Detecting Mother-Father Differences in Spending on Children: A New Approach Using Willingness-to-Pay Elicitation

By Rebecca Dizon-Ross and Seema Jayachandran*

This paper tests whether mothers and fathers differ in their spending on their daughters relative to their sons. We compare mothers' and fathers' willingness to pay (WTP) for specific goods for their children, diverging from the previous literature's approach of comparing the expenditure effects of mothers' versus fathers' income. Our method, which we apply in Uganda, allows us to estimate gender differences and explore mechanisms with greater precision. We find that fathers have a lower WTP for their daughters' human capital than their sons' human capital, whereas mothers do not. We then examine why spending patterns differ between mothers and fathers, e.g., altruism, personal returns to investing in children. We find evidence that altruism plays a role: fathers' WTP for goods that simply bring joy to their daughters is lower than their WTP for such goods for their sons, but mothers' is not.

Do fathers invest less in their daughters than their sons? Are mothers less discriminatory against their daughters? If so, these relationships would be important for policy, as they would imply that improvements in gender equality are self-reinforcing. As women gain more say in household decision-making, household spending on daughters may increase, producing more gender equality in the next generation. This virtuous cycle could help to close the gender gaps in schooling and health care that are pervasive in developing countries (Evans, Akmal and Jakiela, 2021).

In this paper, we examine if and why fathers underspend on their daughters' health and education relative to how mothers spend. We adopt a new approach to measure parents' spending preferences that has higher statistical power than the traditional approach used in the literature: we elicit and compare mothers' and fathers' willingness to pay (WTP) for various goods for their sons and daughters. We conduct the study in Uganda among a sample of 1,084 households with young children in which we interviewed one randomly selected parent or usually both (separately).

We find that fathers have a significantly lower WTP for their daughters' human capital

* Dizon-Ross: University of Chicago, University of Chicago Booth School of Business. 5807 South Woodlawn Avenue. Chicago, IL 60637, rdr@chicagobooth.edu. Jayachandran: Princeton University, 126 Julis Romo Rabinowitz Building. Princeton NJ 08544, jayachandran@princeton.edu. We thank Rebekah Chang, Jake Gosselin, Suanna Oh, Juliana Sánchez Ariza, and Varun Satish for excellent research assistance and Lori Beaman, Kate Orkin, Heather Sarsons, and Alessandra Voena for helpful comments.Jayachandran gratefully acknowledges support from the National Science Foundation (SES-1156941). The study received institutional review board approval from Northwestern University (STU00055491), Stanford University (21651) and the Uganda National Council for Science and Technology (SS3002). The data and code for the study are available through the AEA Data and Code Repository (Dizon-Ross and Jayachandran, 2022*a*)

than their sons' human capital. In contrast, mothers, if anything, have a higher WTP for their daughters' human capital than their sons'. As a result, willingness to spend on daughters is higher among mothers than fathers. While previous papers have hinted that mothers might spend more on daughters while fathers spend more on sons (see Lundberg, 2005 for a review), the previous literature's power to statistically reject equality between male and female caregivers has been limited (e.g., Duflo, 2003).

We then investigate the underlying reason for the mother-father differences. In a review paper on possible mother-father differences in son-daughter preferences published two decades ago, Lundberg (2005) noted that "one central unanswered question for economists is whether these differences emerge from parental preferences ...[or from] real or perceived differences in the returns to parental inputs." Today, we still know remarkably little about this question.

The preference-based explanation is simple: fathers care more about their sons than their daughters, and mothers do not. Alternatively, the differences between mothers and fathers could be due to different returns to investment ("investment-based" explanation). Mothers might expect to be more dependent on support from their children in old age, as women have lower earnings and longer life expectancy than men, on average. This could cause mothers to spend more on daughters than fathers do if they believe, as the majority of our sample does, that daughters are more likely to help support their parents in old age than sons are. Fathers and mothers could also have different beliefs about the returns to human capital investment for boys and girls.

To test between these hypotheses, we examine whether there are similar mother-father son-daughter WTP differences for goods that bring joy to the children but do not add to their human capital: toys and candy (hereafter, "enjoyment goods"). Under an investment-based explanation, one would expect the gaps to be observed for human capital goods but not enjoyment goods. Conversely, the patterns being similar for both types of goods would point to a preference-based explanation.

The results support a preference-based explanation: We find that fathers also have a lower WTP for enjoyment goods for their girls than for their boys, suggesting that they have less altruism or love for their daughters than their sons. Mothers, in contrast, have no lower WTP for enjoyment goods for their girls than for their boys. A second piece of evidence consistent with altruism as the underlying explanation is that the mother-father differences in human capital spending are almost entirely driven by households that say the mother loves the children more than the father does.

Our paper's main contribution is to introduce a new way of testing whether mothers and fathers differ in their spending preferences. The standard approach in the literature is to examine the effects of exogenous changes in female versus male income, asking: does a change in income for mothers lead to larger spending on, say, girls' education, or children's education in general, than the same change in income for fathers? However, children's health and education — and goods assignable to children more generally — typically constitute a modest share of household expenditures. Detecting mother-father differences off a small base is statistically challenging. Thus, while a few studies can reject that women's income and men's income have identical effects (e.g., Lundberg, Pollak and Wales, 1996) or can rule out modest differences between women's and men's spending (e.g., Somville, Almås and Vandewalle, 2020), many studies in this literature are underpowered to compare mothers' and fathers' overall spending on children, let alone whether mother-father patterns differ for sons versus daughters. We are unaware of any prior paper that uses the standard approach and statistically rejects that mothers and fathers spend identically on their sons relative to daughters.[1,2]

Our method has considerably higher statistical power to test if mothers and fathers have the same spending patterns for their sons and daughters. To be concrete, Haushofer and Shapiro (2016) compare large cash grants given to men or women in Kenya and report that the minimum detectable effect (MDE) size gap between male and female recipients for their health and education outcomes is 0.24 to 0.25 standard deviations (SD). The MDE for mother-father differences in child spending in our study, which uses a comparable number of households, is 0.08 SD. The gains with our method come from zeroing in on the expenditure category of interest, namely children's human capital.

Another advantage of using WTP to study parental differences in spending is that it directly elicits parent-specific choices rather than inferring them from household-level choices. In the standard approach, if an increase in women's relative income does not change household spending on children, that might be because women's bargaining power did not increase, in which case one cannot conclude whether or not mothers and fathers have similar spending patterns.

Finally, the approach we use is practical. It could easily be incorporated into studies that are interested in measuring parental spending or son preference. The ability to compare WTP for goods with different attributes or that are relevant for different children offers valuable flexibility to test mechanisms. Indeed, this flexibility makes possible our second contribution to the literature, namely shedding light on *why* mothers and fathers spend differently on their children.

I. Data

The data for the study were collected in Iganga district in eastern Uganda. The sample comprises households with a child in primary school. In the first round of data collection (March-May 2013), we surveyed one randomly-selected parent per household. The randomization means that household and child characteristics are balanced when we compare

[1] For example, Duflo (2003) analyzes the child health effects of pensions in South Africa, and one likely could not reject that there is no difference in the effect of grandmothers' versus grandfathers' income on boys relative to girls. (The paper does not report this test.) Other studies on male-female differences in investment in children include Thomas (1994), Qian (2008), Benhassine et al. (2015), and Akresh, De Walque and Kazianga (2016).

[2] Two related papers use lab experiments to compare mothers' and fathers' allocation of money to their child or to themselves (Cherchye et al., 2021; Ringdal and Sjursen, 2021). These studies were conducted after ours. Ringdal and Sjursen (2021) also find some evidence for gender-concordant patterns, although acknowledge that there might be biased selection into their sample by the gender of the child which contributes to their gender-concordant finding. Another related study is Nikiforidis et al. (2018), who asked a small sample of mothers and fathers visiting a zoo in the US to choose between a boy's or girl's backpack and posed a similar question about a savings bond to an Amazon MTurk sample, with the main finding that choices were gender-concordant.

mothers and fathers. In the second round of data collection (September-October 2013), we returned to a subset of the households, specifically those who also had a child age 3 to 8 years old and surveyed the other parent. Figure 1 provides a visual summary of the sampling strategy.

Sampling

We sampled households with children enrolled in grade 4 to 6 in 40 government primary schools with whom we partnered to offer one of our education goods (practice exams that schools administer for a fee). We began with a listing of eligible households in the participating schools. The first eligibility criterion was that the child lived with both biological parents (94% of children). The second was that, for the current academic term, the child's parents had not paid for all of the practice exams the school offered (70% of children). The rationale was that we would be eliciting WTP at different prices up to but not exceeding the market price, so households already purchasing the good at the market price would generate no variation in WTP.

First survey

For the first survey, we randomly selected whether to interview the mother or the father, stratified on school and whether the household was polygamous. A surveyor visited the home and administered a screening questionnaire to confirm eligibility. The final sample comprises 1084 households that met the eligibility criteria and agreed to participate.

The survey elicited WTP for goods for the focal child enrolled in grade 4-6.[3] If a household also had a younger child aged 3-8, we gathered WTP for one good for that younger child. Hereafter, we refer to these two children as the older and younger child, respectively.

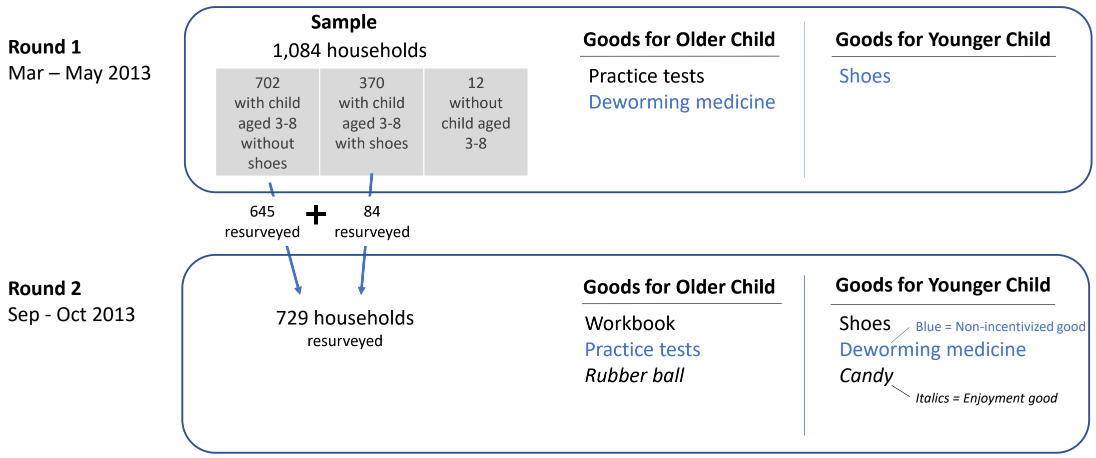
The randomly chosen parent was interviewed with no one else present, besides infants or toddlers. The survey first collected information on household composition, family background, and income and assets for each parent. We then elicited WTP for a set of goods. Finally, we asked questions related to mechanisms, such as expected old age support. The survey took approximately 75 minutes to complete. The participant received 8,000 or 10,000 Ugandan shillings (1 USD $\approx$ 2600 UGX) as compensation for their time and to minimize cash constraints affecting WTP.[4]

Second survey

In the second round of data collection, we revisited 729 of the original 1084 households and surveyed the other parent, using a similar survey structure and content. The reason for revisiting a subsample was budgetary constraints. We focused on the subsample with

[3] In cases of more than one eligible older child in a household, we randomly selected one, and the same for the younger child.

[4] We randomized the payment level to test for cash-on-hand and gift-exchange effects on WTP. Appendix Table A.1 shows that receiving the higher compensation level does not affect WTP for the goods we offered.



Note: The market prices of each good in UGX are as follows. Practice tests: 6,000. Shoes: 2,000 or 2,500 (note that this varied by foot size; we used the child's age as a proxy when choosing the maximum price for the shoe elicitation). Deworming medicine: 4,000. Workbook: 4,500. Rubber ball: 1,500. Candy: 3,000.

FIGURE 1. OVERVIEW OF SAMPLING AND CHILD GOODS FOR BOTH SURVEY ROUNDS

a child aged 3 to 8. Because one of the goods would be shoes for the younger child, we over-sampled households in which the child did not own shoes, according to the parent interviewed in the first survey. We attempted to interview all 702 of these households and completed interviews with 645. We also resurveyed a random subset (84 households) of the other households with a child age 3 to 8. All respondents received 9000 UGX for participating.

Procedure for WTP elicitation

To elicit WTP, we used the Becker-DeGroot-Marschak (BDM) mechanism, asking the respondent if he or she was willing to purchase the good at a series of prices, in declining order from the market price to a price near zero (Becker, DeGroot and Marschak, 1964).[5] The decrement was chosen so that respondents were asked about roughly 12 price levels per good. The respondent was told that after the price questions, one price would be randomly chosen and she would purchase the good from us at that price if and only if her response had been that she wanted to. The surveyor explained the procedure in detail to ensure comprehension, and we also asked debrief questions (such as regret about one's choices) to confirm comprehension. The selection of the randomized price and exchange of money and goods, if applicable, were conducted just after the BDM questions were asked

[5]Recent studies validating the BDM method in developing countries include Berry, Fischer and Guiteras (2020) and Burchardi et al. (2021).

for a good. In the second wave, to increase sample size without increasing study costs, we grouped five of the goods, and first randomly chose one good and then one price level for that good; respondents were informed in advance that a transaction could only occur for one of these five goods; this two-step randomization occurred after the BDM questions for all five goods.

Also to increase sample size, in each wave we asked WTP in a similar but non-incentivized way for additional goods. The surveyor followed the same protocol of showing the actual good to the respondent so that it was concrete, but respondents knew in advance that for these goods, no transaction would take place. Appendix B presents evidence that the non-incentivized WTP elicitation appears to have worked quite similarly to the incentivized WTP elicitation. As a result, we pool incentivized and non-incentivized WTP in our main specifications for statistical power.

Children's goods

We used several criteria when choosing which goods to offer parents. First, we wanted parents to be familiar with the good and its market price; otherwise, based on piloting, variation in the perceived quality and market price would add noise and potentially bias the results. Second, the good should be something that most households value at less than the market price, but place some value on, so that there is variation in WTP. Third, the good should not be particularly appealing to one gender, *within* the categories of human capital or "enjoyment goods" (non-human capital); the goods are intended to represent the broad categories of human capital or enjoyment goods, so while each whole category might be favored more by one gender, we would not want a good that, say, fathers idiosyncratically like more than mothers do or that is more appropriate for girls than boys. In addition to doing extensive preliminary fieldwork to choose goods that met these criteria, we asked questions on the survey to verify our assumptions.

Figure 1 lists each child good, which survey it was collected in, which focal child it was for, and whether its WTP elicitation was incentivized. The market prices of the goods ranged from 1,500 to 6,000 UGX.

Human capital goods

In the first survey, we elicited WTP for three human capital goods. The first, measured in an incentivized way, was practice exams for the older child, administered by the child's school. Schools offer practice exams during and at the end of each of the three terms of the school year, but charge students to participate. Our survey was conducted during Term 1 of 2013, and we sold a bundle of all of the exams for Terms 2 and 3.

The other human capital goods in the first survey were deworming medicine for the older child and, if the household had a child age 3 to 8, rubber-soled shoes for them. [6] Many

[6]In the first survey, we asked about deworming only for those randomized to receive the higher payment for participating in the study. The reason was to justify the higher compensation for some people by their survey being longer.

young children do not wear shoes, and being barefoot is a risk factor for intestinal worms, as well as cuts and injuries. WTP for these two goods was measured in a non-incentivized way.

In the second survey, we elicited WTP for four human capital goods: rubber-soled shoes for the younger child (incentivized); a grade-appropriate math workbook for the older child (incentivized); deworming medicine for the younger child (non-incentivized) and practice exams for the older child for the first two terms of the 2014 school year (non-incentivized).

Enjoyment goods

In the second survey, we also elicited WTP for fun goods for children that are not human-capital enhancing. Both goods were offered for the younger child: a rubber ball (incentivized) and a packet of candy (incentivized).[7]

Each child good was intended for a specific child, and we use the gender of that child to compare how parents spend on sons versus daughters. However, some goods, such as the math workbook, rubber ball, and candy, might be transferable between siblings or shared among them, which would bias us against finding WTP differences by child gender.

Benchmark goods

We elicited WTP for goods used by adults as a benchmark of each respondent's general WTP for goods. We control for the adult good WTP in our regressions to increase precision, as factors such as cash on hand or gift-exchange motives should affect adult good WTP similarly to how they affect child good WTP (Dizon-Ross and Jayachandran, 2022*b*). In principle, controlling for the adult good WTP could also help address any systematic differences in such factors between male and female respondents. In practice, controlling for the adult good WTP does not change our point estimates much, only the standard errors, and we show robustness to excluding the control. We attempted to identify goods that were not gendered. In our preparatory fieldwork, we asked our field team and pilot respondents if they thought potential goods were preferred more by one gender or the other, and we chose goods that did not have that characteristic. In the first survey, the adult good was a metal cup for drinking (3600 UGX) and in the second survey, the two adult goods were a poster (2000 UGX) and a pair of jerry cans (4000 UGX).

Survey questions to test mechanisms

Our strongest test of whether mothers and fathers simply differ in how much they care about their children's well-being is to examine WTP for enjoyment goods, but we also asked direct questions about which parent cared more about the children. We similarly asked direct questions about other potential mechanisms such as perceived returns to investing in human capital.

[7] We also asked about a separate toy for girls (teddy bear) and boys (toy truck) in both surveys (non-incentivized). To make comparisons across boys and girls, it is important to use the same gender-neutral enjoyment good, so we exclude these toys from our analysis.

Summary statistics and balance

Table A.2 presents summary statistics for household and focal child characteristics in the full sample and shows that these characteristics are balanced between the mother and father subsamples. Almost all households in our sample own land, and 25% are polygamous. Older focal children are 12 years old, on average, and younger focal children, almost 6. Appendix C presents additional summary statistics and balance tests.

II. Empirical strategy and results

Spending on children's goods

We begin by testing whether parents collectively spend more on their daughters or sons and whether mothers or fathers spend more on average on their children. To do so, we estimate the following equation:

$$WTP_{ihgc} = \alpha + \beta Daughter_{ihc} + \gamma Mother_{ihc} + \delta X_{ihgc} + \epsilon_{ihgc} \tag{1}$$

where each observation is for parent i in household h asked about a good g for child c. The independent variables of interest are *Daughter*, an indicator for the child being female, and $Mother$, an indicator for the respondent being female. In principle, we could estimate the difference using a single good, but for statistical power and so that the results are less specific to a particular good, we pool several goods. To make WTP comparable across goods, we normalize the WTP for each good by its within-sample standard deviation, with the results robust to other normalizations. The vector of control variables X includes good-by-survey-wave fixed effects, and stratum fixed effects. To increase precision, X also includes WTP for the adult goods. The standard errors allow for clustering within a household.

Column (1) of Table 1 presents the results. There is no statistically significant difference between parents' WTP for goods for their daughters versus their sons and also no statistical difference between mothers' and fathers' WTP overall.

Next, to understand whether daughter-son spending preferences differ across mothers and fathers, we add the interaction term and estimate the following equation:

$$WTP_{ihgc} = \alpha + \beta Daughter_{ihc} + \delta Mother_{ih} \times Daughter_{ihc} + \lambda Mother_{ih} + \gamma X_{ihgc} + \epsilon_{ihgc} \tag{2}$$

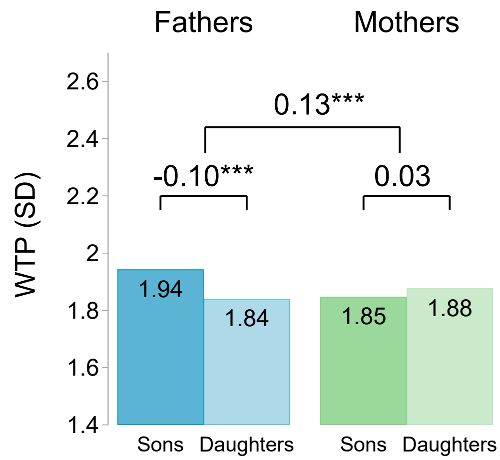
In this case, the coefficient on *Daughter* tells us how much lower fathers' WTP is for daughters than sons, and the coefficient on $Mother \times Daughter$ tells us how different the *Daughter* effect (i.e., the spending on daughters relative to sons) is for mothers than fathers. Finally, the coefficient on $Mother$ represents how much less mothers spend overall (on sons) than fathers.

Table 1—Fathers spend less on girls but mothers do not

	WTP normalized by...					
	SD	SD	Market price	SD	SD	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Daughter	-0.037	-0.102	-0.029	-0.094	-0.065	-0.159
	(0.024)	(0.032)	(0.009)	(0.037)	(0.036)	(0.052)
Mother × Daughter		0.131	0.036	0.142	0.069	0.209
		(0.046)	(0.013)	(0.053)	(0.053)	(0.076)
Mother	-0.029	-0.095	-0.028	-0.066	-0.091	-0.090
	(0.028)	(0.036)	(0.010)	(0.043)	(0.040)	(0.059)
p-val: Mother + Mother × Daughter = 0		0.318	0.420	0.070	0.589	0.032
p-val: Daughter + Mother × Daughter = 0		0.399	0.454	0.204	0.931	0.349
Dep. var. mean father-son	1.943	1.943	0.537	1.943	1.793	2.164
Fixed effects	Stratum	Stratum	Stratum	HH	Stratum	Stratum
Goods included	All	All	All	All	Incentivized	Non-incentivized
Number of observations	6,673	6,673	6,673	6,673	4,000	2,673

Notes: All columns control for survey round, adult WTP, and adult WTP interacted with survey round. Columns 1-3 control for strata and good fixed effects. Column 4 controls for household and good fixed effects. Standard errors are clustered by household.

Column (2) of Table 1 shows that fathers have a lower WTP for goods for their daughters, while mothers do not. Figure 2 displays the same result graphically. In the figure, the left set of bars shows that fathers have a 0.10 SD lower WTP for daughters than sons. (In the table, this corresponds to the coefficient of -0.10 for *Daughter*.) The bars on the right show that, in contrast to fathers, mothers spend similarly on their daughters and sons – a 0.03 SD difference which, if anything, points to spending more on their daughters. (In the table, the 0.03 SD estimate is the sum of the *Daughter* and $Mother \times Daughter$ coefficients, and the p-value at the bottom of the table shows that the difference is not significant.) Finally, one can statistically reject that fathers and mothers have the same boy-girl gap in spending. (This corresponds to the positive and significant 0.13 SD $Mother \times Daughter$ coefficient in the table.)



Note: Figure is based on the regression reported in Table 1, column (2).

Figure 2. Fathers spend less on girls but mothers do not

To help gauge the magnitudes of these boy-girl and mother-father differences, the bottom row of Table 1 reports the mean of WTP for the father-son subsample. Dividing the effect sizes by that mean expresses how large the WTP gaps are in percentage terms. For example, the daughter-son gap for fathers seen in Figure 2 (i.e., the *Daughter* coefficient in Table 1, column (2)) maps to fathers being willing to pay 5% less for goods for their daughters than their sons. Moreover, under certain assumptions, the percent WTP gap can also be interpreted as the percent gap in demand (i.e., in the expected quantity of goods purchased). Thus, fathers also have 5% lower demand for goods for their daughters than sons. This mapping from WTP to overall demand is laid out in a simple framework in Appendix D that assumes there are an array of goods available at different prices.

Similarly, dividing the $Mother \times Daughter$ coefficient by the father-son dependent variable mean shows that that effect represents a 7% change in WTP. Adding that term to the negative 5% *Daughter* effect suggests that mothers have 2% higher demand for goods for their daughters than their sons. Appendix D also discusses how to estimate the percent gaps in expenditures (as opposed to quantity demanded), and finds similar magnitudes. For example, the estimates in Figure 2 map to fathers spending 8% less on their daughters than sons.[8]

The findings shown in Figure 2 and Table 1, column (2) are robust across several different specifications. Table 1, column (3) shows robustness to normalizing the WTP for each good by its market price. The main effect of -0.03 for *Daughter* normalized by the dependent variable mean for fathers and sons again suggests that fathers have 5% lower demand for goods for their daughters than for their sons. Column (4) shows that the results are robust

[8] Appendix D also presents a second way to calculate the implied effect on expenditures that assumes different market prices for the goods we asked about, which yields similar magnitudes.

to including household fixed effects.

Our main findings are based on pooling incentivized and non-incentivized WTP observations. One potential concern with using non-incentivized WTP data is that perhaps mothers are more prone to social desirability bias than fathers (or vice versa), which could cause mothers to have inflated non-incentivized WTP relative to fathers. However, we collected a measure of the respondent's propensity to give socially desirable survey answers (the Marlowe-Crowne scale) in our second survey round and find no differences between mothers and fathers (Crowne and Marlowe, 1960).[9] Moreover, we can test whether our results depend on whether we incentivized the elicitation of WTP. Columns (5) and (6) of Table 1 display the results from estimating equation (1) separately for the goods for which the elicitation was incentivized and non-incentivized, respectively. While these estimates are less precise than the pooled estimate in column (2), the qualitative takeaway is the same in both cases: Fathers have significantly lower spending on daughters than sons (i.e., $Daughter$ is negative), whereas mothers do not (i.e., $Daughter + Mother \times Daughter$ is not negative). We cannot reject that the coefficients in columns (5) and (6) are identical (p-values for equality: 0.16, 0.13, and 0.92 for the $Daughter$, $Mother \times Daughter$, and $Mother$ coefficients, respectively), although the coefficient magnitudes vary between the columns, so we also cannot rule out substantive differences.

Thus, while non-incentivized WTP seems to have performed well, and thus our results are promising for combining non-incentivized WTP with incentivized WTP to identify parents' spending patterns in future work, there is still some uncertainty about how much real stakes matter for the quantitative estimates.

Appendix E presents further sensitivity analyses on the results shown in Figure 2, such as excluding WTP for the adult good as a control variable, and concludes that it is robust across several alternative specifications.

Spending on human capital

The child development implications of parents' spending preferences depend on whether the results above hold for human capital in particular. We estimate equation (2) using human capital goods only and show the results in Figure 3(a), as well as in Column (1) of Appendix Table A.3. Mean WTP for each human capital good is 2,100 UGX, roughly 15% of total per-term per-child educational spending.

Again, fathers have significantly lower WTP and demand for goods for their sons than daughters, with an effect size of -0.11 SD, or -5%. Meanwhile, mothers, if anything, have higher demand for goods for daughters than sons. Mothers spend roughly 0.09 SD (4%) more on daughters than fathers do, and roughly 0.06 SD (3%) less on sons than fathers do, although only the former difference is statistically significant. (See the p-values on $Mother$ + $Mother \times Daughter$ and on $Mother$, respectively, in Column (1) of Appendix Table

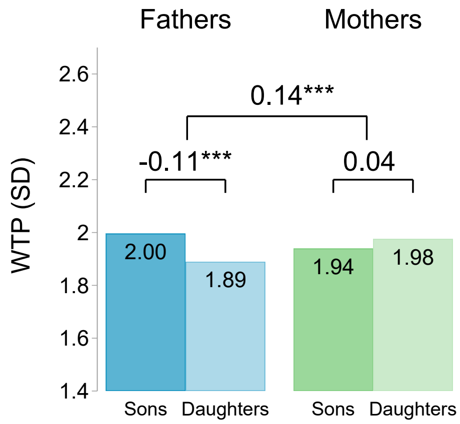
[9]The Marlowe-Crowne measure captures a person's general tendency to given socially desirable survey answers so does not allow us to assess whether mothers or fathers have different levels of social desirability bias when asked about girls in particular. Nonetheless, we view the lack of mother-father differences in the general propensity for social desirability bias as reassuring.

A.3.) The difference in mothers' and fathers' pattern of spending on daughters relative to sons (i.e., the $Mother \times Daughter$ effect) is the highly significant 0.14 SD or 7% gap shown in Figure 3(a).

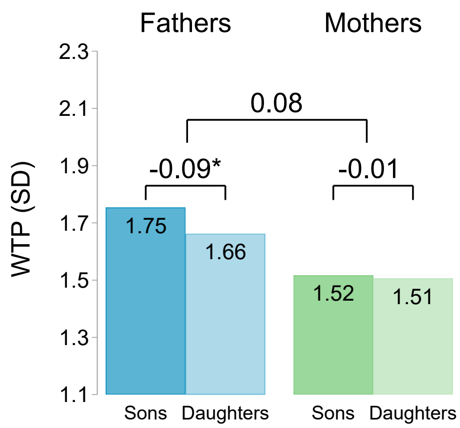
Columns (2) and (3) of Appendix Table A.3 show that the patterns are qualitatively similar for the two subcategories of human capital – education and health – with the magnitude and significance of the coefficients larger for health.

Testing why mothers and fathers spend differently

What underlies these differences between mothers and fathers in their spending on sons versus daughters? We begin by testing whether altruism towards sons relative to daughters plays a role. We examine parents' WTP for goods that are purely for their children's enjoyment and do not build human capital. Differential WTP for those goods by child gender suggests differences in altruism towards sons and daughters.



(a) Human capital



(b) Enjoyment goods

Note: Figures 3(a) and 3(b) present the coefficient estimates from estimating equation 2 using only human capital goods and only enjoyment goods, respectively.

FIGURE 3. FOR BOTH HUMAN CAPITAL AND ENJOYMENT GOODS, FATHERS SPEND LESS ON DAUGHTERS BUT MOTHERS DO NOT

We estimate equation 2 using only the "pure enjoyment" (non-human capital) goods and display the results in Figure 3(b). We also report the result in Appendix Table A.3, column (4). Among fathers, WTP for fun items for daughters is considerably lower than for boys. Comparing the two panels of Figure 3, for both enjoyment goods and human capital goods, fathers have around 0.1 SD lower WTP for their daughters than their sons. The similar pattern for enjoyment goods as for human capital is consistent with fathers having lower altruism for their daughters than sons.[10] Meanwhile, mothers have no lower WTP for their daughters' enjoyment than their sons'. [11]

One concern in interpreting these findings is that parents might see the rubber ball and candy as having other benefits besides joy for their children, and these perceptions might differ by parent or child gender. For example, the items might keep the child busy, and mothers, as the primary caregivers, might value this more than fathers do. Appendix Figures A.1 and A.2 verify that the main benefit of these goods, according to parents, was joy for their children. Keeping the child occupied was a negligible benefit cited for the rubber ball, and, while it was mentioned as a benefit of the candy, there are no significant gender differences in these responses. Moreover, the results are robust to controlling for fixed effects for the parent's main reason for valuing each good (see Appendix Table A.3, column (5)), so they are not driven by parental differences in perceived benefits of the candy and ball.

We conclude from Figure 3 that altruism is likely an important reason for mother-father differences in spending on daughters. At the same time, the evidence leaves scope for other explanations too. The magnitude of the $Mother \times Daughter$ effect for enjoyment goods (0.08 SD) is smaller than the corresponding $Mother \times Daughter$ effect for human capital (0.14 SD), although we cannot reject equality (Appendix Table A.3, column (6), p-value = 0.31). Taking the two effect sizes at face value would imply that altruism explains around half (0.08/0.14) of the $Mother \times Daughter$ effect on human capital. This suggests that a second explanation – such as investment motives – might explain the other half. Consistent with this idea, the difference between the $Mother \times Daughter$ effects for human capital and enjoyment goods stems from mothers spending more on their daughters' human capital than their sons', while spending the same amount on their sons' and daughters' enjoyment goods. This pattern is what one would expect if mothers have stronger investment motives, specifically in daugthers. In our survey, 55% of parents stated that adult daughters support their parents more than adult sons do, compared to only 20% believing that sons provide more support. (The remaining 25% believed sons and daughters provide equal support.) In addition, mothers were more likely than fathers to state that they were the parent who would receive the most support from their adult children.

A variant of the investment story is that mothers perceive the returns to female human capital (relative to male human capital) to be higher than men do. However, as shown in Appendix Table A.4, across several questions about parents' beliefs about the value of schooling for boys and girls, there are no meaningful differences between mothers and

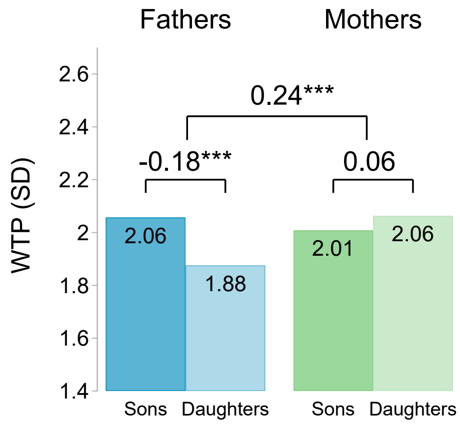
[10]We cannot reject that fathers have the daughter-son gap in spending for human capital and enjoyment goods. We estimate equation 2 using both types of goods, adding in interactions of all of the regressors with a dummy for the good being an enjoyment good. The regression is reported in Appendix Table A.3 , column (6), and the relevant coefficient is for $Daughter \times Enjoyment\ good$.

[11]Mothers also have markedly lower spending than fathers on enjoyment goods for both sons and daughters. While this could reflect mothers having lower altruism overall towards their children, it could also stem from mothers not believing in spending on "frivolous" goods or fathers wanting to be the "fun" parent. It could also reflect women having less control over household income than their husbands, as average income is 3 times as high for fathers as for mothers, and 73% of female and 92% of male respondents say that the man does more of the household spending than the woman does. Notably, however, these other factors (i.e., mothers' different attitudes towards frivolous goods or mothers' different income levels) should not differ between sons and daughters.

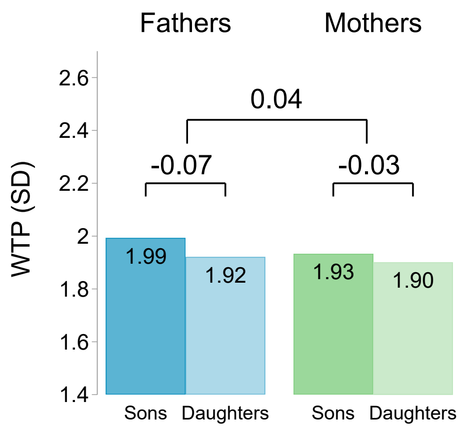
fathers.

As our final analysis, we present a second test of the hypothesis that altruism underlies the mother-father differences in discrimination against daughters. We conduct heterogeneity analysis using a proxy for whether the mother is the more altruistic parent within the household. We estimate equation (2) for human capital goods, separately in two subsamples of households. The first is the roughly 50% of households where both parents identify the mother as the parent who cares more about the children. The second is the other 50% of households, in which either only one or neither of the parents identified the mother as the more loving parent.[12] While we did not ask about which parent cared more about their girls in particular, in a context where boys might be more universally beloved, caring more about children in general may also proxy for caring more about female children.

[12]We can only construct this variable for households where we surveyed both parents, so this analyses only includes these households.



(a) Mother loves more = 1



(b) Mother loves more = 0

Note: Figures 4(a) and 4(b) present the coefficient estimates from estimating equation 2 using human capital goods, for the subsample of households that (a) say the mother loves the children more than the father does and (b) say they care about them equally or the father cares about them more.

FIGURE 4. GENDERED SPENDING PATTERNS ARE DRIVEN BY HOUSEHOLDS THAT SAY THE MOTHER LOVES THE CHILDREN MORE

Figure 4 presents the results. The pattern that fathers spend less on daughters is much stronger and only statistically significant among households in which the mother is described as the more altruistic parent (panel (a)). Fathers' lower spending on daughters is nearly three times as large for these households compared to the rest of the sample (-0.18 versus -0.07). In addition, mothers' relatively greater WTP for daughters' human capital (the $Mother \times Daughter$ effect) is much larger in the mother-loves-the-children-more households than in other households (0.24 versus 0.04). This difference across subsamples is statistically significant at the 10% level. (Appendix Table A.3, column 7 reports the pooled regression that enables this statistical test.) Thus, this second test also suggests that altruism underlies the mother-father differences in son preference.

III. Conclusion

We revisit the classic question in family economics of whether mothers and fathers spend differently on children, using a different approach than past studies: we elicit each parent's WTP for goods for their children. The advantages of this approach, compared to using exogenous changes in women's and men's income, are statistical power and the ability to choose goods with attributes that enable one to test mechanisms. We apply this method in rural Uganda.

We find that fathers but not mothers spend less on daughters than sons. We then investigate why that is. Specifically, we test between a preference-based explanation, in which mothers care about daughters relatively more than fathers do, and an investment-based explanation, in which mothers enjoy a higher financial return on investment in daughters. We find support for the preference-based explanation. A key test examines parents' WTP for goods that bring joy to the children but do not build their human capital. We find similar patterns for these enjoyment goods as we did for human capital, consistent with fathers' lower altruism toward their daughters playing an important role in spending differences.

Our investigation leaves unanswered many questions about what underlies parental spending differences. We do not consider all possible explanations. For example, gendered norms or specialization could also play a role.[13] We also do not explore the deep causes of preference differences. For example, the literature in sociology and psychology has proposed that preference differences could stem from parents identifying more closely with same-gender children (Belsky, 1979; Nikiforidis et al., 2018). Our results are consistent with men and women both having same-gender favoritism. If mothers and fathers had equal financial resources, such favoritism would cancel out. However, because men control more resources than women do, daughters end up disadvantaged. Continuing to explore the reasons for parental differences in spending is a rich area for further research, and using WTP elicitation as a research design could aid in this exploration.

[13]Doepke and Tertilt (2019) present and test a model in which mothers spend more on children's human capital because each parent specializes in providing different public goods for the household.

REFERENCES

Akresh, Richard, Damien De Walque, and Harounan Kazianga. 2016. "Evidence from a Randomized Evaluation of the Household Welfare Impacts of Conditional and Unconditional Cash Transfers Given to Mothers or Fathers." *World Bank Policy Research Working Paper*, , (7730).

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak. 1964. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science*, 9(3): 226–232.

Belsky, Jay. 1979. "Mother–Father–Infant Interaction: A Naturalistic Observational Study." *Developmental Psychology*, 15(6): 601.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen. 2015. "Turning a Shove Into a Nudge? A "Labeled Cash Transfer" for Education." *American Economic Journal: Economic Policy*, 7(3): 86–125.

Berry, James, Greg Fischer, and Raymond Guiteras. 2020. "Eliciting and Utilizing Willingness to Pay: Evidence From Field Trials in Northern Ghana." *Journal of Political Economy*, 128(4): 1436–1473.

Burchardi, Konrad, Jonathan De Quidt, Selim Gulesci, Benedetta Lerva, and Stefano Tripodi. 2021. "Testing Willingness to Pay Elicitation Mechanisms in the Field: Evidence from Uganda." *Journal of Development Economics*, 102701.

Cherchye, Laurens, Pierre-André Chiappori, Bram De Rock, Charlotte Ringdal, and Frederic Vermeulen. 2021. "Feed the Children." CEPR Discussion Paper No. DP16482.

Crowne, D., and D. Marlowe. 1960. "A new scale of social desirability independent of psychopathology." *Journal of Consulting Psychology*, 24(4): 349–354.

Dizon-Ross, Rebecca, and Seema Jayachandran. 2022*a*. "Data and code for: Detecting Mother-Father Differences in Spending on Children: A New Approach Using Willingness-to-Pay Elicitation." *American Economic Association* [publisher]. Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/E179162V2.

Dizon-Ross, Rebecca, and Seema Jayachandran. 2022*b*. "Improving Willingness-to-Pay Elicitation by Including a Benchmark Good." *American Economic Association Papers and Proceedings*, 122(5): 551–555.

Doepke, Matthias, and Michele Tertilt. 2019. "Does Female Empowerment Promote Economic Development?" *Journal of Economic Growth*, 24(4): 309–343.

Duflo, Esther. 2003. "Grandmothers and Granddaughters: Old Age Pension and Intra-Household Allocation in South Africa." *World Bank Economic Review*, 17(1).

Evans, David K, Maryam Akmal, and Pamela Jakiela. 2021. "Gender gaps in education: The long view." *IZA Journal of Development and Migration*, 12(1).

Haushofer, Johannes, and Jeremy Shapiro. 2016. "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya." *Quarterly Journal of Economics*, 131(4): 1973–2042.

Lundberg, Shelly. 2005. "Sons, Daughters, and Parental Behaviour." *Oxford Review of Economic Policy*, 21(3): 340–356.

Lundberg, Shelly J., Robert A. Pollak, and Terence J. Wales. 1996. "Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit." *Journal of Human Resources*, 32(4): 463–480.

Nikiforidis, Lambrianos, Kristina M Durante, Joseph P Redden, and Vladas Griskevicius. 2018. "Do Mothers Spend More on Daughters While Fathers Spend More on Sons?" *Journal of Consumer Psychology*, 28(1): 149–156.

Qian, Nancy. 2008. "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance." *Quarterly Journal of Economics*, 123(3): 1251–1285.

Ringdal, Charlotte, and Ingrid Hoem Sjursen. 2021. "Household Bargaining and Spending on Children: Experimental Evidence From Tanzania." *Economica*, 88(350): 430–455.

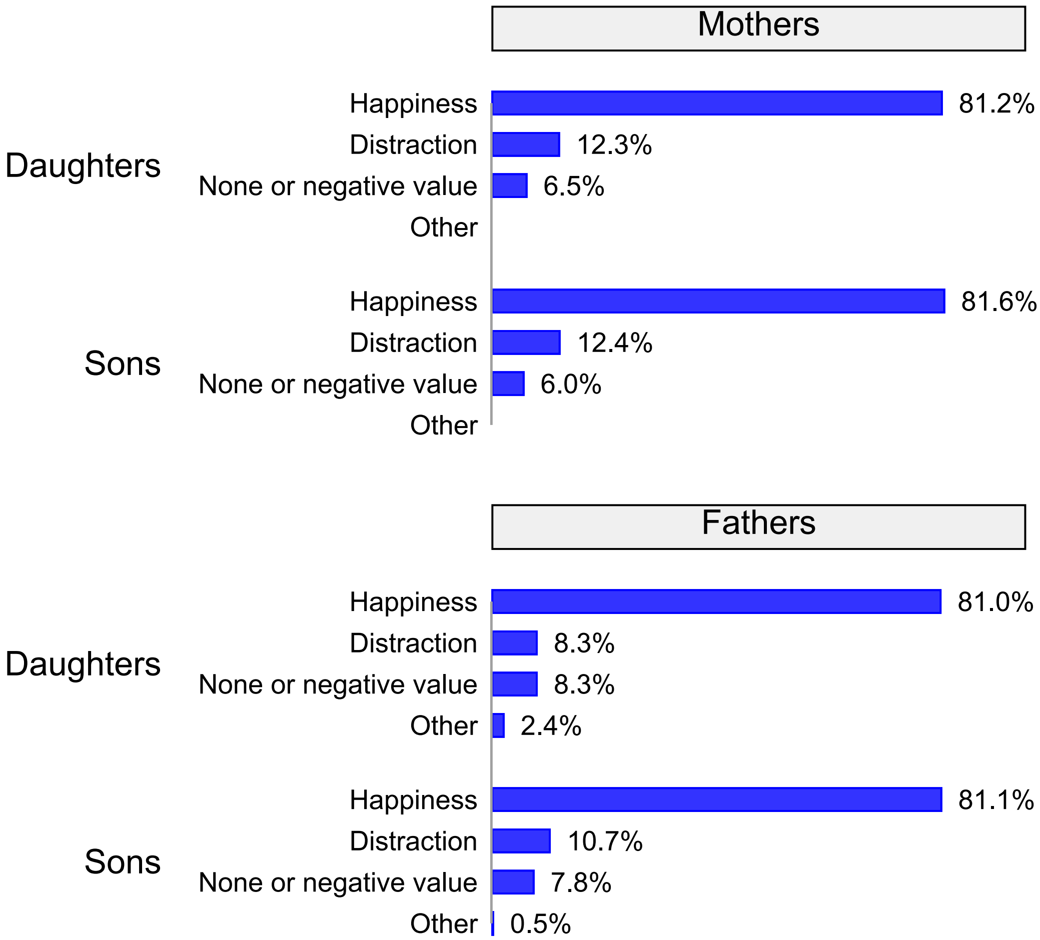
Somville, Vincent, Ingvild Almås, and Lore Vandewalle. 2020. "The Effect of Gender-Targeted Transfers: Experimental Evidence From India." *NHH Dept. of Economics Discussion Paper*, , (16).

Thomas, Duncan. 1994. "Like Father, Like Son, Like Mother, Like Daughter: Parental Education and Child Health." *Journal of Human Resources*, 29(4): 950–988.

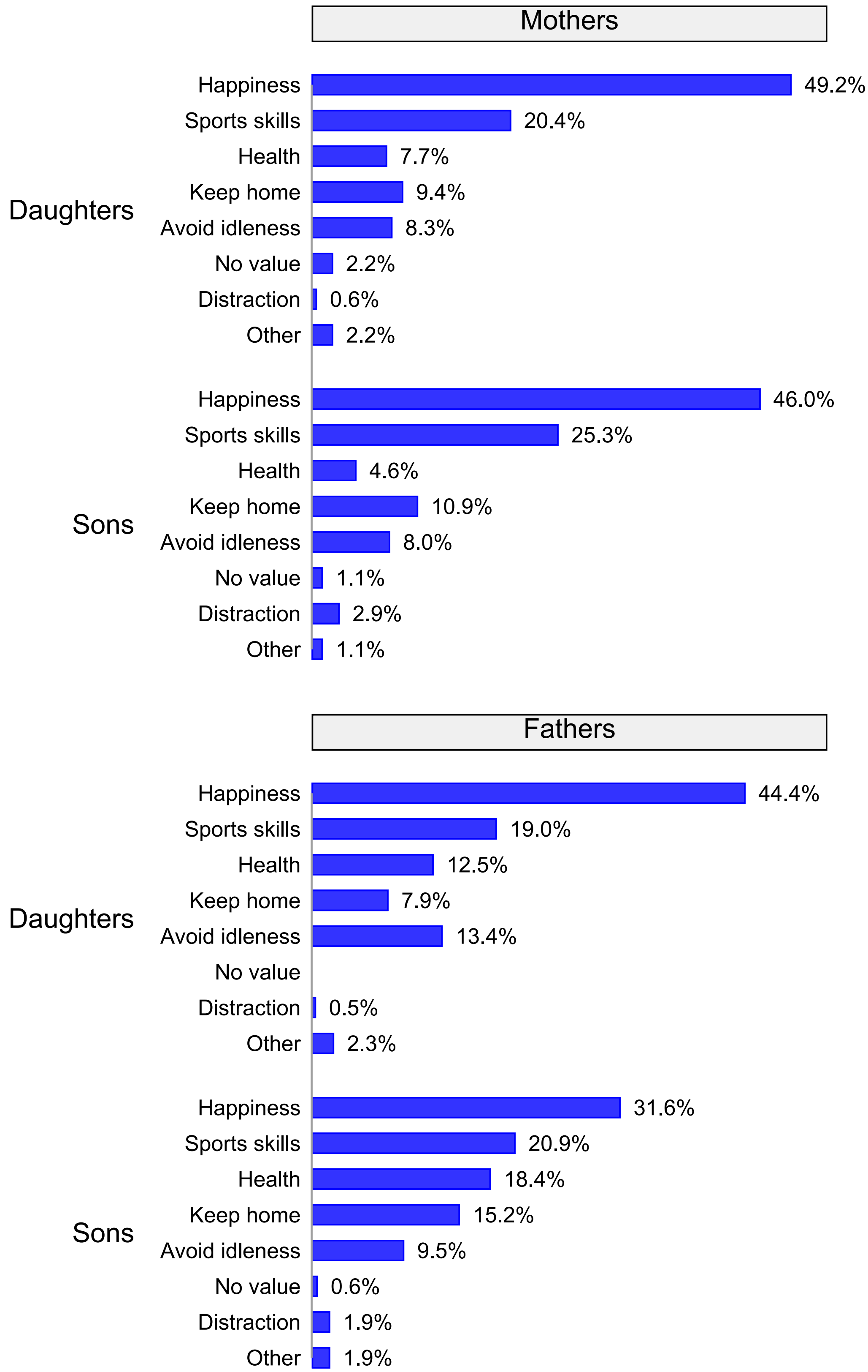
Online Appendix

Detecting Mother-Father Differences in Spending on Children: A New Approach Using Willingness-to-Pay Elicitation

Rebecca Dizon-Ross and Seema Jayachandran



Appendix Figure A.1. Parents' statements of the candy's primary value, by child and parent gender



Appendix Figure A.2. Parents' statements of the rubber ball's primary value, by child and parent gender

Appendix Table A.1—Compensation for study participation does not affect WTP

	Adult good (cup) WTP	Tests WTP	Shoes WTP
	(1)	(2)	(3)
Received higher payment	-0.036	0.004	-0.001
	(0.063)	(0.063)	(0.083)
Observations	1084	1084	680

Notes: Analysis uses first survey round, in which compensation was randomized. The omitted group received 8,000 UGX for participation. The higher payment amount was 10,000 UGX; the payment level was randomized. The analysis omits deworming medicine WTP, as it was elicited only for the higher-payment group. All regressions include strata fixed effects. Standard errors are heteroskedasticity-robust.

Appendix Table A.2—Household and child characteristics are balanced across mother and father samples

Variable	Mothers (1)	Fathers (2)	Standardized diff (3)
Panel A: Household characteristics			
Number of children	9.119 [3.101]	9.025 [2.967]	0.031
Number of cattle	1.006 [1.338]	1.069 [1.372]	-0.046
Number of motos	0.041 [0.199]	0.042 [0.200]	-0.005
Number of rooms	3.014 [1.197]	3.047 [1.189]	-0.028
Owns land	0.914 [0.280]	0.934 [0.248]	-0.076
Polygamous	0.247 [0.434]	0.248 [0.434]	-0.002
Panel B: Focal child characteristics			
Older focal child male	0.544 [0.498]	0.538 [0.499]	0.012
Older focal child age	11.958 [1.949]	12.047 [2.034]	-0.045
Weekly study hours	4.875 [5.305]	5.396 [9.155]	-0.070
Older focal child school performance	3.175 [0.944]	3.233 [0.939]	-0.062
Believes older focal child will support parents more than other children	0.480 [0.500]	0.495 [0.500]	-0.030
Has younger focal child	0.782 [0.413]	0.796 [0.403]	-0.034
Younger focal child male	0.553 [0.498]	0.552 [0.498]	0.002
Younger focal child age	5.733 [1.783]	5.784 [1.824]	-0.028
Younger focal child in school	0.625 [0.485]	0.675 [0.469]	-0.105
Younger focal child grade	1.488 [1.044]	1.538 [1.073]	-0.047
Number of observations	900	913	
Joint p-value	0.495		

Notes: In the regression to test for joint orthogonality, we impute missing values with the sample mean and include missing flags. We also control for survey round and strata fixed effects, to match the main specification. Standard errors are clustered at the household level. The unit of observation is a household-parent.

Appendix Table A.3—Regressions underlying Figures 3 and 4 + results by human capital type

	Goods included:						
	Human Capital (1)	Education (2)	Health (3)	Enjoyment (4)	Enjoyment (5)	All (6)	Human Capital (7)
Daughter	-0.106	-0.081	-0.145	-0.091	-0.085	-0.106	-0.072
	(0.036)	(0.051)	(0.052)	(0.051)	(0.051)	(0.037)	(0.059)
Mother × Daughter	0.143	0.085	0.197	0.081	0.074	0.143	0.040
	(0.051)	(0.070)	(0.075)	(0.081)	(0.081)	(0.051)	(0.085)
Mother	-0.057	-0.064	-0.043	-0.237	-0.230	-0.057	-0.060
	(0.039)	(0.052)	(0.052)	(0.062)	(0.062)	(0.040)	(0.066)
Mother × Daughter × Covariate						-0.062	0.196
						(0.086)	(0.111)
Daughter × Covariate						0.015	-0.110
						(0.057)	(0.079)
Mother × Covariate						-0.180	0.011
						(0.065)	(0.085)
Main perceived benefit X good fixed effects	No	No	No	No	Yes	No	No
Covariate used in interactions						Enjoyment good	Mothers love more
p-val: Mother + Mother × Daughter = 0	0.025	0.639	0.006	0.016	0.016		
p-val: Daughter + Mother × Daughter = 0	0.325	0.930	0.323	0.866	0.859		
p-val: Mother × Daughter + Mother × Daughter × Cov. = 0						0.312	0.001
p-val: Daughter + Daughter × Cov. = 0						0.070	0.001
Dep. var. mean father-son	1.996	1.618	2.299	1.754	1.754	1.943	2.028
Number of observations	5,215	2,542	2,673	1,458	1,458	6,673	4,640

Notes: The dependent variable is WTP for the good. Columns 1,2,3, and 7 control for strata and good fixed effects, survey round, adult WTP, adult WTP interacted with survey round, and all previous controls interacted with the covariate of interest. Columns 4 to 6 control for strata and good fixed effects, survey round, adult WTP, and adult WTP interacted with survey round. Column 5 additionally controls for fixed effects for the parent's main perceived benefit of the good × good fixed effects.The column 4 regression is the analog of Figure 3(b), and includes only the two enjoyment goods (rubber ball and candy). Column 7 is limited to the households with observations from both the mother and the father.

Appendix Table A.4—Summary statistics on mothers' and fathers' beliefs about education

Variable	Full sample (1)	Mothers (2)	Fathers (3)	Standardized diff (4)
Agree: It is useless to send girls to secondary school since they will marry	0.041 [0.198]	0.037 [0.188]	0.045 [0.207]	-0.040
Agree: Even boys who will become farmers will be better at farming if they have gone to school.	0.905 [0.293]	0.890 [0.313]	0.920 [0.271]	-0.102
Would make son finish primary if they wanted to quit	0.881 [0.323]	0.879 [0.326]	0.884 [0.321]	-0.015
[If yes to above] Would make son finish O levels if they wanted to quit	0.943 [0.231]	0.955 [0.208]	0.932 [0.252]	0.100
Would make daughter finish primary if they wanted to quit	0.855 [0.352]	0.847 [0.361]	0.864 [0.343]	-0.048
[If yes to above] Would make daughter finish O levels if they wanted to quit	0.945 [0.229]	0.942 [0.233]	0.947 [0.225]	-0.022
Number of observations	1813	900	913	

Notes: All variables are observed in both survey rounds, except the one reported in the first row, which is only available in the first round.

Validation of non-incentivized WTP

To assess the performance of our non-incentivized WTP elicitation, we examine WTP for practice tests, which is a good we asked the sample about in both an incentivized and non-incentivized manner. In the first round, the WTP elicitation for tests was incentivized and in the second round it was not. In addition, we asked several survey questions in both rounds that might predict demand for tests, such as perceived quality of the tests and spending on educational inputs, which we can use as potential predictors. We also use household and child characteristics as potential predictors.

Using the households surveyed in both rounds, we first use LASSO to identify the primary predictors of non-incentivized WTP and of incentivized WTP. We then use OLS to test for differences in the relationships between non-incentivized WTP and incentivized WTP and their primary predictors. To be able to conduct valid inference in the second step, we randomly split our sample in half and use one half to fit the predictive LASSO model and the second half for OLS inference on the predictive coefficients.

The evidence suggests that non-incentivized WTP performs well. We find that LASSO identifies the same predictors for non-incentivized and incentivized WTP, and that the predictive coefficients are similar. In addition, with OLS, we are unable to reject equality in the predictive relationships. We first show the LASSO results and then the OLS.

Table B.1 displays the coefficients from using LASSO in the first half of the sample to identify the two most informative predictors of incentivized WTP (column 1) and non-incentivized WTP (column 2). Notably, the table shows that LASSO selects the same primary predictors for both incentivized and non-incentivized WTP. In addition, the predictors it chooses are both intuitive and sensible: WTP for the adult good, and an indicator for the parent thinking that the tests are higher-quality than the tests offered by the child's school. The fact that LASSO picks the same predictors for both incentivized and non-incentivized WTP – and with similar predictive coefficients – is evidence that non-incentivized WTP performed well.

Appendix Table B.1—LASSO chooses the same predictors of incentivized WTP and non-incentivized WTP for tests

Variable	Dep. var.: WTP for tests	
	Incentivized	Non-incentivized
	(1)	(2)
Number of children	.	.
Number of cattle	.	.
Number of motos	.	.
Number of rooms	.	.
Owns land	.	.
Polygamous	.	.
Assets PCA	.	.
Adult Good WTP	0.420	0.445
Believes child is very likely to attend school	.	.
Expect child to finish primary	.	.
Would spend more on child's education than other parent	.	.
Would spend less on child's education than other parent	.	.
Tests more useful than classes	.	.
Tests more useful than workbook	.	.
Believes tests are higher quality than those school offers	0.117	0.082
Food fees from school	.	.
Uniform fees from school	.	.
Textbook fees from school	.	.
Spending on non-school books	.	.
Spending on extra lessons/coaching	.	.
Spending on education outside school	.	.
Supplemental expenses	.	.
Total spending on education	.	.
Total spending on education (log)	.	.
Child is male	.	.
Child age	.	.
Has younger focal child	.	.
Younger focal child male	.	.
Younger focal child age	.	.
Number of observations	364	364

Notes: Columns show coefficients from LASSO regressions that regress the WTP for tests on all of the predictors listed in the rows. Column (1) uses incentivized WTP for tests f as the dependent variable, and column (2) uses non-incentivized WTP for tests as the dependent variable. "." means that the LASSO coefficient is 0.

Next, we use the other half of the sample to conduct statistical inference on whether the relationship between WTP for tests and the two primary predictors chosen by LASSO differs for incentivized and non-incentivized WTP. We regress WTP for tests, pooled across incentivized and non-incentivized observations, on the two primary predictors and their interactions with whether a WTP observation was gathered in an incentivized or non-incentivized manner.[1] Table B.2 shows the results.

Reassuringly, neither interaction term is significant. Note that, while there is no significant difference in the predictors of WTP between incentivized and non-incentivized WTP, average incentivized WTP for the tests is lower than average non-incentivized WTP. However, this marginally significant effect is difficult to interpret: the incentivized and non-incentivized WTP were collected in different time periods (i.e., different survey waves, which occurred during different school years), and so we cannot distinguish whether this negative main effect simply reflects a time effect.

[1]We estimate the following regression:

(A1)

$$\begin{aligned} WTP^{test}_{ihc} =& \alpha + \beta AdultGoodWTP_{ihc} + \gamma AdultGoodWTP_{ihc} \times NonIncentivized_{ihc} + \\ & \lambda HiQuality_{ihc} + \delta HighQuality_{ihc} \times NonIncentivized_{ihc} + \nu NonIncentivized_{ihc} + \varepsilon_{ihc} \end{aligned}$$

where WTP^{test}_{ihc} is parent i in household h's WTP, either incentivized or non-incentivized, for tests for child c; $AdultGoodWTP_{ihc}$ is that same parent's WTP for the adult good; $HiQuality_{ihc}$ is an indicator for the parent thinking the tests were higher quality than the school's regular offering; and $NonIncentivized_{ihc}$ is an indicator that WTP was gathered in a non-incentivized way.

Appendix Table B.2—No significant difference in predictive coefficients for incentivized WTP and non-incentivized WTP

	Tests WTP (1)
Adult good WTP	0.502 (0.049)
Adult good WTP $\times$ Non-incentivized	0.087 (0.073)
Believes tests high quality	0.158 (0.095)
Believes tests high quality $\times$ Non-incentivized	-0.036 (0.132)
Non-incentivized	0.188 (0.148)
Number of observations	730

Notes: The dependent variable is the WTP for tests, pooled across incentivized and non-incentivized elicitations. "Believes test high quality" is an indicator for the respondent thinking the tests are higher quality than those offered by their child's school. The regression additionally controls for missing flags for adult good WTP and high-quality tests, and missing flags interacted with the non-incentivized binary. Standard errors clustered at household level.

Summary statistics and balance tests

This section describes several tests to confirm balance. For 729 households, we have surveys of both the mother and father. For 355 additional households that we did not revisit in the second round, we have data for one randomly-chosen parent. To verify that the randomization yielded balance, Appendix Table A.2 conducts an omnibus balance test between the mother and father subsamples. We fail to reject the null of joint orthogonality of all variables (p-value 0.50). Following **?**, we also calculate the difference between groups divided by the pooled standard deviation. These standardized differences are all far below the rule-of-thumb "cutoff" for good balance of 0.25 SD. Table C.1 shows that, in addition, within the subsamples of male focal children and of female focal children, mothers and fathers have balanced characteristics.

While we randomized the gender of the parent within each household, we did not randomize child gender. Reassuringly, Table C.2 shows that household and parent characteristics are nevertheless similar between the girl and boy subsamples. An omnibus balance test fails to reject the null that they are identical, and all standardized differences between the two samples are far below 0.25 SD. Appendix Table C.3 shows that there is also boy-girl balance within the subsamples of mothers and of fathers.

A parent's gender is bundled with other individual characteristics, such as earnings, and child gender is similarly bundled with other traits. Tables C.4 and C.5 summarize the personal characteristics of mothers and fathers, and of daughters and sons, respectively. Mothers are younger and have less income than fathers, on average. In contrast, daughters and sons have similar characteristics, such as age and school performance.[2]

[2]We have relatively more older focal children among girls (standardized difference of 0.185), likely due to sampling error. We show that all of our results are robust to controlling for whether the child is the younger focal child in parallel to how child gender enters the regressions.

Appendix Table C.1—Household and child characteristics are balanced across mothers and fathers within the daughter sample and within the son sample

Variable	Daughters			Sons		
	Mothers (1)	Fathers (2)	Std. diff (3)	Mothers (4)	Fathers (5)	Std. diff (6)
Panel A: Household characteristics						
Number of children	9.031 [3.058]	8.979 [2.917]	0.017	9.239 [3.050]	9.141 [2.963]	0.033
Number of cattle	0.937 [1.274]	0.995 [1.304]	-0.044	1.044 [1.352]	1.080 [1.368]	-0.027
Number of motos	0.040 [0.195]	0.035 [0.185]	0.026	0.043 [0.202]	0.043 [0.202]	0.000
Number of rooms	3.015 [1.164]	3.055 [1.160]	-0.034	2.941 [1.185]	3.021 [1.206]	-0.068
Owns land	0.912 [0.284]	0.938 [0.242]	-0.099	0.920 [0.272]	0.932 [0.252]	-0.046
Polygamous	0.242 [0.432]	0.247 [0.435]	-0.011	0.249 [0.436]	0.253 [0.438]	-0.009
Panel B: Focal child characteristics						
Older focal child male	0.820 [0.385]	0.821 [0.383]	-0.002	0.260 [0.439]	0.259 [0.438]	0.002
Older focal child age	11.827 [1.906]	11.918 [2.043]	-0.046	12.058 [1.982]	12.108 [2.003]	-0.025
Believes older focal child will support parents more than other children	0.475 [0.500]	0.497 [0.500]	-0.044	0.492 [0.500]	0.486 [0.500]	0.012
Has younger focal child	0.861 [0.346]	0.880 [0.325]	-0.059	0.895 [0.307]	0.893 [0.309]	0.006
Younger focal child male	0.335 [0.434]	0.327 [0.437]	0.017	0.771 [0.388]	0.775 [0.385]	-0.009
Younger focal child age	5.698 [1.632]	5.807 [1.713]	-0.064	5.774 [1.708]	5.756 [1.721]	0.011
Younger focal child in school	0.637 [0.410]	0.687 [0.395]	-0.122	0.625 [0.424]	0.650 [0.412]	-0.061
Younger focal child grade	1.501 [0.690]	1.545 [0.854]	-0.060	1.503 [0.733]	1.507 [0.657]	-0.005
Number of observations	805	817		799	823	
Joint p-value	0.576			0.968		

Notes: In the regression to test for joint orthogonality, we impute missing values with the sample mean and include missing flags. We also control for survey round and strata fixed effects, to match the main specification. Standard errors are clustered at the household level. The unit of observation is a household-parent.

Appendix Table C.2—Household and parent characteristics are balanced across daughter and son samples

Variable	Daughters (1)	Sons (2)	Standardized diff (3)
Panel A: Household characteristics			
Number of children	9.005	9.189	-0.061
	[2.987]	[3.005]	
Number of cattle	0.966	1.062	-0.072
	[1.289]	[1.360]	
Number of motos	0.038	0.043	-0.026
	[0.190]	[0.202]	
Number of rooms	3.035	2.982	0.045
	[1.162]	[1.196]	
Owns land	0.925	0.926	-0.004
	[0.264]	[0.262]	
Polygamous	0.245	0.251	-0.014
	[0.433]	[0.437]	
Panel B: Parent characteristics			
Parent age	39.426	40.076	-0.026
	[24.740]	[24.583]	
Has some education	0.813	0.799	0.035
	[0.390]	[0.400]	
Income (10000s UGX)	56.980	55.853	0.009
	[126.025]	[124.525]	
Number of observations	1622	1622	
Joint p-value	0.284		

Notes: In the regression to test for joint orthogonality, we impute missing values with the sample mean and include missing flags. We also control for survey round and strata fixed effects, to match the main specification. Standard errors are clustered at the household level. The unit of observation is a household-parent-focal child.

Appendix Table C.3—Household characteristics are balanced across daughters and sons within the mother sample and within the father sample

Variable	Mothers			Fathers		
	Daughters (1)	Sons (2)	Std. diff (3)	Daughters (4)	Sons (5)	Std. diff (6)
Panel A: Household characteristics						
Number of children	9.031 [3.058]	9.239 [3.050]	-0.069	8.979 [2.917]	9.141 [2.963]	-0.054
Number of cattle	0.937 [1.274]	1.044 [1.352]	-0.081	0.995 [1.304]	1.080 [1.368]	-0.064
Number of motos	0.040 [0.195]	0.043 [0.202]	-0.015	0.035 [0.185]	0.043 [0.202]	-0.041
Number of rooms	3.015 [1.164]	2.941 [1.185]	0.063	3.055 [1.160]	3.021 [1.206]	0.029
Owns land	0.912 [0.284]	0.920 [0.272]	-0.030	0.938 [0.242]	0.932 [0.252]	0.023
Polygamous	0.242 [0.432]	0.249 [0.436]	-0.016	0.247 [0.435]	0.253 [0.438]	-0.014
Panel B: Parent characteristics						
Age	34.900 [33.229]	35.671 [33.383]	-0.031	43.886 [9.374]	44.352 [8.526]	-0.019
Has some education	0.763 [0.426]	0.731 [0.444]	0.081	0.863 [0.344]	0.865 [0.340]	-0.005
Income (10000s UGX)	27.776 [85.183]	28.273 [82.761]	-0.004	85.755 [150.766]	82.629 [149.908]	0.025
Number of observations	805	799		817	823	
Joint p-value	0.242			0.716		

Notes: In the regression to test for joint orthogonality, we impute missing values with the sample mean and include missing flags. We also control for survey round and strata fixed effects, to match the main specification. Standard errors are clustered at the household level. The unit of observation is a household-parent-focal child.

Appendix Table C.4—Mothers and fathers have different characteristics

Variable	Mothers (1)	Fathers (2)	Standardized diff (3)
Age	35.784 [31.608]	44.650 [9.258]	-0.375
Has some education	0.743 [0.437]	0.865 [0.342]	-0.307
Income (10000s UGX)	29.536 [89.637]	88.185 [155.526]	-0.450
Number of observations	899	913	

Notes: The unit of observation is a household-parent.

Appendix Table C.5—Summary statistics: Boys and girls have similar characteristics

Variable	Daughters (1)	Sons (2)	Standardized diff (3)
Panel A: All focal children			
Older focal child	0.605	0.513	0.185
	[0.489]	[0.500]	
Panel B: Older focal children			
Child age	11.873	12.083	-0.106
	[1.980]	[1.994]	
Grade	4.655	4.620	0.046
	[0.773]	[0.733]	
Weekly study hours	5.226	4.960	0.039
	[7.422]	[6.073]	
School performance	3.208	3.189	0.020
	[0.948]	[0.944]	
Will support parents more than other children	0.486	0.489	-0.006
	[0.500]	[0.500]	
Panel C: Younger focal children			
Child age	5.752	5.766	-0.008
	[1.794]	[1.813]	
In school	0.667	0.634	0.069
	[0.472]	[0.482]	
Grade	1.533	1.495	0.036
	[1.110]	[1.004]	
Number of observations	1622	1622	

Notes: The unit of observation is a household-parent-focal child.

Interpreting the magnitudes of the estimates

In this subsection, we set up a simple framework to elucidate how the gender gaps in WTP that we estimate map to gender gaps in purchases and spending. Let each individual i have an underlying willingness pay parameter wtp_i. We assume that a person's willingness to pay for a specific good, wtp_{ig}, is wtp_i scaled by a good-specific constant, β_g. That is, $wtp_{ig} \equiv \beta_g wtp_i$.[3] We assume that wtp_i is drawn from a bounded distribution with a minimum of 0 and a standard deviation of 1. (Note that all of our wtp_{ig} observations in practice already have a minimum of 0.)

With these assumptions, to recover the latent distribution of wtp_i from our data on wtp_{ig}, we normalize wtp_{ig} by its standard deviation in our sample. When we do this for each good, we can stack them in a regression, having essentially partialled out the good-specific β_g parameters.

We want to use our empirical estimate of the wtp distribution to compare how much different subpopulations (e.g., fathers with daughters, or mothers with sons) would purchase of some good g'. The quantity purchased and expenditures depend on the market price of the good. We outline two approaches below: one that assesses purchases and expenditures under a range of potential prices that are distributed uniformly, and one that assesses them under specific assumed price points. We first describe the approaches and then present the results.

Below, without loss of generality, we assume that $\beta_{g'} = 1$, so $wtp_{ig'} = wtp_i$.[4] Hence, we drop the g' subscript for simplicity going forward.

[3]For simplicity, we are ignoring the error term. More precisely, $wtp_{ig} \equiv \beta_g wtp_i + \varepsilon_{ig}$, where ε_g has mean 0, is independent of wtp, and has a standard deviation that is proportional to β_g (so that the distribution of wtp has the same coefficient of variation for all goods). Note that the addition of an error term means that wtp_{ig} can in practice be negative. We assume that this will be rare in practice and ignore going forward for simplicity.

[4]This is without loss of generality since assuming a different $\beta_{g'}$ is isomorphic to assuming a different price distribution.

Purchases and expenditures under uniform prices.

We assume that the potential market prices for good g', p, are realizations of a random variable P that is distributed uniformly between 0 and $\bar{P}$, with $\bar{P}$ greater than or equal to the maximum value of WTP.

Consider multiple subpopulations s that we want to compare. Let wtp_{is} represent willingness to pay of member i of subpopulation s for good g', and let $x_{is}(p) = 1\{p < wtp_{is}\}$ be an indicator for whether that person would purchase the good at price realization p. Note that, by construction, $wtp_{is} \leq \bar{P}$.

The expected proportion of price realizations for which individual i from subpopulation s would purchase the good is then $x_{is} \equiv E_P[x_{is}(p)] = wtp_{is}/\bar{P}$.[5] That is, they purchase the good if the price is between 0 and their willingness to pay, and this occurs with probability $wtp_{is}/\bar{P}$ given the uniform distribution. The expected share of the subpopulation who would purchase the good (averaged across potential price realizations) is $E_X[x_{is}] = E_{WTP}[wtp_{is}]/\bar{P}$, where the expectations are taken across all individuals i in subpopulation s.

The ratio of expected purchases between two subpopulations, $s = m, f$, is: $\frac{E_{X_m}[x_{im}]}{E_{X_f}[x_{jf}]} = \frac{E_{WTP_m}[wtp_{im}]/\bar{P}}{E_{WTP_f}[wtp_{jf}]/\bar{P}} = \frac{E_{WTP_m}[wtp_{im}]}{E_{WTP_f}[wtp_{jf}]}$. Thus, under these assumptions, the ratio of average standardized WTP for two subpopulations gives us an estimate of the ratio of goods purchased between those groups. Hence, in our regression analysis, the point estimates for WTP (in a specification where WTP has been standardized across the population by the good-specific standard deviation) can be divided by the mean in the reference group to obtain an estimate of the the percent difference in expected purchases across those two subpopulations (e.g., fathers with sons versus fathers with

[5]To see this more rigorously: $E_P[x_{is}(p)] = E[1\{p < wtp_{is}\}] = \int_0^{\bar{P}} 1\{p < wtp_{is}\}/\bar{P}dp = \int_0^{wtp_{is}} dp/\bar{P} = wtp_{is}/\bar{P}$.

daughters).

The above discussion regards the number of goods purchased, not expenditures. To estimate expenditures, we have $E_P[x_{is}(p)p] = E_P[1\{p < wtp_{is}\}p] = \int_0^{\bar{P}} 1\{p < wtp_{is}\}p/\bar{P}dp = \int_0^{wtp_{is}} p/\bar{P}dp = wtp_{is}^2/(2\bar{P})$. The ratio of expected expenditures between two subpopulations, $s = m, f$, is: $\frac{E_{X_m}[x_{im}(p)p]}{E_{X_f}[x_{jf}(p)p]} = \frac{E_{WTP_m}[wtp_{im}^2]/2\bar{P}}{E_{WTP_f}[wtp_{jf}^2]/2\bar{P}} = \frac{E_{WTP_m}[wtp_{im}^2]}{E_{WTP_f}[wtp_{jf}^2]}$. Hence to estimate differences across populations in expenditures, we use standardized willingness to pay squared as the dependent variable in the regression.

Purchases and expenditures for specific price points.

An alternate way to think about purchases and expenditures is to imagine that the potential good g' we are considering has a specific price point, p^*, where we could define p^* as, say, the 20th percentile of the wtp distribution or the 60th percentile. In that case, purchases become $x_{is}(p^*) = 1\{wtp_{is} > p^*\}$ and expenditures become $x_{is}(p^*)p^* = 1\{wtp_{is} > p^*\}p^*$. As a result, we have $\frac{E_{X_m}[x_{im}(p^*)]}{E_{X_f}[x_{jf}(p^*)]} = \frac{E_{WTP_m}[1\{wtp_{im}>p^*\}]}{E_{WTP_f}[1\{wtp_{jf}>p^*\}]}$ for demand and $\frac{E_{X_m}[x_{im}(p^*)p^*]}{E_{X_f}[x_{jf}(p^*)p^*]} = \frac{E_{WTP_m}[1\{wtp_{im}>p^*\}]}{E_{WTP_f}[1\{wtp_{jf}>p^*\}]}$ for expenditures. That is, the estimate for the difference in both purchases and expenditures across two subpopulations for a given price level p^* is the ratio across the subpopulations of the average of $1\{wtp_{is} > p^*\}$. Hence to estimate differences between subpopulations in purchases and expenditures for fixed price points, we use $1\{wtp_{is} > p^*\}$ as the dependent variable in the regression. To explore a range of potential prices, we set p^* at the 20th, 40th, 60th, and 80th percentile of the wtp distribution.

Results

Table D.1 shows the results of estimating equation 2 using the following as dependent variables: standardized WTP (column 1; this is our main specification, repeated for reference), standardized WTP squared (column 2), and indicators for whether standardized WTP is above the 20th, 40th, 60th, and 80th percentiles of the distribution in our sample (columns 3-6). Dividing the *Daughter* coefficients in each of these regressions by the dependent variable mean for fathers and sons (shown in a bottom row of the table) quantifies how much lower, in percent terms, father's demand (columns 1 and 2-6) and/or spending (columns 2 - 6) is for their daughters than their sons. The different columns assume either uniformly distributed potential prices (columns 1 and 2) or specific price points (columns 3-6). Although we lose some power in the percentile specifications in columns (3) - (6) due to not using all of the underlying variation in the data, the high-level take-away is similar across all columns. Across all specifications, the *Daughter* effect is meaningful in percent terms, with a median of 6%. The estimates from columns (1) and (2), which capture percent changes in demand and spending, respectively, under uniform prices are 5% and 8%. The estimates in columns (3) through (6), which allow us to calculate percent changes in both demand and spending at specific price points, are 3%, 5%, 6%, and 16%. (The observed increase in the percent effect across price points is almost mechanical, as a lower price point corresponds to a higher base level of spending.) Similarly, we can normalize the $Daughter \times Mother$ coefficient by the dependent variable mean for fathers and sons to understand how much smaller the *Daughter* effect is for mothers than fathers, in percentage terms. The estimates have a median of 7% and range across columns 1-6 from 5% (column 5) to 13% (column 6).

Appendix Table D.1—Results across the potential price distribution

	WTP specification					
	Standardized (1)	Standardized squared (2)	$\geq$ 20th percentile (3)	$\geq$ 40th percentile (4)	$\geq$ 60th percentile (5)	$\geq$ 80th percentile (6)
Daughter	-0.102	-0.382	-0.027	-0.034	-0.024	-0.034
	(0.032)	(0.132)	(0.013)	(0.016)	(0.015)	(0.013)
Mother $\times$ Daughter	0.131	0.372	0.054	0.062	0.020	0.029
	(0.046)	(0.187)	(0.019)	(0.022)	(0.021)	(0.018)
Mother	-0.095	-0.140	-0.056	-0.053	-0.014	-0.001
	(0.036)	(0.148)	(0.015)	(0.017)	(0.016)	(0.014)
p-val: Mother + Mother $\times$ Daughter = 0	0.318	0.098	0.875	0.610	0.691	0.023
p-val: Daughter + Mother $\times$ Daughter = 0	0.399	0.942	0.064	0.087	0.786	0.722
Dep. var. mean father-son	1.943	5.088	0.834	0.638	0.423	0.219
Number of observations	6,673	6,673	6,673	6,673	6,673	6,673

Notes: The dependent variable in the first column is standardized WTP (i.e., WTP normalized by the good-level standard deviation), as in column (2) of Table 1. The dependent variable in the second column is standardized WTP squared. The dependent variable in columns (3), (4), (5), and (6) are indicators that standardized WTP is at least as large as the 20th, 40th, 60th, and 80th percentiles, respectively, of the standardized WTP distribution for the goods included in the regression.

Robustness of main results

Since daughters are marginally more likely than sons to be the older focal child, Appendix Table E.1 shows that the results are robust to controlling for an indicator that the child is the younger focal child in parallel to how *Daughter* enters the regression. In addition, Appendix Table E.2 shows that our results are robust to excluding the control for the respondent's WTP for the adult good; the main change is that the standard errors are about 30% larger when we omit this control variable.

Appendix Table E.3 repeats these robustness checks for the results corresponding to the two panels of Figure 3.

Appendix Table E.1—Robustness of Table 1 to controlling for *Mother* × *Younger child good*

	WTP normalized by...					
	SD	SD	Market price	SD	SD	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Daughter	-0.037	-0.104	-0.029	-0.099	-0.065	-0.164
	(0.024)	(0.032)	(0.009)	(0.037)	(0.036)	(0.052)
Mother × Daughter		0.135	0.037	0.150	0.067	0.218
		(0.047)	(0.013)	(0.053)	(0.053)	(0.076)
Mother	-0.029	-0.115	-0.032	-0.101	-0.083	-0.144
	(0.028)	(0.039)	(0.012)	(0.046)	(0.043)	(0.067)
p-val: Mother + Mother × Daughter = 0		0.588	0.661	0.259	0.707	0.224
p-val: Daughter + Mother × Daughter = 0		0.367	0.428	0.177	0.950	0.308
Dep. var. mean father-son	1.943	1.943	0.537	1.943	1.793	2.164
Fixed effects	Stratum	Stratum	Stratum	HH	Stratum	Stratum
Goods included	All	All	All	All	Incentivized	Non-incentivized
Number of observations	6,673	6,673	6,673	6,673	4,000	2,673

Notes: All columns control for survey round, adult WTP, and adult WTP interacted with survey round. Columns 1-3 control for strata and good fixed effects; Column 4 controls for household and good fixed effects. Standard errors are clustered by household.

Appendix Table E.2—Robustness of Table 1 to omitting control for adult good WTP

	WTP normalized by...					
	SD	SD	Market price	SD	SD	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Daughter	-0.064	-0.111	-0.029	-0.077	-0.082	-0.163
	(0.033)	(0.045)	(0.009)	(0.039)	(0.052)	(0.057)
Mother × Daughter		0.095	0.036	0.113	0.039	0.179
		(0.058)	(0.013)	(0.059)	(0.069)	(0.079)
Mother	-0.041	-0.089	-0.028	-0.053	-0.094	-0.081
	(0.036)	(0.047)	(0.010)	(0.049)	(0.056)	(0.062)
p-val: Mother + Mother × Daughter = 0		0.887	0.420	0.216	0.317	0.092
p-val: Daughter + Mother × Daughter = 0		0.705	0.454	0.374	0.413	0.781
Dep. var. mean father-son	1.943	1.943	0.537	1.943	1.793	2.164
Fixed effects	Stratum	Stratum	Stratum	HH	Stratum	Stratum
Goods included	All	All	All	All	Incentivized	Non-incentivized
Number of observations	6,673	6,673	6,673	6,673	4,000	2,673

Notes: All columns control for survey round, adult WTP, and adult WTP interacted with survey round. Columns 1-3 control for strata and good fixed effects; Column 4 controls for household and good fixed effects. Standard errors are clustered by household.

Appendix Table E.3—Robustness checks for Figure 3

	Human capital goods:				Enjoyment goods:		
WTP normalized by	Market price (1)	SD (2)	SD (3)	SD (4)	Market price (5)	SD (6)	SD (7)
Daughter	-0.029	-0.110	-0.120	-0.095	-0.028	-0.087	-0.087
	(0.010)	(0.037)	(0.045)	(0.045)	(0.015)	(0.051)	(0.072)
Mother × Daughter	0.039	0.152	0.114	0.149	0.023	0.073	0.024
	(0.014)	(0.051)	(0.059)	(0.061)	(0.024)	(0.081)	(0.103)
Mother	-0.017	-0.098	-0.054	-0.043	-0.070	-0.199	-0.218
	(0.011)	(0.043)	(0.046)	(0.048)	(0.018)	(0.073)	(0.085)
Extra control variables		Mother × YoungerChild		Household FEs		Mother × YoungerChild	
Excluded control variables			Adult good WTP				Adult good WTP
p-val: Mother + Mother × Daughter = 0	0.038	0.183	0.184	0.020	0.017	0.081	0.022
p-val: Daughter + Mother × Daughter = 0	0.356	0.269	0.884	0.246	0.812	0.831	0.403
Dep. var. mean father-son	0.543	1.996	1.996	1.996	0.518	1.754	1.754
Number of observations	5,215	5,215	5,215	5,215	1,458	1,458	1,458

Notes: All columns control for strata and good fixed effects, and survey round. All columns except columns 3 and 7, also control for adult WTP and adult WTP interacted with survey round. Standard errors are clustered by household.