Big Sisters

Pamela Jakiela, Owen Ozier, Lia Fernald, and Heather Knauer*

June 29, 2020

Abstract

We model household investments in young children when parents and older siblings both contribute to human capital accumulation. The empirical pattern we observe is consistent with a model where: (i) older siblings' investments in young children contribute to their human capital accumulation, (ii) both parents and older children perceive lower returns to investing in older girls than in older boys. We estimate the impact of having an older sister (rather than an older brother) on early childhood development in a sample of rural Kenyan households with otherwise similar family structures. Older sibling gender is not related to household structure, subsequent birth spacing, or other observable characteristics, so we treat the presence of an older girl (as opposed to an older boy) as plausibly exogenous. Having an older sister improves younger siblings' vocabulary and fine motor skills by more than 0.1 standard deviations.

JEL codes: O12, J13, J16, D13

Keywords: sisters, girls, girl effect, girl power, Family Care Indicators, early childhood, human capital, household structure, parental investments, natural experiment

*Jakiela: Center for Global Development, BREAD, and IZA, email: pjakiela@cgdev.org; Ozier: World Bank Development Research Group, BREAD, and IZA, email: oozier@worldbank.org; Fernald: University of California at Berkeley, email: fernald@berkeley.edu; Knauer: University of Michigan, email: heather.knauer@gmail.com. Orazio Attanasio, Maya Eden, Dave Evans, Deon Filmer, Joe Henrich, Amanda Glassman, Vikram Maheshri, Ania Ozier, Justin Sandefur, and Dana Schmidt provided helpful comments. We are grateful to William Blackmon, Julian Duggan, Laura Kincaide, and Elyse Thulin for excellent research assistance. This research was funded by Echidna Giving and by the Strategic Impact Evaluation Fund and the Early Learning Partnership at the World Bank. All errors are our own.

1 Introduction

Investments in early childhood are a critical determinant of later life outcomes, and stimulating activities — for example, shared reading and infant-directed speech — are an important way that older family members invest in young children (Knudsen, Heckman, Cameron, and Shonkoff 2006, Grantham-McGregor *et al.* 2007, Almond and Currie 2011, Walker *et al.* 2011, Aizer and Cunha 2012). Underinvestment in early childhood is an acute problem in low- and middle-income countries (LMICs), where an estimated 43 percent of children are at risk of failing to meet their developmental potential because of inadequate nutrition and cognitive stimulation (Black *et al.* 2017). A growing interdisciplinary literature examines the causes and consequences of parental investments in young children in LMICs, seeking to identify interventions that can change parenting practices to improve developmental outcomes in children and increase incomes in adulthood (cf. Gertler *et al.* 2014, Black *et al.* 2017, Andrew *et al.* 2018). However, parents are not the only caregivers in most societies — in many low-income contexts, much of that work is done by older siblings, particularly sisters (Weisner and Gallimore 1977, Lancy 2015). Though this pattern is well-documented in the anthropology literature, older siblings' role in childrearing is often ignored in academic and policy discussions of investments in early childhood.[1]

We model older siblings' contributions to the human capital accumulation of young children. Our model extends existing work in economics by incorporating several insights from anthropology and psychology. First, older children do much of the childcare in many LMIC settings, and the quality of their caregiving practices impacts the human capital accumulation of their younger siblings (Weisner and Gallimore 1977, Maynard 2002, Maynard

[1]The Family Care Indicators, one of the most widely used measures of early childhood stimulation, does not record stimulating activities carried out by older children and adolescents (Hamadani *et al.* 2010, Kariger *et al.* 2012). A recent systematic review of 466 impact evaluations of early childhood development interventions in LMICs found that only four measured indirect effects on older siblings in middle childhood or adolescence (Evans, Knauer, and Jakiela 2020). This tendency to ignore the caretaking role of older siblings is at least partially attributable to the widespread perception that parental stimulation is more beneficial to young children than stimulation by older children (though it may also entail a higher opportunity cost). In a published response to Weisner and Gallimore (1977), Brian Sutton-Smith argued that "Maximal personal and social development of infants is produced by the mother (or caretaker) who interacts with them in a variety of stimulating and playful ways. Unfortunately the intelligence to do this with ever more exciting contingencies is simply not present in child caretakers." (Weisner and Gallimore, 1977, p. 184).

and Tovote 2010, Lancy 2015). Second, older siblings involved in caregiving make active tradeoffs, deciding how much to invest in their own human capital and how much to invest in their younger siblings (Brody 2004, Bock 2010). Third, even when older children are less effective caretakers than parents (Ellis and Rogoff 1982), it may be optimal to delegate some childrearing to older siblings when the opportunity cost of their time is low relative to adults (Chick 2010, Lancy and Grove 2010). We extend a simple model of parental investments in children to consider the direct contributions of older siblings and the tradeoffs that arise when siblings and parents share caregiving responsibilities in this way. We show that parental investments in the youngest household members cannot be interpreted as measures of parental preferences in such settings. Because parents' and siblings' investments are substitutes, having an older child who is an effective caregiver allows parents to increase their labor supply and substitute to older children without compromising young children's human capital.

In most societies where older children play a substantial caretaking role, older sisters do more childcare than older brothers (Weisner and Gallimore 1977, Lancy 2015). Our model demonstrates that this can occur because older sisters are more effective caregivers than older brothers, or because households perceive a lower return to investing in the human capital of older girls than older boys. In either case, young children with an older sister (rather than an older brother) are likely to benefit, receiving more cognitive stimulation as a result. However, a treatment effect of having an older sister can also arise in a standard model, where older children do not contribute to human capital accumulation in their younger siblings. If older siblings' investments in young children's human capital are not productive, any treatment effect of having an older sister must be driven by parental investments: parents would invest more in young children when they have an older girl rather than an older boy because they have less incentive to invest in their older child. In contrast, when older siblings contribute to the development of young children's human capital, parents with an older girl may actually invest *less* time in their young children — because they have an effective substitute. Our model illustrates that any treatment effect of older sisters (on child development) that is not explained by differential investments by parents can be attributed to — and provides evidence of — the contributions of older siblings.

To test the empirical predictions of our model, we estimate the impact of older sisters on early childhood development in a sample of rural Kenyan households that have one or two young children aged three to six and exactly one older sibling aged seven to 14.[2] In this sample, we show that the gender of the older sibling is unrelated to household or community characteristics — and hence plausibly exogenous.[3] We find that young children with one older sister experience significantly more cognitive stimulation than those with one older brother. This pattern results from increased stimulation by older sisters, not by parents. Our model suggests that this empirical pattern will arise when both parents and older siblings perceive a gender gap in the return to investing in older children's human capital, and parents know that stimulating activities with older siblings increase the youngest children's human capital.

Differential patterns of household investment translate into meaningful impacts on child development. An aggregate index of language and motor development is 0.1 standard deviations higher when a young child's older sibling is a sister and not a brother. The magnitude of this difference is commensurate with that between children of primary-educated and secondary-educated mothers. Impacts on fine motor skill development are concentrated in the bottom half of the distribution, but impacts on language skills are not. Our results suggest that older siblings play an important role in shaping younger children's human

[2]Reviewing ethnographic evidence from 50 traditional societies, Rogoff *et al.* (1975) report that the typical age at which societies begin assigning older children childcare responsibilities is between five and seven years old. Weisner and Gallimore (1977) report that (Kenyan) Luhya girls aged 6–8 years old spent 60 percent of their waking hours looking after younger children, though this caretaking was often under the explicit or implicit supervision of nearby adults. Similarly-aged boys (6–8 years old) and younger girls (aged 3–5) spent about half as much time caring for small children. Apoko (1967) and Lijembe (1967) also relate how in neighboring Acholi communities (where the role is called the *lapidi*) as well as in Idakho communities, young girls are usually tasked with caring for infants; if there is no appropriately-aged older sister, the task may fall to an older brother.

[3]An extensive literature treats the sex composition of children as a source of exogenous variation (cf. Angrist and Evans 1998, Washington 2008, Glynn and Sen 2015). However, the assumptions required for such estimates to identify causal impacts are unlikely to hold in general (Bisbee, Dehejia, Pop-Eleches and Samii 2017, Clarke 2018). In the Unites States, Dahl and Moretti (2008) show that having a firstborn daughter increases the likelihood of parental separation. In India, existing evidence suggests that son preference influences birth spacing and total fertility (Clark 2000, Jayachandran and Kuziemko 2011), so households with a firstborn son may not be comparable to households with a firstborn daughter.

capital in this context, and that optimizing households are well aware of this fact.

Though economic models of investments in children typically focus on parents' investment decisions, several recent papers have highlighted the important role played by older sisters in LMIC settings. In Turkey, Alsan (2017) shows that a nationwide vaccination campaign targeting young children improved literacy and educational attainment among adolescent girls — who are often forced to stay home tending sick younger siblings. In Mozambique, Martinez, Naudeau and Pereira (2017) find that the construction of new community-based preschools increased the likelihood that older children had ever attended school, decreased their childcare hours, and increased the amount of time spent on school work. In Pakistan, Qureshi (2018) demonstrates that increasing older girls' educational attainment has positive impacts on the literacy and numeracy of younger brothers. In Kenya, Ozier (2018) shows that infants and toddlers whose older siblings were exposed to a school-based deworming program saw improvements in cognitive development, and that gains were larger among children with more older sisters. In Brazil, Attanasio *et al.* (2019) find that access to publicly-provided daycare increased employment and income among older sisters (aged 15 and above). These papers highlight the special alloparenting role played by older girls, showing that it has empirical implications for both the girls themselves and their younger siblings.

Our work is related to several broader strands of literature. First, we contribute to the growing body of work on early childhood development in low-income settings (Heckman 2007, Almond and Currie 2011, Black *et al.* 2017), specifically research analyzing the human capital production function (cf. Cunha, Heckman and Schennach 2010) and recent evaluations of interventions intended to change parenting practices (cf. Chang *et al.* 2015, Weisleder *et al.* 2017, Hamadani *et al.* 2019, Attanasio *et al.* 2020, Doyle 2020). Second, our work relates to the wider literature on gender norms affecting children in LMICs (cf. Dhar, Jain and Jayachandran 2018), particularly work on factors constraining girls' education (cf. Kremer, Miguel and Thornton 2009, Baird, McIntosh and Özler 2011, Jensen 2012) and on the differential chore and carework responsibilities of male and female children (cf.

Edmonds 2006, Montgomery 2010).[4] Finally, our work builds on literature in demography, economics, political science, and sociology using quasi-experimental variation in the sex composition of children to identify the impact of daughters on parents' attitudes and behaviors (cf. Dahl and Moretti 2008, Washington 2008, Glynn and Sen 2015).

The rest of this paper is organized as follows. Section 2 presents our model of familial investments in young children and our empirical tests of the model. Section 3 describes our study setting and data set. Section 4 presents our empirical results, and Section 5 concludes.

2 Conceptual Framework

2.1 A Simple Model of Parental Investment

We first consider a standard conceptual framework in which stimulating activities performed by older siblings do not contribute to the human capital accumulation of the youngest family members. In this setting, only parents can intentionally invest in the human capital of young children, and any effect of older siblings is explained by changes in parental investment.

Consider a household comprising a parent, an older (school-aged) child, and a younger (not yet school-aged) child. Parental investments increase child ability, leading to higher incomes (or greater overall welfare) in adulthood. The parent divides their time between household production and investing in their two children. The household utility function is given by

$$U = y\left(L_p\right) + \tilde{h}_o(p_o) + \tilde{h}_y\left(p_y\right) \tag{1}$$

subject to the constraint

$$L_p = 1 - p_o - p_y. \tag{2}$$

$L_p \geq 0$ is the amount of time allocated to home production, and $y\left(\cdot\right)$ is a strictly concave

[4] Beyond the large literature on gender differences in schooling around the world (cf. Psaki, McCarthy and Mensch 2018, Evans, Akmal and Jakiela 2020a), there is of course a rich literature on gender differences in adult behavior (cf. Pitt, Rosenzweig and Hassan 2012, Alesina, Giuliano and Nunn 2013), but those topics are not the emphasis of the present paper.

production function satisfying Inada conditions. Let $k \in \{o, y\}$ index children within the household, indicating whether a child is the older or younger sibling. p_k is the parent's investment of time in the human capital of child k, and $\tilde{h}_k(\cdot)$ is a strictly concave human capital production function satisfying Inada conditions.[5] To model gender gaps, we let

$$\tilde{h}_k(p_k) = \lambda_k^z h_k(p_k) \tag{3}$$

for $k \in \{o, y\}$ and $z \in \{G, B\}$, where z indicates child k's gender.[6] Thus, age-specific human capital production functions are identical up to a parameter characterizing the relative returns by gender (as perceived by the parent).[7] Inada conditions guarantee an interior optimum characterized by the first-order condition:

$$y'(1 - p_o^* - p_y^*) = \lambda_o^z h_o'(p_o^*) = \lambda_y^z h_y'(p_y). \tag{4}$$

Three results are immediately apparent. First, a younger sibling's human capital only depends on the gender of the older sibling if there are gender differences in the human capital production function (as perceived by the parent): if $\lambda_o^G = \lambda_o^B$ and $\lambda_y^G = \lambda_y^B$, parental investments in human capital do not depend on the gender of either child. Second, if parents prefer boys — or, equivalently, if the returns to investments in human capital are systematically lower for girls than for boys at all ages — parents will invest less in girls and more in boys (conditional on child age). So, if $\lambda_o^G < \lambda_o^B$ and $\lambda_y^G < \lambda_y^B$, parents will invest

[5] Many models of human capital formation divide childhood into multiple periods (cf. Heckman 2007). We abstract from the intertemporal dynamics of investment in a particular child to focus on the intra-household process of building young children's human capital. An extension to our model would allow for consideration of dynamic effects in setting where older children contribute to the production of younger children's human capital.

[6] Because z always appears a superscript on a parameter that is also indexed by a subscript k, we omit the subscript k (on z) to simplify notation.

[7] In this framework, lower objective returns — for example, gender differences in the return to schooling — are equivalent to lower subjective parental valuation of (objective) returns. For example, in a patrilocal society, parents' private return to educating a daughter may be low because the return is captured by the girl's husband's family. Alternatively, parents who simply prefer boys might place more weight on their sons' future income and wellbeing (relative to daughters' future welfare). The utility weights λ_k^z reflect both objective and subjective factors influencing parents' perceptions of the return to investing in a child's human capital.

less in older girls than in older boys, they will invest less in younger girls than in younger boys, and — conditional on the gender of the younger child — they will invest more in young children with an older sister than in young children with an older brother. Finally, if $\lambda_o^G < \lambda_o^B$ and $\lambda_y^G = \lambda_y^B$, parental investments in both children depend on the gender of the older sibling: if the older sibling is a girl, parents will invest less in her and more in her younger sibling — irrespective of the gender of the younger child. Thus, we would expect to see a treatment effect of older sibling gender on the developmental outcomes of young children, and this effect would be driven by differences in parental investments in those children. If the marginal return to investing in older girls is relatively low (i.e. when one assumes that $\lambda_o^G < \lambda_o^B$ and $\lambda_y^G = \lambda_y^B$), parents with older girls have more time available to invest in their younger children. However, because $\lambda_y^G = \lambda_y^B$, parents would not invest more in young boys than in young girls (on average, holding the gender of the older sibling constant).

2.2 The Contributions of Older Siblings

We now extend the model to consider the contributions of older children in a framework that characterizes the active tradeoffs made by both parents and older siblings. Again, we consider a unitary household comprising a parent, an older child, and a younger child, but now we allow the actions of the older child to influence both their own human capital accumulation and the human capital of their younger sibling. Familial (rather than parental) investments in children increase child ability, leading to higher adult welfare. The parent divides their time between household production and investing in their two children, and the older child divides their time between schoolwork (i.e. investing in their own human capital) and engaging in stimulating activities with the younger sibling.

The younger child's human capital depends only on investments by older family members — since preschool-aged children do not make active choices (e.g. how hard to work in school) that increase human capital. The younger child's human capital production function is given by

$$\tilde{h}_y(I_y) = \lambda_y^z h_y(I_y) \tag{5}$$

where $h_y(\cdot)$ is assumed to be an increasing, concave function that satisfies Inada conditions. $I_y = p_y + \delta_o^z o_y$ where p_y is the parent's investment in the younger child, o_y is the older sibling's investment in her younger sibling, and $\delta_o^z < 1$ is a quality parameter indexing the productivity (with respect to human capital production) of investments made by an older sibling of gender z (relative to investments by adults in the household). Later, we will refer to I_y and its component $\delta_o^z o_y$ as *stimulation*, to distinguish them from the underlying *investment*, o_y. Thus, parents and older siblings are assumed to be perfect substitutes in the production of younger children's human capital — though parents may have an absolute advantage.[8]

Older children invest in their own human capital by exerting effort in school, and they also benefit from investments (in them) made by the adults in their household. The older sibling's human capital production function is given by

$$\begin{aligned}\tilde{h}_o\left(E_o, p_o\right) &= \lambda_y^z h_o\left(E_o, p_o\right) \\ &= \lambda_y^z\left[h_{o\to o}(E_o) + h_{p\to o}(p_o)\right]\end{aligned} \tag{6}$$

where E_o is the child's level of investment in their own human capital (e.g. in schoolwork) and p_o is the parent's investment in the older child. Thus, $h_o\left(E_o, p_o\right)$ is assumed to be additively separable in child and adult investments in human capital.[9] Both $h_{o\to o}(E_o)$ and $h_{p\to o}(p_o)$ are increasing and concave functions satisfying Inada conditions.

The parent divides one unit of time between household production and stimulating her

[8]When δ_o^G and δ_o^B are sufficiently small, investments made by the older sibling do not improve the younger siblings' human capital, and the model reduces to the version considered in Section 2.1 — as we discuss further below.

[9]Cases where child effort and parental investment are complements have an intuitive appeal — for example, if parents assist school-aged children with their homework. However, such complementaries allow for the possibility of multiple equilibria. For this reason, we focus our analysis on the determinants of investments in younger children and simplify the rest of the environment as much as possible.

children (as in Section 2.1). Household utility is given by:

$$U = y\left(L_p\right) + \lambda_o^z h_{o\to o}(E_o) + \lambda_o^z h_{p\to o}(p_o) + \lambda_y^z h_y\left(I_y\right) \tag{7}$$

where

$$L_p = 1 - p_o - p_y, \tag{8}$$

$$E_o = 1 - o_y, \tag{9}$$

and

$$I_y = p_y + \delta_o^z o_y. \tag{10}$$

We assume a unitary household at this point in our exposition, deferring relaxation of this assumption until Section 2.3. If an interior solution (p_o^*, p_y^*, o_y^*) exists, the following are true at the optimum: first, parents equate the marginal product of labor with the marginal product of additional time invested in each child by setting

$$y'(1 - p_o^* - p_y^*) = \lambda_o^z h'_{p\to o}(p_o^*) = \lambda_y^z h'_y(p_y^* + \delta_o^z o_y^*), \tag{11}$$

and second, the older child equates the marginal product of investing in their own human capital with the marginal product of investing in their younger sibling by setting

$$\lambda_o^z h'_{o\to o}(1 - o_y^*) = \delta_o^z \lambda_y^z h'_y(p_y^* + \delta_o^z o_y^*). \tag{12}$$

Two corner solutions are also possible: at the optimum, either o_y^* or p_y^* (but not both) might be equal to 0. If older children are sufficiently proficient at stimulating their younger siblings, parents may delegate this task to them by setting $p_y^* = 0$. On the other hand, when older children's investments in their younger siblings are sufficiently unproductive, (i.e. when δ_o^G and δ_o^B are sufficiently small), older siblings will devote all their time to building their own human capital by setting o_y^* to 0 and E_o to 1. For the rest of this exposition, we focus on the interior solution.[10]

When $\delta_o^G = \delta_o^B = 0$, the model reduces to simple case described in Section 2.1.[11] Reviewing those predictions: if human capital production functions do not differ by gender (i.e. if $\lambda_k^G = \lambda_k^B$ for $k \in \{o, y\}$), we would not expect gender gaps in parental investment or treatment effects of older sibling gender; when parents favor boys over girls (i.e. when $\lambda_y^G < \lambda_y^B$ and $\lambda_o^G < \lambda_o^B$), they invest more in boys than girls at all ages, and they also invest more in the younger siblings of older girls; finally, when $\lambda_y^G = \lambda_y^B$ and $\lambda_o^G < \lambda_o^B$, parents do not invest more in younger boys than in younger girls on average, but they invest more in the younger siblings of older girls — because they perceive a low return to spending time building the human capital of school-aged girls; thus, in summary, there is a treatment effect of older siblings that is mediated by parental investments.

2.2.1 Gender Differences in Productivity

We now characterize behavior when older children can improve their younger siblings' human capital by engaging in stimulating activities with them. In our framework, there are two reasons older girls might stimulate their younger siblings more than older boys. First, girls might be better at producing younger siblings' human capital with a given level of (time) investment ($\delta_o^G > \delta_o^B$). Alternatively, older boys and and girls might be equally good at caring for younger siblings, but the return to human capital might be lower for (older) girls than for (older) boys ($\lambda_k^G < \lambda_k^B$).[12] We have already considered

[10]Beyond the the interior solution that is our emphasis here, one might note that a sufficient (but not necessary) condition under which this latter corner solution occurs is when $\delta_o^z = 0$. One way to guarantee that the corner solution is inapplicable is to assume that $h'_{o \to o}(E_o) \to 0$ as $E_o \to 1$. Relaxing this assumption does not change our analysis substantively, so we entertain the corner solutions no further here.

[11]Specifically, Equation 4 is the special case of Equation 11 that occurs when $\delta_o^G = \delta_o^B = 0$.

[12]It is apparent that one could extend the model to introduce other reasons that girls might spend more time stimulating younger siblings. In a model of occupational specialization, girls who expect to specialize in home production might see a high return to the development of home-specific human capital (such as child-rearing skills). Alternatively, one could introduce social norms that make it costly for boys or girls to engage in behaviors commonly associated with the opposite gender (see the model presented in Jakiela and Ozier (2019) for a simple example). Many of these theoretical extensions yield predictions that are identical to those derived here. Indeed, the δ_o^z parameter captures some of these social norm effects in a simplified way if we interpret as a measure of the amount of stimulation (for example, singing or storytelling) an older sibling engages in per unit of time spent caring for a younger sibling. If stimulating activities are perceived as feminine because they are often done by mothers, older brothers may be less likely to engage in such

the implications of the latter possibility, letting $\lambda_k^G < \lambda_k^B$, in the special case when δ_o^G and δ_o^B are both equal to 0. We now consider the first of these possibilities: the consequences of gender differences in δ_o^z, the relative productivity of older siblings' investments in young children's human capital (compared to the parents' investments), when human capital production functions do not differ by gender. Specifically, let $\delta_o^G > \delta_o^B$, $\lambda_o^G = \lambda_o^B = \bar{\lambda}_o$, and $\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$. Let $p_o^*(\delta_o^z, \bar{\lambda}_o, \bar{\lambda}_y)$ denote the optimal level of parental investment in an older child of gender z (given our assumptions about δ_o^z and λ_k^z). Define $p_y^*(\delta_o^z, \bar{\lambda}_o, \bar{\lambda}_y), o_y^*(\delta_o^z, \bar{\lambda}_o, \bar{\lambda}_y), L_p^*(\delta_o^z, \bar{\lambda}_o, \bar{\lambda}_y), E_o^*(\delta_o^z, \bar{\lambda}_o, \bar{\lambda}_y)$, and $I_y^*(\delta_o^z, \bar{\lambda}_o, \bar{\lambda}_y)$ analogously for $z \in \{G, B\}$.[13]

In Proposition 1, we show that when older sisters are more productive than older brothers (when it comes to improving younger siblings' human capital), children with older sisters receive more stimulation overall. However, parents with an older daughter substitute away from investing their time in early childhood stimulation because their older child is a good substitute, investing more in the older child's human capital and increasing their own their own labor supply in consequence. Impacts on older siblings' time allocation are ambiguous and depend on the functional forms of the production functions, but the overall quantity of stimulation by older siblings ($\delta_o^z o_y^*$) is higher when the older sibling is female.

Proposition 1. *Let $\delta_o^G > \delta_o^B > 0$, and further assume δ_o^B is sufficiently far above zero to guarantee that $o_y^*(\delta_o^B, \lambda_o^B, \lambda_y^z) > 0$ and $o_y^*(\delta_o^B, \lambda_o^B, \lambda_y^z) > 0$. Let $\lambda_o^G = \lambda_o^B = \bar{\lambda}_o$ $\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$. The following are true:*

i. $I_y^*\left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) > I_y^*\left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$,

ii. $p_o^*\left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) > p_o^*\left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$,

iii. $p_y^*\left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) < p_y^*\left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$,

iv. $\delta_o^G o_y^*\left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) > \delta_o^B o_y^*\left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$, *and*

socially costly behaviors.

[13] An equilibrium is fully characterized by p_y^*, p_o^*, and o_y^*. The optimal L_p^*, E_o^*, and I_y^* are then defined by Equations 11 and 12.

v. $L_p^*\left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) > L_p^*\left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$.

Proof. See Appendix. □

When girls are more productive caregivers than boys, young children benefit from having an older sister: they receive more stimulation from their older sibling and more stimulation overall. Parents also benefit because older girls provide more effective support at home. As a result, gender differences in older children's effectiveness as caregivers translate into empirical predictions about both younger siblings' development and parents' responses.[14] Optimizing parents delegate more early childhood stimulation to more competent sibling caregivers, substituting toward other activities that cannot be done by their older children. Thus, if older girls are more effective caregivers than older boys, parents will appear to favor older girls by investing more in their human capital — but this appearance is deceptive because it results from gender differences in children's productivity rather than parents' preferences. This highlights the importance of explicitly modeling the human capital production function within the household, and accounting for the role that older children play in shaping younger children's human capital.

2.2.2 Gender Differences in the Returns to Human Capital

Thus far, we have seen that a treatment effect of having an older sister could be explained by two different mechanisms: either a gender gap in the return to human capital investment among older children when older siblings do not contribute to building younger children's human capital, or a gender gap in productivity where older sisters are better than older brothers at improving their younger siblings' human capital. We've considered each mechanism in isolation, and seen that the models make divergent predictions about parental investments in young children. Here, we extend the model to characterize household behavior when the returns to human capital investments in school-aged children differ by

[14] We consider the case of gender differences, but the proof is equally valid if other observable factors (e.g. older sibling age) generate systematic differences in δ_o^z.

gender *and* older siblings contribute to the development of human capital in young children by engaging in stimulating activities with them.

In Proposition 2, we show that when returns to human capital are lower for girls than for boys and older boys and girls are equally efficient at improving younger children's human capital, children with older sisters receive more stimulation overall. When returns to older siblings' human capital differ by gender, parents invest less in older sisters (relative to older brothers). However, when older children contribute to human capital accumulation in their younger siblings in this context, older sisters also invest less in their own human capital and more in the human capital of their younger siblings — breaking the link between the gender gap in the returns to investing in older children's human capital and parental investments in younger children.

Let $\delta_o^G = \delta_o^B = \bar{\delta}_o > 0$, $\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$, and $\lambda_o^G < \lambda_o^B$. Paralleling our analysis in Section 2.2.1, we let $p_o^*(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y)$ denote the optimal level of parental investment in an older girl under these assumptions, and we define $p_y^*(\bar{\delta}_o, \lambda_o^z, \bar{\lambda}_y)$, $o_y^*(\bar{\delta}_o, \lambda_o^z, \bar{\lambda}_y)$, $L_p^*(\bar{\delta}_o, \lambda_o^z, \bar{\lambda}_y)$, $E_o^*(\bar{\delta}_o, \lambda_o^z, \bar{\lambda}_y)$, and $I_y^*(\bar{\delta}_o, \lambda_o^z, \bar{\lambda}_y)$ analogously for $z \in \{G, B\}$.

Proposition 2. *Let $\lambda_o^G < \lambda_o^B$, and let $\lambda_y^G = \lambda_y^B$ and $\delta_y^G = \delta_y^B > 0$, and assume $\bar{\delta}_o$ is sufficiently far above zero to guarantee that $o_y^*(\bar{\delta}_o, \lambda_o^G, \lambda_y^z) > 0$ and $o_y^*(\bar{\delta}_o, \lambda_o^B, \lambda_y^z) > 0$. The following are true:*

i. $I_y^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) > I_y^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

ii. $p_o^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) < p_o^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

iii. $o_y^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) > o_y^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

iv. $\delta_o^G o_y^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) > \delta_o^B o_y^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

iii. $E_o^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) V E_o^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

and

vi. $L_p^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) > L_p^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$.

Proof. See Appendix. □

Proposition 2 highlights the importance of older siblings' investments in young children — even when the treatment effect of older sisters is driven by gender differences in the return to education as opposed to productivity. As discussed in Section 2.1, when $\delta_o^G = \delta_o^B = 0$, parents respond to gender gaps in the return to parental investment in older children by investing more in their younger children. Incorporating older siblings into the model as optimizing agents changes this prediction because older girls also invest less in themselves — and more in their younger siblings — than older boys.

We have considered the $\delta_o^G = \delta_o^B > 0$ case, but results are similar when δ_o^G and δ_o^B are not exactly equal. If older girls are substantially more effective caregivers than older boys (i.e. if δ_o^G is substantially larger than δ_o^B), the gender gap in sibling productivity will be more important than the gender gap in the return to schooling, so parents of older girls will invest less in their young children than parents of older boys. The opposite is true if older boys are substantially more effective sibling caregivers than older girls. In both cases, any gender gap in the returns to investing in older children's human capital may offset the effect of the gender gap in productivity. The key insight is that when older siblings make tradeoffs between their own human capital and that of their younger siblings, the effect of older sibling gender on parental investments in younger siblings is ambiguous because older sisters facing a lower return to investing in their own human capital invest more in the human capital of their siblings.

2.3 Extensions to the Model

In much of our analysis, we have assumed that parents are not inherently prejudiced against girls. If gender differences were driven by parental bias, we would expect parents to invest less in older girls than in older boys, and we would also expect them to invest less in younger girls than in younger boys. In our framework, $\lambda_y^G < \lambda_y^B$ implies a lower optimal level of familial investment in young girls than in young boys — a prediction that is testable in our data.

We have also assumed a unitary household that can be represented by a single utility function. However, if parents perceive a low return to investing in the human capital of older girls (relative to older boys) but older girls do not, that is, that older girls have a utility function identical to the parent's utility function except for the value of λ_y^G, the unitary household assumption may be inappropriate. This will tend to shift older girls toward investing more in their own human capital relative to the parental optimum; parents will partially offset this by investing more in the younger siblings of an older girl than they would under the unitary model — though the overall treatment effect of older sibling gender on parental investments in young children remains ambiguous. This channel can only matter when older siblings contribute to the human capital of the young children. If they did not, older siblings would invest all their time in building their own human capital irrespective of the gender gap in the returns to schooling.

2.4 Summary of Predictions

Table 1 summarizes the theoretical predictions that we will test empirically.[15] As discussed below, our data set includes information on the amount of early childhood stimulation done by parents, by older siblings, and by other individuals. The model summarizes predictions about three outcomes: p_y^*, the amount of parental stimulation of young children, $\delta_o^z o_y^*$; the amount of stimulation done by the older sibling; and I_y^*, the total amount of early childhood stimulation experienced by the youngest family members. When either $\delta_o^G > \delta_o^B$ or $\lambda_o^G < \lambda_o^B$, young children with an older sister will receive more stimulation than young children with an older brother. Where this increase in overall stimulation comes from provides information about the underlying parameter values. When $\delta_o^G = \delta_o^B = 0$, older siblings' investments are not productive, so they do not invest time in stimulating their younger siblings. Hence, any overall impact of an older sister is mediated by parental investments. On the other hand, when $\delta_o^G > \delta_o^B > 0$ and $\lambda_o^G = \lambda_o^B$, the treatment effect of an older sister results from the fact that older sisters are more productive caregivers, and

[15]The full set if theoretical predictions is presented in Table A1.

they do more stimulation of their younger siblings than older brothers. Parents respond to this by investing less in their younger children and more in their older children. Finally, when $\lambda_o^G < \lambda_o^B$ and $\delta_o^G \geq \delta_o^B > 0$, both mechansims are at play. Young children receive more stimulation from older sisters than older brothers. Because of this, parents may invest either more or less in their younger children. The fact that $\lambda_o^G < \lambda_o^B$ pushes them toward investing less in their older daughters and more in their younger children. However, older daughters also invest less in themselves and more in their younger siblings, lowering the marginal return to parental investments in young children. Thus, the overall impact on parental investments in young children cannot be signed when both mechanisms are at work.[16]

3 Data

Our sample includes data on 552 households from 73 rural communities in western Kenya. Data were collected during the baseline survey that preceded a pre-literacy intervention (Jakiela, Ozier, Fernald and Knauer 2020). Households living within 750 meters of the local government primary school were included in the sample if they had children between three and six years old. Here, we restrict attention to those households which also had exactly one older child between the ages of seven and 14. Our treatment of interest is an indicator for having an older child who is female. In this restricted sample of households, having an older child who is female is uncorrelated with a range of covariates, as we discuss further below.

Our data set includes information on household and parental characteristics (e.g. household assets and mother's education) as well as multiple measures of child development and familial investments in young children. We consider two main developmental outcomes that can be measured in preschool-aged children: vocabulary and fine motor skills. Both are measured through direct child assessment.

[16]This is true for any values of λ_o^G and λ_o^B such that $\lambda_o^G \geq \lambda_o^B > 0$.

Our vocabulary index combines includes three sub-scales: expressive vocabulary and receptive vocabulary in English (one of Kenya's national language and the primary language of instruction at upper levels of primary school) and Luo (a Nilotic language that is the mother tongue of all of the children in our sample). Receptive vocabulary is the ability to understand words, while expressive vocabulary is the ability to produce words — for example, to identify familiar objects. Children begin developing receptive vocabulary before they begin to express themselves through speech (Fernald, Prado, Kariger and Raikes 2017). To measure receptive vocabulary in English and Luo, we adapted items from the British Picture Vocabulary Scale, a version of the Peabody Picture Vocabulary Test suitable for speakers of British or Commonwealth English (Dunn and Dunn 1997; Dunn, Dunn and Styles 2009; Knauer et al. 2019b). We assessed expressive vocabulary through a 37-item assessment developed and validated for the EMERGE study (Knauer, Kariger, Jakiela, Ozier and Fernald 2019b).

We assessed children's fine motor skills using a subset of items from the Malawi Developmental Assessment Test (Gladstone et al. 2010). Specifically, the survey included six questions from the MDAT fine motor sub-scale that showed high predictive power (in terms of other development outcomes) in a pilot study (Knauer et al. 2019a). The items measure young children's ability to build simple structures (e.g. a tower) with blocks and to use a pencil to make elementary drawings (e.g. a circle). Both vocabulary and fine motor indices are converted into age-normalized z-scores. We then average the individual (z-score) components to construct an overall measure of child development.

To understand the mechanisms through which sibling gender impacts child development, we collected data on early childhood stimulation using an expanded version of the Family Care Indicators (FCI) questionnaire (Hamadani et al. 2010, Kariger et al. 2012). The FCI asks about six types of stimulating activities: for example, reading, singing, storytelling, and physical play. We expanded this set to include additional stimulating activities more appropriate for slightly older children: for instance, teaching a child letters or English words (Knauer et al. 2019a). Based on extensive piloting, we also expanded the questionnaire to better capture the full range of family members who engage in early childhood stimulation. While the original instrument asks about stimulation by a child's mother, father, and by other adults, we also ask about stimulating activities by older sisters, older brothers, and grandparents. Summary statistics on who engages in early childhood stimulation are shown in Appendix Figure A1. On average, older sisters engage in more stimulating activities with young children than any other household members. Though older sisters do the most, even older brothers play an important role: older brothers do more than fathers or grandparents.

4 Analysis

4.1 Empirical Strategy

To estimate the impact of big sisters on child development, we assume that child gender is plausibly exogenous.[17] We estimate the regression equation:

$$Y_i = \alpha + \beta Sister_i + \varepsilon_i \tag{13}$$

where $Sister_i$ is an indicator equal to one if the older sibling in household i is female. Parents cannot control the sex of any given child, and households in our study area have little access to sex-selection technologies — so gender is not explicitly endogenous. Nevertheless, our estimates of the treatment effect of older sisters will be biased if $Sister_i$ is correlated with any (observed or unobserved) covariates that also predict outcomes. For example, if fathers were less likely to be present in households with an older girl (Dahl and Moretti 2008) and fathers' presence had a direct effect on child development outcomes, $\hat{\beta}$ would not capture the causal impact of having an older sister on child development. We test for this by comparing the observable characteristics of households with and without an older sister.

Summary statistics comparing households with an older sister to households with an older brother are presented in Table 2. Households are broadly similar in terms of family structure, parental characteristics, and living conditions; and younger children are similar

[17] See Washington (2008) for a similar estimation approach.

in terms of gender, age, and school enrollment. Older sisters and older brothers are also similar in age, suggesting comparable patterns of fertility and birth spacing in the two types of households. Since families with older sisters and older brothers look similar in terms of observable characteristics, we treat the gender of the older child as plausibly exogenous in our subsequent analysis.

4.2 The Impact of Big Sisters on Child Development

Kernel density estimates of our early childhood development index are presented in Figure 1. Negative z-scores are more common among young children with older brothers, and z-scores are more concentrated about zero among children with older sisters. The density functions are quite similar for z-scores above one. Thus, the graphical evidence suggests that poor early childhood development outcomes are less common in families with an older child who is female.

Regression estimates of the impact of older sisters on younger siblings' development are reported in Table 3. Having an older sister rather than an older brother has a large and statistically significant effects on their younger siblings. Estimates of Equation 13 suggest that young children with an older sister score 0.129 standard deviations higher on our aggregate measure of early childhood development (p-value 0.035). In specifications that include controls for child gender, age (fixed effects for child age in months), mother's education, and an index of household assets, the estimated impact of big sisters rises to 0.141 (p-value 0.023). Both coefficients are large in magnitude and developmentally meaningful. For comparison, the estimated effect is roughly equivalent to the difference in development between children whose mothers' completed secondary school and those whose mothers' only completed primary.[18]

Quantile regressions of the early childhood development index on the indicator for having an older sister rather than an older brother are summarized in Figure 2. We estimate

[18] In OLS specifications including controls for child gender, age (fixed effects for child age in months), mother's education, and an index of household wealth, the coefficient on maternal education (years of schooling) is 0.031 (p-value 0.042).

one regression for every quantile between 0.02 and 0.98. The pattern suggests that impacts are largest at the bottom of the distribution. For the lowest quantiles, the estimated treatment effects are large but imprecisely estimated. For quantiles between about 0.2 and 0.5, estimated treatment effects are positive and confidence intervals typically exclude zero. Above the median, estimated treatment effects are closer to zero and never statistically significant. Thus, results from quantile regressions formalize the evidence from the kernel density estimates: the treatment effects of sisters appear to be concentrated on the bottom half of the distribution.

In Panel B of Table 3, we decompose the underlying elements of the early childhood development index, estimating the treatment effect of big sisters on young children's vocabulary and fine motor skills. Results show that having an older sister leads to improvements in both outcomes. In specifications including controls (child gender, child age, mother's education, and an index of household assets), having an older sister as opposed to an older brother is associated with a 0.130 standard deviation increase in vocabulary (p-value 0.042) and a 0.151 standard deviation increase in fine motor skills (p-value 0.063).

In Figures 3 and 4, we present quantile regressions of the impact of older sisters on vocabulary and fine motor skills. In both cases, the largest point estimates occur near the bottom of the distribution. However, impacts on the quantiles of fine motor skills are precise zeros in the top half of the distribution, while estimated impacts on quantiles below the median are substantially larger and often statistically significant. Thus, having an older sister appears to improve fine motor skills, but only below the median. In contrast, estimated impacts on vocabulary skills are consistently small and positive, though they are rarely statistically significant above the 20^{th} percentile.

4.3 The Impact of Big Sisters on Investments in Young Children

As discussed in Section 2, there are several different reasons that younger siblings might benefit from having an older sister. One possibility is that older sisters are more effective than older brothers at improving children's younger siblings' human capital. Alternatively, older girls and their parents might believe that the returns to investing in their human capital are relatively low (as compared with similarly aged boys). If older siblings' investments in young children are not productive, parents who invest less in their older girls will invest more in their youngest children. On the other hand, if older children contribute to the human capital accumulation of their younger siblings, older girls will invest less in themselves and more in their younger siblings — and the impact of older sibling gender on parental investment will be ambiguous.

We test the predictions of the model using data on early childhood stimulation — both the overall amount of stimulation received by each young child, and the amount of stimulation done by different family members (e.g. the mother, the father, the older siblings, etc.). Estimates of the impact of having an older daughter on the overall level of early childhood stimulation are reported in Panel C of Table 3. We find large and statistically significant impacts of older sisters on the level of early childhood stimulation a child experiences. Having an older sister increases the number of different stimulating activities (out of 12) over the three days prior to the survey by between 0.637 (without controls, p-value 0.006) and 0.703 (with controls, p-value 0.002). Among households where the older child is male, the mean number of stimulating activities if 5.147; hence, the estimated treatment effect of having an older sister represents more than a ten percentage point increase in early childhood stimulation. From a theoretical perspective, this suggests that we can rule out the possibility that both $\delta_o^G = \delta_o^B$ and $\lambda_o^G = \lambda_o^B$ — since we observe a clear treatment effect of older sibling gender on early childhood stimulation.

Next, we test whether parents are the main channel of impact. In Figure 5, we summarize the estimated treatment effect of having an older sister on 10 different outcome variables. First, we present the treatment effect on the overall level of early childhood stimulation (replicating the specification in Panel C of Table 3 that was discussed above). Then, we present treatment effects on the amount of early childhood stimulation done by parents and the amount done by siblings. Finally, we present treatment effects on the amount of early childhood stimulation done mothers, by fathers, by sisters, by brothers, by grandmothers, by grandfathers, and by other individuals. Having an older sister does not impact the amount of stimulation young children receive from their parents, nor does it impact the amount received from the mother or father specifically, from either grandparent, or from others. All estimated coefficients are relatively precise zeros. Instead, having an older sister leads to a significant increase in the amount of stimulation received from siblings. Having an older sister increases the amount of stimulation done by sisters and decreases the amount done by brothers, but the positive impact on stimulation by sisters is larger — leading to a positive impact on the overall level of sibling stimulation.[19]

Seen through the lens of our theoretical model, this pattern of empirical results suggests two things. First, older siblings contribute to the human capital accumulation of their young siblings, and both parents and older siblings know it. If this were not the case, we would expect a treatment effect of older siblings on parental investments in young children. Second, both parents and older siblings perceive lower returns to investments in older girls' human capital relative to older boys. If this were not the case — and the treatment effect of older sisters was driven entirely by gender differences in the productivity of human capital investments by sisters vs. brothers — we would expect a negative treatment effect of older sisters on stimulation (of young children) by parents. We do not observe this, suggesting that part of the effect is attributable to gender gaps in the returns to human capital investments among older children, which lead older sisters to invest less in themselves and more in their young siblings while discouraging parents from ramping up their investments in older girls.

5 Conclusion

Older sisters have a positive and significant impact on their younger siblings' development. Our results are not consistent with a model in which parents of older sisters invest more in young children only because of a relaxed budget constraint associated with lower perceived

[19]Since our measure also captures stimulation by adult siblings, the level of stimulation by older sisters is not zero in households where the older child is male.

labor market returns to investment in the older sister. Instead, our results suggest that older siblings and parents *both* contribute to the development of young children's human capital, and that households know the productive value of these contributions. Importantly, the empirical patterns we observe can only arise if both parents and older children see a lower return to human capital investments in older girls, relative to older boys. Older sisters invest less in their own human capital than older brothers, so they invest more in their younger siblings. This changes the marginal utility of parental investments, so parents of older girls may or may not invest more in their youngest progeny than parents of older boys.

Our results highlight the critical importance of older children (both sisters and brothers) in child rearing in developing country contexts. In our sample, siblings do more cognitive stimulation than any other household member — but their role is typically ignored in models of household investments in children and policy discussions about early childhood. Our results suggest that evaluations of early childhood interventions are unlikely to fully capture effects on households if they do not take account of older siblings; likewise, evaluations targeting older children should explicitly consider impacts on younger siblings in critical stages of child development. Siblings, particularly sisters, play an important role in shaping the developmental trajectories of young children.

References

Aizer, Anna and Flavio Cunha, "The Production of Human Capital: Endowments, Investments and Fertility," Technical Report, National Bureau of Economic Research 2012.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the origins of gender roles: Women and the plough," *The Quarterly Journal of Economics*, 2013, *128* (2), 469–530.

Almond, Douglas and Janet Currie, "Human Capital Development Before Age Five," in David Card and Orley Ashenfelter, eds., *Handbook of Labor Economics*, Vol. 4, Elsevier, 2011, pp. 1315–1486.

Alsan, Marcella, "The Gendered Spillover Effect of Young Children's Health on Human Capital: Evidence from Turkey," Technical Report, National Bureau of Economic Research 2017.

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

Angrist, Joshua D. and William N. Evans, "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, 1998, *88* (3), 450–477.

Apoko, Anna, "At Home in the Village: Growing Up in Acholi," in Lorene K. Fox, ed., *East African Childhood*, Oxford University Press, 1967, pp. 45–75.

Attanasio, Orazio, Ricardo Paes de Barro, Pedro Carneiro, David K. Evans, Lycia Lima, Pedro Olinto, and Norbert Schady, "Public Childcare, Child Development, and Labor Market Outcomes," Technical Report 2019.

___, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina, "Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia," *American Economic Review*, 2020, *110* (1), 48–85.

Baird, Sarah, Craig McIntosh, and Berk Özler, "Cash or condition? Evidence from a cash transfer experiment," *The Quarterly journal of economics*, 2011, *126* (4), 1709–1753.

Bisbee, James, Rajeev Dehejia, Cristian Pop-Eleches, and Cyrus Samii, "Local Instruments, Global Extrapolation: External Validity of the Labor Supply-Fertility Local Average Treatment Effect," *Journal of Labor Economics*, 2017, *35* (S1), S99–S147.

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

Bock, John, "An Evolutionary Perspective on Learning in Social, Cultural, and Ecological Context," in David F. Lancy, John Bock, and Suzanne Gaskins, eds., *The Anthropology of Learning in Childhood*, Lanham, MD: AltaMira Press, 2010, pp. 11–34.

Brody, Gene H., "Siblings' Direct and Indirect Contributions to Child Development," *Current Directions in Psychological Science*, 2004, *13* (3), 124–126.

Chang, Susan M, Sally M Grantham-McGregor, Christine A Powell, Marcos Vera-Hernández, Florencia Lopez-Boo, Helen Baker-Henningham, and Susan P Walker, "Integrating a Parenting Intervention with Routine Primary Health Care: A Cluster Randomized Trial," *Pediatrics*, 2015, *136* (2), 272–280.

Chick, Gary, "Work, Play, and Learning," in David F. Lancy, John Bock, and Suzanne Gaskins, eds., *The Anthropology of Learning in Childhood*, Lanham, MD: AltaMira Press, 2010, pp. 119–143.

Clark, Shelley, "Son Preference and Sex Composition of Children: Evidence from India," *Demography*, 2000, *37* (1), 95–108.

Clarke, Damian, "Children and Their Parents: A Review of Fertility and Causality," *Journal of Economic Surveys*, 2018, *32* (2), 518–540.

Cunha, Flavio, James J Heckman, and Susanne M Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

Dahl, Gordon B. and Enrico Moretti, "The Demand for Sons," *Review of Economic Studies*, 2008, *75* (4), 1085–1120.

Dhar, Diva, Tarun Jain, and Seema Jayachandran, "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India," Technical Report, National Bureau of Economic Research 2018.

Doyle, Orla, "The First 2,000 Days and Child Skills," *Journal of Political Economy*, 2020, *128* (6), 2067–2122.

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

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

Edmonds, Eric V, "Understanding sibling differences in child labor," *Journal of Population Economics*, 2006, *19* (4), 795–821.

Ellis, Shari and Barbara Rogoff, "The Strategies and Efficacy of Child versus Adult Teachers," *Child Development*, 1982, *53*, 730–735.

Evans, David K, Maryam Akmal, and Pamela Jakiela, "Gender Gaps in Education: The Long View," Center for Global Development Working Paper 523 2020.

Evans, David K., Pamela Jakiela, and Heather Knauer, "The Impact of Early Childhood Development Interventions on Mothers and Others," working paper 2020.

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

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M Chang, and Sally Grantham-McGregor, "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica," *Science*, 2014, *344* (6187), 998–1001.

Gladstone, M., G. A. Lancaster, E. Umar, M. Nyirenda, E. Kayira, N. R. van den Broek, and R. L. Smyth, "The Malawi Developmental Assessment Tool (MDAT): The Creation, Validation, and Reliability of a Tool to Assess Child Development in Rural African Settings," *PLoS Medicine*, 2010, *7* (5), http://doi.org/10.1371/journal.pmed.1000273.

Glynn, Adam N and Maya Sen, "Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Women's Issues?," *American Journal of Political Science*, 2015, *59* (1), 37–54.

Grantham-McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, Barbara Strupp, International Child Development Steering Group et al., "Developmental potential in the first 5 years for children in developing countries," *The Lancet*, 2007, *369* (9555), 60–70.

Hamadani, J. D., F. Tofail, A. Hilalay, S. N. Huda, P. Engle, and S. M. Grantham-McGregor, "Use of Family Care Indicators and Their Relationship with Child Development in Bangladesh," *Journal of Health, Population, and Nutrition*, 2010, *28* (1), 23–33.

Hamadani, Jena D, Syeda F Mehrin, Fahmida Tofail, Mohammad I Hasan, Syed N Huda, Helen Baker-Henningham, Deborah Ridout, and Sally Grantham-McGregor, "Integrating an Early Childhood Development Programme into Bangladeshi Primary Health-Care Services: an Open-Label, Cluster-Randomised Controlled Trial," *The Lancet Global Health*, 2019, *7* (3), e366–e375.

Heckman, James J., "The Economics, Technology, and Neuroscience of Human Capability Formation," *Proceedings of the national Academy of Sciences*, 2007, *104* (33), 13250–13255.

Jakiela, Pamela and Owen Ozier, "Gendered Language," 2019.

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

Jayachandran, Seema and Ilyana Kuziemko, "Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India," *The Quarterly journal of economics*, 2011, *126* (3), 1485–1538.

Jensen, Robert, "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India," *The Quarterly Journal of Economics*, 2012, *127* (2), 753–792.

Kariger, P., E. A. Frongillo, P. Engle, P. M. Britto, S. M. Sywulka, and P. Menon, "Indicators of Family Care for Development for Use in Multicountry Surveys," *Journal of Health, Population, and Nutrition*, 2012, *30* (4), 472–486.

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

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

Knudsen, Eric I, James J Heckman, Judy L Cameron, and Jack P Shonkoff, "Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce," *Proceedings of the National Academy of Sciences*, 2006, *103* (27), 10155–10162.

Kremer, Michael, Edward Miguel, and Rebecca Thornton, "Incentives to learn," *The Review of Economics and Statistics*, 2009, *91* (3), 437–456.

Lancy, David F., *The Anthropology of Childhood: Cherubs, Chattel, Changelings*, Cambridge, UK: Cambridge University Press, 2015.

—— **and M. Annette Grove**, "The Role of Adults in Children's Learning," in David F. Lancy, John Bock, and Suzanne Gaskins, eds., *The Anthropology of Learning in Childhood*, Lanham, MD: AltaMira Press, 2010, pp. 145–179.

Lijembe, Joseph A., "The Valley Between: A Muluyia's Story," in Lorene K. Fox, ed., *East African Childhood*, Oxford University Press, 1967, pp. 1–41.

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

Maynard, Ashley E., "Cultural Teaching: The Development of Teaching Skills in Maya Sibling Interactions," *Child Development*, 2002, *73* (3), 969–982.

——— **and Katrin E. Tovote**, "Learning from Other Children," in David F. Lancy, John Bock, and Suzanne Gaskins, eds., *The Anthropology of Learning in Childhood*, Lanham, MD: AltaMira Press, 2010, pp. 181–205.

Montgomery, Heather, "Learning Gender Roles," in David F. Lancy, John Bock, and Suzanne Gaskins, eds., *The Anthropology of Learning in Childhood*, Lanham, MD: AltaMira Press, 2010, pp. 287–305.

Ozier, Owen, "Exploiting Externalities to Estimate the Long-Term Effects of Early Childhood Deworming," *American Economic Journal: Applied Economics*, 2018, *10* (3), 235–62.

Pitt, Mark M, Mark R Rosenzweig, and Mohammad Nazmul Hassan, "Human capital investment and the gender division of labor in a brawn-based economy," *American Economic Review*, 2012, *102* (7), 3531–60.

Psaki, Stephanie R, Katharine J McCarthy, and Barbara S Mensch, "Measuring gender equality in education: Lessons from trends in 43 countries," *Population and Development Review*, 2018, *44* (1), 117–142.

Qureshi, Javaeria A, "Additional Returns to Investing in Girls' Education: Impact on Younger Sibling Human Capital," *Economic Journal*, 2018, *128* (616), 3285–3319.

Rogoff, Barbara, Martha Julia Sellers, Sergio Pirotta, Nathan Fox, and Sheldon H. White, "Age of Assignment of Roles and Responsibilities to Children," *Human Development*, 1975, *18*, 353–369.

Walker, Susan P., Theodore D. Wachs, Sally Grantham-McGregor, Maureen M. Black, Charles A. Nelson, Sandra L. Huffman, Helen Baker-Henningham, Susan M. Chang, Jena D. Hamadani, Betsy Lozoff, Julie M. Meeks Gardner, Christine A. Powell, Atif Rahman, and Linda Richter, "Inequality in Early Childhood: Risk and Protective Factors for Early Child Development," *Lancet*, 2011, *6736* (11), 60555–2.

Washington, Ebonya L, "Female Socialization: How Daughters Affect Their Legislator Fathers," *American Economic Review*, 2008, *98* (1), 311–32.

Weisleder, A, Denise S.R. Mazzuchelli, Aline Sá Lopez, Walfrido Duarte Neto, Carolyn Brockmeyer Cates, Hosana Alves Gonçalves, Rochele Paz Fonseca, João Oliveira, and Alan L. Mendelsohn, "Reading Aloud and Child Development: A Cluster-Randomized Trial in Brazil," *Pediatrics*, 2017, *141* (1), e20170723.

Weisner, Thomas S. and Ronald Gallimore, "My Brother's Keeper: Child and Sibling Caretaking," *Current Anthropology*, 1977, *18* (2), 169–190.

Figure 1: Kernel Density Estimates of Early Childhood Development Indices

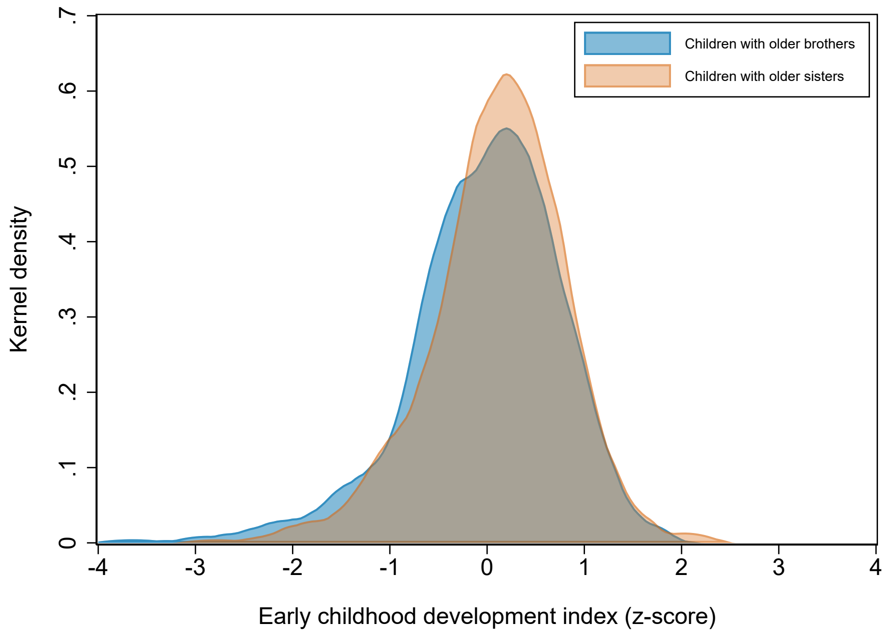

Figure 2: Quantile Regressions of the Impact of Sisters on Early Childhood Development

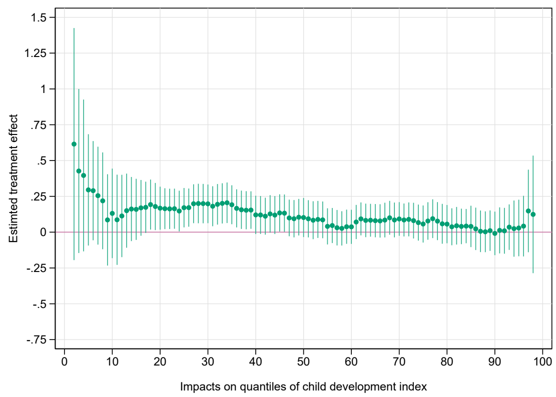

Figure 3: Quantile Regressions of the Impact of Sisters on Vocabulary

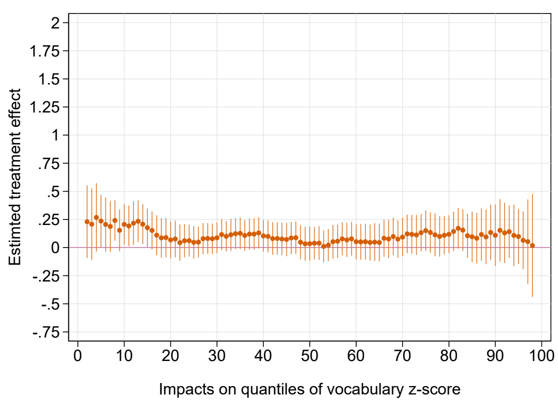

Figure 4: Quantile Regressions of the Impact of Sisters on Fine Motor Skills

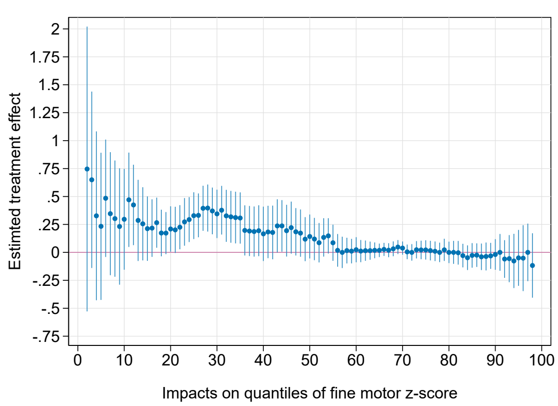

Figure 5: Decomposing the Impact of Having a Sister on Early Childhood Stimulation

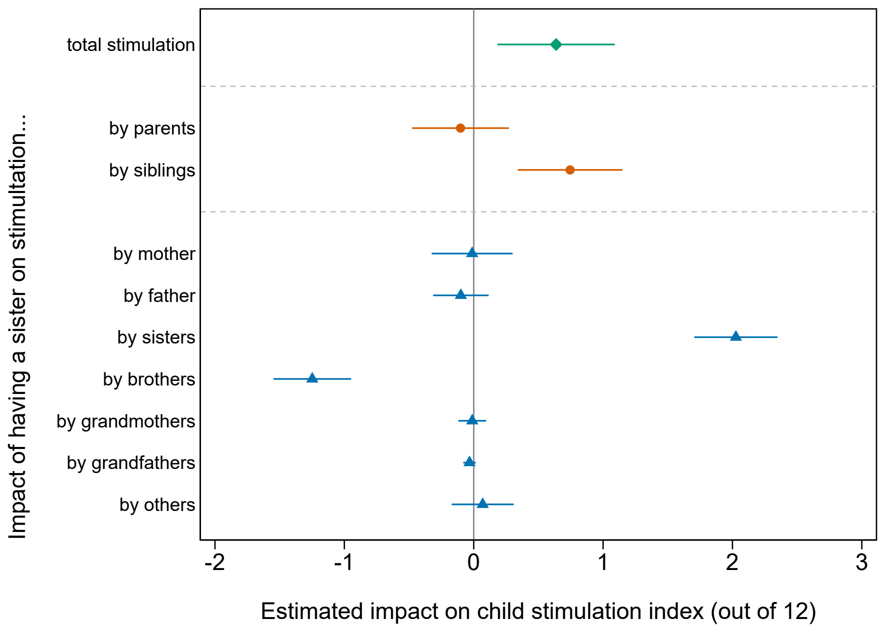

Table 1: Testable Predictions of the Theoretical Model when $\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$

Assumptions:	p_y^*	$\delta_o^z o_y^*$	I_y^*
No gender differences			
$\delta_o^G = \delta_o^B \geq 0$			
$\lambda_o^G = \lambda_o^B = \bar{\lambda}_o$	$p_y^G = p_y^B$	$\delta_o^G o_y^G = \delta_o^B o_y^B$	$I_y^G I_y^B$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$			
Differential returns, zero sibling productivity			
$\delta_o^G = \delta_o^B = 0$			
$\lambda_o^G < \lambda_o^B$	$p_y^G > p_y^B$	$\delta_o^G o_y^G = \delta_o^B o_y^B = 0$	$I_y^G > I_y^B$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$			
Equal returns, differential sibling productivity			
$\delta_o^G > \delta_o^B > 0$			
$\lambda_o^G = \lambda_o^B = \bar{\lambda}_o$	$p_y^G < p_y^B$	$\delta_o^G o_y^G > \delta_o^B o_y^B$	$I_y^G > I_y^B$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$			
Differential returns, equal (positive) sibling productivity			
$\delta_o^G = \delta_o^B > 0$			
$\lambda_o^G < \lambda_o^B$	–	$\delta_o^G o_y^G > \delta_o^B o_y^B$	$I_y^G > I_y^B$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$			
Differential returns, differential (positive) sibling productivity			
$\delta_o^G > \delta_o^B > 0$			
$\lambda_o^G < \lambda_o^B$	–	$\delta_o^G o_y^G > \delta_o^B o_y^B$	$I_y^G > I_y^B$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$			

Table 2: Summary Statistics by Older Sibling Gender

Older sibling is a...	Sister		Brother		Difference	
	Mean	S.D.	Mean	S.D.	Diff.	S.E.
Child is male	0.48	0.50	0.53	0.50	-0.05	0.04
Child age (in months)	59.70	13.72	60.46	14.14	-0.76	0.90
Child is enrolled in school	0.90	0.30	0.87	0.33	0.03	0.02
Older sibling age	9.53	2.16	9.51	2.19	0.02	0.20
Caregiver is child's mother	0.84	0.37	0.84	0.37	0.00	0.03
Caregiver is child's father	0.01	0.08	0.00	0.05	0.00	0.00
Caregiver is child's grandmother	0.12	0.33	0.13	0.34	-0.01	0.03
Caregiver illiterate	0.48	0.50	0.54	0.50	-0.06	0.05
Child's mother is alive	0.96	0.19	0.97	0.17	-0.01	0.01
Mother's age	30.50	7.03	30.44	6.90	0.06	0.60
Mother is Luo	0.95	0.21	0.95	0.22	0.01	0.02
Mother's education in years	7.88	2.39	8.02	2.42	-0.15	0.21
Father unknown or deceased	0.24	0.43	0.19	0.39	0.05	0.04
Mother and father married	0.82	0.38	0.82	0.39	0.01	0.04
Father's age	39.41	9.43	38.37	9.22	1.04	0.87
Father is Luo	0.99	0.10	0.98	0.14	0.01	0.01
Father's education in years	8.67	2.76	8.89	2.59	-0.23	0.25
Number of young children	0.38	0.48	0.47	0.50	-0.09^{*}	0.05
Has cement floor	0.15	0.36	0.16	0.37	-0.01	0.03
Has iron roof	0.98	0.13	0.99	0.12	-0.00	0.01
Has latrine or toilet	0.81	0.40	0.78	0.42	0.03	0.04
Has solar power	0.39	0.49	0.44	0.50	-0.04	0.04
Distance to primary school (in meters)	438.64	185.28	428.14	156.11	10.50	15.35
Observations	352		347			

Statistical significance: ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

Table 3: Impacts of Big Sisters on Early Childhood Development

		No Controls		W/ Controls	
	Mean	Coef.	S.E.	Coef.	S.E.
Panel A. Summary Measures of Younger Siblings' Development					
Child development index (z-score)	-0.022	0.129**	0.061	0.141**	0.062
Child is stunted	0.128	-0.059**	0.024	-0.047*	0.026
Panel B. Components of Child Development Index					
Child vocabulary (z-score)	-0.015	0.108*	0.064	0.130**	0.064
Fine motor skills (z-score)	-0.028	0.149*	0.078	0.151*	0.081
Panel C. Household Investments in Young Children					
Early childhood stimulation index (out of 12)	5.147	0.637***	0.231	0.703***	0.225
Child given Vitamin A supplement	0.481	0.084**	0.042	0.100**	0.042
Child given iron supplement	0.037	0.005	0.017	0.011	0.016
Child plays with store-bought toys	0.380	0.014	0.043	0.014	0.043

OLS coefficients reported. Robust standard errors clustered at the household level. The mean indicates the average variable of each outcome variable among households with a single male child between the ages of 7 and 14; the OLS coefficient estimates denote the treatment effect of having a older sister rather than an older brother (in the eligible age range). The specification with controls includes child age (fixed effects for age in months), child gender, mother's education, household size, and an index of household assets. Statistical significance: ***, **, and * indicate significance at the 1, 5, and 10 percent levels, respectively.

A Online Appendix: not for print publication

A.1 Mathematical Appendix

A.1.1 Proof of Proposition 1.

Statement of Proposition 1. Let $\delta_o^G > \delta_o^B > 0$, and further assume δ_o^B is sufficiently far above zero to guarantee that $o_y^*(\delta_o^B, \lambda_o^B, \lambda_y^z) > 0$ and $o_y^*(\delta_o^B, \lambda_o^B, \lambda_y^z) > 0$ (so, older brothers allocate a strictly positive amount of time to engaging in stimulating activities with their younger siblings). Let $\lambda_o^G = \lambda_o^B = \bar{\lambda}_o$ $\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$. Then, the following are true:

i. $I_y^* \left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) > I_y^* \left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$,

ii. $p_o^* \left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) > p_o^* \left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$,

iii. $p_y^* \left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) < p_y^* \left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$,

iv. $\delta_o^G o_y^* \left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) > \delta_o^B o_y^* \left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$, and

v. $L_p^* \left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right) > L_p^* \left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$.

Notation. To simplify notation within the proof, we omit the arguments of the quantities agents are maximizing over. We use I_y^G to denote $I_y^* \left(\delta_o^G, \bar{\lambda}_o, \bar{\lambda}_y\right)$ and I_y^B to denote $I_y^* \left(\delta_o^B, \bar{\lambda}_o, \bar{\lambda}_y\right)$. For $z \in \{G, B\}$, p_o^z, p_y^z, o_y^z, L_p^z, and E_o^z are defined analogously. The arguments are unnecessary within the proof because we have explicitly stated our assumptions regarding the values of δ_o^z, λ_o^z, and λ_y^z above. Within the proof, we use * (e.g. in I_y^*) in comparative statics analysis to indicate the optimal value defined as a function of δ, not the optimum at a specific value of δ such as δ_o^G or δ_o^B.

Step 1. An increase in δ_o^z leads to an increase in I_y^*, so $I_y^G > I_y^B$. Assume not: assume an increase in δ_o^z leads to either a decrease or no change in $I_y^* = p_y^* + \delta_o^z o_y^*$.

First, consider the possibility that an increase in δ_o^z leads to a decrease in I_y^* and thus an increase in $\bar{\lambda}_y h_y'(p_y^* + \delta_o^z o_y^*)$. By Equation 11, this implies an increase in both $y'(1 - p_o^* - p_y^*)$ and $\bar{\lambda}_o h_{p\to o}'(p_o^*)$. The latter implies a decrease in p_o^* since $h_{p\to o}(\cdot)$ is strictly concave. By a similar argument, the former implies an increase in $p_o^* + p_y^*$; since we've already shown that p_o^* must decrease, p_y^* must increase. So, if an increase in δ_o^z leads to a decrease in $\delta_o^z o_y^* + p_y^*$, it implies an increase in p_y^*; thus, the decrease in $\delta_o^z o_y^* + p_y^*$ must come from an increase in o_y^*.

An increase in δ_o^z must also lead to either an increase in $h_{o\to o}'(1 - o_y^*)$ or a decrease in $h_y'(\delta_o^z o_y^* + p_y^*)$ (or both) if Equation 12 is to hold. Since we started from the assumption that $\delta_o^z o_y^* + p_y^*$ decreases, $h_y'(\delta_o^z o_y^* + p_y^*)$ must increase. Hence, Equation 12 can only hold if $h_{o\to o}'(1-o_y^*)$ increases. However, we have already shown that o_y^* must decrease, so $1-o_y^*$ and $h_{o\to o}(1-o_y^*)$ must increase — leading to a decrease in $h_{o\to o}'(1-o_y^*)$. Thus, the assumption that an increase in δ_o^z leads to an decrease in I_y^* leads to a contradiction.

Next, consider the possibility that an increase in δ_o^z leads to no change in I_y^*. This means that there is no change in $h_y'(\delta_o^z o_y^* + p_y^*)$. There is consequently no change in either p_o^* or p_y^* (by Equation 11). Since there is no change in p_y^*, o_y^* must decrease to offset the increase in δ_o^z (keeping I_y^* constant). This implies an decrease in $h_{o\to o}'(1-o_y^*)$. However, Equation 12 requires an *increase* in $h_{o\to o}'(1-o_y^*)$ to offset the increase in δ_o^z — since $h_y'(\delta o_y^* + p_y^*)$ and $\bar{\lambda}_y$ do not change. So $h_{o\to o}'(1-o_y^*)$ must increase and decrease simultaneously — a contradiction.

Step 2. $I_y^G > I_y^B$ implies $p_o^G > p_o^B$. This follows directly from Equation 11 since $h_{p\to o}(\cdot)$ and $h_y(\cdot)$ are concave.

Step 3. $I_y^G > I_y^B$ and $p_o^G > p_o^B$ together imply $L_p^G > L_p^B$ and $p_y^G < p_y^B$. Since $h_y(\cdot)$ and $y(\cdot)$ are both strictly concave, the increase in I_y^* means that $y'(1-p_o^*-p_y^*)$ must decrease if Equation 11 is to hold. Hence, $L_p^G > L_p^B$ must increase. We have already shown that $p_o^G > p_o^B$. Since $L_p^z = 1 - p_o^z - p_y^z$, $L_p^G > L_p^B$ and $p_o^G > p_o^B$ together imply $p_y^G < p_y^B$.

□

A.1.2 Proof of Proposition 2.

Statement of Proposition 2. Let $\lambda_o^G < \lambda_o^B$, $\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$, and $\delta_o^G = \delta_o^B = \bar{\delta}_o > 0$. Then, the following are true:

i. $I^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) > I^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

ii. $L_p^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) > L_p^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

iii. $p_o^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) < p_o^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

iv. $o_y^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) > o_y^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$,

v. $\delta_o^G o_y^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) > \delta_o^B o_y^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$, and

vi. $E_o^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right) < E_o^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$.

Notation. To simplify notation within the proof, we omit the arguments of the quantities agents are maximizing over. We use I_y^G to denote $I_y^*\left(\bar{\delta}_o, \lambda_o^G, \bar{\lambda}_y\right)$ and I_y^B to denote $I_y^*\left(\bar{\delta}_o, \lambda_o^B, \bar{\lambda}_y\right)$. For $z \in \{G, B\}$, p_o^z, p_y^z, o_y^z, L_p^z, and E_o^z are defined analogously. The arguments are unnecessary within the proof because we have explicitly stated our assumptions regarding the values of δ_o^z, λ_o^z, and λ_y^z. Within the proof, we use * (e.g. in I_y^*) to indicate the optimal value defined as a function of δ, not the optimum at a specific value of λ_o such as λ_o^G or λ_o^B.

Step 1. A decrease in λ_o leads to a decrease in p_o^*.

Assume not: assume a decrease in λ_o leads to either a increase or no change in p_o^*.

By Equation 11, $\lambda_o = y'(1 - p_o^* - p_y^*)/h'_{p\to o}(p_o^*)$ (Equation 11). Hence, a decrease in λ_o means that either $y'(1-p_o^*-p_y^*)$ must decrease or $h'_{p\to o}(p_o^*)$ must increase. Because $h_{p\to o}(\cdot)$ is concave, $h'_{p\to o}(p_o^*)$ can only increase if p_o^* decreases. So, for λ_o to decrease without a decrease in p_o^*, $y'(1 - p_o^* - p_y^*)$ must decrease — and for this to happen without a decrease in p_o^*, p_y^* must decrease. So, if λ_o decreases, p_y^* must decrease.

By Equation 11, $\lambda_o = \bar{\lambda}_y h'_y(\bar{\delta}_o o_y^* + p_y^*)/h'_{p\to o}(p_o^*)$. So, if λ_o decreases and p_o^* does not, $h'_y(\bar{\delta}_o o_y^* + p_y^*)$ must decrease (since $\bar{\lambda}_y$ does not change). Since $h'_y(\cdot)$ is concave, this implies an increase in either o_y^* or p_y^*. Above, we demonstrated that p_y^* must decrease (if λ_o decreases and p_o^* does not), so o_y^* must increase.

Combining Equation 11 and Equation 12, we see that

$$\frac{h'_{p\to o}(p_o^*)}{h'_{o\to o}(1 - o_y^*)} = \frac{1}{\delta}. \tag{14}$$

Since o_y^* must increase and $\bar{\delta}_o$ does not change, we see that p_o^* must decrease — though we have assumed that it does not. Thus, starting from the assumption that p_o^* does not decline leads to a contradiction. Hence, a decrease in λ_o implies a decrease in p_o^*.

Step 2. The decrease in p_o^* implies a decrease in E_o^* and an increase in o_y^*.

This follows directly from Equation 14 and the definition of E_o^*.

Step 3. The decrease in p_o^* implies an increase in I_y^* and L_p^*.

We proceed by contradiction. We have already shown that o_y^* must increase. As a consequence, if we assume that I_y^* does not increase, then p_y^* must decrease. Since we have already shown that p_o^* must decrease, this means that $1 - (p_o^* + p_y^*)$ must increase, and (by concavity) $y'(1 - p_o^* - p_y^*)$ must decrease. Note, however, that if I_y^* does not increase, then $h_y'(\bar{\delta}_o o_y^* + p_y^*)$ cannot decrease and (as a result) Equation 11 cannot hold. This is a contradiction. So, I_y^* must increase, and (by Equation 11) L_p^* must increase as well.

Table A1: Summary of Theoretical Predictions when $\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$

Assumptions:	p_o^*	p_y^*	L_p^*	o_y^*	$\delta_o^z o_y^*$	E_o^*	I_y^*
$\delta_o^G = \delta_o^B = 0$							$I_y^G = I_y^B$
$\lambda_o^G = \lambda_o^B = \bar{\lambda}_o$	$p_o^G = p_o^B$	$p_y^G = p_y^B$	$L_p^G = L_p^B$	$o_y^G = o_y^B = 0$	$\delta_o^G o_y^G = \delta_o^B o_y^B$	$E_o^G = E_o^B = 1$	$I_y^G = p_y^G$
$\delta_o^G = \delta_o^B = 0$							$I_y^G > I_y^B$
$\lambda_o^G < \lambda_o^B$	$p_o^G < p_o^B$	$p_y^G > p_y^B$	$L_p^G > L_p^B$	$o_y^G = o_y^B = 0$	$\delta_o^G o_y^G = \delta_o^B o_y^B = 0$	$E_o^G = E_o^B = 1$	$I_y^G = p_y^G$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$							$I_y^B = p_y^B$
$\delta_o^G > \delta_o^B > 0$							
$\lambda_o^G = \lambda_o^B = \bar{\lambda}_o$	$p_o^G > p_o^B$	$p_y^G < p_y^B$	$L_p^G > L_p^B$	–	$\delta_o^G o_y^G > \delta_o^B o_y^B$	–	$I_y^G > I_y^B$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$							
$1 \geq \delta_o^G = \delta_o^B > 0$							
$\lambda_o^G < \lambda_o^B$	$p_o^G < p_o^B$	–	$L_p^G > L_p^B$	$o_y^G > o_y^B$	$\delta_o^G o_y^G > \delta_o^B o_y^B$	$E_o^G < E_o^B$	$I_y^G > I_y^B$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$							
$\delta_o^G > \delta_o^B > 0$							
$\lambda_o^G < \lambda_o^B$	–	–	$L_p^G > L_p^B$	–	$\delta_o^G o_y^G > \delta_o^B o_y^B$	–	$I_y^G > I_y^B$
$\lambda_y^G = \lambda_y^B = \bar{\lambda}_y$							

Figure A1: Who Engages in Cognitively Stimulating Activities with Young Children