I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17228

Free Schooling Reverses Sibling Rivalry

João R. Ferreira
Wayne Aaron Sandholtz

AUGUST 2024

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17228

Free Schooling Reverses Sibling Rivalry

João R. Ferreira
Nova School of Business and Economics and Universidade NOVA de Lisboa

Wayne Aaron Sandholtz
Nova School of Business and Economics, NOVAFRICA, CESifo and IZA

AUGUST 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Free Schooling Reverses Sibling Rivalry*

We use administrative data to measure sibling spillovers on academic performance before and after Tanzania's introduction of Free Secondary Education (FSE). Prior to FSE, students whose older siblings narrowly passed the secondary school entrance exam were less likely to go to secondary school themselves; with FSE, the effect became positive. Negative spillovers in the pre-reform period were concentrated in poorer regions; positive spillovers in the post-reform period were largest for lower-performing younger siblings. This suggests that FSE alleviated financial constraints, allowing families to distribute educational investments more equitably rather than concentrating resources on high-performing children.

JEL Classification: D10, I20, I25, J13

Keywords: sibling spillovers, educational achievement, resource constraints, high-stakes exams

Corresponding author:
Wayne Aaron Sandholtz
Nova School of Business and Economics
R. da Holanda 1
2775-405 Carcavelos
Portugal
E-mail: wayne.sandholtz@novasbe.pt

* For helpful comments and discussions we thank Krzysztof Karbownik, Supreet Kaur, Gaurav Khanna, Jeremy Magruder, Meera Mahadevan, Ted Miguel, Marianne Page, Javaeria Qureshi, Sally Sadoff, Tom Vogl, and seminar participants at LESE: Lisbon Economics and Statistics of Education conference. All errors are our own.

1 Introduction

Can free public education affect the way households allocate human capital investments across siblings? Theoretically, parents face a tension between "compensating" investments in lower-ability children vs. "reinforcing" investments in higher-ability children. Studies from the developing world document the existence of parental reinforcement, often denoted "sibling rivalry:" parents invest more in the human capital of higher-ability children, generating negative sibling spillovers (Vogl, 2013; Jensen & Miller, 2017; Morduch, 2000; Akresh, Bagby, de Walque, & Kazianga, 2012; Rosenzweig & Schultz, 1982).[1] In the developed world, however, sibling spillovers in academic achievement are generally positive, suggesting that resource constraints may partly drive the sibling rivalry observed in poorer settings (Qureshi, 2018b; Karbownik & Ozek, 2019; Figlio, Karbownik, & Özek, 2023). Households in both high- and low-income countries express preferences for equalizing investments across siblings (Berry, Dizon-Ross, & Jagnani, 2024; Adhvaryu & Nyshadham, 2016; Behrman, Pollak, & Taubman, 1982). By relaxing resource constraints, free public education policies may enable households to allocate human capital investments more equitably among children rather than "picking a winner" (Banerjee & Duflo, 2012).

There is little existing evidence on how policy affects sibling spillovers. While large literatures on sibling spillovers exist for the developed world — where school access is near-universal — and for the developing world — where access is more limited — no studies have yet managed to measure spillovers in the context of a shock to school access. Exogenous variation permitting the identification of sibling spillovers is rare, and coincident policy changes in access to schooling are even rarer. Some RCT evidence from the developing world has shown that increased resources (in the form of cash transfers or scholarships) can have null or even negative effects on educational outcomes for recipients' siblings (Barrera-Osorio, Bertrand, Linden, & Perez-Calle, 2008; Duflo, Dupas, & Kremer, 2017). However, it is not obvious whether effects would be similar for large-scale government programs for increasing school access (Bold, Kimenyi, Mwabu, Ng'ang'a, & Sandefur, 2018).

In this paper, we use administrative data to identify sibling spillovers in education before and after the introduction of a nationwide Free Secondary Education policy (FSE) in Tanzania. Tanzania offers an ideal context for examining how household investments across siblings respond to increased education access, for three reasons. First is the policy itself: the 2016 abolition of lower secondary school fees (primary school fees were abolished over a decade earlier). Secondary school is the frontier of free public education in much of the developing world: about half of countries in sub-Saharan Africa offered fee-free

[1] This is consistent with the theoretical prediction of Becker and Tomes (1976), who first formally articulated this tension.

secondary school as of March 2023 (Garlick, 2019; Gruijters, Abango, & Casely-Hayford, 2023). Second, there is evidence that in Tanzania school fees were a binding constraint on many households' educational investments prior to the policy, and that household educational expenditure in this context is responsive to a positive resource shock (Sandholtz, 2024; Burchardi, de Quidt, Gulesci, & Sulaiman, 2024). Third, the binding exam score threshold for secondary school entrance provides exogenous variation in secondary school entry.

To measure sibling spillovers, we exploit a discontinuity at the exam score required for admission to public secondary school. We draw upon the universe of administrative data on the results of the Primary School Leaving Exam (PSLE). This exam is high-stakes: a pass is required to continue in the public school system, and failing is associated with a higher likelihood of moving from primary school into "mining, grazing, and family activities" instead of secondary instruction (Kippenberg, 2014). Our data include multiple cohorts of exam takers before and after the introduction of FSE. We link students by last name within wards to identify pairs of likely siblings. Our sample consists of sibling pairs in which the older sibling scored at or near the passing threshold on the exam. We compare the educational outcomes of younger siblings in these pairs whose older sibling narrowly did vs. did not qualify for entry into public secondary school. We measure the "intent-to-treat" effect of having a sibling pass the qualifying exam, rather than using exam passing as an instrument for secondary school entry.[2]

We find that sibling spillovers are significant and negative prior to the introduction of the FSE reform — but significant and positive afterward. Prior to the reform, an older sibling who marginally passed the exam *reduced* their younger siblings' likelihood of transitioning to secondary school by about 4% (.09 percentage points) – consistent with models of sibling rivalry. After secondary school fees were abolished, older siblings' exam passage *increased* their younger siblings' transition rates by 1% (.07 percentage points). For comparison, these effect sizes represent a meaningful fraction of the descriptive difference in transition rates between places with above- vs. below-median poverty (4.7 pp in the pre-reform period).[3] For a comparison in the literature, a Tanzanian conditional cash transfer program of 3600 Tanzanian shillings per child per month (USD $4 in 2024) raised children's likelihood of ever having attended school by about 4 percentage points (Evans, Gale, & Kosec, 2023).

We show suggestive evidence that sibling rivalry is driven by financial constraints, which are partly alleviated by FSE. Negative pre-reform spillovers are strongest — and positive post-reform spillovers

[2] If passing the exam is interpreted by parents as a signal of student ability, this could affect investments in subsequent children independently of whether the older sibling ends up attending secondary school or not, which would violate the exclusion restriction.

[3] We proxy for poverty using the fraction of households with thatched roofs of grass or leaves – see Section 4.

weakest – in districts with higher poverty rates. We also show in the appendix that positive post-reform spillovers accrue mostly to smaller families. Examining heterogeneity by gender, we find that negative spillovers in the pre-period are driven by sibling pairs involving boys (with no significant effect on younger sisters of girls). This may be because families appear to invest less in girls' education in general: prior to FSE, pass rates on the secondary school entrance exam were 11 percent (7 pp) lower for girls than for boys. Post reform, these gender differences in spillover effects disappear.

We also find evidence consistent with parental preferences for equalization of opportunity across siblings. For cohorts of younger students who sat the qualifying exam after the introduction of FSE, we observe a measure of *ex-ante* student ability: 4th grade exam scores (available data do not extend back far enough to observe this measure for pre-reform cohorts). We find that positive sibling spillovers on pass rates and transition rates are concentrated among lower-ability students. We interpret this as consistent with parents responding to positive older-sibling information shocks by reallocating resources to "compensate" struggling younger siblings.

Our main contribution is to provide empirical evidence that expansions of public school access can have a meaningful impact on the intra-household distribution of human capital investments. Much of the prior work on the demand for schooling focuses on forces largely outside the state's control, including shocks to agricultural productivity and labor market opportunities (Foster & Rosenzweig, 1996; Shah & Steinberg, 2017; Oster & Steinberg, 2013; Jensen, 2012; Atkin, 2016; Glewwe & Jacoby, 2004; Jensen & Miller, 2017). Other work has shown that state-provided school inputs can even crowd out educational inputs from households such as books (Das et al., 2013). Here we show that demand-side human capital investments can complement supply-side policy implemented at scale. Our findings also contribute to the debate over the expansion of (expensive and arguably regressive) free schooling policies, adding another dimension to the argument that a full accounting of the costs and benefits of these policies should consider their effects on siblings (Qureshi, 2018a; Brudevold-Newman, 2021; Crawfurd, 2024).

We also advance the literature on sibling spillovers by helping to reconcile divergent estimates from the developed and developing world. In poorer countries, where resource constraints are more likely to bind, much of the existing literature documents the existence of sibling rivalry. Using Tanzanian household survey data, Morduch (2000) shows a positive correlation between a child's educational attainment and the share of his siblings who are girls, consistent with "rivalry for scarce resources in which parents favor sons." Our causal results (from the pre-reform period) support this interpretation. Akresh et al. (2012) and Giannola (2023) show similarly that parents invest more in high-ability children in Burkina

Faso and India. These negative spillovers are often gendered: Vogl (2013) shows that it is common in sub-Saharan Africa for early marriage and dropout to create competition among sisters, while Shrestha and Palaniswamy (2017) show in Nepal that men's educational opportunities crowd out human capital investment in their sisters. Meanwhile, nearly all studies on developed countries find evidence of positive sibling spillovers (Qureshi, 2018b; Karbownik & Ozek, 2019; Figlio et al., 2023; Zang, Tan, & Cook, 2023).[4] Ours is the first study to examine how sibling spillovers change over time in a given country, and our study period coincides with an important policy for expanding access to public education. After instituting FSE, Tanzania's pattern of sibling spillovers began to resemble that of richer countries. While any number of factors which differ between rich and poor countries might drive the contrasting findings in the literature, our results suggest an important role for the ability of fee-free public education to alleviate financial constraints.

Finally, we contribute more broadly to the literature on sibling spillovers by introducing a novel source of variation in older sibling school attendance: discontinuities in qualifying exam scores. Existing work on sibling spillovers has exploited variation in siblings' college-going (Altmejd et al., 2021), disabilities (Black et al., 2021), school starting age rules (Karbownik & Ozek, 2019; Zang et al., 2023), grade retention policies (Figlio et al., 2023), teen pregnancy (Heissel, 2021), peer quality (Nicoletti & Rabe, 2019), and teacher quality (Qureshi, 2018b). Much of this literature comes from developed countries where secondary schooling is not subject to an entrance exam. However, secondary school qualifying exams with binding thresholds are still common in many parts of the developing world, providing an important potential source of variation in siblings' school attendance that can be used for measuring sibling spillovers in other developing contexts.

2 Context and Data

Prior to the implementation of FSE in 2016, the vast majority of Tanzanian students finished primary school but failed to finish secondary school. Primary school is compulsory from the age of 7 and lasts for seven years (Standards 1-7). These are followed by four years of non-mandatory lower secondary instruction (Forms 1-4) and 2 years of upper secondary (Forms 5-6). Public schools are dominant, comprising 93% of all primary schools and 79% of all secondary schools as of 2016. Net enrollment rates in 2016 stood

[4] While this pattern in the sign of spillovers in the literature is strong, it is not exact: Qureshi (2018a) and Lindskog (2013) find evidence of positive spillovers in Pakistan and Ethiopia, respectively. Meanwhile de Gendre (2022) identifies small negative spillovers of higher class rank on younger siblings' test scores in the Netherlands, though she also finds that higher-ranked older siblings cause greater parental investment in younger siblings.

at 84% for primary but only 24% for secondary, in which Tanzania trails the average completion rate in Sub-Saharan Africa by approximately 10 percentage points (World Bank, 2022). However, returns to secondary school in Tanzania have been estimated at 15% per year, suggesting that many more students could benefit from staying in school longer (Montenegro & Patrinos, 2014).

2.1 Introduction of Free Secondary Education (FSE)

In November 2015, the Tanzanian government approved a Free Secondary Education policy (FSE). Under the new policy, public schooling — free at the primary level since 2001 — would become fee-free through the first four years of secondary school. This entailed the abolition of school fees and examination fees, as well as some school equipment charges. Prior to FSE, yearly secondary school fees were about TZS 20,000 per pupil (equivalent to about USD $25 in 2024 at PPP) (Oxford Business Group, 2019; World Bank, 2022). For comparison, average household consumption expenditure in Tanzania around this time was just over 400,000 TZS (Ministry of Finance and Planning - Poverty Eradication Division (MoFP- PED) [Tanzania Mainland] and National Bureau of Statistics (NBS), 2019). Descriptive evidence suggests that these fees had been binding constraints on many students' transition to secondary school; transition rates rose sharply after the policy's implementation, even though students were still required to pass the qualifying exam in order to enroll in secondary school. Evidence from household surveys suggests strongly that the policy succeeded in its most direct aim: the amount of money households reported spending on school fees for their children in secondary public schools fell dramatically after the policy came into effect (Sandholtz, 2024).

2.2 Administrative data on national standardized exams

The National Examinations Council of Tanzania (NECTA) administers a series of national examinations that determine whether and how students may continue their studies. These tests are graded centrally — not by each school's own teachers — thereby alleviating concerns of systematic manipulation of scores. We use the full universe of administrative data from two of these exams: the Primary School Leaving Examination (PSLE) and the Form Two National Assessment (FTNA).

Primary School Leaving Examination (PSLE): This exam is administered at the end of the seventh and final year of primary school, in September. It is not required, but in practice well over 95% of pupils enrolled in Standard 7 sit the exam. It is a high-stakes test, functioning as a qualifying exam for continuing in the public education system: a passing grade is required to enroll in secondary (i.e., non-vocational) instruction at a government-run school. (Private schools are not required to consider PSLE scores for admission but in practice many do.) Therefore in this paper we use interchangeably the terms "PSLE" and "secondary school entrance exam" or "qualifying exam."

Five subjects are covered in the PSLE: English, Mathematics, Swahili, Social Studies, and Science, with each subject accounting for one-fifth of the possible exam marks (i.e., 50 out of 250 total points). NECTA does not make students' precise marks available, but does provide the associated letter grades for each subject and the overall test. The marks for each subject map onto letter grades according to the following scale: 0-9 marks correspond to an E, 10-19 to a D, 20-29 to a C, 30-39 to a B, and 40-50 to an A. Regarding the overall score, 200 marks is the threshold for an A, whereas B's, C's, and D's correspond to scores of at least 150, 100, and 50 respectively. The passing threshold for the exam is an overall score of 100 marks — a C average. Our data therefore provide a sharp measure of which students passed the PSLE, and a noisy measure of their exam marks (overall and by subject). Our analysis covers 7 PSLE cohorts (2013-2019) totaling 6,068,930 pupils.

Form Two National Assessment (FTNA): Two years after enrolling in secondary education, pupils sit this exam in order to determine whether they can proceed to Form 3. Because administrative data on transition to secondary school is not available at the individual level, we use students' participation in this exam as a proxy for transition to secondary school. Because NECTA had not yet adopted unique student identifiers in the exam data during the time period we study, we match students across exams using their names. About 98% of PSLE takers' names are unique within their cohort.

We proxy for whether a PSLE taker in year t transitioned to secondary school using a binary variable for whether their full name appeared in the list of FTNA takers in year $t + 2$. (We drop the 2% of students with non-unique names, following Sandholtz (2024) and Sandholtz, Gibson, and Crawfurd (2024).) We consider this to be a conservative measure of secondary transition for a number of reasons. First, students who enrolled in secondary school but dropped out before the end of the second year will be missed. Second, students who sat both exams but used differently-spelled names (e.g. middle name vs. middle initial) will be missed. Third, students who skipped or repeated a year will be missed.[5] We confirm that passing the PSLE is a strongly binding constraint on secondary transition: only about 1% of those who fail appear as FTNA takers two years later (Figure A.1). (See cohort-wise summary statistics for these exams

[5]In practice, less than 1% (5%) of PSLE takers in year t match to an FTNA taker's name from year $t + 1$ ($t + 3$), suggesting that normal grade progression is common.

in Table A.1.)

3 Design

Causally identifying sibling spillovers is challenging. Siblings are a type of peer, whose behavior may affect each other mutually. Identifying the effect of one sibling's educational outcomes on those of another requires overcoming the reflection problem using a credible source of exogenous variation (Manski, 1993).

Our strategy for overcoming this problem entails comparing the younger siblings of older students who narrowly passed or failed the PSLE to qualify for secondary school. This allows us to bypass the problem of simultaneity in peer effects by measuring non-contemporaneous outcomes: we are interested in how older siblings' exam passage affects the performance of younger siblings at least one year later. It also affords us a source of quasi-random variation in the older sibling's educational attainment. To carry out this strategy, we must first identify pairs of siblings, and then identify which older siblings are near the qualifying exam passing threshold.

3.1 Sibling Matching

The NECTA administrative data does not contain family identifiers, but we can proxy for them by combining information from students' names with information about their locations.[6] We link students into family groups using shared last names within wards (Swahili: *kata*), the third administrative subdivision. Naming conventions in Tanzania dictate that children typically take their father's last name, meaning that siblings (children of the same father) share a last name. To reduce false positives, we consider a set of potential 'older siblings' as PSLE takers whose last name is unique within their PSLE cohort $\times$ ward (constituting 48% of PSLE takers).[7] (We define a student's ward as the ward of the school in which she sat the PSLE.) We then match these older siblings to the set of individuals in more recent PSLE cohorts who share that last name in the same ward: the 'younger siblings'. We discard family groups consisting of more than six matches (one per year) as likely false positives due to common names. This procedure yields around 1.54 million older-younger sibling pairs. These exclusions mean our analysis is not representative of the entire country – students with more common last names are underrepresented – but permits greater confidence in the precision of sibling matching.

[6]See Cruz, Labonne, and Querubin (2017) for a similar example of inferring family connections through naming patterns.

[7]Due to its high linguistic and ethnic diversity, Tanzania shows a high degree of surname variability. Out of approximately 6 million PSLE takers, we find over 300,000 different last names, of which over 170,000 appear more than once.

We explore alternative sibling matching procedures in Appendix B. We show that our results are qualitatively similar when we match siblings at the school (rather than ward) level, and when we match siblings on middle and last names (rather than just on last names). Our preferred specifications look at outcomes on younger siblings but take as the unit of observation the sibling pair, considering all matched pairs which meet the above criteria. This means that some individuals may appear more than once, for example as the younger sibling to multiple older siblings. Our results are robust to enforcing that each younger sibling appear only once (by excluding all but the closest-spaced sibling pair of each younger sibling).

3.2 Probability of Passing the PSLE

Our identification strategy rests on the assumption that older siblings who barely pass the PSLE are fundamentally similar *ex-ante* to those who barely fail it, meaning that the determination of which of these students have the opportunity to attend secondary school is as good as randomly assigned. Our strategy therefore shares the intuition of a regression discontinuity design, with the important difference that in the absence of precise numerical PSLE scores, there is no clear running variable.

Instead, we use the information on subject letter grades to identify older siblings with grade combinations which correspond to scores near the passing threshold. We first compute the probability of passing the PSLE conditional on each of the 2000 distinct observed combinations of the five subject grades. Many of these are unambiguous: students whose five subject grades are all equal to or higher than C have a 'passing probability' of 1, and pupils receiving any combination of D's and E's have a passing probability of 0. Many other grade combinations, however, reflect scores near enough to the passing threshold that they include both passers and failers. For instance, a pair of students who earned the same combination of C's and D's could still end up on different sides of the passing threshold. Figure A.2 presents the distribution of PSLE score combinations and implied probabilities of passing the exam.

We consider older siblings near the passing threshold to be those whose grade combinations are not determinate of passing status; i.e., those with a passing probability in the interval of]0, 1[. We regard these students near the cutoff as quasi-randomly assigned to the treatment (qualifying for secondary school), employing the idea of local randomization (Cattaneo, Titiunik, & Vazquez-Bare, 2017). Among these students near the passing threshold, passing the PSLE has a large and statistically significant effect on a student's own likelihood of making the transition to secondary school: about 25 percentage points in the pre-reform period, and 43 percentage points after the reform's implementation. Both before and after the reform, students near the passing threshold who fail the PSLE have transition rates of 2% or less. We show in the appendix that these direct effects of passing the PSLE on one's own transition probability – as well as our main results on younger siblings – are robust to narrowing the interval of passing probability (Table A.2, Table A.3).

3.3 Econometric specification

Our estimates of sibling spillover effects in educational achievement are obtained by estimating equations of the form:

$$y_i = \beta PSLEpass_i + \gamma_i + \rho_i + \tau_i + \varepsilon_i \tag{1}$$

where y is one of our binary outcomes of interest: transition to secondary school, passing the PSLE, and achieving a high score on that exam (A or B). $PSLEpass_i$ is a dummy variable that takes value one if the older sibling in pair i passed the PSLE, and zero otherwise; β is the parameter of interest. In order to compare the younger siblings of older siblings who are as similar as possible, we employ a set of extremely granular fixed effects: γ_i captures the fixed effect of the older sibling's subject-specific PSLE score combination, while ρ_i and τ_i capture school and cohort fixed effects, respectively. Our main analyses estimate this equation separately for pre-FSE and post-FSE cohorts. In later robustness analyses, we also extend equation (1) by including controls for the sex of each sibling and the age gap between them. Standard errors are clustered by (older sibling's) score combination (Lee & Card, 2008).

We also show estimates from a dynamic version of equation 1:

$$y_{it} = \sum_{t=2014}^{2019} \beta_t PSLEpass_{it} \times \tau_t + \gamma_i + \rho_i + \tau_t + \varepsilon_{it} \tag{2}$$

As mentioned above, we restrict our primary analysis sample to sibling pairs in which the older sibling's PSLE grade combination could have corresponded to either a passing or failing numerical score (i.e., an implied "probability of passing" strictly between zero and one). Our identification assumption is that — conditional on the older sibling's PSLE subject letter grades, school, and cohort — her passing the PSLE is as good as random.

We test this assumption by showing that characteristics do not differ significantly between the younger siblings of PSLE passers and failers near the threshold (see Table A.4 for more detail).

4 Results

Our main results can be found in Table 1. As outcomes of interest, we consider three complementary margins of educational achievement: enrolling in secondary school (*Transition*), passing the PSLE, and receiving a high score (i.e., an A or B average) on that exam. We find that before FSE, a student whose older sibling passed the exam was 0.9 percentage points less likely to enroll in secondary school than a pupil with a comparable older sibling who did not pass the exam. Moreover, she faced a 1.2 p.p. lower probability of passing the PSLE herself (though we find no effect on the likelihood of obtaining an A or a B – a "high score"). However, after the introduction of FSE, students whose older sibling passed the exam became 0.7 percentage points *more* likely to enroll in secondary schooling, 0.4 percentage points more likely to pass the PSLE, and 0.8 percentage points more likely to obtain a high score compared to their peers whose older sibling failed the exam. Table A.5 estimates these effects in a more structured way, confirming that the effect of older siblings passing the exam on each of these younger sibling outcomes is statistically different across the two periods.

Table 1: Spillover effects from older siblings' achievement

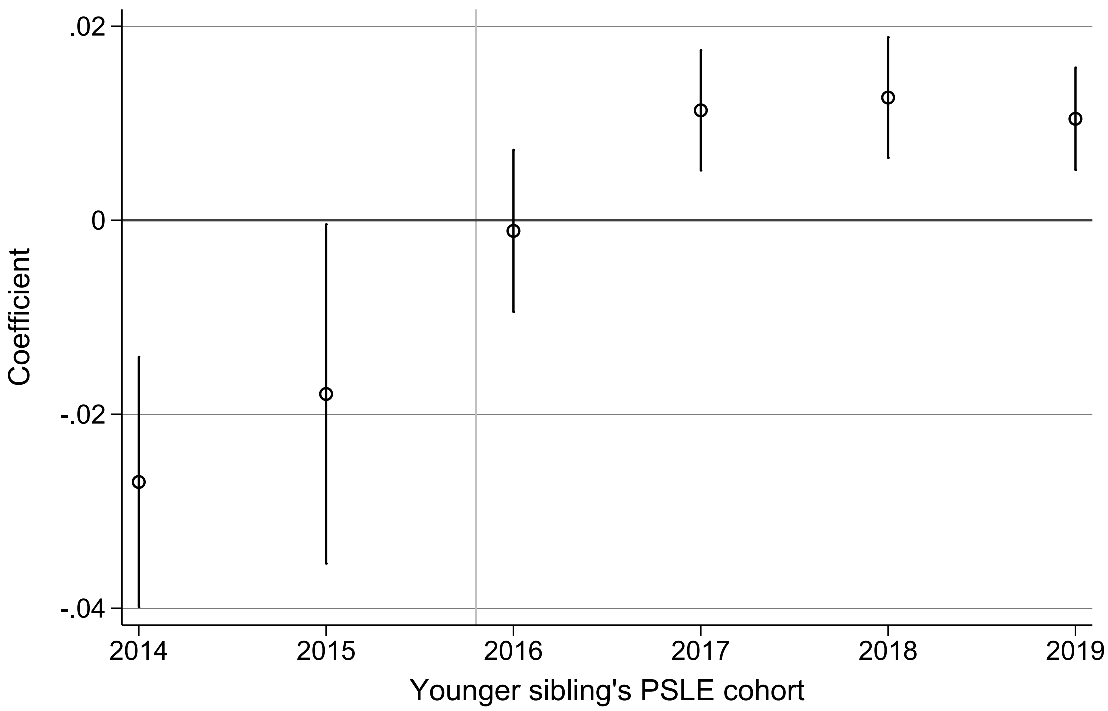
	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.009**	-0.012***	0.004	0.007***	0.004**	0.008***
	(0.005)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
N	84,035	84,730	84,730	517,117	521,738	521,738
Mean dep. var.	0.239	0.622	0.161	0.608	0.754	0.250
R^2	0.292	0.333	0.385	0.176	0.171	0.230
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

To illustrate the temporal changes in the sign of the sibling spillover effects, we plot coefficients from the treatment $\times$ year dummy interactions in (2), with secondary transition as the outcome variable in Figure 1. This shows the estimated spillover effect for each cohort of younger siblings. The estimates are negative (around -2 percentage points) for younger siblings that sat the PSLE in 2014 and 2015, but positive (just over 1 percentage point) for the 2017, 2018, and 2019 PSLE cohorts. We find no significant effect in

2016; this may reflect the fact that the policy was announced only in late 2015 (after that year's cohort had taken the PSLE), leaving little time for adaptation on either the supply or the demand side prior to the policy's intended start in January of 2016. See the first column of Table A.6 for detailed estimates.

Figure 1: Effect of older sibling's qualification for secondary school
on younger sibling's secondary school attendance, by cohort



Notes: Coefficients and 95% confidence intervals from estimates of Equation 2 regressing younger sibling's secondary transition on a dummy for whether their older sibling passed the secondary school qualification exam, conditional on being near the threshold.

4.1 Robustness

Further analyses in the appendix confirm the robustness of these results to the inclusion of controls and to alternative choices over matching and estimation. Results are virtually unchanged when we include controls for (both siblings') gender and the age gap between siblings in the model (see Table A.7). When considering a smaller passing probability interval, the effects remain similar in magnitude (though less precise), as shown in Table A.3. Restricting the sample to only the closest-spaced sibling pair for each younger sibling (the 'most recent' PSLE-taker among one's older siblings) also yields results which are quantitatively very similar to our main findings both before and after the FSE policy was in effect (see

Table A.8). Appendix B shows that our results are broadly similar when using alternate methods for matching sibling pairs.

Finally, there is evidence that Tanzanian schools in this period sometimes strategically excluded low-performing students from being allowed even to sit the PSLE (Cilliers, Mbiti, & Zeitlin, 2021). To test whether our results might be partly explained by changes to the composition of which younger siblings sit the exam, we test whether an older sibling marginally passing the PSLE has an effect on the probability of having any younger sibling sit the exam. We find no evidence in favor of this hypothesis (see Appendix A.1).

4.2 Evidence on mechanisms

We present suggestive evidence that the effects we measure are consistent with FSE alleviating families' resource constraints in a way that allowed them to "compensate" rather than "reinforce" ability differences between children. We do this by testing for heterogeneity in treatment effects by poverty, gender, and students' *ex-ante* ability.[8]

4.2.1 Heterogeneity by poverty

FSE directly reduced the cost of secondary education. It did not reduce the cost all the way to zero – even if the abolition of school fees were perfectly enforced, school fees represented only a part of the cost of school attendance, including the opportunity cost. But its effect was large enough to show up in household survey data as a huge reduction in reported school fees (as well as overall spending) for students in public secondary school (Sandholtz, 2024). For families with a marginal ability to send students to secondary school, this may have been sufficient to change the optimal strategy from "quality" to "quantity" — in other words, to invest in various siblings' education rather than concentrating resources on one promising potential "winner." For families who were still far from being able to send many kids to secondary school even after the abolition of school fees, FSE would have no expected effect on the allocation of educational resources among children. Finding that negative sibling spillovers are concentrated among poorer families – and that positive spillovers after the reform are concentrated among richer ones – would be consistent with the narrative of FSE changing the sign of sibling spillovers by alleviating resource constraints for

[8] We test an alternative mechanism, school selection, in Appendix A.2. In particular, we check whether parents were more or less likely to send their younger children to a different or a 'better' school (as measured by its *ex-ante* PSLE pass rate) when the older sibling passed the PSLE. We find virtually no evidence in favor of the 'school selection' alternative channel, apart from a very small (0.4 p.p.) increase in the post-reform probability of sending a younger child to the same school her older sibling attended when the latter passed the PSLE.

families on the margin. There is recent empirical evidence that in this context saving is difficult, and financial constraints are binding for household educational investments (Carroll et al., 2023; Burchardi et al., 2024).

While our administrative data do not provide family-level measures of socioeconomic status, we can proxy for families' income levels using census data at the district level. We use data on the proportion of homes with grass- or leaf-thatched roofs, reported for each district in the 2012 Population and Housing Census, as a proxy for local poverty. This is a common poverty measure, used for example by means-tested programs elsewhere in the region (Egger, Haushofer, Miguel, Niehaus, & Walker, 2022). Specifically, we code the upper half of the matched sample (as ordered by this variable) as living in 'poorer districts'. We then estimate equation (1) with an additional interaction term between $PSLEpass_i$ and this poverty indicator (the latter variable is not separately included in the estimated model since it is time-invariant for each district and thus collinear with school fixed effects). Results can be found in Table 2.

We find that, before the introduction of FSE, negative sibling spillovers on secondary transition were essentially driven by pupils in poorer districts. Moreover, post-reform spillover effects on both the transition rate and the PSLE pass rate are null for students from poorer districts, while positive in richer parts of the country. While district-level measures of poverty can capture household-level socioeconomic well-being only coarsely, the pattern we observe suggests that a crucial mechanism for the effects we observe overall may be FSE's alleviation of resource constraints. Results remain qualitatively similar when using an alternative measure of poverty based on the fraction of iron-sheet roofs (Table A.9).

Table 2: Heterogeneity by proxy for district-level poverty

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.008	-0.010**	0.013***	0.010***	0.004**	0.010***
	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)
Older sibling passed x Poorer district	-0.008*	-0.007	-0.017***	-0.006**	-0.001	-0.003
	(0.005)	(0.005)	(0.005)	(0.003)	(0.002)	(0.002)
Passed + Passed x Poorer	-0.016***	-0.017***	-0.004	0.004	0.004	0.007***
	(0.005)	(0.005)	(0.006)	(0.002)	(0.002)	(0.002)
N	78,893	79,535	79,535	483,923	488,197	488,197
Mean dep. var.	0.239	0.623	0.161	0.610	0.757	0.252
Mean dep. var. (Richer)	0.259	0.662	0.173	0.650	0.778	0.265
R^2	0.292	0.332	0.384	0.177	0.171	0.231
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. 'Poorer district' is proxied by (being in the upper half of) the proportion of households living in homes where grass/leaves were used as a roof-building material, measured as of the 2012 Population and Housing Census. A 'Poorer district' variable is not included as it would be collinear with school FE. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

A related question is how these effects vary by family size (which in our context, we can only measure as the number of siblings sitting the PSLE during the window we observe). Resource constraints are likely to be more binding when there are more siblings among whom resources may potentially be divided. We show in Table A.10 that the positive sibling spillovers we estimate in the post-reform period are limited to smaller families.[9]

4.2.2 Heterogeneity by gender

We next examine how sibling spillovers vary by gender. In previous descriptive work on sibling rivalry using household survey data from Tanzania, Morduch (2000) shows that children with a greater share of female siblings have higher educational attainment, which he interprets as consistent with a parental preference to invest in sons when resources are scarce. The administrative data we use are descriptively consonant with this interpretation – girls are 8 percentage points less likely to pass the PSLE than boys in the pre-reform period, a gap which narrows to a still-significant 3 percentage points after the reform's introduction. Previous literature has also studied how gender differences may arise for cultural reasons which intersect with economic incentives: for instance, if boys and girls have different responsibilities

[9]The existence of differential effects in urban and rural areas could be of related interest. To test this possibility, we code each ward as 'urban', 'mixed', or 'rural' according to its Census classification. We do not find evidence of differently-sized spillover effects in urban wards as compared to non-urban regions (see Table A.11).

and are thus expected to perform different (gender-specific) tasks at home, sibling spillovers may be more favorable when a boy's older sibling is a girl, and vice-versa, than for same-sex siblings (Lindskog, 2013).

To explore this possibility, we regress our outcomes of interest on an augmented set of independent variables: in particular, we expand equation (1) to include dummies for each sibling's gender, as well as a series of interaction terms between our main independent variable ('Older sibling passed the PSLE'), 'Female,' and 'Female older sibling.' For additional detail regarding these estimations, see Table A.12.

Figure 2 presents the estimated spillovers by gender pair combination. (*Transition* is the dependent variable. See columns (1) and (4) of Table A.12 for detailed results.) We find that negative spillovers pre-FSE appear only for gender pairs involving boys, consistent with boys being the relevant margin of sibling competition for educational resources. In the post-FSE period, positive spillovers appear to affect sibling pairs of all genders more or less equally. This further supports the idea that FSE altered patterns of sibling spillovers by alleviating financial constraints.

Figure 2: Sibling spillover effects on secondary transition, by gender

Notes: Coefficients from regressing younger siblings' secondary school transition on a triple interaction of younger sibling's sex, older sibling's sex, and older sibling's PSLE pass result (conditional on being near the threshold). Confidence intervals at 95% and 90% are shown.

4.2.3 Heterogeneity by Prior Academic Achievement

Finally, we test whether the students who benefit most from their older siblings' academic achievement are those younger siblings with high or low *ex-ante* measures of ability. This is vital for assessing whether parents practice a strategy of "compensating" or "reinforcing."

In order to have a measure of younger siblings' prior academic achievement, we link our PSLE data set to results from the earliest national standardized test students take in primary school: the Standard Four National Assessment (SFNA), taken at the end of students' fourth year of primary school. The SFNA is taken three years before the PSLE; unlike the latter, it does not determine pupils' opportunities of further schooling. But it does provide an *ex-ante* measure of student ability. Unfortunately, SFNA information is only available from 2015 onward, so we are only able to consider younger siblings who sat the PSLE in the post-FSE period – 2018 and 2019 (i.e., those who were part of the 2015 and 2016 SFNA cohorts, respectively). We mirror the matching method used to link PSLE to FTNA, limiting attention to nationally-unique full name matches between individuals who sat the PSLE in year t and pupils who took the SFNA in year $t-3$. We are able to match just short of 50% of our (2018/9) preferred sample to a prior SFNA score.

We estimate an expanded version of equation (1) that includes a binary indicator for whether younger sibling i achieved a low grade (C or worse) on the SFNA, and interacts that indicator with $PSLEpass$, our measure of older siblings' academic achievement.

Table 3: Heterogeneity by younger siblings' prior academic achievement

	(1) Transition	(2) PSLE Pass	(3) High score
Older sibling passed the PSLE	-0.017***	-0.010*	0.019***
	(0.006)	(0.005)	(0.006)
Low ability	-0.220***	-0.183***	-0.359***
	(0.005)	(0.004)	(0.007)
Older sibling passed x Low ability	0.027***	0.025***	-0.018**
	(0.006)	(0.005)	(0.007)
Passed + Passed x Low ability	0.010**	0.016***	0.002
	(0.004)	(0.003)	(0.003)
N	124,310	124,348	124,348
Mean dep. var.	0.639	0.786	0.297
Mean dep. var. (Low ability)	0.593	0.745	0.208
R^2	0.292	0.285	0.405
Older sibling's score FE	Yes	Yes	Yes
School FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Notes: Standard errors are clustered by grade combination. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. 'Low ability' refers to an average grade below B (A-E scale) on the Standard Four National Assessment, an exam taken three years before the PSLE. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. We only have information for the 2015 and 2016 SFNA, which we link to the 2018 and 2019 PSLE cohorts. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table 3 shows that individuals with lower ability benefit more from exposure to an older sibling's achievement: positive spillover effects on the probabilities of primary-to-secondary school transition and of passing the PSLE are driven by students who achieved a grade of C or lower on the SFNA. These results suggest that, post-FSE, sibling spillovers mitigate preexisting differences in ability, which may reflect parents' desire to equalize educational outcomes and/or investments across children.[10]

5 Conclusion

In this paper, we study sibling spillover effects in education before and after a major policy reform: the nationwide abolition of public secondary school tuition fees in Tanzania (Free Secondary Education, or FSE). We draw on the universe of administrative data on national standardized test results. We compare the outcomes of the younger siblings of students who narrowly passed vs. failed the secondary school qualifying exam. Before FSE was implemented, we identify negative spillovers of an older sibling's educational attainment on younger siblings' Primary School Leaving Examination pass rates and secondary enrollment rates. After the introduction of FSE, the sign on both these effects was reversed. These results are robust to a wide variety of alternative estimation choices.

We present suggestive evidence that this change in the sign of sibling spillover effects was due to FSE's alleviation of families' financial constraints. We also show evidence that in the post-reform period, parents preferred to compensate lower-ability children rather than reinforce the advantages of higher-ability children.

Our results show that expansions in public schooling access — which lower the cost of educating children — may allow families on the margin to distribute human capital investments more equitably among children. Existing literature on sibling spillovers tends to find positive effects in developed countries and negative effects in many (though not all) developing countries. Our paper may help reconcile these divergent results, especially in light of existing empirical evidence on the context-specific nature of findings in the peer effects literature more broadly (Sacerdote, 2014). These findings are relevant for assessing the ongoing expansion of access to secondary education in the developing world.

References

Adhvaryu, A., & Nyshadham, A. (2016). Endowments at Birth and Parents' Investments in Children. *The Economic Journal*, *126*(593), 781–820. Retrieved 2024-06-10, from https://www.jstor.org/stable/24738175 (Publisher: [Royal Economic Society, Wiley])

Akresh, R., Bagby, E., de Walque, D., & Kazianga, H. (2012). Child Ability and Household Human Capital Investment Decisions in Burkina Faso. *Economic Development and Cultural Change*, *61*(1), 157–186. Retrieved 2024-05-16, from https://www.journals.uchicago.edu/doi/abs/10.1086/666953 (Publisher: The University of Chicago Press) doi: 10.1086/666953

[10]This also implies at least a minimal level of parental knowledge about children's relative abilities, despite evidence that parents' beliefs about their children's abilities are often inaccurate (Dizon-Ross, 2019; Cunha, Elo, & Culhane, 2020).

Altmejd, A., Barrios-Fernández, A., Drlje, M., Goodman, J., Hurwitz, M., Kovac, D., ... Smith, J. (2021). O brother, where start thou? Sibling spillovers on college and major choice in four countries. *The Quarterly Journal of Economics*, *136*(3), 1831–1886.

Atkin, D. (2016). Endogenous Skill Acquisition and Export Manufacturing in Mexico. *American Economic Review*, *106*(8), 2046–2085. Retrieved 2024-06-18, from https://www.aeaweb.org/articles?id=10.1257/aer.20120901 doi: 10.1257/aer.20120901

Banerjee, A. V., & Duflo, E. (2012). *Poor economics: a radical rethinking of the way to fight global poverty* (Paperback ed.). New York: PublicAffairs.

Barrera-Osorio, F., Bertrand, M., Linden, L. L., & Perez-Calle, F. (2008). Conditional cash transfers in education: design features, peer and sibling effects Evidence from a randomized experiment in Colombia. *NBER Working Paper no. 13890*.

Becker, G. S., & Tomes, N. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, *84*(4), S143–S162. Retrieved 2024-06-05, from https://www.jstor.org/stable/1831106 (Publisher: University of Chicago Press)

Behrman, J. R., Pollak, R. A., & Taubman, P. (1982). Parental Preferences and Provision for Progeny. *Journal of Political Economy*, *90*(1), 52–73. Retrieved 2024-06-10, from https://www.jstor.org/stable/1831229 (Publisher: University of Chicago Press)

Berry, J., Dizon-Ross, R., & Jagnani, M. (2024). Not Playing Favorites: Parents and the Value of Equal Opportunity.

Black, S. E., Breining, S., Figlio, D. N., Guryan, J., Karbownik, K., Nielsen, H. S., ... Simonsen, M. (2021). Sibling Spillovers. *The Economic Journal*, *131*(633), 101–128. Retrieved 2024-06-08, from https://academic.oup.com/ej/article/131/633/101/5890836 doi: 10.1093/ej/ueaa074

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., & Sandefur, J. (2018). Experimental evidence on scaling up education reforms in Kenya. *Journal of Public Economics*, *168*, 1–20. Retrieved from https://doi.org/10.1016/j.jpubeco.2018.08.007 doi: 10.1016/j.jpubeco.2018.08.007

Brudevold-Newman, A. (2021). Expanding access to secondary education: Evidence from a fee reduction and capacity expansion policy in Kenya. *Economics of Education Review*, *83*, 102127.

Burchardi, K., de Quidt, J., Gulesci, S., & Sulaiman, M. (2024). Borrowing Constraints and Demand for Remedial Education: Evidence from Tanzania. *The Economic Journal*, ueae024. Retrieved 2024-06-18, from https://doi.org/10.1093/ej/ueae024 doi: 10.1093/ej/ueae024

Carroll, P. P., Myamba, F., Nielson, D. L., Price, J., Roessler, P., & Sandholtz, W. A. (2023). Priming the pump: Can paying interest upfront increase savings?

Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2017). Comparing inference approaches for rd designs: A reexamination of the effect of head start on child mortality. *Journal of Policy Analysis and Management*, *36*(3), 643–681.

Cilliers, J., Mbiti, I. M., & Zeitlin, A. (2021). Can Public Rankings Improve School Performance? *Journal of Human Resources*, 0119–9969R1. doi: 10.3368/jhr.56.3.0119-9969r1

Crawfurd, L. (2024). Feasibility first: Expanding access before fixing learning. *International Journal of Educational Development*, *104*(November 2023), 102949. Retrieved from https://doi.org/10.1016/j.ijedudev.2023.102949 doi: 10.1016/j.ijedudev.2023.102949

Cruz, C., Labonne, J., & Querubin, P. (2017). Politician family networks and electoral outcomes: Evidence from the Philippines. *American Economic Review*, *107*(10), 3006–37.

Cunha, F., Elo, I., & Culhane, J. (2020). Maternal subjective expectations about the technology of skill formation predict investments in children one year later. *Journal of Econometrics*.

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., & Sundararaman, V. (2013). School Inputs, Household Substitution, and Test Scores. *American Economic Journal: Applied Economics*, *5*(2), 29–57.

de Gendre, A. (2022). *Class rank and sibling spillover effects* (Working Paper). University of Melbourne.

Dizon-Ross, R. (2019). Parents' Beliefs about Their Children's Academic Ability: Implications for Educational Investments. *American Economic Review*, *109*(8), 2728–2765. Retrieved 2024-06-08, from https://pubs.aeaweb.org/doi/10.1257/aer.20171172 doi: 10.1257/aer.20171172

Duflo, E., Dupas, P., & Kremer, M. (2017). The Impact of Free Secondary Education: Experimental Evidence from Ghana. (1254167), 1–92. Retrieved from https://web.stanford.edu/{~}pdupas/DDK{_}GhanaScholarships.pdf

Egger, D., Haushofer, J., Miguel, E., Niehaus, P., & Walker, M. (2022). General Equilibrium Effects of Cash Transfers: Experimental Evidence From Kenya. *Econometrica*, *90*(6), 2603–2643. Retrieved 2024-05-08, from https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA17945 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA17945) doi: 10.3982/ECTA17945

Evans, D. K., Gale, C., & Kosec, K. (2023, February). The educational impacts of cash transfers in Tanzania. *Economics of Education Review*, *92*, 102332. Retrieved 2024-07-12, from https://www.sciencedirect.com/science/article/pii/S0272775722001054 doi: 10.1016/j.econedurev.2022.102332

Figlio, D. N., Karbownik, K., & Özek, U. (2023). *Sibling Spillovers May Enhance the Efficacy of Targeted School Policies* [Working Paper]. National Bureau of Economic Research. Retrieved 2024-06-08, from https://www.nber.org/papers/w31406 doi: 10.3386/w31406

Foster, A. D., & Rosenzweig, M. R. (1996). Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution. *The American Economic Review*, *86*(4), 931–953. Retrieved 2024-06-18, from https://www.jstor.org/stable/2118312 (Publisher: American Economic Association)

Garlick, R. (2019). Quality-Quantity Tradeoffs in Pricing Public Secondary Education. *SSRN Electronic Journal*, 1–66. doi: 10.2139/ssrn.3453239

Giannola, M. (2023). Parental Investments and Intra-Household Inequality in Child Human Capital: Evidence from a Survey Experiment. *The Economic Journal*(16727), 1–57.

Glewwe, P., & Jacoby, H. G. (2004). Economic growth and the demand for education: is there a wealth effect? *Journal of Development Economics*, *74*(1), 33–51. Retrieved 2024-06-18, from https://www.sciencedirect.com/science/article/pii/S0304387803001792 doi: 10.1016/j.jdeveco.2003.12.003

Gruijters, R., Abango, M. A., & Casely-Hayford, L. (2023). Secondary School Fee Abolition in Sub-Saharan Africa: Taking Stock of the Evidence.

Heissel, J. A. (2021). Teen Fertility and Siblings' Outcomes: Evidence of Family Spillovers Using Matched Samples. *Journal of Human Resources*, *56*(1), 40–72. Retrieved 2024-06-13, from https://jhr.uwpress.org/content/56/1/40 (Publisher: University of Wisconsin Press Section: Articles) doi: 10.3368/jhr.56.1.0218-9341R2

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *Quarterly Journal of Economics*, *127*(2), 753–792. Retrieved from https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjs002 doi: 10.1093/qje/qjs002

Jensen, R., & Miller, N. H. (2017). *Keepin' 'em Down on the Farm: Migration and Strategic Investment in Children's Schooling* [Working Paper]. National Bureau of Economic Research. Retrieved 2024-06-13, from https://www.nber.org/papers/w23122 doi: 10.3386/w23122

Karbownik, K., & Ozek, U. (2019). Setting a good example? Examining sibling spillovers in educational achievement using a regression discontinuity design. *NBER Working Paper no. 26411*.

Kippenberg, J. (2014). *Tanzania: Let's tackle the primary school leaving exam. Daily News (Tanzania)*. (May 13. Last accessed on December 14, 2022. Retrieved from https://www.hrw.org/news/2014/05/13/tanzania-lets-tackle-primary-school-leaving-exam)

Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, *142*(2), 655-674.

Lindskog, A. (2013). The effect of siblings' education on school-entry in the Ethiopian highlands. *Economics of Education Review*, *34*, 45–68.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531–542.

Ministry of Finance and Planning - Poverty Eradication Division (MoFP- PED) [Tanzania Mainland] and National Bureau of Statistics (NBS). (2019, dec). *Tanzania mainland household budget survey 2017-18, key indicators report.* Retrieved from https://www.nbs.go.tz/nbs/takwimu/hbs/2017_18_HBS_Key

_Indicators_Report_Engl.pdf (Dodoma, Tanzania. Retrieved 16 July 2024.)
Montenegro, C. E., & Patrinos, H. A. (2014). Comparable estimates of returns to schooling around the world. *World Bank policy research working paper*(7020).
Morduch, J. (2000). Sibling rivalry in Africa. *American Economic Review*, *90*(2), 405–409.
Nicoletti, C., & Rabe, B. (2019). Sibling spillover effects in school achievement. *Journal of Applied Econometrics*, *34*(4), 482–501.
Oster, E., & Steinberg, B. M. (2013). Do IT service centers promote school enrollment? Evidence from India. *Journal of Development Economics*, *104*, 123–135. Retrieved 2024-06-18, from https://www.sciencedirect.com/science/article/pii/S0304387813000837 doi: 10.1016/j.jdeveco.2013.05.006
Oxford Business Group. (2019). *Enrolment rates rise in Tanzania with free secondary education.* (Last accessed on December 4, 2022. Retrieved from https://oxfordbusinessgroup.com/overview/smart-adaptations-introduction-free-secondary-education-has-boosted-enrolment-rates-while-private)
Pew Research Center. (2012, dec). *Table: Religious composition by country.* PDF. Retrieved from https://assets.pewresearch.org/wp-content/uploads/sites/11/2012/12/globalReligion-tables.pdf (The Global Religious Landscape. Pew Research Center. 18 December 2012. p. 50. Retrieved 16 July 2024.)
Qureshi, J. A. (2018a). Additional returns to investing in girls' education: Impact on younger sibling human capital. *The Economic Journal*, *128*(616), 3285–3319.
Qureshi, J. A. (2018b). Siblings, teachers, and spillovers on academic achievement. *Journal of Human Resources*, *53*(1), 272–297.
Rosenzweig, M. R., & Schultz, T. P. (1982). Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India. *The American Economic Review*, *72*(4), 803–815. Retrieved 2024-06-18, from https://www.jstor.org/stable/1810018 (Publisher: American Economic Association)
Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: two steps forward? *Annual Review of Economics*, *6*(1), 253–272.
Sandholtz, W. A. (2024). Secondary school access raises primary school achievement [Working Paper].
Sandholtz, W. A., Gibson, C., & Crawfurd, L. (2024). Catholic schools add value: Evidence from lesotho, tanzania, and uganda. *Working Paper*.
Shah, M., & Steinberg, B. M. (2017). Drought of Opportunities: Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital. *Journal of Political Economy*, *125*(2), 527–561. doi: 10.1086/690828
Shrestha, S. A., & Palaniswamy, N. (2017). Sibling rivalry and gender gap: intrahousehold substitution of male and female educational investments from male migration prospects. *Journal of Population Economics*, *30*(4), 1355–1380.
Vogl, T. S. (2013). Marriage Institutions and Sibling Competition: Evidence from South Asia*. *The Quarterly Journal of Economics*, *128*(3), 1017–1072. Retrieved 2024-06-14, from https://academic.oup.com/qje/article/128/3/1017/1850653 doi: 10.1093/qje/qjt011
World Bank. (2022). *World development indicators.* (Data retrieved from World Development Indicators, https://data.worldbank.org/indicator/PA.NUS.PPP?locations=TZ, https://data.worldbank.org/indicator/SP.POP.0014.TO, https://data.worldbank.org/indicator/SP.POP.1519.MA, https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?locations=TZ, https://data.worldbank.org/indicator/SE.SEC.CMPT.LO.ZS?locations=ZG-TZ, and https://data.worldbank.org/indicator/SP.POP.1519.FE)
Zang, E., Tan, P. L., & Cook, P. J. (2023). Sibling Spillovers: Having an Academically Successful Older Sibling May Be More Important for Children in Disadvantaged Families. *American Journal of Sociology*, *128*(5), 1529–1571. Retrieved 2024-06-08, from https://www.journals.uchicago.edu/doi/full/10.1086/724723 (Publisher: The University of Chicago Press) doi: 10.1086/724723

A Supplementary Figures and Tables

Figure A.1: PSLE score and transition to secondary school

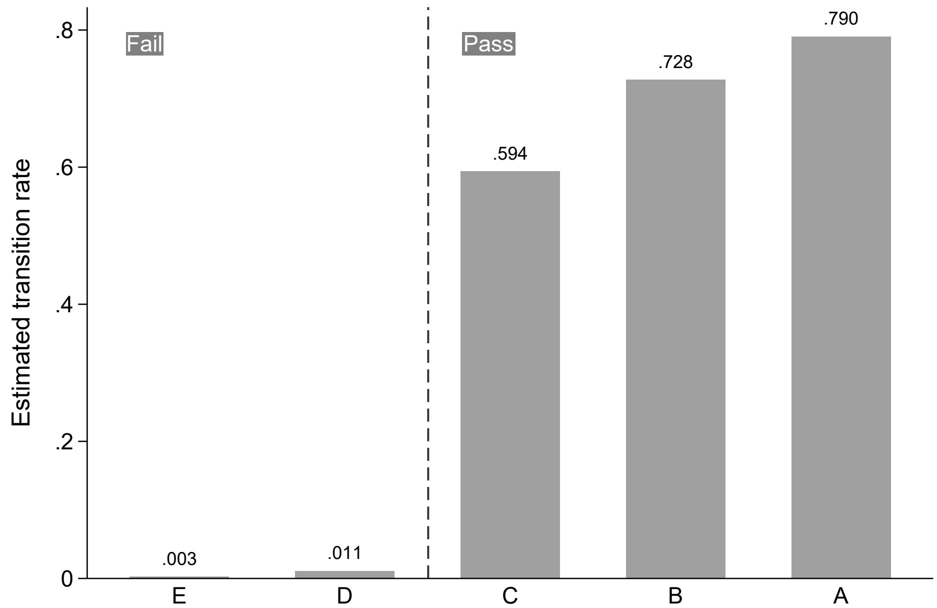


Figure A.2: PSLE letter grade combination and probability of passing

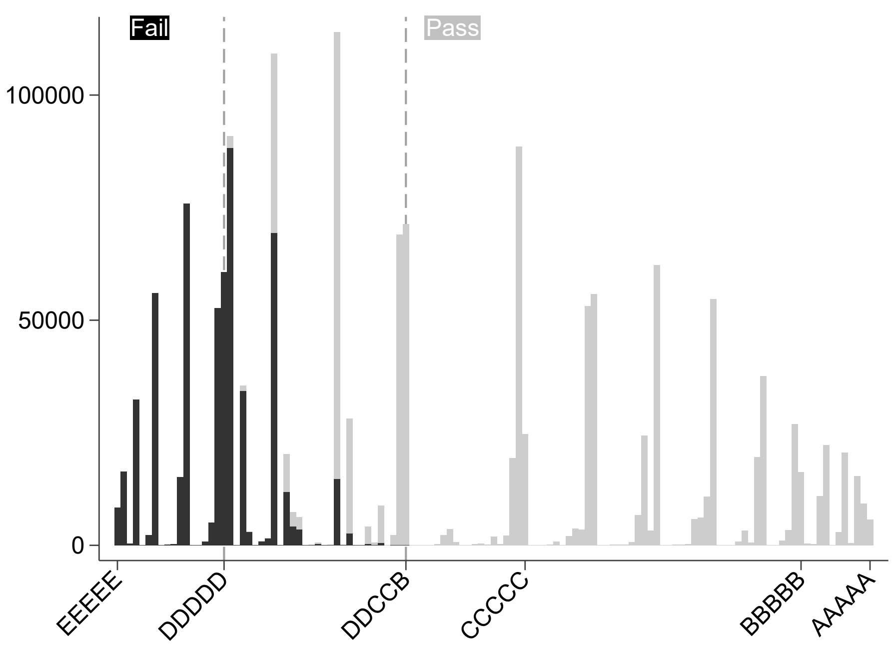


Table A.1: Summary statistics

	2013	2014	2015	2016	2017	2018	2019
Primary schools							
Number of schools	15,656	15,867	16,096	16,350	16,575	16,826	17,047
Public schools (proportion)	0.962	0.964	0.961	0.955	0.952	0.945	0.940
PSLE							
Number of exam takers	867,983	808,085	775,273	795,740	916,885	957,893	947,071
Female (proportion)	0.525	0.532	0.534	0.531	0.528	0.525	0.524
Pass rate (proportion)	0.493	0.559	0.668	0.698	0.722	0.766	0.803
FTNA							
Transition rate (proportion)	0.161	0.177	0.270	0.453	0.620	0.640	0.654
Preferred sample							
Number of older siblings	74,830	76,419	53,468	51,779	36,544	16,310	—
Number of younger siblings	—	27,810	55,893	73,715	104,672	125,631	135,415
Urban (proportion)	—	0.187	0.186	0.196	0.182	0.181	0.185

Notes: the proxy for primary-to-secondary-school transition (obtained using FTNA data) is explained in Subsection 2.2. Our preferred sample is created through ward-level sibling matching, as described in Section 3. 'Number of older [younger] siblings' is a count of individuals, not sibships.

Table A.2: Effect of passing PSLE on student's own secondary transition probability

	Pre-FSE		Post-FSE	
Passed PSLE	0.250***	0.266***	0.433***	0.426***
	(0.007)	(0.015)	(0.011)	(0.017)
N	26,213	3,617	198,942	44,981
Mean, PSLE pass = 0	0.009	0.013	0.014	0.020
Pass prob. window	]0,1[	]0.3,0.7[	]0,1[	]0.3,0.7[
R^2	0.433	0.550	0.400	0.478
PSLE score FE	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. Outcome measure – student's own transition to secondary school – is only measured for students whose full name is unique across the country within their cohort. All regressions include school and grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each student. The unit of observation is a unique 'Pass prob. window' the sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.3: Spillover effects from older siblings' achievement; passing probability interval]0.3, 0.7[

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.007	-0.022***	0.000	0.004**	0.002	0.009***
	(0.011)	(0.005)	(0.007)	(0.002)	(0.002)	(0.002)
N	16,695	16,848	16,848	120,160	121,259	121,259
Mean dep. var.	0.237	0.627	0.149	0.606	0.751	0.244
R^2	0.422	0.444	0.480	0.243	0.235	0.290
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0.3 and 0.7. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.4: Balance: "Effect" of older siblings' achievement on younger siblings' characteristics

	Pre-FSE			Post-FSE				
	Female	Muslim (inferred)	Name length	Female	Muslim (inferred)	Name length	Not missing SFNA	Low grade SFNA
Old sib pass PSLE	-0.009	0.002	-0.047	-0.001	-0.001	0.007	0.004	-0.004
	(0.006)	(0.005)	(0.041)	(0.002)	(0.002)	(0.014)	(0.003)	(0.003)
N	85,221	85,221	85,221	523,447	523,447	523,447	319,746	124,348
Mean dep. var.	0.544	0.197	20.949	0.532	0.197	20.921	0.393	0.768
R^2	0.161	0.318	0.249	0.047	0.217	0.136	0.238	0.374
Old sib score FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. 'Muslim (inferred)' means the student had at least one name from a list of common Muslim names. 'Not missing SFNA' is defined for younger siblings in PSLE cohorts 2018 and 2019 (the start of data availability), and indicates that the student was able to be matched to a Standard Four National Assessment result from the fourth year of primary school. 'Low Grade SFNA' is only defined for students who are matched to an SFNA score, and means the student scored a C or lower on the SFNA. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.4 compares time invariant characteristics of students whose older sibling was on one side or the other of the PSLE passing threshold. While our administrative data do not provide much data about individual characteristics, we are able to make some inferences. The data does provide information about students' sex. The data also provides student names, which we can use to infer religious affiliation by comparing them against a list of common Muslim names. This conservative inference method identifies

20% of students as having Muslim names in both the pre- and the post- period. (About 35 % of Tanzania's population is Muslim (Pew Research Center, 2012).) We can also test for differences in the number of characters in a student's name – this is important because our main outcome measure, Transition to secondary school, relies on matching students by name. Finally, and perhaps most importantly, we are able to measure younger students' *ex-ante* ability using their results from the Standard Four National Assessment (SFNA), taken in the fourth year of primary school. These results are only available from 2015 onward (corresponding in our data to students who ended up sitting the PSLE in 2018 and 2019). We first test for differences in whether a student can be matched to an SFNA score at all, then for differences in scores for those who match. As expected, we find no significant differences across the older sibling's PSLE passing threshold on any of these characteristics, suggesting that variation in exam passing near the threshold can be considered quasi-random, and that older siblings passing the PSLE does not alter the composition of younger siblings who sit the exam.

Table A.5: Spillover effects from older siblings' achievement, interacted with pre-post dummy

	(1) Transition	(2) PSLE Pass	(3) High score
Older sibling passed PSLE	-0.021***	0.015***	0.004*
	(0.008)	(0.005)	(0.002)
Older sibling passed PSLE × Post-FSE	0.030***	-0.015***	0.005**
	(0.008)	(0.005)	(0.002)
Passed + Passed x Post-FSE	0.009***	0.000	0.009***
	(0.002)	(0.002)	(0.001)
N	603,143	608,450	608,450
Mean dep. var.	0.555	0.736	0.238
R^2	0.215	0.172	0.220
Older sibling's score FE	Yes	Yes	Yes
School FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Post-FSE' is a dummy for whether the year is 2016 or later. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.6: Spillover effects from older siblings' achievement (year-by-year)

	(1)	(2)	(3)
Matching:	Ward level	School level	Middle name
Younger sibling's PSLE cohort = 2015	0.093***	0.095***	0.098***
	(0.004)	(0.005)	(0.008)
Younger sibling's PSLE cohort = 2016	0.261***	0.259***	0.264***
	(0.006)	(0.005)	(0.008)
Younger sibling's PSLE cohort = 2017	0.423***	0.409***	0.392***
	(0.008)	(0.009)	(0.008)
Younger sibling's PSLE cohort = 2018	0.450***	0.438***	0.416***
	(0.007)	(0.009)	(0.007)
Younger sibling's PSLE cohort = 2019	0.465***	0.457***	0.436***
	(0.007)	(0.009)	(0.008)
Older sibling passed the PSLE x (Younger's cohort = 2014)	-0.027***	-0.021**	-0.027***
	(0.007)	(0.009)	(0.009)
Older sibling passed the PSLE x (Younger's cohort = 2015)	-0.018**	-0.020*	-0.007
	(0.009)	(0.010)	(0.006)
Older sibling passed the PSLE x (Younger's cohort = 2016)	-0.001	0.002	-0.000
	(0.004)	(0.004)	(0.006)
Older sibling passed the PSLE x (Younger's cohort = 2017)	0.011***	0.014***	0.014***
	(0.003)	(0.003)	(0.005)
Older sibling passed the PSLE x (Younger's cohort = 2018)	0.013***	0.011***	0.012***
	(0.003)	(0.003)	(0.005)
Older sibling passed the PSLE x (Younger's cohort = 2019)	0.010***	0.009**	0.008
	(0.003)	(0.003)	(0.005)
N	603143	668300	329487
Mean dep. var.	0.555	0.539	0.568
R^2	0.215	0.201	0.200
Older sibling's score FE	Yes	Yes	Yes
School FE	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition', the dependent variable, is only measured for students whose full name is unique across the country within their cohort. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.7: Spillover effects from older siblings' achievement (with controls)

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.010**	-0.013***	0.004	0.008***	0.004***	0.008***
	(0.005)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Female older sibling	-0.001	-0.001	0.006**	-0.002	0.001	0.004***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)
Female	-0.012***	-0.078***	-0.069***	0.011***	-0.029***	-0.081***
	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)	(0.001)
Age gap	-0.002	0.000	0.002	0.004***	0.003***	0.003***
	(0.004)	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)
N	84,035	84,730	84,730	517,117	521,738	521,738
Mean dep. var.	0.239	0.622	0.161	0.608	0.754	0.250
R^2	0.292	0.338	0.393	0.176	0.172	0.238
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. 'Female' and 'Female older sibling' are dummies for the gender of each sibling. 'Age gap' is proxied by the difference in years between siblings' PSLE cohorts. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.8: Spillover effects from older siblings' achievement (Most recent older sibling)

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.011**	-0.013***	0.006	0.005**	0.002	0.005**
	(0.005)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
N	77,965	78,585	78,585	330,736	333,335	333,335
Mean dep. var.	0.236	0.619	0.158	0.597	0.745	0.239
R^2	0.290	0.334	0.383	0.184	0.177	0.234
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes the most recent older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.9: Heterogeneity by proxy for district-level poverty

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.016***	-0.013**	-0.006	0.003	0.003	0.007***
	(0.006)	(0.006)	(0.005)	(0.002)	(0.002)	(0.002)
Older sibling passed x Richer district	0.008	0.000	0.020***	0.007**	0.003	0.003
	(0.006)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)
Passed + Passed x Richer	-0.009	-0.013***	0.013***	0.010***	0.005**	0.010***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.002)	(0.002)
N	78,893	79,535	79,535	483,923	488,197	488,197
Mean dep. var.	0.239	0.623	0.161	0.610	0.757	0.252
Mean dep. var. (Poorer)	0.221	0.571	0.147	0.553	0.719	0.231
R^2	0.292	0.332	0.384	0.177	0.171	0.231
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. 'Richer district' is proxied by (being in the upper half of) the proportion of households living in homes where iron sheet was used as a roof-building material, measured as of the 2012 Population and Housing Census. A 'Richer district' variable is not included as it would be collinear with school FE. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.10: Spillover effects from older siblings' achievement (Heterogeneity by sibship size)

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.003	0.001	0.016*	0.018***	0.015***	0.016***
	(0.007)	(0.009)	(0.008)	(0.004)	(0.004)	(0.003)
Sibship size	0.002***	-0.001	0.002	0.001	-0.000	0.001***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)
Older sibling passed x Sibship size	-0.002	-0.004*	-0.004	-0.003***	-0.003***	-0.002***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
N	84,035	84,730	84,730	517,117	521,738	521,738
Mean dep. var.	0.239	0.622	0.161	0.608	0.754	0.250
Mean sibship size	3.242	3.248	3.248	3.296	3.302	3.302
R^2	0.292	0.333	0.385	0.176	0.171	0.230
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.11: Spillover effects from older siblings' achievement (urban/rural heterogeneity)

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.009*	-0.010**	0.001	0.008***	0.004**	0.007***
	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)
Older sibling passed x Urban	-0.001	0.001	0.011	-0.002	-0.004*	0.003
	(0.006)	(0.005)	(0.008)	(0.003)	(0.002)	(0.003)
Passed + Passed x Urban	-0.010	-0.010**	0.012	0.006*	0.001	0.010***
	(0.006)	(0.005)	(0.007)	(0.003)	(0.002)	(0.004)
N	80,240	80,892	80,892	494,341	498,685	498,685
Mean dep. var.	0.239	0.622	0.162	0.607	0.755	0.252
Mean dep. var. (Urban)	0.259	0.738	0.221	0.750	0.848	0.330
R^2	0.291	0.333	0.388	0.175	0.170	0.230
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. Since school fixed effects are always included in our models and given that we classify wards according to the 2012 Census definition (i.e., ward status is time-invariant), we do not include a dummy for ward status ('Urban') in the estimated equations. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

Table A.12: Spillover effects from older siblings' achievement (Heterogeneity by older sibling's gender)

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Old sib passed PSLE	-0.020**	-0.017**	0.007	0.006	0.004	0.009**
	(0.008)	(0.007)	(0.005)	(0.004)	(0.003)	(0.004)
Female	-0.018***	-0.082***	-0.059***	0.002	-0.037***	-0.077***
	(0.006)	(0.007)	(0.006)	(0.004)	(0.004)	(0.004)
Old sib passed PSLE × Female	0.005	0.003	-0.007	0.002	0.002	-0.008*
	(0.009)	(0.011)	(0.007)	(0.005)	(0.006)	(0.005)
Old sib female	-0.006	0.003	0.010**	-0.008**	-0.003	0.002
	(0.007)	(0.006)	(0.004)	(0.004)	(0.004)	(0.003)
Old sib passed PSLE × Old sib female	0.003	-0.014	0.005	0.000	-0.004	0.004
	(0.010)	(0.010)	(0.006)	(0.005)	(0.005)	(0.005)
Female × Old sib female	-0.001	-0.010	-0.008	0.011	0.009	-0.002
	(0.008)	(0.009)	(0.005)	(0.007)	(0.006)	(0.005)
Old sib passed PSLE × Female × Old sib female	0.020	0.033**	-0.006	0.001	0.005	0.004
	(0.014)	(0.015)	(0.008)	(0.009)	(0.007)	(0.006)
Boys with older brothers	-0.020**	-0.017**	0.007	0.006	0.004	0.009**
	(0.008)	(0.007)	(0.005)	(0.004)	(0.003)	(0.004)
Boys with older sisters	-0.015**	-0.014*	-0.000	0.008**	0.006	0.001
	(0.006)	(0.008)	(0.007)	(0.004)	(0.004)	(0.003)
Girls with older brothers	-0.017***	-0.031***	0.012**	0.006*	-0.000	0.013***
	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)	(0.002)
Girls with older sisters	0.008	0.005	-0.002	0.010***	0.007***	0.009***
	(0.006)	(0.006)	(0.005)	(0.002)	(0.002)	(0.002)
N	84,035	84,730	84,730	517,117	521,738	521,738
Mean dep. var.	0.239	0.622	0.161	0.608	0.754	0.250
R^2	0.292	0.339	0.393	0.176	0.172	0.238
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

A.1 Extensive Margin

Table A.13 tests whether passing the PSLE affects the likelihood that younger siblings sit the PSLE. This could be a concern given evidence of schools strategically excluding some poorly performing students from sitting the exam (Cilliers et al., 2021). To analyze this possibility, we examine the sample of PSLE takers from years 2013-2018 whose last name is unique within ward×cohort. We match each cohort of these students by last name within ward to all younger cohorts of PSLE takers whose last name is unique within ward×cohort (2014-2019). (This is to limit/minimize false positives.) We then create outcome variables from these merges: a dummy for whether anyone with the same last name sat the PSLE in the same ward in the following year; a dummy for whether anyone with the same last name sat the PSLE in the same ward in any subsequent year; and the number of people with the same last name who sat the PSLE in the same ward in all subsequent years. As before, we regress these outcomes on a dummy for exam passage, considering only students with grade combinations indicating ambiguous pass status, and including fixed effects for year, school, and grade combination.

Table A.13: Extensive margin: effect of passing PSLE on whether younger siblings sit PSLE

	Pre-FSE			Post-FSE		
	Sibling sat PSLE in following yr	Any younger sib sat PSLE	Num. younger sibs sat PSLE	Sibling sat PSLE in following yr	Any younger sib sat PSLE	Num. younger sibs sat PSLE
Pass PSLE	-0.002	-0.001	-0.002	-0.000	-0.000	0.002
	(0.002)	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)
N	510,012	510,012	510,012	405,802	405,802	405,802
Mean dep. var.	0.138	0.458	0.724	0.150	0.274	0.324
R^2	0.043	0.079	0.084	0.052	0.097	0.110
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. Unit of observation is a PSLE taker. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the student got an A or B average on the PSLE, and zero otherwise. All regressions include grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort. The sample includes all students from year 2013-2018 whose last name is unique within their ward × cohort, whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

We find no evidence that passing the PSLE on the margin changed the likelihood of younger siblings taking the PSLE.

A.2 School selection

Table A.14 tests whether an older sibling passing the PSLE affects the quality of schools at which younger students sit the PSLE. Existing literature suggests this could be a meaningful mechanism in some contexts; school quality appears to have been an important channel for the sibling spillovers measured by Figlio et al. (2023).

We create three measures of households' selection into higher-quality schools: 1) A dummy for whether the younger sibling sat the PSLE at the same school as the older siblings; 2) the *ex-ante* PSLE pass rate of the younger sibling's school (measured in 2013, prior to any of our sample of younger siblings sitting the exam to avoid endogeneity)[11]; and 3) a dummy for whether the younger sibling sat the PSLE at a higher-*ex-ante*-pass-rate school than her older sibling.

[11] School-level correlations between pass rate percentile (within the national distribution) from one year to the next are above .5 for all years in our data.

Table A.14: School selection: effect of passing PSLE on quality of schools at which younger siblings sit PSLE

	Pre-FSE			Post-FSE		
	Different school	Ex-ante PSLE pass rate	School pass rate higher than older sib's	Different school	Ex-ante PSLE pass rate	School pass rate higher than older sib's
Older sibling passed the PSLE	0.010	0.001	0.002	-0.004**	-0.001	-0.001
	(0.007)	(0.001)	(0.006)	(0.002)	(0.001)	(0.002)
N	84,999	82,003	81,939	521,583	436,585	433,651
Mean dep. var.	0.450	0.488	0.237	0.482	0.487	0.232
R^2	0.303	0.719	0.336	0.231	0.675	0.250
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
Older sibling's school FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. All regressions include older sibling's school and grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

We find little evidence that an older sibling passing the PSLE changed school selection for younger siblings, either before or after the reform. In the post period, parents appear slightly more likely to send younger siblings to the same school as an older sibling in response to the older sibling (marginally) passing the PSLE at that school, which may be interpreted as positive school selection. The effect is small, however: less than one percent (0.4 percentage points from a base of 48%).

B Alternative Sibling Matching

Besides our preferred (ward-level) sibling matching procedure, we test two alternative methods. First, we repeat the matching procedure but search for within-school matches (instead of within-ward). Additionally, we perform the initial matching algorithm but look for last- *and* middle-name matches in the same ward, exploiting a Tanzanian naming convention by which children of the same father have the same middle and last names. While this method has the potential to minimize mismatches, it is only possible for sibling pairs in which both siblings' full name (first, middle, last) is present in the data; many students' records feature only a middle initial. This matching method therefore reduces the sample size dramatically.

Figures B.1 and B.2 show the estimated (spillover) effect of older sibling's qualification for secondary school on younger sibling's secondary school attendance (i.e., *Transition*), by cohort, under the school-level and middle-name matching procedures, respectively. The evolution of the coefficients over time is remarkably similar to that displayed in Figure 1 (for a more detailed analysis, compare the estimates in the second and third columns of Table A.6 with those in the first column).

Meanwhile, Tables B.1 and B.2 present the aggregated (i.e., pre- and post-FSE) estimates under the two alternative matching methods – these results compare with those in Table 1 under our preferred procedure. Once again, our estimates are quantitatively similar under any of these alternative matching methodologies.

The similarity in headline results under different procedures is reassuring for two reasons: first, they are shown to be robust to different reasonable matching assumptions; second, it suggests that even if sibling mismatches were not negligible in our sample, there is no evidence that they significantly attenuate our estimates.

B.1 School-level Matching

Figure B.1: Effect of older sibling's qualification for secondary school on younger sibling's secondary school attendance, by cohort (School-level Matching)

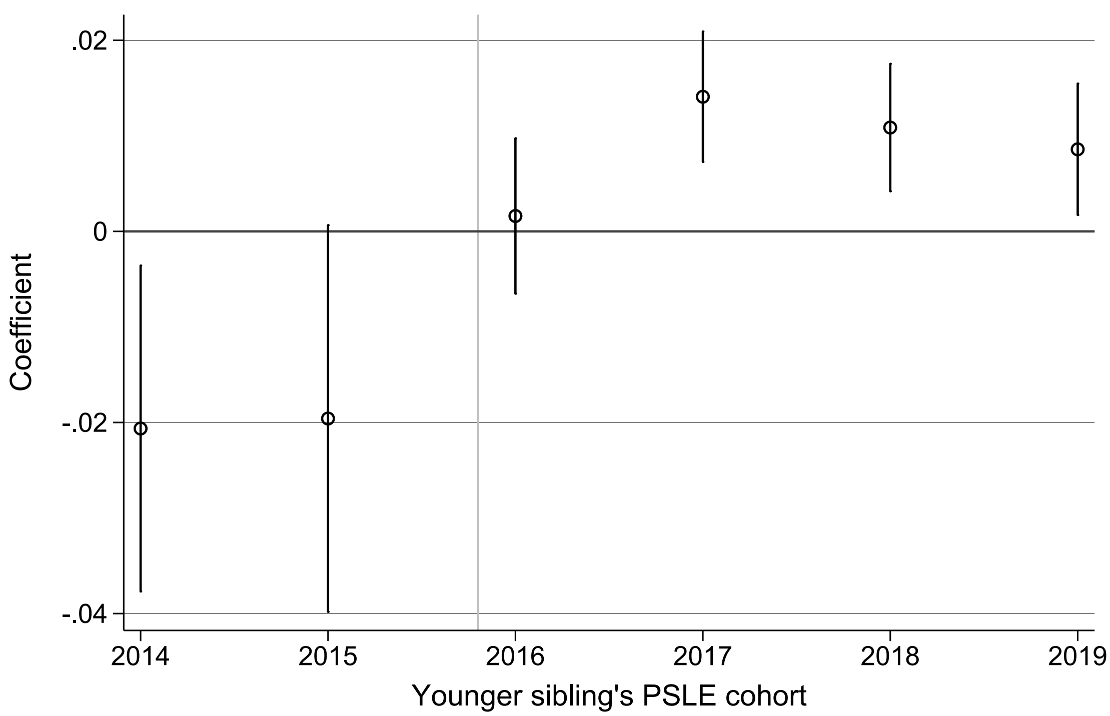


Table B.1: Spillover effects from older siblings' achievement (School-level Matching)

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	0.002	-0.002	0.007***	0.006***	0.006***	0.007***
	(0.005)	(0.006)	(0.003)	(0.002)	(0.002)	(0.002)
N	99,494	101,567	101,567	567,211	580,166	580,166
Mean dep. var.	0.238	0.609	0.142	0.592	0.739	0.224
R^2	0.273	0.309	0.333	0.161	0.157	0.183
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. All regressions include school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01

B.2 Middle and Last Name Matching

Figure B.2: Effect of older sibling's qualification for secondary school
on younger sibling's secondary school attendance, by cohort (Middle and Last Name Matching)

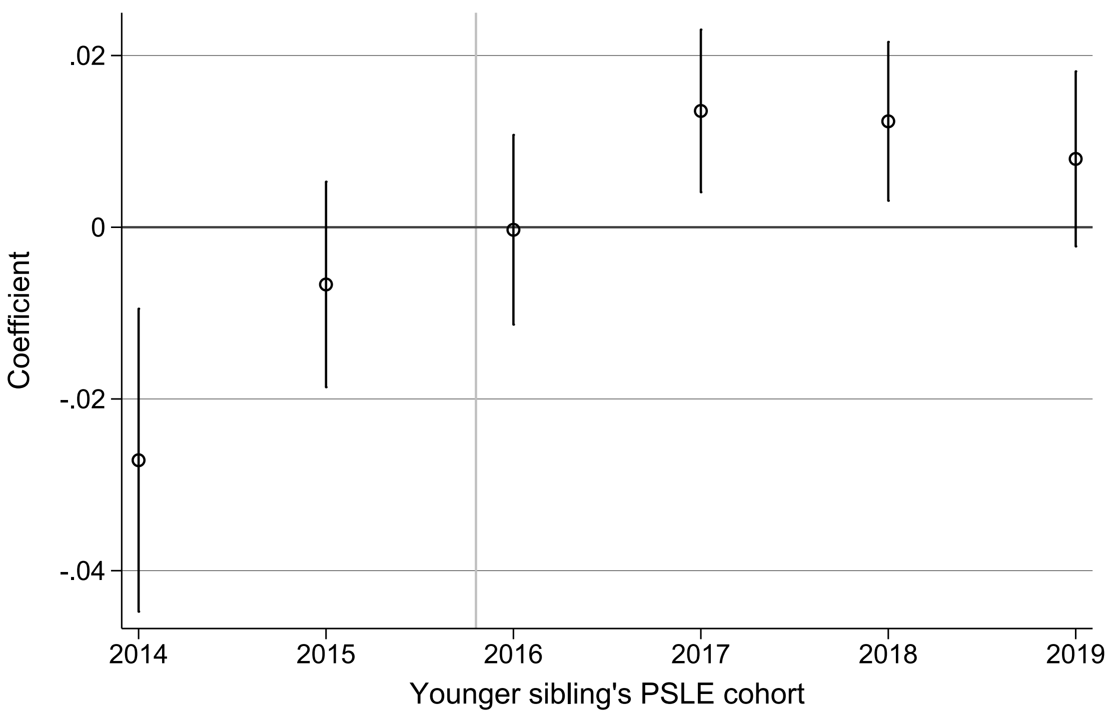


Table B.2: Spillover effects from older siblings' achievement (Middle and Last Name Matching)

	Pre-FSE			Post-FSE		
	(1) Transition	(2) PSLE Pass	(3) High score	(4) Transition	(5) PSLE Pass	(6) High score
Older sibling passed the PSLE	-0.009	-0.006	-0.001	0.009**	0.006**	0.007***
	(0.008)	(0.007)	(0.004)	(0.004)	(0.003)	(0.002)
N	36,798	38,737	38,737	289,442	303,243	303,243
Mean dep. var.	0.283	0.624	0.145	0.608	0.746	0.228
R^2	0.358	0.385	0.420	0.176	0.170	0.219
Older sibling's score FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered by grade combination in parentheses. 'Transition' is only measured for students whose full name is unique across the country within their cohort. 'High score' is a binary variable equal to one if the younger sibling got an A or B average on the PSLE, and zero otherwise. All regressions include (younger sibling's) school and (older sibling's) grade combination fixed effects. 'Year FE' denotes binary indicators for the PSLE cohort of each younger sibling. The sample includes all older siblings whose grade combination was associated with a probability of passing the PSLE between — and not including — 0 and 1. Significance levels: * <0.1, ** <0.05, *** <0.01