VA approach if we adequately control for μ in the production function, or for FE if differencing eliminates μ from the production function.

the first time, test scores that are *both* comparable over time and obtained regardless of school attendance. Prior panel data analysis did not have the benefit of repeated measures with both features. Similarly, only three prior papers implemented IV: Gunnarsson et al. (2006), Bezerra et al. (2009), Dumas (2012). They all suffer from weak instruments, while Young Lives provides a richer set of instruments.

While our simple model divides child time into work, school and leisure, the Young Lives data contain more detailed time use categories. Of particular interest is that it splits work time into market/farm work and household chores. We are interested in whether these different types of work have different effects on child development. For example, if all that each type of work does is detract from school/study time, we would expect both to have the same negative effect on child cognitive development. On the other hand, market/farm work might be worse because it is more tiring, or better because the content of work teaches useful skills. It is also possible that different types of work have different effects on math versus verbal skills (e.g., types of work that involve interaction with adults might improve verbal skills). Furthermore, the content of work and school may differ by county and child age.

Thus, in Section 5 we discuss how we generalize (1) to allow for a more detailed set of child time inputs that distinguishes between market/farm work and household chores. We also explain how we let the parameters of (1) differ flexibly by country/age.

4. The Young Lives data

Young Lives collects data on two cohorts of children from four countries: Ethiopia, India (Andhra Pradesh and Telangana), Peru, and Vietnam. The "Older Cohort" (OC) was born in 1994/95 while the "Younger Cohort" (YC) was born in 2001/02. They were surveyed in 2002, 2006, 2009, and 2013. Thus, the YC is observed from birth through ages 12/13, while the OC is observed from ages 8/9 through ages 19/20.

The sample contains about 12,000 children, consisting of 2000 YC and 1000 OC children from each country.[10] One hundred YC and 50 OC children were randomly sampled from 20 sites in each country, selected to represent its demographic diversity, with a pro-poor bias.[11] Young Lives collect detailed information about each target child, as well as detailed data on household economic circumstances and demographics.

All data were collected through home visits by interviewers; these include interviews with the person who knows most about the socioeconomic circumstances of the household, and the primary caregiver of the target child. The interviewer also does a direct assessment of the target child. At older ages, the target child was also interviewed. In each round, an intense effort was made to find children who had moved since the prior survey. As a result, attrition from 2002 to 2013 is very low—only 3.6% in the YC and 8.3% in the OC.[12]

---

[10]In Peru, only 714 OC children were enrolled due to capacity constraints during the first round.

[11]With the exception of Peru, samples were not selected to be statistically representative. However, subsequent comparisons to nationally representative surveys in each country have shown that the data reflect the diversity of children in the two cohorts across a wide number of variables. See technical notes at www.younglives.org.uk.

[12]These figures increase to 5.5% and 9.1% if we include mortality.

TABLE 1. Analysis sample sizes by country, round and outcome variable.

	Ethiopia			India			Peru			Vietnam		
	Math	Verbal	N	Math	Verbal	N	Math	Verbal	N	Math	Verbal	N
R3, 2009 (YC, 8)	1240	488	1885	1672	1382	1931	1846	1738	1942	1701	1640	1961
R4, 2013 (YC, 12)	1799	714	1867	1857	1560	1912	1828	1750	1902	1840	1807	1918
R3, 2009 (OC, 15)	924	362	974	931	859	976	653	627	676	918	910	976
R4, 2013 (OC, 19)	851	NA	905	920	NA	951	575	NA	635	775	NA	882

Cognitive ability test scores were not collected in the first round, so we restrict our analysis to the 2nd, 3rd, and 4th rounds. The YC children were on average age 5, 8, and 12 in rounds 2, 3, and 4, respectively; the OC children were 12, 15, and 19. We control for lagged test scores in our models, so the round two data are only used to form these lags. Combining the two cohorts, we can study outcomes at ages 8, 12, 15, and 19 in rounds 3 and 4. Verbal test scores were not collected at age 19.

Table 1 shows total (N) and analysis sample size for each country at each age. In most cases, the analysis sample is slightly smaller than N due to missing values for test scores, time use and/or controls. But in some cases, we lose more data for two main reasons.

First, we focus on the largest language group in each country, as the verbal skill measure is language specific. In Ethiopia, this is Amharic, which is only a quarter of the full sample. We also lose 20% of the India sample for a similar reason. Second, time-use data were only collected at ages 5+. Not all children had turned 5 by the 2nd round, so we lose 10 to 20% of the age 8 observations for the YC due to missing lagged scores.

4.1 *Child skill assessments*

The first key feature of Young Lives that makes it ideal for our analysis is the high-quality of the cognitive skill measures. Math and verbal skills were tested *regardless* of whether a child was in school, allowing us to avoid the selection bias present in prior studies. As we discuss in Appendix B, state-of-the-art item response theory (IRT) methods were used by the Young Lives team to create tests that are comparable across ages and countries.

Math tests were administered to the young cohort at ages 5, 8, and 12 and the older cohort at 12, 15, and 19. Math skills at age 5 were assessed by the Cognitive Development Assessment (CDA), designed by the International Association for Evaluation of Educational Achievement to assess cognitive development of young children (Montie, Xiang, and Schweinhart (2006)).[13] After age 5, math skills are assessed using paper-based tests designed by education experts on the Young Lives team. In order to cover the wide range of math proficiency across countries and ages, the tests contain items of highly varied levels of difficulty (see Appendix B).

---

[13]The CDA tests children's understanding of concepts such as few, most, half, many, equal, and pairs with statements such as "Point to the plate that has fewer cupcakes."

Verbal skills are measured by the Peabody Picture Vocabulary Test (PPVT).[14] The PPVT was administered to the younger cohort at ages 5, 8, and 12 and the older cohort at ages 12 and 15. In the PPVT, the child is asked to select a picture that best gives the meaning of a stimulus word presented orally by the examiner.[15]

4.2 *Children's time-use*

The second feature that makes Young Lives ideal for our analysis is the data on child time use, collected starting in round 2, when the YC were 5 and OC were 12. Young Lives measures time spent on eight activities in a "typical weekday": (1) sleep; (2) caring for other household members; (3) domestic tasks (e.g., fetching water or firewood, cleaning, cooking, washing); (4) tasks on the family farm, cattle herding, other family business; (5) paid work and other activities outside the home; (6) school (including travel time); (7) study outside of school (at home, extra tuition); (8) play time and leisure. We combine categories (2) and (3) into a single category that we call "domestic chores," and we combine categories (4) and (5) into a single category we call "market/farm work."

Table 2 presents summary statistics for time allocation. Work at ages 8 and 12 is most prevalent in Ethiopia, where children average 4 hours of total work per day (domestic chores plus market/farm work). Peru is next with total hours of 1.63 per day at age 8 and 2.37 at age 12. Work time at young ages is lowest in India—0.58 hours per day at age 8 and 1.15 hours at age 12. Combined school/study time at ages 8 and 12 follows the reverse pattern. It is highest in India (about 9 1/2 hours per day), and lowest in Ethiopia (6 to 7 hours per day). Interestingly, school time in Vietnam is lower than in any other country, but study time is much higher.

In all four countries, children at ages 8, 12, and 15 spend (on average) more time on domestic chores than market/farm work, typically by a wide margin. But at age 19 time spent on market/farm work is greater. Combined work time increases sharply in India, Peru, and Vietnam at age 19, catching up to the levels in Ethiopia.

Leisure time is highest in Vietnam (5 to 6 hours per day depending on age), and next highest in India (4 to 5 hours per day). In Peru, it varies from 3 1/3 to 4 1/4 hours per day, and it is lowest in Ethiopia, where it varies from 3 to 4 1/3 hours. In all four countries, sleep time is 9 to 10 hours at age 8 and drops to about 8 1/3 hours at age 19.

Table 3 summarizes work time along the extensive and intensive margins. At age 8, the fraction of children who do some form of work is 38% in India, 54% in Vietnam, 76% in Peru to a high of 93% in Ethiopia. By age 12, the large majority of children do some work in every country. Domestic chores are more common than market/farm work in all countries and at all ages. At age 8, the fraction of children engaged in market/farm work ranges from low single digits in India and Vietnam to 16% in Peru and 40% in Ethiopia. But by age 15, about *a quarter* to *a half of* children are engaged in market/farm work.

---

[14]The PPVT-III (Dunn, Dunn, and American Guidance Service (1997)) was adapted for administration in Vietnam, Ethiopia, and India. In Peru, the Spanish version PPVT-R adapted for Latin America was used (Dunn, Padilla, Lugo, and Dunn (1986)).

[15]Several studies find a strong positive correlation between PPVT and commonly-used intelligence tests, such as the Wechsler and McCarthy Scales (Campbell (1998), Campbell, Bell, and Keith (2001), Gray, Plante, Vance, and Henrichsen (1999).

TABLE 2. Child time allocation—mean hours of each activity in a "typical day."

	Sleep	Domestic Chores*	Market/Farm Work*	All Work	At School	Studying at Home	Play/Leisure
Ethiopia							
Age 8	9.63	2.47	1.47	3.94	5.07	1.05	4.31
Age 12	9.25	2.43	1.63	4.06	5.63	1.47	3.59
Age 15	8.66	3.16	1.70	4.86	5.62	1.88	2.98
Age 19	8.50	2.65	3.26	5.91	3.72	1.68	4.20
India							
Age 8	9.15	0.56	0.03	0.58	7.69	1.84	4.74
Age 12	8.95	0.97	0.18	1.15	8.06	1.85	3.99
Age 15	8.27	1.68	1.47	3.16	6.50	2.04	4.03
Age 19	8.24	2.59	3.16	5.75	3.78	1.21	5.02
Peru							
Age 8	9.97	1.37	0.25	1.63	6.21	1.94	4.25
Age 12	9.51	1.87	0.51	2.37	6.24	2.07	3.80
Age 15	9.15	2.18	1.07	3.25	6.11	2.16	3.33
Age 19	8.47	2.62	3.93	6.55	3.70	1.54	3.74
Vietnam							
Age 8	9.72	0.80	0.10	0.89	4.96	2.84	5.59
Age 12	8.97	1.36	0.42	1.78	5.46	2.71	5.08
Age 15	8.69	1.62	1.44	3.06	4.24	3.05	4.96
Age 19	8.28	1.83	4.16	5.99	2.83	1.25	5.64

Note: Domestic chores include caring for others in the household (e.g., younger children or ill household members) and domestic tasks (such as fetching water, firewood, cleaning, cooking, and washing); Market/farm work includes tasks on the family farm, cattle herding, other family business (not just farming), as well as paid work or activities outside of the household for someone not in the household.

Looking at the column of Table 3 headed "Both," we can see that the large majority of children who do market/farm work also do chores (consistent with MICS data in Edmonds (2007)). But less than half of children who do domestic chores also do in market/farm work.

Conditional on working, children aged 8–12 work about 4 hours per day in Ethiopia, and roughly 2 hours per day in the other three countries. Ethiopia stands out as the country with highest level of child work, with 90% of children working at all ages, and mean hours (conditional on working) at least 4 1/4 per day, even at age 8.

Figure 1 describes how school time varies with total hours spent on market/farm work or household chores. The results are from a multivariate fractional polynomial model, including controls for country, child age, and household demographics (see Appendix C for details). As the figure shows, it is possible to sustain up to 3 hours of work or chores time with only modest reductions in school hours.[16] Beyond that, school hours fall more quickly. Also, we see that market/farm work crowds out school time more rapidly than chores.

[16]Similarly, using data from UNICEF's MICS project, Edmonds (2009) finds work hours can increase up to a threshold with little effect on school attendance.

TABLE 3. Prevalence and intensity of chores and work by age and country.

	Incidence of Work (%)				Hours Conditional on Work		
	Work any	Domestic Chores (any)	Market and Farm Work (any)	Both	Domestic Chores	Market/Farm Work	Total (if any Work)
Ethiopia							
Age 8	93	85	40	33	2.89	3.67	4.25
Age 12	94	85	48	39	2.87	3.41	4.33
Age 15	99	92	45	38	3.45	3.77	4.92
Age 19	91	73	53	34	3.64	6.20	6.50
India							
Age 8	38	38	01	01	1.48	2.39	1.53
Age 12	69	68	04	03	1.43	4.37	1.65
Age 15	77	72	23	19	2.33	6.29	4.12
Age 19	88	83	41	36	3.14	7.65	6.55
Peru							
Age 8	76	75	16	14	1.84	1.57	2.12
Age 12	90	88	25	23	2.12	2.02	2.65
Age 15	91	89	33	30	2.46	3.28	3.56
Age 19	91	74	53	35	3.55	7.47	7.19
Vietnam							
Age 8	54	52	04	03	1.53	2.30	1.66
Age 12	77	74	21	18	1.83	2.06	2.32
Age 15	93	87	33	26	1.86	4.41	3.27
Age 19	93	80	57	44	2.28	7.33	6.44

Note: Domestic chores include caring for others in the household (e.g., younger children or ill household members) and domestic tasks (such as fetching water, firewood, cleaning, cooking and washing); Market/farm work includes tasks on the family farm, cattle herding, other family business (not just farming), as well as paid work or activities outside of the household for someone not in the household.

4.3 *Other inputs and descriptive statistics*

Our simple model of Section 3 abstracted from goods inputs into child development, but it is important to control for them as well. As in all prior work, we do not observe goods directly, so we proxy for them using measures of household resources.[17] These are a household wealth index and parents' education, which proxy for permanent income, the child's height-for-age z-score, which is a good proxy for nutrition inputs, and the number of siblings and relatives in the household, which influence resources available for the target child. Good data on current income is only available for Ethiopia, Peru, and India in round 3, which correspond to age 8 for the YC and 15 for the OC. But in these cases we find income is insignificant in our models once we control for our resource measures, so we exclude it from our main specifications.[18]

---

[17]As we discussed in Section 3, child work can be an important source of income. Hence, child labor bans may harm children by reducing household resources (Basu (1999), Jacoby and Skoufias (1997)).

[18]Given the volatility of current income, as well as measurement error, our resource measures are likely to be better proxies for goods inputs, as they capture permanent income. Furthermore, height-for-age is a very good proxy for nutrition inputs, which is presumably very highly correlated with overall goods inputs.
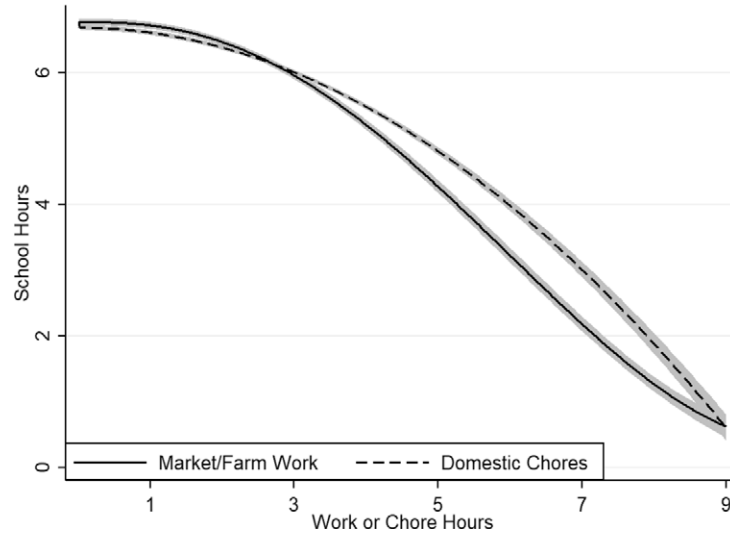
FIGURE 1. Evolution of school hours with market/farm work and domestic chores. *Note*: Figure 1 shows the predicted level of school hours conditional on market/farm hours or hours of domestic chores. We combine data for all countries and ages. The results are from a multivariate fractional polynomial model, including controls for country, child age, and household demographics. See Appendix C for details. The curve for market/farm work holds chores fixed, and vice versa. 95% confident intervals are shaded.

Appendix D reports summary statistics for the control variables (in round 2 in 2006). Parental education is much lower in Ethiopia and India than in Peru and Vietnam. About half of mothers in Ethiopia and India have no schooling. Ethiopian children tend to live in larger households with more siblings. The Ethiopia, India, and Vietnam samples are mostly rural, while Peru is more urban. In all countries, the average child is at least a standard deviation shorter than a typical healthy child of their age, reflecting the pro-poor bias of the survey.

5. ESTIMATION: VALUE ADDED AND FIXED EFFECTS APPROACHES

The child cognitive ability production function that we take to the data is a substantial generalization of Eq. (1) in the simple expository model of Section 3. To fix ideas, let the production function for *measured* skill, given by the test score Y of child i at age a, be

$$Y_{ia} = F_a\big(X_i(a), U_i(a), \mu_i, m_{ia}\big), \tag{2}$$

where $X_i(a) = (x_{i1}, \ldots, x_{ia})$ and $U_i(a) = (u_{i1}, \ldots, u_{ia})$ are vectors that contain the history of observed and unobserved inputs into child development, while μ_i is child i's innate ability (or latent skill), and m_{ia} captures transitory measurement error in the skill test. Assuming for simplicity we observe a child for two periods, and the production function is linear, we have:

$$Y_{i1} = x'_{i,1}\beta_1^1 + u'_{i,1}\gamma_1^1 + \rho_1\mu_i + m_{i1}, \tag{3}$$

$$Y_{i2} = x'_{i,2}\boldsymbol{\beta}_2^2 + x'_{i,1}\boldsymbol{\beta}_1^2 + u'_{i,2}\boldsymbol{\gamma}_2^2 + u'_{i,1}\boldsymbol{\gamma}_1^2 + \rho_2\mu_i + m_{i2}, \tag{4}$$

where $\boldsymbol{\beta}_t^a$ is the effect of an age t input on ability at age a, and similarly for $\boldsymbol{\gamma}_t^a$.

A great challenge in estimating this model is omitted variables. We do not observe μ_i or (u_{i1}, u_{i2}) as no data set can contain a complete history of relevant inputs. An endogeneity problem arises if these unobserved inputs are correlated with the observed inputs. There are two main approaches to deal with this problem using panel-data: First, under the assumption that $\rho_2 = \rho_1 = \rho$, a first-difference (FD) transform will eliminate latent skill, as in

$$Y_{i2} - Y_{i1} = x'_{i,2}\boldsymbol{\beta}_2^2 - x'_{i,1}(\boldsymbol{\beta}_1^1 - \boldsymbol{\beta}_1^2) + (u'_{i,2}\boldsymbol{\gamma}_2^2 - u'_{i,1}(\boldsymbol{\gamma}_1^1 - \boldsymbol{\gamma}_1^2) + m_{i2} - m_{i1}). \tag{5}$$

Thus, estimation of a first-difference model, or fixed effects (FE) if $T > 2$, gives consistent estimates of effects of *observed* inputs—time-use variables in our case—if: (i) the effects of *unobserved* inputs are time invariant (i.e., $u'_{i,2}\boldsymbol{\gamma}_2^2 = u'_{i,1}(\boldsymbol{\gamma}_1^1 - \boldsymbol{\gamma}_1^2)$) so they drop out of (5), or if (ii) observed inputs are uncorrelated with the non-zero part of $u'_{i,2}\boldsymbol{\gamma}_2^2 - u'_{i,1}(\boldsymbol{\gamma}_1^1 - \boldsymbol{\gamma}_1^2)$.

Alternatively, one may use a value added (VA) specification to attempt to control for unobserved ability and other unobserved inputs. The standard VA model has the form:

$$Y_{i2} = \rho Y_{i,1} + x'_{i2}\boldsymbol{\beta}_2^2 + (u'_{i,2}\boldsymbol{\gamma}_2^2 + \varepsilon_{i2}), \tag{6}$$

where we make explicit the unobserved inputs in the error. Substituting for Y_{i1} we obtain

$$Y_{i2} = x'_{i2}\boldsymbol{\beta}_2^2 + x'_{i,1}\rho\boldsymbol{\beta}_1^1 + u'_{i,1}\rho\boldsymbol{\gamma}_1^1 + \rho\rho_1\mu_i + \rho m_{i1} + u'_{i,2}\boldsymbol{\gamma}_2^2 + \varepsilon_{i2}. \tag{7}$$

Note that (7) is equivalent to (4) if (i) the effect of the latent skill endowment depreciates at the rate ρ, so $\rho_2 = \rho\rho_1$, (ii) the effect of lagged inputs (both observed and unobserved) also depreciate at rate ρ, so that $\boldsymbol{\beta}_1^2 = \rho\boldsymbol{\beta}_1^1$, and $\boldsymbol{\gamma}_1^2 = \rho\boldsymbol{\gamma}_1^1$, and (iii) $\varepsilon_{i2} \equiv m_{i2} - \rho m_{i,1}$. Thus, in a VA model, the lagged test score controls for latent ability and lagged inputs provided that the common depreciation rate assumption holds.[19]

The error term in the VA model in (6) contains unmeasured current period inputs, u_{i2}. So does the error term in the FE model in (5) unless $u'_{i,2}\boldsymbol{\gamma}_2^2 = u'_{i,1}(\boldsymbol{\gamma}_1^1 - \boldsymbol{\gamma}_1^2)$. So in each case, endogeneity arises if u_{i2} is correlated with measured current inputs (x_{i2}).[20] One way to deal with this problem is IV, but we defer discussion of that approach until Section 7. In the VA and FE approaches, we include measures of family background and nontime inputs in our models in an attempt to control for u_{ia} as well as

---

[19]With more than two periods, conditions (i) and (ii) generalize to $\rho_a = \rho^a \ \forall a$, and $\boldsymbol{\beta}_{a-k}^a = \rho^k \boldsymbol{\beta} \ \forall a$, and $\boldsymbol{\gamma}_{a-k}^a = \rho^k \boldsymbol{\gamma} \ \forall a$. That is, the VA model assumes effects of lagged inputs $x_{i,a-k}$ and the initial skill endowment μ_i on Y_{ia} all depreciate at rate ρ. Of course, as $\rho \to 1$ the VA and first difference (or FE) models are equivalent.

[20]For example, suppose the nutrition input into child development is unmeasured. A positive shock to the wage of child labor may cause an increase in child work that is exogenous in the sense it is uncorrelated with ability, but it may also improve the nutrition input because household income increases. Thus, we may underestimate the *ceteris paribus* negative effect of child work because it coincides with improved nutrition. But we provide evidence that our resource measures are good enough to allay this concern; see Section 4.3.

possible. [21, 22] Letting $\boldsymbol{b}_{ia}$ denote this vector of controls, and extending our model to a periods, we rewrite (6) as

$$Y_{ia} = \rho Y_{i,a-1} + \boldsymbol{x}_{ia}'\boldsymbol{\beta}_a^a + \boldsymbol{b}_{ia}'\boldsymbol{\delta}_{ia} + \varepsilon_{ia}. \tag{8}$$

As Todd and Wolpin (2003) stress, the FE and VA models give consistent estimates under different conditions. The FE model assumes latent child ability and other unobserved inputs have constant effects over time, so differencing or demeaning eliminates them, while the VA approach assumes their effects depreciate at a common rate over time, so lagged test scores control for them. We adopt VA as our preferred approach, as specification tests in Todd and Wolpin (2007) and Fiorini and Keane (2014) favor it, as do the validation studies noted in the Introduction. But we also estimate FE models as a robustness check.

Finally, we present our estimating equation. Following Todd and Wolpin (2007), we relax the VA model's common depreciation rate assumption by estimating an "extended" VA model that adds lagged inputs $\boldsymbol{x}_{i,a-1}$ to (8). We also extend the VA model by allowing the depreciation rate of the lagged score and lagged inputs to vary by age. Then we have

$$Y_{ia} = \rho_a Y_{i,a-1} + \boldsymbol{x}_{ia}'\boldsymbol{\beta}_a^a + \boldsymbol{x}_{i,a-1}'\boldsymbol{\pi}_a + \boldsymbol{b}_{ia}'\boldsymbol{\delta}_{ia} + \varepsilon_{ia}. \tag{9}$$

We estimate Eq. (9) separately for children of each age in each country.

In our specification of (9), the time-varying inputs $\boldsymbol{x}_{ia}$ are the vector of child time uses: time spent on market/farm work, domestic chores, school, study, leisure, and sleep. The control variables $\boldsymbol{b}_{ia}$ are characteristics of the *child* (age, gender, religion, ethnicity, height-for-age z-score), the *parents* (father's and mother's age and education, whether both parents are present), and the *household* (urban/rural, wealth, siblings, and other adults present).

Note that the elements of $\boldsymbol{b}_{ia}$ either vary deterministically over time (age), change little over time, or are time-invariant. Thus, we do not include $\boldsymbol{b}_{i,a-1}$ in (9). The full set of control variables included in the main models are listed in Appendix D.

We also estimated VA models like (9) with height-for-age as the dependent variable; see Appendix E. We find children's time use is insignificant in this regression, while our household resource proxies are highly significant with expected signs. This implies time use is uncorrelated with unmeasured nutrition inputs once we control for our resource measures.

---

[21]Formally, the VA model gives consistent estimates of the effects of *current* observed inputs if (i) and (ii) hold, $Y_{i,a-1}$ is predetermined and observed current inputs x_{ia} are uncorrelated with unobserved current inputs u_{ia}.

[22]Measurement error in the lagged test score is also a potential problem for the VA model. The error term ε_{ia} in (8) includes—$\rho m_{i,a-1}$. So ε_{ia} is negatively correlated with measurement error in the lagged test score. This biases $\hat{\rho}$ downward, and biases $\hat{\boldsymbol{\beta}}$ in an ambiguous direction. We find instrumenting for the lagged score using a different test shifts $\hat{\rho}$ upward by 50% or more, but has little impact on estimated effects of the time inputs.

6. Effects of child work on child development—disaggregate VA results

We present our results in four parts. Here, we estimate the impact of child work on child development at the country/age level using the VA approach. Sections 7 and 8 report results using FE or IV methods. Then in Section 9, we return to the VA approach and focus on heterogeneity in effects of work by country, age, level of hours, skill, and gender.

6.1 *Child work and child development: Results by country/age*

First, we estimate the cognitive ability production function in (9) separately by country/age. For each of the four countries we estimate math equations for ages 8, 12, 15, and 19, and verbal equations for ages 8, 12, and 15. This gives 28 regressions (16 for math, 12 for verbal), so we cannot present all estimates. Instead, Figures 2 and 3 summarize our main results *graphically* by plotting the coefficients on the key variables of interest: hours of domestic chores and market/farm work. We also report 95% confidence intervals around the point estimates. We focus on effects of current time use (see Appendix F for discussion of the lagged test score and lagged time inputs, and Appendix K for the resource measures).

We begin by looking at simple models where time use enters linearly. Thus, effects that we estimate here should be interpreted as effects at the margin for a typical child. Later, in Section 9, we study whether effects of child work differ by the level of hours or ability.

To identify the model, we must choose an omitted time use category. We then obtain estimates of the effect of child work *relative* to that category. While the choice of omitted category has no substantive importance, it has practical importance for proper interpretation of the estimates. For example, if we omit leisure, the coefficient on chores is interpreted as the effect of increasing time spent on chores by one hour while reducing leisure by one hour. Thus, a simple way to visualize the effect of child work relative to each alternative time-use category is to omit each category in turn and plot the coefficients on domestic chores and market/farm work in each case. We do this in the three panels of Figures 2 and 3.

The results presented in Figures 2 and 3 provide no clear evidence that time spent working is less productive for children's math and verbal ability than time spent in leisure. In contrast, there is strong evidence that child work is detrimental for the development of math and verbal skills if it crowds out school time or study time. In interpreting the evidence in the figures, we will abstain from commenting on age or country-specific coefficients, and instead focus on overall patterns. The large number of effects we estimate creates a high probability of Type I error, so we do not want to read too much into significance of individual estimates.

For instance, In Figure 2 Panel A, where leisure is the omitted category, we find that for math scores, the effect of an extra hour spent on domestic chores instead of leisure is significantly negative (5% level) in only one of the 16 cases. Similarly, only 2 out of the 16 estimates of the effect of time spent on market/farm work relative to leisure are

Panel A – Leisure Omitted

Panel B – School Omitted

Panel C – Study Omitted
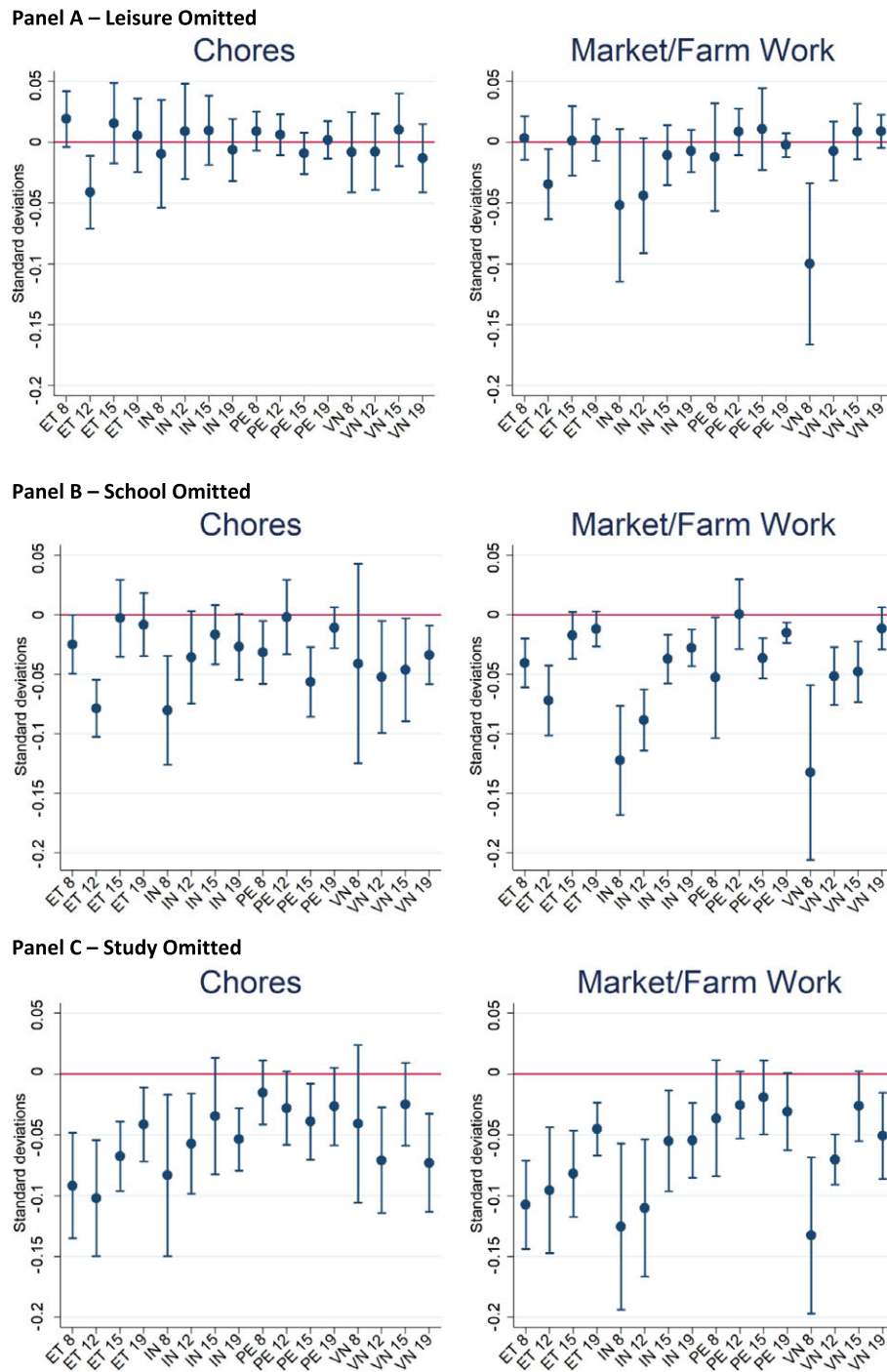
Figure 2. Effect of child work on math scores.

Panel A – Leisure Omitted

Panel B – School Omitted
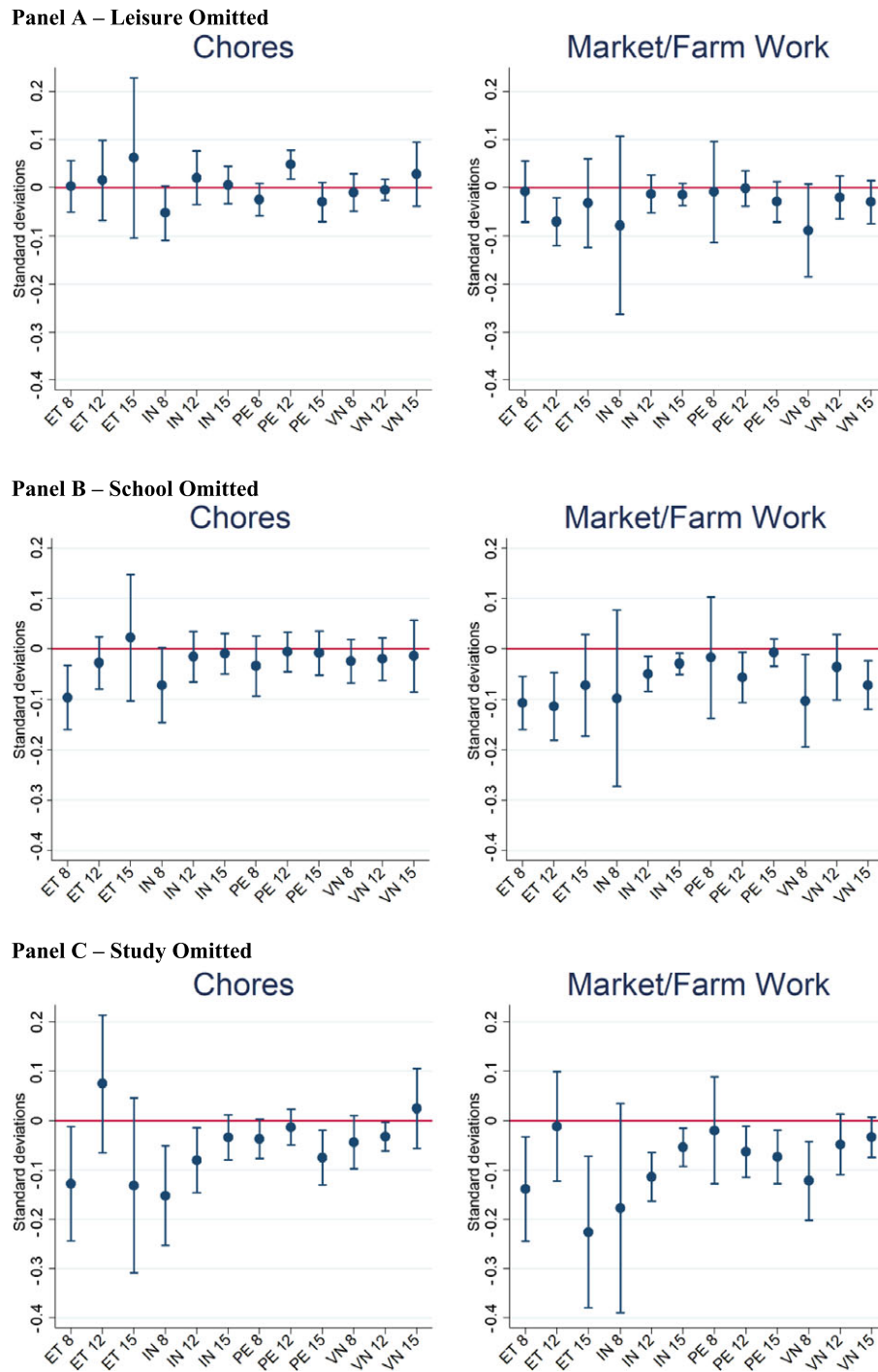

Panel C – Study Omitted

Figure 3. Effect of child work on verbal scores.

significantly negative.[23] Almost all coefficients are small in magnitude, and 15 of the 32 are the "wrong" sign (i.e., the point estimate of the child work effect is actually positive).

Figure 3 Panel A shows similar results for verbal development. The effect of chores relative to leisure is significantly negative (at the 10% level) in only 1 of the 12 cases, and in 7 cases the estimates are the "wrong" sign. Similarly, only 1 of 12 estimates of effects of market/farm work relative to leisure is significantly negative at the 5% level. Thus, we find no evidence for the proposition that child work is detrimental for cognitive development, either math or verbal, *provided* that work only crowds out leisure time.

In contrast, we find strong evidence that child work is detrimental for development of math and verbal skills if it crowds out school time. In Figure 2 Panel B, where school is the omitted category, we find that, for math scores, 12 of the 16 estimates of the effect of an extra hour spent on market/farm work instead of school are significantly negative at the 5% level. Similarly, for verbal skills (see Figure 3(B)), all 12-point estimates of the effect of market/farm work relative to school are negative, and 7 are significant (4 at the 1% level).

We also find clear evidence that time spent on domestic chores is detrimental for child development if it crowds out school. For math scores all point estimates are negative, and 8 of 16 are significant (see Figure 2 Panel B). And the median point estimate for the effect of chores is very close to what we found for market work (both are roughly −0.04 standard deviations per hour). But for verbal scores our results are much weaker. While almost all point estimates are negative, only a couple are significant (see Figure 3, Panel B), and the median point estimate is small.

Finally, in Panel C of Figures 2 and 3 we show results with study time as the omitted category. Here, there is a clear pattern of mostly significant negative effects of both chores and market/farm work on both math and verbal scores. In fact, our estimates of the negative effects of work time relative to study time are more consistently significant than were our estimates for work relative to school. This suggests that study is more productive for child development than school time *per se* in the LMIC context. This is not surprising given the evidence of low school quality in many LMIC's (see Glewwe and Muralidharan (2016)).[24]

To put the effect sizes in Figures 2 and 3 in some context, the median point estimate implies an extra hour of market/farm work relative to time at school reduces math scores by 0.04 standard deviations, and verbal scores by 0.065 standard deviations. Recall from Table 3 that the average child engaged in market/farm work in Ethiopia works about 3.5 hours per day at ages 8, 12, and 15. The median point estimate (−0.04) implies such an increase in work relative to school would lower math scores at a given age by 0.14 standard deviations. This is equivalent to 16% of the average gain in math scores

---

[23]Given we are reporting 32 estimates, it would not be the least bit surprising to find 3 false positives due to type I error under the (true) null hypothesis that the effect of child work is no different from that of leisure.

[24]In Appendix G, Figure G1, we show results where study is an omitted category so we can focus on the effect of school relative to study. In the large majority of cases, the point estimates imply time spent studying is more productive than time spent in school. But the difference is not usually significant. Combining the math and verbal results, we find that in 4 out of 28 cases the difference is significant at the 5% or 1% level.

between age 8 and 12, and 25% of the average gain between age 15 and 19. This effect is magnified by the dynamics of the model.

It is interesting that our estimates of the negative effects of domestic chores on child outcomes are in most cases very similar to those for market/farm work.[25] Some prior work finds market/farm work has a more adverse impact on child development (see Bezerra et al. (2009), Buonomo Zabaleta (2011), Emerson, Ponczek, and Portela Souza (2017), Gunnarsson, Orazem, and Sánchez (2006), Orazem and Gunnarsson (2004)). We explore this contrast in the next subsection.

6.2 *Comparison to "status quo" results*

We have estimated effects of time spent working *relative* to specific alternative activities. In contrast, the typical approach in the child labor literature—due to lack of data—is to simply estimate the effect of work time without controlling for other time uses.[26] We call this the "status quo" approach. Here, we replicate the "status quo" type of analysis by estimating the model in Eq. (9) but including only measures of time spent on domestic chores or on market/farm work, omitting any controls for other uses of time.

The results are reported in Figure 4. These imply that hours spent on market/farm work have an unambiguous negative effect on math and verbal skills. In contrast, there is no evidence that hours spent on domestic chores have a detrimental effect on verbal skills. There is some weak evidence that hours spent on domestic chores are detrimental for math skills (i.e., 4 of 16 estimated effects are significantly negative), but the point estimates are generally much smaller in magnitude than we see for market/farm work.

Thus, application of the "status quo" analysis to our data yields the conclusion, similar to that drawn in several previous studies, that market/farm work is harmful for child development, while time spent on domestic chores is either not harmful, or much less so. In contrast, our results in Figures 2 and 3 imply that *neither* time spent on domestic chores nor market/farm work is harmful for child development if they substitute for leisure time. But time spent on domestic chores and market/farm work are *both* harmful for development—and roughly equally so in most cases—if they substitute for school time or study time.[27]

The key reason the "status quo" approach implies that market/farm work is more harmful than chores is that time spent on market/farm work is associated with greater reductions in school time than time spent on domestic chores. This is shown in Figure 1.

7. Fixed effects results

Here, we investigate whether our results are robust to using a fixed effects (FE) approach to deal with unobserved ability and inputs, instead of value added (VA). The FE model

---

[25]In Appendix G, Figure G2, we make domestic chores the omitted time-use category and focus on the impact of market/farm work relative to chores. We find little evidence that work is worse for development than chores.

[26]Examples of this type of analysis are Akabayashi and Psacharopoulos (1999), Bezerra et al. (2009), Buonomo Zabaleta (2011), Heady (2003), and Ray and Lancaster (2005).

[27]Some "status quo" studies include an indicator for whether a child works, rather than actual work hours. In Appendix H, we show this generates similar results.
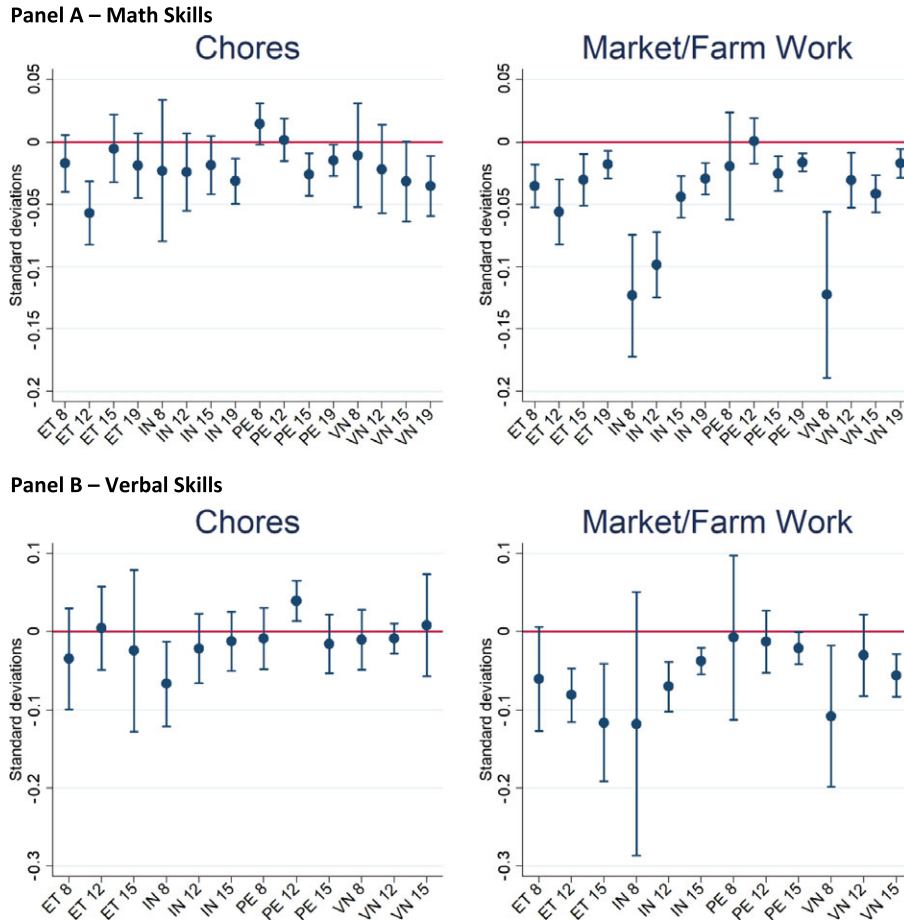
Panel A – Math Skills

Panel B – Verbal Skills

FIGURE 4. "Status quo" analysis of impact of work hours on math and verbal scores. *Note*: The coefficients are from regressions that include only hours of market/farm work or chores, without controls for other time uses. This "status quo" approach contrasts with our main models that control for time allocated to six possible time-use categories that make up a complete 24 hours.

is obtained by generalizing Eq. (5) to a periods, and assuming our control variables $\boldsymbol{b}_{ia}$ capture the influence of unobserved inputs:

$$Y_{ia} - Y_{i,a-1} = \boldsymbol{x}'_{i,a}\boldsymbol{\beta}^a_a - \boldsymbol{x}'_{i,a-1}\boldsymbol{\beta}^{a*}_{a-1} + \boldsymbol{b}'_{ia}\boldsymbol{\delta}_{ia} + \varepsilon^*_{ia}, \tag{10}$$

where $\boldsymbol{\beta}^{a*}_{a-1} \equiv (\boldsymbol{\beta}^{a-1}_{a-1} - \boldsymbol{\beta}^a_{a-1})$ and $\varepsilon^*_{ia} \equiv (m_{ia} - m_{i,a-1})$.[28] While the VA model assumes unobserved ability is captured by the lagged outcome measure, in the FE model the effect of latent ability is differenced out under the assumption that its effect on child outcomes is age invariant. Large differences between the VA and FE estimates would suggest our results are vulnerable to the specific way in which we control for unobserved ability.

Figure 5 reports the FE results for math. The FE model is less efficient than VA, for the usual reason that it relies purely on within subject variation (differencing out all between

---

[28]This is analogous to how we derived our VA estimating equation Eq. (9) from the model in (6).
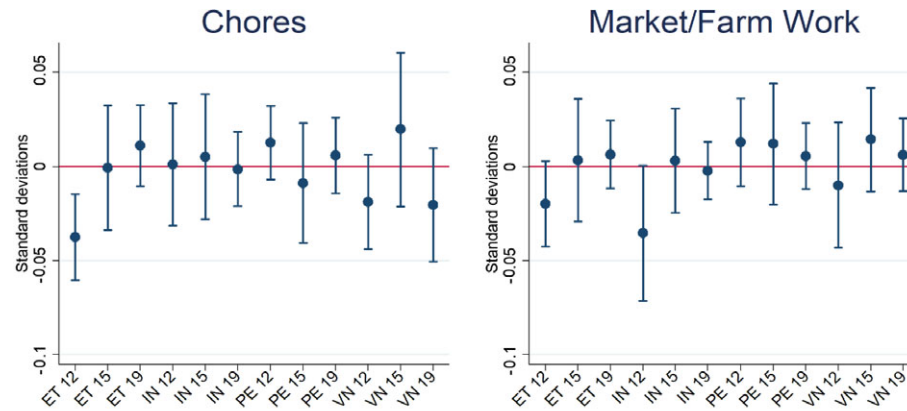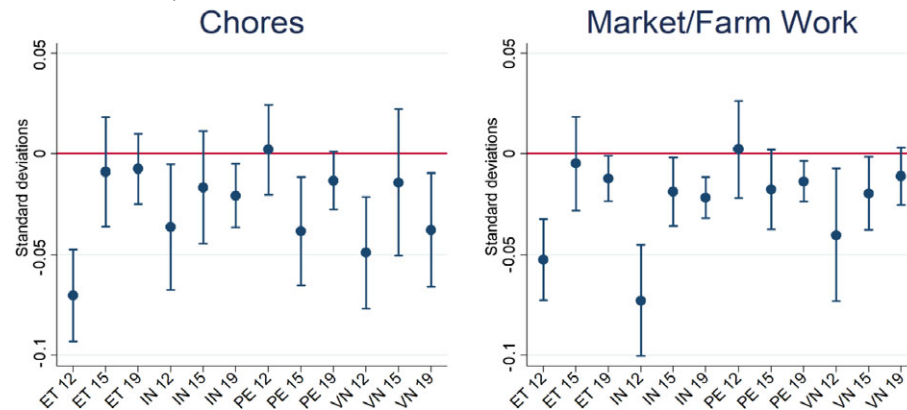
A. Leisure omitted

B. School/Study Time omitted

FIGURE 5. Fixed effects results for math scores.

subject variation). For this reason, we merge the school and study categories into a single "school/study" category. Also, we cannot obtain fixed effects estimates for age 8 because the CDA math test taken at age 5 is not exactly comparable to math tests at later ages.

As before, we find no evidence that chores or market/farm work are detrimental for child math skills if they substitute for leisure time (Figure 5, Panel A). However, both chores and market/farm work do have detrimental effects on math ability if they substitute for school/study time (Figure 5, Panel B). Thus, our FE results for math are qualitatively similar to the VA results. The FE results for verbal scores (available on request) are also very similar.

It is comforting that our main results are not sensitive to the specific way we control for latent ability and unobserved inputs. The similarity of the VA and FE results, despite the fact that these estimators give consistent estimates under somewhat different assumptions, gives us added confidence in the robustness of our main conclusions.

Appendix I presents additional discussion of the role of unobservables in our analysis. We find OLS estimates exaggerate the negative impact of work and chores on math and verbal scores by a factor of 2 to 3.

8. Instrumental variables results

We are aware of only three prior studies that attempt to infer the effect of child work on cognitive test scores using IV. Gunnarsson, Orazem, and Sánchez (2006) instrument for child work using mandatory school entry/exit ages across nine Latin American countries. Bezerra, Kassouf, and Arends-Kuenning (2009) use the unskilled wage in the state of residence in Brazil. And Dumas (2012) uses rainfall (children work less in droughts) and commuting time to school in Senegal. The results are contradictory, as the former two papers find negative effects of child work, while the latter finds positive effects. All three studies find 2SLS estimates of effects of work on cognitive ability that are, surprisingly, several times *greater* in magnitude than OLS estimates. But the 2SLS estimates are also very imprecise, suggesting weak instruments are an issue.

Our challenge in finding appropriate instruments is greater, as we seek to infer the effect of a child's entire vector of time use, so we need to instrument for several endogenous variables.[29] To mitigate the problem, we combine school and study into one category (which we treat as the omitted category). We then need to instrument for four endogenous variables: hours spent on market/farm work, chores, leisure, and sleep. A key challenge is to find instruments that generate independent variation in all four, in particular, in work versus chores.

As a guide to finding good instruments, we examined prior papers that instrument for child work to estimate its effect on schooling: Boozer and Suri (2001) use regional rainfall in Ghana, while Beegle, Dehejia, and Gatti (2009) use the price of rice at the community level in Vietnam. In the Nicaraguan context, Buonomo Zabaleta (2011) uses a large set of instruments including weather shocks, wage rates, distance to school, and access to electricity and water.[30] We also consulted reduced form studies of how shocks affect child labor: Beegle, Dehejia, and Gatti (2006) find that crop loss shocks increase child work and chores, and reduce school attendance in Tanzania. Trinh, Posso, and Feeny (2020) find that below average rainfall reduces child work and chores in Vietnam.

Based on this prior work, we decided to construct instruments based on agricultural prices, local wages and prices, rainfall and weather shocks, and distance to school. More specifically, for each household/year we construct an agricultural price index where specific crop price indices are weighted by the quantity of each crop the household grows. We also construct a livestock output price index based on the price of each animal's output (e.g., price of eggs for chickens), weighted by the quantity of each type of animal the household holds.

We construct community wage indices using local wage rates for men and women in several occupations reported in Young Lives, and taking the first two principle components of those series. Similarly, we construct a consumer price index by taking local

---

[29]Notably, Dumas (2012) attempted to estimate the effect of both work hours and school but was not successful in finding strong instruments for school.

[30]Buonomo Zabaleta (2011) also uses school quality, household assets, household head education, household size, and number of young children. We chose not to use these sort of variables as excluded instruments as we suspect they may be correlated with household resources, or enter the test score production function directly.

prices for a range of common goods reported in Young Lives, and taking the first two principle components of those series.

Next, we linked Young Lives to local rainfall data for the 2 and 12 month preinterview periods by using satellite data (see Appendix J). We take deviations from 20-year means to construct rainfall shocks. These are interacted with interview month to allow different effects of rainfall at different times of year. We also use frost shock indicators from the Young Lives data.

Additional instruments are distance to school and an index of local services and infrastructure. This includes such things as community access to electricity, drinkable water, and children's playgrounds. We assumed this would be be useful in predicting time devoted to chores and leisure.

Our instruments can be classified into those that vary at the household level; namely, the agricultural and livestock price indices, school distance and type, rainfall in the months prior to the interview date, frost shocks, interview month, and whether the household is agricultural—and those that are at the community level—local wage and consumer price indices, and the community service/infrastructure index.

We interacted the community level instruments and rainfall with several child and household chacteristics to let them differentially affect different types of children/households. These are child gender, number of older siblings, number of brothers/sisters, time to school, household wealth, urban/rural, the agricultural and livestock price indices, frost shocks, and month of interview. Including interactions, we have 120 to 130 instruments (see Appendix J for a complete list). The number varies slightly by country/age, as interviews were spread over fewer months in some cases.

First, we assess whether our instruments are sufficiently strong to make 2SLS feasible, using first-stage F-statistics for weak instruments. We have four endogenous variables, so we use the Sanderson and Windmeijer (2016) F-statistic designed to assess *independent* variation in each endogenous variable generated by the instruments. Given the large instrument set, the many instrument problem may be important, so we consider both 2SLS on the full instrument set, and results using LASSO to select a subset of strongest instruments (see details below).

The results show our instruments are reasonably strong for Ethiopia at ages 8 and 12. Other cases are borderline at best, so we do not consider them further. Presumably, the reason we succeed for Ethiopia is that almost all child work there is agricultural, and our instruments are specifically designed to predict work in an agricultural context. Also, child work is much more common in Ethiopia than the other countries (see Section 4.2), so we are not trying to predict a rare event.

We report the IV results for age 8 and 12 math scores in Ethiopia in the top panel of Table 4. The first three columns report results from VA and OLS (with and without controls) for comparision purposes. The next three columns report 2SLS using the full instrument set, followed by 2SLS and optimal (2-step) GMM using the subset of instruments chosen by LASSO. The last three columns also report IV results, except here we *both* control for the lagged test score and instrument for time use, thus combining the IV and VA approaches.

TABLE 4. Instrumental variable results for math scores.

	Linear Regression			IV (Lagged Test Excluded)			IV (Lagged Test Included)		
	OLS (No Controls)	OLS	VA	2SLS	2SLS (Lasso)	GMM (Lasso)	2SLS	2SLS (Lasso)	GMM (Lasso)
Ethiopia Age 8									
Market/Farm Work									
Estimate (β)	−0.135	−0.058	−0.055	−0.082	−0.086	−0.077	−0.083	−0.101	−0.091
Std. Err.	0.006	0.007	0.008	0.011	0.015	0.014	0.011	0.015	0.015
F-Statistic				6.855		12.035	6.907		13.453
Chores									
Estimate (β)	−0.125	−0.053	−0.038	−0.085	−0.110	−0.103	−0.084	−0.120	−0.115
Std. Err.	0.007	0.007	0.011	0.015	0.020	0.020	0.015	0.022	0.022
F-Statistic				4.606		9.661	4.365		10.310
Leisure									
Estimate (β)	−0.089	−0.056	−0.057	−0.049	−0.045	−0.036	−0.044	−0.037	−0.035
Std. Err.	0.006	0.005	0.007	0.010	0.012	0.012	0.010	0.012	0.012
F-Statistic				4.086		7.965	3.956		10.946
Instruments (#)				122		24	122		20
Ethiopia Age 12									
Market/Farm Work									
Estimate (β)	−0.196	−0.115	−0.078	−0.165	−0.169	−0.149	−0.101	−0.099	−0.076
Std. Err.	0.009	0.011	0.015	0.023	0.036	0.035	0.022	0.033	0.032
F-Statistic				7.331		18.209	6.004		16.048
Chores									
Estimate (β)	−0.226	−0.121	−0.084	−0.214	−0.214	−0.166	−0.141	−0.121	−0.074
Std. Err.	0.012	0.012	0.012	0.028	0.037	0.035	0.026	0.034	0.032
F-Statistic				5.230		19.378	4.651		16.571
Leisure									
Estimate (β)	−0.112	−0.080	−0.044	−0.079	−0.036	−0.022	−0.039	0.013	0.023
Std. Err.	0.011	0.010	0.011	0.021	0.035	0.034	0.019	0.029	0.028
F-Statistic				4.022		9.884	4.463		10.234
Instruments (#)				131		19	131		20

Note: Omitted time category is school and study in all regressions. The first stage F-statistic is calculated according to Sanderson and Windmeijer (2016). The optimal regularization parameter in Lasso was picked according to the minimal BIC.

For Ethiopia at age 8, if we implement 2SLS using the full set of 122 instruments, the coefficients on market/farm work and chores are -0.083 and -0.088, respectively. This implies an extra hour of work or chores that detracts from school/study time reduces math scores by roughly 0.085 standard deviations. The standard errors are 0.011 and 0.15, respectively, so the point estimates are precise despite instrumenting. But the SW first-stage F-statistics for work and chores are 6.86 and 4.61, respectively. These values may raise some concern that the large instrument set could cause a many/weak instrument problem.[31]

To address this concern, we reduce the instrument set by using LASSO regression to select the best instruments. We use the Bayes information criterion (BIC) to assess the fit of alternative first-stage models, adjusting for number of parameters.[32] To ensure LASSO does not choose instruments that merely give the best fit to each endogenous variable *individually*, but rather to maximize their *independent* variation, we use a multistep approach: First, use LASSO to chose a subset of instruments that best explain the child work variable. Second, use LASSO to choose a subset of instruments that best explain the *residual* variation in the chores variable (not explained by instruments selected in step 1). Repeat for leisure and sleep.

Adopting this procedure, we chose a subset of 24 instruments, listed in Appendix J Table J3.[33] While it is hard to completely describe the first stage results, the most prominent patterns are that higher agricultural prices shift time from chores to market/farm work (especially for boys), distance to school increases market/farm work (but not chores), rainfall reduces market/farm work (as expected), and higher prices for staples increase both chores and market/farm work (with the effect on work being greater for boys and children with older siblings). Using these 24 variables as the excluded instruments in 2SLS, the SW F-statistics for work and chores improve markedly to 12.0 and 9.7, respectively.[34] The point estimates for work and chores become slightly more negative, and standard errors increase by only 30%, so the estimates remain precise. In contrast to the IV papers in the prior literature (see above), our instruments give much more precise estimates of child labor effects.

The next column of Table 4 reports optimal (2-step) GMM-LASSO estimates which are similar to the 2SLS-LASSO estimates. The last three columns report the same models except we add the lagged test score as an additional control, thus combining the IV and VA approaches to dealing with latent child ability. This has little impact on the IV results.

The GMM estimates for Ethiopia at age 8, that control for lagged test score and use the 20 instruments chosen by LASSO, give coefficients on market/farm work and chores

---

[31]Stock–Yogo critical values for weak instrument F-statistics in the case of multiple endogenous variables and instrument sets larger than 100 are not available. A common rule-of-thumb suggests an F-stat of 10 or more.

[32]Of course, we force LASSO to to retain all 2nd stage control variables as instruments in the first stage. LASSO then choses *additional* variables from the set of *excluded* instruments. See Appendix J for details on LASSO.

[33]The table lists the instruments that generate independent variation in each endogenous variable.

[34]With four endogenous variables and 24 instruments, the Stock–Yogo critical value for rejecting the null that IV bias relative to OLS is greater than 10% is 11.4. The Stock–Yogo critical values for rejecting the null of IV bias greater than 30%, 20%, 10%, or 5% of the OLS bias are 4.4, 6.2, 11.4, and 21.4, respectively.

of −0.091 and −0.115, and these are precisely estimated (the standard errors are 0.015 and 0.022, respectively). In one way, these results are consistent with our value added (VA) results, in that effects of work and chores are not significantly different. Substituting an extra hour of either for school/study reduces math test scores by roughly a tenth of standard deviation.

However, the IV estimates are roughly twice as large as the VA estimates. There are (at least) two plausible explanations: The first is measurement error in the time use variables, which would bias estimated effects of work/chores toward zero. But we are skeptical of this story, because if we use child reports of time use to instrument for parent reports it has little impact on our VA results.[35] A second possibility is that there is heterogeneity in the effects of child work, and that the children at the margin—who are shifted in and out of work/chores by the variation in the instruments—are relatively more adversely affected.

Note that while the IV estimates exceed VA, they are smaller than what we obtain in a simple OLS regression of test scores on the time use vector (with no controls). Thus, we find it is relatively lower skill children who are selected into work (biasing OLS downward).

The bottom panel of Table 4 reports IV results for math scores at age 12 in Ethiopia. Our instruments are stronger here, as the SW F-statistics for those selected by LASSO exceed 16 for both chores and for work. The most prominent patterns in the first stage for age 12 are that higher livestock prices shift time from chores to market/farm work (especially for boys), better local services reduce chores, and increase market/farm work for boys, higher staples' prices reduce chores in urban areas, and higher local wages reduce chores (more so for boys and in urban areas) while increasing market/farm work for boys and in urban areas.

Using these instruments, the GMM estimates that control for the lagged test score give coefficients on market/farm work and chores of −0.076 and −0.074, respectively. These estimates are very similar to the value added (VA) results (see column 3 of Table 4), both in magnitude and in the sense that effects of work and chores are not significantly different. The estimated effects at age 12 are slightly smaller than what we found at age 8. They imply that substituting an extra hour of either work or chores for school/study reduces math test scores by roughly 0.075 standard deviations.

Next, we consider the results for verbal scores. As instrumenting leads to a loss in precision, we decided to increase sample size by pooling language groups as discussed in Appendix J. The top panel of Table 5 reports IV results for verbal scores in Ethiopia at age 8. The GMM estimates in the last column—that control for lagged test score and use the 20 instruments chosen by LASSO—give coefficients on market/farm work and chores of −0.243 and −0.178 at age 8, with standard errors of 0.032 and 0.061, respectively. We cannot reject that these effects are equal, so it appears that substituting an extra hour of either work or chores for school/study reduces math test scores by roughly two-tenths of a standard deviation. This is twice as large as the effect we found for math scores. Again, we find that at age 8 the IV estimates are roughly twice as large as the VA

---

[35]Child reports of their own time use are available at ages 12 and 15 only.

TABLE 5. Instrumental variable results for verbal scores.

	Linear Regression			IV (Lagged Test Excluded)			IV (Lagged Test Included)		
	OLS (No Controls)	OLS	VA	2SLS	2SLS (Lasso)	GMM (Lasso)	2SLS	2SLS (Lasso)	GMM (Lasso)
Ethiopia Age 8									
Market/Farm Work									
Estimate (β)	−0.271	−0.130	−0.120	−0.182	−0.216	−0.235	−0.170	−0.222	−0.243
Std. Err.	0.012	0.014	0.011	0.024	0.032	0.031	0.024	0.033	0.032
F-Statistic				11.463		15.034	11.966		16.468
Chores									
Estimate (β)	−0.253	−0.081	−0.081	−0.169	−0.182	−0.193	−0.121	−0.154	−0.178
Std. Err.	0.014	0.016	0.021	0.037	0.058	0.057	0.036	0.063	0.061
F-Statistic				3.957		6.651	3.986		6.732
Leisure									
Estimate (β)	−0.189	−0.106	−0.090	−0.162	−0.206	−0.207	−0.152	−0.207	−0.208
Std. Err.	0.012	0.011	0.014	0.022	0.031	0.030	0.024	0.037	0.037
F-Statistic				4.094		6.572	4.027		6.620
Instruments (#)				122		24	122		20
Ethiopia Age 12									
Market/Farm Work									
Estimate (β)	−0.297	−0.150	−0.110	−0.189	−0.165	−0.168	−0.098	−0.071	−0.048
Std. Err.	0.014	0.015	0.021	0.034	0.059	0.059	0.032	0.054	0.054
F-Statistic				8.221		15.525	8.421		13.905
Chores									
Estimate (β)	−0.212	−0.051	−0.020	−0.132	−0.166	−0.163	−0.017	−0.159	−0.137
Std. Err.	0.018	0.018	0.022	0.042	0.074	0.074	0.040	0.073	0.072
F-Statistic				5.151		12.148	5.119		8.726
Leisure									
Estimate (β)	−0.122	−0.041	−0.007	−0.100	−0.214	−0.194	0.000	−0.040	−0.029
Std. Err.	0.017	0.015	0.025	0.032	0.055	0.055	0.029	0.054	0.054
F-Statistic				4.701		12.562	6.704		8.633
Instruments (#)				131		18	131		20

Note: Omitted time category is school and study in all regressions. The first stage F-statistic is calculated according to Sanderson and Windmeijer (2016). The optimal regularization parameter in Lasso was picked according to the minimal BIC.

estimates. The age 12 results are fairly similar, but effect sizes are smaller, less precisely estimated, and less stable across specifications.

Finally, Tables 4 and 5 also report the coefficients on leisure. Our VA results imply that child work only has a negative impact if it substitutes for school/study (not leisure). Our IV results for verbal scores are consistent with this hypothesis. But our IV results for math imply that child work has adverse effects relative to leisure in Ethiopia at both ages 8 and 12.

Overall, the IV results are similar to VA in that work and chores have similar negative effects relative to school/study. But they differ from VA in that they imply larger negative effects of both work and chores, both relative to school/study and, for math, relative to leisure. This may suggest that work/chores are particularly detrimental for "marginal" children pushed in and out of child labor by the variation of the instruments. Nevertheless, the IV estimates of adverse child labor effects remain smaller than OLS with no controls, implying negative selection into work based on ability.

9. Heterogeneity in effects of child work

A weakness of the existing literature on child work is that it looks at specific contexts, driven by limited data availability. But the nature of child work differs across contexts, and it may have very different effects at different ages (Attanasio, Meghir, Nix, and Salvati (2017), Del Boca, Monfardini, and Nicoletti (2017)). Young Lives surveys children from four LMIC countries with diverse economic contexts, at ages from childhood to late adolescence. This gives us an opportunity to study how effects of child work vary by context. We consider five dimensions: country, age, level of hours, skill, and gender. As we will see, we find evidence of heterogenous effects in all five dimensions.

In Section 6, we ran separate regressions by age/country to avoid the risk of bias due to pooling inappropriately. In this section, we pool the data across either the country or age dimension (or both) to gain efficiency, enabling us to draw more reliable inferences about heterogeneous effects. Pooling also lets us study age/country specific effects in a way that is less susceptible to Type I error, as the number of estimated parameters is reduced.

9.1 *Pooled VA results*

We start by checking if our main results are robust to pooling. Recall that only our math measure is comparable across age-groups *and* countries, while verbal scores are only comparable across age but not language. Thus, our analysis in this section is mostly restricted to math skills. Table 6 column 1 presents estimates of Eq. (9) obtained by pooling across both age and country dimensions. This model includes age and country fixed effects, but all slope coefficients are assumed homogeneous. Study time is the omitted time-use category.

The pooled results are very similar to the disaggregated results from Section 6. The coefficients on chores and market/farm work are both -0.058, with standard errors of

TABLE 6. Pooled models for math scores (study omitted): geographic variation.

	Pooled (1)	Ethiopia (2)	India (3)	Peru (4)	Vietnam (5)
School (hrs/day)	−0.018	−0.064	−0.015	−0.004	−0.003
	(0.007)	(0.014)	(0.009)	(0.009)	(0.011)
Chores (hrs/day)	−0.058	−0.084	−0.055	−0.028	−0.057
	(0.006)	(0.008)	(0.011)	(0.008)	(0.011)
Market/farm work (hrs/day)	−0.058	−0.088	−0.059	−0.034	−0.043
	(0.006)	(0.009)	(0.009)	(0.008)	(0.009)
Leisure (hrs/day)	−0.055	−0.085	−0.054	−0.034	−0.048
	(0.005)	(0.009)	(0.010)	(0.006)	(0.008)
Adjusted R^2	0.703	0.627	0.621	0.702	0.619
Sample size	20,330	4814	5380	4902	5234
F-test for equality of chores and leisure coefficients:					
F-statistic	0.61	0.03	0.00	2.07	0.91
p-value	0.44	0.86	0.96	0.16	0.35
F-test for equality of market/farm work and leisure coefficients:					
F-statistic	0.78	0.14	0.59	0.00	0.56
p-value	0.38	0.71	0.45	0.97	0.46

Note: Study time is the omitted time-use category. The pooled regression includes country and age fixed effects. The country-specific regressions include age fixed effects. Standard errors are in parentheses.

0.006. Thus, our estimates imply an extra hour of *either* market/farm work or chores that crowds out study time reduces math scores by 0.06 standard deviations.[36]

Furthermore, the coefficient on leisure time is nearly identical to those on chores and market/farm work. Thus, our pooled results strongly support our conclusion in Section 6 that market/farm work and chores are equally detrimental for child development if they substitute for study time, while neither is detrimental if it only substitutes for leisure.[37]

The coefficient on school time in the pooled model is small (−0.018) but significant (standard error 0.007). This supports our result in Figures 2 and 3 that school is slightly less productive for child development than study.

Reassuringly then, our pooled results for math are very similar to our disaggregate results. It appears that any misspecification due to pooling is not serious enough to alter our main findings.[38] Appendix K reports estimates of the household resource variables in the pooled model. As expected, household wealth and parents' education have a strong positive effect on test scores, as does the height-for-age z-score, which proxies for nutrition inputs. Having more siblings reduces test scores, as it means less resources for the target child.

---

[36]The point estimates in the pooled model are very similar to the median (across ages and countries) of the point estimates from the disaggregated models in Figure 2, Panel C, which are −0.055 for work and −0.048 for chores.

[37]Our results also imply that study time is more productive for child cognitive ability than leisure.

[38]Pooled fixed effects results (available on request) are also very similar.

9.2 *Differences by country*

A key aim of the Young Lives study was to investigate child development in *different* LMIC settings. This was the rationale behind selecting four countries from four major regions of the developing world, with diverse socioeconomic characteristics. Child work in such different contexts is likely to mean different things. This motivates an investigation of how the effects of child work differ between the study countries.

Our quantitative data do not contain detailed information on the nature of work, but qualitative data that was also collected as part of Young Lives reveals differences across countries/contexts.[39] For example, in Ethiopia, India, and Vietnam most child work outside the household is linked to agricultural tasks, while in Peru it is informal trade (Morrow and Boyden (2018)). Recall that the Peru sample is much more urban.

Table 6, columns 2–5, reports results by country. Consistent with our general strategy in this section, we pool across ages to increase efficiency and reduce the number of country-specific parameters that we estimate down to a manageable number. Given that we pool across ages, we do include age fixed effects.

Our results imply that child work is most detrimental (relative to study) in Ethiopia. It is least detrimental in Peru, where the coefficients on both market/farm work and chores are only about 1/3 as large. India and Vietnam lie in between.[40] Ethiopia is the poorest country in the sample, and the children work quite a bit more than in the other three countries, especially at ages 8 and 12 (see Tables 2 and 3). Households are more reliant on child work in Ethiopia than in the other countries, which may translate into children having to do more physically demanding tasks. Child work may also be more detrimental at the margin simply because the typical child works much higher hours in Ethiopia.

However, as we saw in Table 2, children in Peru tend to work more than in Vietnam or India, especially at ages 8 and 12. So there is no simple relationship whereby the negative impact of child work is greater at higher levels of hours. The relatively benign impact of work in Peru is more likely explained by the different nature of work.

In all four countries, we cannot reject the hypothesis that effects of market/farm work, chores, and leisure (relative to study) are all equal. Thus, our country level results support our main finding that work activities are only detrimental for child development if they detract from school/study time (not just leisure time).

We also find that school time in Ethiopia is less productive than study. But there is no evidence that school is less productive than study in other countries.[41]

---

[39]Young Lives conducted in-depth qualitative research on a sub-sample of 200 children, selected to reflect main ethnic/caste groups, rural and urban sites, and poor and less-poor communities (see www.younglives.org.uk).

[40]Formally testing for differences across countries, we find effects in Ethiopia are significantly larger than in the other three countries, but the differences between India, Peru, and Vietnam are not statistically significant.

[41]Singh (2015, 2020) presents evidence that Ethiopia and India have lower school productivity than Vietnam and Peru. But we only find evidence of this for Ethiopia and not India.

TABLE 7. Pooled models for math scores (study omitted): child characteristics.

	Age 8 (1)	Age 12 (2)	Age 15 (3)	Age 19 (4)	Female (5)	Male (6)
School (hrs/day)	−0.021	−0.014	−0.018	−0.024	0.014	−0.004
	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)
Chores (hrs/day)	−0.051	−0.063	−0.056	−0.043	−0.049	−0.040
	(0.011)	(0.009)	(0.008)	(0.007)	(0.008)	(0.009)
Market/farm work (hrs/day)	−0.073	−0.075	−0.050	−0.042	−0.039	−0.071
	(0.012)	(0.010)	(0.008)	(0.007)	(0.008)	(0.008)
Leisure (hrs/day)	−0.062	−0.054	−0.055	−0.040	−0.048	−0.057
	(0.010)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)
Adjusted R^2	0.620	0.646	0.668	0.657	0.661	0.657
Sample size	6459	7324	3426	3121	9833	10,497
F-test for equality of chores and leisure coefficients:						
F-statistic	3.02	1.82	0.03	0.19	0.05	8.08
p-value	0.09	0.18	0.87	0.67	0.82	0.01
F-test for equality of market/farm work and leisure coefficients:						
F-statistic	1.72	8.23	0.60	0.16	3.02	10.00
p-value	0.19	0.01	0.44	0.69	0.09	0.00

Note: Study time is the omitted time-use category. The age-specific regressions include country fixed effects. The gender-specific regressions include country and age fixed effects. Standard errors are in parentheses.

9.3 *Differences by child age*

Table 7 columns 1–4 presents results by age. Here, we pool across the four countries, and include country fixed effects. Looking at the age pattern of the coefficients, there is no clear pattern of the effect of chores differing by age. However, the point estimates imply that market/farm work has more detrimental effects on child development at ages 8 and 12 than at ages 15 and 18. The estimated effects of market/farm work (relative to study) are about 50% greater at the younger ages, and the differences are statistically significant.

There are two distinct explanations for why child work may have a more negative effect on child development at younger ages. First, as Cunha and Heckman (2008) argue, investments in child cognitive development may be more productive at younger ages. Then work that crowds out school/study time by any given number of hours will have a larger negative impact if it occurs at younger ages.

Second, a given amount of work time may be more tiring for younger children. In fact, reports in the qualitative data suggest that some of the work that children are typically expected to do can be especially physically demanding and tiring for young children. As an example, in Ethiopia and India young children's responsibilities can include operation of irrigation pumps on the family farm.[42] This has to be done at night when electricity supply is reliable and requires speed and strength to control and direct fast-moving water. Children report being tired at school as a result (Morrow, Tafere, and Vennam (2014)). The qualitative data also reveal substantial age differences in the type of work that children do.

---

[42]Recall that we categorize work on the family farm as part of "market/farm work" rather than chores.
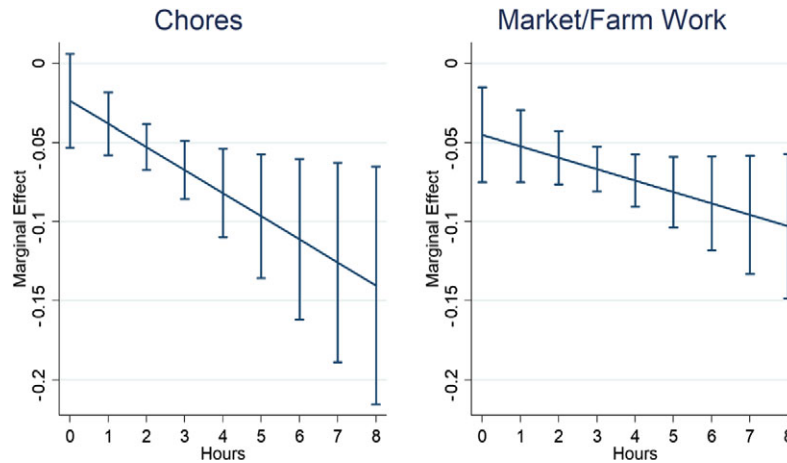
FIGURE 6. Nonlinear effects of time on chores and market/farm work at age 12. *Note*: The figure shows marginal effects of an additional hour of chores or market/farm work on math scores at different levels of hours. We present the results for age 12 only, as we did not find evidence of nonlinearity at later ages.

9.4 *Differences by level of hours*

It seems intuitive that the detrimental effect of child work may increase nonlinearly with hours. The qualitative data provide evidence that, beyond a certain threshold, work starts to impinge on the children's ability to be productive while studying and going to school. In particular, children often report being too tired to learn and pay attention at school because of long work hours, and gradually falling behind in their studies as a result (Zharkevich, Roest, and Thi Thanh Huong (2016)). In this section, we test for nonlinearities in the effects of child work.

Specifically, we estimate models that include quadratics in time spent on chores and market/farm work. At age 8, our estimates are very imprecise, as few children work high levels of hours, making nonlinearities very difficult to identify. At ages 15 and 19, we do obtain precise estimates, but we find no clear evidence that the detrimental effect of work increases *more than linearly* with the level of hours.

However, we do find evidence of nonlinearity at age 12, consistent with the evidence above that child work is more detrimental for younger children. Figure 6 reports age 12 results for math scores, plotting the marginal effects of hours of work against the level of hours. We see that negative effects of chores and market/farm work both increase with the level of hours spent in these activities. For example, at zero hours of chores, the marginal effect of the first hour of chores (substituting for study) is only -0.025. But at 3 hours the marginal effect of an additional hour is three times as large.[43]

In Appendix L, we adopt a different approach to testing for nonlinearities. We pool the data across countries and ages as in Table 6 column (1), and also combine

---

[43]Estimation of nonlinearities in the effect of work *at a point in time* does not require tests scores that are consistent across ages/countries, so we also obtained results for verbal scores at age 12 (not reported). The point estimates turn more negative as hours increase, but the changes are not significant. Thus, we find no clear evidence that the adverse effects of either work or chores increase with the level of hours. It

school/study into a single category, in order to gain efficiency. We then consider models that add quadratics in school/study time, hours of work or chores, or all three. We find these quadratic terms are not significant, and the estimated effects of school and work time are little affected by their inclusion. So our main results are not sensitive to allowing for nonlinearities in hours.

9.5 *Differences by skill level*

Here, we examine whether effects of time use differ by skill level. To this end, we combine school/study into one category, and interact time use with the lagged math ability test score. Results are reported in Appendix M. Column (1), which combines school/study without including interactions, confirms our result that school/study is beneficial relative to all alternative times uses, while work, chores, and leisure are equivalent. Column (2) interacts school/study time with the lagged math ability test. Interestingly, the interaction is significant and *negative*. This implies additional school time would be more beneficial for relatively low ability children. In column (3), we interact all time use variables with the lagged math score. The results imply that child work is less detrimental (relative to school) for higher ability children. These results are consistent with a model where school and ability are substitutes in the production of cognitive ability, while ability complements learning from work and chores.

9.6 *Differences by gender*

In Table 7, columns 5 and 6, we present separate results by gender. Here, we pool across ages and countries, and include age and country fixed effects. The results clearly show that market/farm work is relatively more detrimental for boys. The effect size is 80% larger for boys than girls (-0.071 vs. -0.039) and the difference is highly statistically significant.[44] We also find that, for boys, time spent on market/farm work has a significantly more negative effect than time spent in leisure.

The finding that market/farm work is more detrimental for boys is notable, as boys are more likely to engage in market/farm work outside the home than girls. In our data, boys spend 1.47 hours per day (on average) on market/farm work compared to 0.84 for girls. But girls spend 2.20 hours per day on chores, compared to only 1.28 hours for boys.

Consistent with the time-use data, the qualitative data suggest that in most study areas girls have more limited opportunities to engage in market/farm work than boys, due to concerns about their security, and a desire to prepare them for their adult roles (i.e., domestic chores). These restrictions mean boys tend to undertake the more physically demanding and risky market/farm work across the four countries. For example, Morrow and Vennam (2012) note that, in paddy operation in India, more physically demanding tasks like ploughing, sowing, weeding, watering/irrigating, and thrashing are more likely

---

is possible work activities stimulate language ability (Patrinos and Psacharopoulos (1997)), mitigating its negative effects. But we leave further investigation of the mechanism behind this result for future research.

[44]If we look only at children aged 8 to 12, we find the negative effects of market/farm work increase to -0.093 for boys (s.e. $= 0.010$) and to -0.058 for girls (s.e. $= 0.011$). Results available on request.

to be done by boys. Hence, an extra hour of market/farm work is likely to be more tiring for boys than girls.

However, a key implication of our results is that, as girls spend almost twice as much time on chores as boys, an exclusive focus on market/farm work, based on the incorrect "status quo" result that chores are not detrimental for development, is *de facto* discriminatory against girls. Indeed, the point estimates in Table 7 imply chores are more detrimental for girls than market/farm work, and that chores are more detrimental for girls than boys. Thus, an exclusive focus on reducing market/farm work ignores the potential benefit of policies to shift girls' time from chores to school/study.

10. Conclusion

Prior work on the effect of child labor on child development has been hampered by serious data limitations, leading to ambiguous conclusions. Our analysis based on Young Lives, a new multicountry longitudinal cohort study, overcomes many of these problems. For the first time in this context, Young Lives collects skill measures regardless of whether children are in school, avoiding an important selection bias in prior studies. It provides state-of-the-art cognitive test scores that are longitudinally consistent, allowing use of panel data methods to control for latent ability. And it collects data on children's complete time budget, allowing us to estimate effects of child work relative to *specific* alternative time uses.

Using Young Lives, we have estimated child cognitive ability productions using both value added and fixed effects methods. Across four countries (India, Vietnam, Peru, and Ethiopia) and a wide age range (8, 12, 15, and 19), we find that time spent on domestic chores or market/farm work are both detrimental for cognitive development if they crowd out school and/or study time. But there is little evidence that time spent on work or chores is detrimental for cognitive development if it only crowds out leisure.

Our approach yields different findings from the typical approach in prior work of estimating effects of child work *without* controls for time in other activities. Had we done that, we would have found, as in several prior studies, that time spent on market/farm work is harmful for child development, while time spent on chores is not. This would be consistent with the current policy focus on reducing children's time in market/farm work *per se*. Instead, we conclude that, to encourage child development, policy should focus on reducing *both* chore and work time, and to shifting that time to school and study.

A weakness of the existing literature on child labor is that effects are studied for specific contexts, driven by limited data availability. The Young Lives data let us study how effects of child work differ across contexts defined by country, age, level of hours, ability, and gender. Our results imply child work is most detrimental in Ethiopia and least detrimental in Peru. This may be due to the different nature of work, with children in Ethiopia mostly engaged in farm work, while those in Peru are mostly engaged in informal trade.

We also find that market/farm work is more detrimental for younger children (ages 8 and 12) than older children (ages 15 and 18). In contrast, there is no pattern in the effect of chores by age. So, a policy focus on reducing work for young children may be appropriate.

We also find that work is relatively less detrimental for higher ability children. This implies a complementarity whereby higher ability children are able to learn from work.

And we find that time spent on market/farm work (rather than school/study) is relatively more detrimental for boys, while the reverse is true for girls. Thus, an exclusive focus on reducing market/farm work (based on the misconception that chores are benign) ignores the potential benefit of policies to shift girls' time from chores to school/study.

We also obtain instrumental variable results using agricultural prices, wages, and weather to instrument for child time-use. This is very challenging as we need to instrument for four endogenous time use variables.[45] We succeed for Ethiopia at ages 8 and 12, where our instruments are strong because child labor is very prevalent and mostly agricultural.

The IV results are consistent with value added and fixed effects in that we find market/farm work and chores have similar negative effects relative to school/study. But the IV results differ in that they imply *larger* negative effects of work and chores, both relative to school/study and, for math, relative to leisure. This may be because work/chores are particularly detrimental for "marginal" children pushed in and out of child labor by the variation of the instruments. Nevertheless, the IV estimates of adverse child labor effects remain smaller than what we obtain by running OLS on the time-use vector with no controls, implying negative selection into work based on ability.

In sum, the answer to the question "Does child work harm cognitive development?" is that it depends on whether work crowds out school/study. Our results imply policies aimed at reducing child labor will do little to enhance human capital formation unless they promote school attendance. One possible way to achieve both objectives is conditional cash transfers; see Skoufias and Parker (2001), Fiszbein and Schady (2009), Todd and Wolpin (2006).[46]

Our results on the importance of school/study time complement a growing literature that uses time use data and panel data methods to assess the impact of children's time use on child development in developed country contexts (see Cano, Perales, and Baxter (2019), Del Boca, Monfardini, and Nicoletti (2017), Del Bono, Francesconi, Kelly, and Sacker (2016), Fiorini and Keane (2014), Funk and Kemper (2016), Hsin and Felfe (2014), Hull (2017)).[47] This literature generally finds that education-intensive uses of child time are most productive for the development of cognitive skills. Our findings extend this result to a range of low to middle income country contexts.

---

[45]We are aware of only three prior studies that attempt to infer the effect of child work on cognitive test scores using IV; namely, Gunnarsson, Orazem, and Sánchez (2006), Bezerra, Kassouf, and Arends-Kuenning (2009), and Dumas (2012). But these studies do not attempt to instrument for the complete vector of child time use, and they all suffer from weak instruments.

[46]According to Todd and Wolpin's model, the Progresa program in Mexico, which provided families with cash transfers conditional on children staying in school, was effective at shifting children from work to school. In contrast, their model implies that a child labor ban would reduce work but have almost no effect on schooling.

[47]For structural work using time diary data, see Del Boca, Flinn, and Wiswall (2014) and Verriest (2020).

References

Akabayashi, Hideo and George Psacharopoulos (1999), "The trade-off between child labour and human capital formation: A Tanzanian case study." *The Journal of Development Studies*, 35 (5), 120–140. [428, 444]

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters (2013), "Explaining charter school effectiveness." *American Economic Journal: Applied Economics*, 5 (4), 1–27. [426]

Assad, Ragui, Deborah Levison, and Hai-Ang Dang (2010), "How much work is too much? Effect of child work on schooling: The case of Egypt." In *Child Labor and the Transition Between School and Work* (Randall Akee, Eric Edmonds, and Kostantinos Tatsiramos, eds.), 53–97, Emerald, Bingley. [428]

Attanasio, Orazio, Costas Meghir, Emily Nix, and Francesca Salvati (2017), "Human capital growth and poverty: Evidence from Ethiopia and Peru." *Review of Economic Dynamics*, 25, 234–259. [453]

Basu, Kaushik (1999), "Child labor: Cause, consequence, and cure, with remarks on international labor standards." *Journal of Economic Literature*, 37 (3), 1083–1119. [436]

Becker, Gary (1960), *An Economic Analysis of Fertility. Demographic and Economic Change in Developed Countries*. Princeton University Press, Princeton. [429]

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti (2006), "Child labor and agricultural shocks." *Journal of Development Economics*, 81, 80–96. [447]

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti (2009), "Why should we care about child labor?: The education, labor market, and health consequences of child labor." *Journal of Human Resources*, 44 (4), 871–889. [428, 447]

Bezerra, Marcio E. G., Ana L. Kassouf, and Mary Arends-Kuenning (2009), "The impact of child labor and school quality on academic achievement in Brazil." IZA Discussion Paper 4062. Available at http://ftp.iza.org/dp4062.pdf. [428, 432, 444, 447, 460]

Boozer, Michael and Tavneet Suri (2001), "Child labor and schooling decisions in Ghana." Working paper, Yale University. [428, 447]

Buonomo Zabaleta, Mariela (2011), "The impact of child labor on schooling outcomes in Nicaragua." *Economics of Education Review*, 30 (6), 1527–1539. [428, 444, 447]

Campbell, Jonathan (1998), "Book review: Peabody picture vocabulary test, third edition." *Journal of Psychoeducational Assessment*, 16 (4), 334–338. [434]

Campbell, Jonathan M., Stephen K. Bell, and Lori K. Keith (2001), "Concurrent validity of the peabody picture vocabulary test-third edition as an intelligence and achievement screener for low SES African American children." *Assessment*, 8 (1), 85–94. [434]

Cano, Tomás, Francisco Perales, and Janeen Baxter (2019), "A matter of time: Father involvement and child cognitive outcomes." *Journal of Marriage and the Family*, 81 (1), 164–184. [460]

Cigno, Alessandro and Furio C. Rosati (2005), *The Economics of Child Labour*. Oxford University Press. [428]

Cunha, Flavio and James Heckman (2008), "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation." *Journal of Human Resources*, 43 (4), 738–782. [456]

Das, Jishnu and Tristan Zajonc (2010), "India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement." *Journal of Development Economics*, 92 (2), 175–187. [428]

Del Boca, Daniela, Chris Flinn, and Matthew Wiswall (2014), "Household choices and child development." *Review of Economic Studies*, 81 (2), 137–185. [460]

Del Boca, Daniela, Chiara Monfardini, and Cheti Nicoletti (2017), "Parental and child time investments and the cognitive development of adolescents." *Journal of Labor Economics*, 35 (2), 565–608. [453, 460]

Del Bono, Emilia, Marco Francesconi, Yvonne Kelly, and Amanda Sacker (2016), "Early maternal time investment and early child outcomes." *Economic Journal*, 126, F96–F135. [460]

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger (2014), "School choice, school quality, and postsecondary attainment." *The American Economic Review*, 104 (3), 991–1013. [426]

Dumas, Christelle (2012), "Does work impede child learning? The case of Senegal." *Economic Development and Cultural Change*, 60 (4), 773–793. [428, 431, 432, 447, 460]

Dunn, Lloyd, Eligio Padilla, Delia Lugo, and Leota Dunn (1986), "Manual Del Examinador Para El Test de Vocabularo En Imagenes Peabody: Adaptacion Hispanoamericana." Circle Pines (Minnesota): AGS, American Guidance Service. [434]

Dunn, Lloyd M., Leota M. Dunn, and American Guidance Service (1997), *A Examiner's Manual for the PPVT-III, Peabody Picture Vocabulary Test*, Third edition. AGS. [434]

Edmonds, Eric V. (2007), "Child labor." In *Handbook of Development Economics*, Vol. 4, 3607–3709, Elsevier. [428, 429, 435]

Edmonds, Eric V. (2009), "Defining child labour: A review of the definitions of child labour in policy research." Working Paper. International Labour Office, International Programme, on the Elimination of Child Labour (IPEC)—Geneva. [435]

Emerson, Patrick M., Vladimir Ponczek, and André Portela Souza (2017), "Child labor and learning." *Economic Development and Cultural Change*, 65 (2), 265–296. [428, 431, 444]

Fiorini, Mario and Michael Keane (2014), "How the allocation of children's time affects cognitive and noncognitive development." *Journal of Labor Economics*, 32 (4), 787–836. [426, 439, 460]

Fiszbein, Ariel and Norbert R. Schady (2009), "Conditional cash transfers: Reducing present and future poverty." World Bank Publications. [460]

Funk, Peter and Thorsten Kemper (2016), "Leisure and learning activities and their effects on child skill development." Economics Working Paper #85, University of Cologne. [460]

Glewwe, Paul (2002), "Schools and skills in developing countries: Education policies and socioeconomic outcomes." *Journal of Economic Literature*, 40 (2), 436–482. [428]

Glewwe, Paul and Karthik Muralidharan (2016), "Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications." In *Handbook of the Economics of Education*, Vol. 5 (Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, eds.), 653–743, Elsevier. [443]

Gray, Shelley, Elena Plante, Rebecca Vance, and Mary Henrichsen (1999), "The diagnostic accuracy of four vocabulary tests administered to preschool-age children." *Language, Speech, and Hearing Services in Schools*, 30 (2), 196–206. [434]

Guarino, Cassandra M., Mark D. Reckase, and Jeffrey M. Wooldridge (2014), "Can value-added measures of teacher performance be trusted?" *Education Finance and Policy*, 10 (1), 117–156. [426]

Gunnarsson, Victoria, Peter F. Orazem, and Mario A. Sánchez (2006), "Child labor and school achievement in Latin America." *The World Bank Economic Review*, 20 (1), 31–54. [428, 429, 432, 444, 447, 460]

Heady, Christopher (2003), "The effect of child labor on learning achievement." *World Development*, 31 (2), 385–398. [428, 444]

Hsin, Amy and Christina Felfe (2014), "When does time matter? Maternal employment, children's time with parents, and child development." *Demography*, 51, 1867–1894. [460]

Hull, Marie (2017), "The time-varying role of the family in student time use and achievement." *IZA Journal of Labor Economics*, 6 (10), 1–20. [460]

International Labour Organization (2017a), Global Estimates of Child Labour: Results and Trends, 2012–2016. [427]

International Labour Organization (2017b), *Ending Child Labour by 2025: A Review of Policies and Programmes*. [427]

Jacoby, Hanan G. and Emmanuel Skoufias (1997), "Risk, financial markets, and human capital in a developing country." *The Review of Economic Studies*, 64 (3), 311–335. [436]

Keane, Michael, Sonya Krutikova, Timothy Neal (2022), "Supplement to 'Child work and cognitive development: Results from four low to middle income countries'." *Quantitative Economics Supplemental Material*, 13, https://doi.org/10.3982/QE1745. [430]

Keane, Michael and Kenneth Wolpin (1997), "The career decisions of young men." *Journal of Political Economy*, 105 (3), 473–522. [430]

Montie, Jeanne E., Zongping Xiang, and Lawrence J. Schweinhart (2006), "Preschool experience in 10 countries: Cognitive and language performance at age 7." *Early Childhood Research Quarterly*, 21 (3), 313–331. [433]

Morrow, Virginia and Jo Boyden (2018), *Responding to Children's Work: Evidence From the Young Lives Study in Ethiopia, India, Peru and Vietnam*. Young Lives, Oxford. Summative Report. [455]

Morrow, Virginia, Yisak Tafere, and Uma Vennam (2014), "Changes to rural children's use of time: Evidence from Ethiopia and Andhra Pradesh, India." In *Growing up in Poverty: Findings From Young Lives* (Michael Bourdillon and Jo Boyden, eds.). Palgrave, Basingstoke. [456]

Morrow, Virginia and Uma Vennam (2012), "Children's responses to risk in agricultural work in Andhra Pradesh, India." *Development in Practice*, 22 (4), 549–561. [458]

Muralidharan, Karthik and Sundararaman Venkatesh (2013), "Contract teachers: Experimental evidence from India." NBER Working Paper No. 19440. [426]

Orazem, Peter and L. Victoria Gunnarsson (2004), "Child labour, school attendance and performance: A review." Iowa State University Dept. of Economics WP # 04001. [444]

Patrinos, Harry A. and George Psacharopoulos (1997), "Family size, schooling and child labor in Peru—an empirical analysis." *Journal of Population Economics*, 10 (4), 387–405. [428, 458]

Ravallion, Martin and Quentin Wodon (2000), "Does child labour displace schooling? Evidence on behavioural responses to an enrollment subsidy." *The Economic Journal*, 110 (462), C158–C175. [428]

Ray, Ranjan and Geoffrey Lancaster (2005), "The impact of children's work on schooling: Multi-country evidence." *International Labour Review*, 144 (2), 189–210. [428, 444]

Sanderson, Eleanor and Frank Windmeijer (2016), "A weak instrument F-test in linear IV models with multiple endogenous variables." *Journal of Econometrics*, 190, 212–221. [448, 449, 452]

Sedlacek, Guilherme, Suzanne Duryea, Nadeem Ilahi, and Masaru Sasaki (2009), "Child labor, schooling and poverty in Latin America." In *Child Labor and Education in Latin America* (Peter F. Orazem, Guilherme Sedlacek, and Zafiris Tzannatos, eds.). Palgrave Macmillan, New York. [428]

Singh, Abhijeet (2015), "Private school effects in urban and rural India: Panel estimates at primary and secondary school ages." *Journal of Development Economics*, 113, 16–32. [455]

Singh, Abhijeet (2020), "Learning more with every year: School year productivity and international learning divergence." *Journal of the European Economic Association*, 18 (4), 1770–1813. [455]

Skoufias, Emmanuel and Susan Parker (2001), "Conditional cash transfers and their impact on child work and schooling: Evidence from the progresa program in Mexico." *Economia*, 2 (1), 45–86. [460]

Todd, Petra and Kenneth Wolpin (2003), "On the specification and estimation of the production function for cognitive achievement." *Economic Journal.*, 113 (485), F3–F33. [430, 439]

Todd, Petra and Kenneth Wolpin (2006), "Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility." *American Economic Review*, 96 (5), 1384–1417. [460]

Todd, Petra and Kenneth Wolpin (2007), "The production of cognitive achievement in children: Home, school, and racial test score gaps." *Journal of Human Capital*, 1 (1), 91–136. [426, 430, 439]

Trinh, Trong-Anh, Alberto Posso, and Simon Feeny (2020), "Child labor and rainfall deviation: Panel data evidence from rural Vietnam." *The Developing Economies*, 58 (1), 53–76. [447]

Verriest, Ewout (2020), "Household labor supply, child development and childcare policies in the United States." SSRN Research Paper 3330051. Available Online: http://dx.doi.org/10.2139/ssrn.3330051. [460]

Zharkevich, Ina, Jennifer Roest, and Vu Thi Thanh Huong (2016), "Gendered trajectories through school, work and marriage in Vietnam." Young Lives Working Paper 156. [457]