Impacts of Secondary School in Kenya after Two Decades

Michelle Layvant, Edward Miguel,
Owen Ozier, and Michael Walker*

March 23, 2026

Abstract

A crucial 8th grade standardized test provides a discontinuity that causes otherwise similar students to have different probabilities of attending and completing secondary school in Kenya. Following survey respondents for two decades after that test, we provide evidence of the lasting effects of secondary school attendance and completion. Secondary school decreases total fertility at almost every year afterward; it durably increases general cognitive function; and it leaves aggregate earnings and per-capita consumption unchanged on average.

JEL codes: I21, J31, O12, O15.

*Michelle Layvant is a PhD candidate at UC Berkeley; email mlayvant@berkeley.edu. Edward Miguel is a Professor of Economics at UC Berkeley, and holds affiliations with BREAD, J-PAL, and NBER; email emiguel@berkeley.edu. Owen Ozier is an Associate Professor of Economics at Williams College, and holds affiliations with BREAD, IZA, and J-PAL; email: owen.ozier@williams.edu. Michael Walker is a Research Economist at UC Berkeley; email mwwalker@berkeley.edu. We gratefully acknowledge funding from the U.S. National Institutes of Health (NIH) and the U.S. National Science Foundation (NSF).

1 Introduction

Twenty-five years ago, the Millennium Development Goals pushed policymakers around the world to achieve, among other things, universal access to primary education. Now, as part of the Sustainable Development Goals, that focus has shifted to secondary schooling (Duflo et al., 2024). Across Africa, secondary school has become increasingly available at low cost or no direct cost to those who enroll (Garlick, 2019; Ferreira and Sandholtz, 2024). In our study location—Kenya—just as free primary education featured in major political campaigns and policy discussions two decades ago, free secondary education has been important in more recent years (Ohba, 2011; Brudevold-Newman, 2021). Yet despite its importance, credible evidence on the long-run impacts of secondary schooling in lower-income settings remains scarce relative to what we know about primary education.

We study a sharp assignment rule at the end of primary school. The Kenya Certificate of Primary Education (KCPE) exam historically governed access to lower-cost government secondary places; pupils scoring just above the relevant threshold were much more likely to enroll in—and complete—four years of secondary school than pupils just below. Ozier (2018) first analyzed this cutoff in early adulthood and found large increases in secondary participation and completion at the threshold, higher scores on survey-administered reasoning and vocabulary tests, and reductions in teen pregnancy and early fertility. We revisit the same design and extend the horizon by roughly 10–15 additional years, using new rounds of data of the Kenya Life Panel Survey (KLPS), a longitudinal study that follows a large cohort from adolescence into adulthood with rich modules on schooling, fertility, labor markets, and cognition.

We link administrative KCPE scores to KLPS using a modernized matching approach, yielding about 1,000 additional matches and a substantially larger analytic sample. The expanded data allow us to examine outcomes that were previously unavailable: (i) mid-life cognition, measured with a battery of 20+ assessments covering memory, executive function, language, orientation, and visuospatial skills; and (ii) later-life socioeconomic and demo-

graphic outcomes, including fertility, earnings, and per-capita household consumption.

We re-estimate the same regression-discontinuity design. The first stage remains strong two decades later: at the cutoff, the probability of completing secondary rises by about 14 percentage points, which translates into roughly 0.6 additional years of schooling on average.

Our results show three main patterns. First, crossing the KCPE cutoff raises mid-life cognitive performance: memory increases by about 0.25 SD (statistically significant at the five percent level) and a general cognition factor by roughly 0.19 SD (significant at the ten percent level), with estimates for executive function, language, and visuospatial skills positive but imprecise. Second, fertility declines meaningfully over the long run: cumulative births are about 0.5 lower nearly two decades after the cutoff, an effect concentrated on the intensive margin among those with any children, while the extensive margin (any child) changes little (though the vast majority of respondents have children); earlier rounds of data on stated fertility preferences also move in the same direction. Third, economic resources are broadly unchanged: total individual earnings and per-capita household consumption show no detectable effects, although the composition of income shifts modestly—lower farming profits alongside offsetting, imprecise movements in wage and self-employment income.

This paper contributes to three main strands of literature. First, we extend the causal literature linking schooling to adult cognition into a low-income African setting. In high-income contexts, quasi-experimental changes to compulsory schooling show lasting cognitive gains: cohorts affected by England's 1947 school-leaving reform perform better on memory and executive tasks years later, particularly among men (Banks and Mazzonna, 2012); using cross-state and cohort variation in U.S. compulsory-schooling laws, Glymour et al. (2008) find that respondents with more schooling perform better on episodic memory decades later, with little effect on orientation or attention; and multi-country European reforms imply roughly 0.2 SD improvements in memory per additional year of education, with some protection against decline in verbal fluency (Schneeweis et al., 2014). By contrast, causal evidence from low-income settings remains scarce, and there is an active debate about how much

effective instruction students receive in African schools. In this paper, we measure mid-life cognition with a battery of 20+ assessments spanning memory, orientation, executive function, language, and visuospatial skills, allowing us to identify domain-specific effects of crossing the KCPE cutoff. Scaling the reduced-form by the first stage, the implied effect is near the upper end of the high-income range – about 0.3 SD higher general cognition per additional year of schooling – with the largest gains in memory.

Second, we contribute to the literature on education and fertility. Ozier (2018) found reductions in teen pregnancy and shifts in early fertility among those near the threshold. We extend the time horizon substantially, tracing cumulative fertility into the 30s, showing that the divergence opens early and continues to widen over adulthood. We decompose this change into a near-zero extensive margin effect and a sizable intensive-margin reduction among those with any births. These patterns are consistent with experimental evidence from Ghana's secondary-school scholarship program, which raised secondary school completion and reduced female fertility 12–15 years later (Duflo et al., 2021, 2024). They also align with quasi-experimental results in other low- and middle-income settings: Nigeria's Universal Primary Education reform implies roughly 0.26 fewer early births per extra year of female schooling (Osili and Long, 2008), and Indonesia's school-construction program delayed marriage and reduced early fertility (Breierova and Duflo, 2004).

Finally, we contribute to reproducibility by replicating and extending core results from Ozier (2018). Using an expanded KLPS sample and a new record-linkage procedure that yields over 1,000 additional matches to the administrative test score data, we recover the same first-stage schooling patterns and early fertility results, and show that they persist under modern RD estimators. By demonstrating that the original findings hold in later waves and under alternative specifications—while adding long-run outcomes—we strengthen the credibility and reproducibility of the research (Vilhuber, 2023).

In the sections that follow, we provide background on schooling in Kenya; the data; our approach to estimation; our results; and some discussion as we conclude.

2 Background: Secondary Schooling in Kenya

Kenya's cohorts in this study were educated under the 8–4–4 system: eight years of primary school followed by four years of secondary school and a terminal exam (KCSE). At the end of primary, pupils sat the Kenya Certificate of Primary Education (KCPE), a high-stakes exam used to allocate scarce, lower-cost places in public secondary schools. Admission into the first year of secondary school (known as Form 1) was organized through a centralized selection that prioritized KCPE totals and rationed places by school type—national, provincial/county, and district/sub-county—and by boarding versus day status. Higher-scoring pupils were funneled toward more competitive boarding schools, while most students enrolled in local day schools.

Effective admission thresholds were not uniform: they varied by school competitiveness and, in practice, by gender. In our setting, girls often faced slightly lower effective cutoffs than boys, consistent with administrative selection patterns and capacity constraints at the time. Families still bore meaningful costs (uniforms, transport, boarding levies), but KCPE-based placement sharply increased the probability of both entry and completion of the four-year cycle (Form 1 - Form 4) for pupils just above relevant thresholds.

Two policy developments help situate these cohorts. First, before 2008, tuition and levies at public secondary schools were substantial relative to family resources, and secondary enrollment was far from universal. Second, starting in 2008, the government introduced Free (Day) Secondary Education grants that covered tuition in public schools; in later years, policymakers articulated a "100 percent transition" goal from primary to secondary. Our cohorts sat KCPE before these policies fully took hold, when access remained capacity- and score-constrained and the KCPE threshold was especially salient for secondary progression.

These institutions motivate our empirical strategy. Because selection probabilities jump at school- and gender-specific KCPE totals, crossing the threshold generates localized changes in the likelihood of starting and completing secondary school—even though no single national cutoff applies to all pupils. In what follows, we treat KCPE totals as the running variable,

center them at empirically relevant thresholds (separately by gender), focus on first-attempt scores, and interpret our estimates as reduced-form effects of crossing the KCPE margin on subsequent outcomes.

3 Data

3.1 Kenya Life Panel Survey

The primary data source for this study is the Kenya Life Panel Survey (KLPS), a long-term longitudinal survey that has followed a representative cohort of children from rural western Kenya since the late 1990s. The KLPS was launched in 2003 to track roughly 7,500 individuals who were enrolled in grades 2 through 7 in 75 primary schools in Busia District in 1998. These schools were the site of the Primary School Deworming Program (PSDP), a large-scale randomized health intervention that provided free deworming treatment to schoolchildren in phases between 1998 and 2001 (Miguel and Kremer, 2004). While subsequent KLPS rounds expanded the sample to include respondents from two additional randomized interventions—the Girls' Scholarship Program (GSP) (Kremer et al., 2009) and a vocational training voucher and cash grant program launched in 2009—our analysis focuses on the original PSDP cohort, since administrative Kenya Certificate of Primary Education (KCPE) test score data (which define the regression discontinuity design) were only collected for pupils in PSDP schools.

To date, four survey rounds have been completed (2003–2005, 2007–2009, 2011–2014, and 2017–2021). Across these waves, the KLPS has collected numerous data on education, labor market outcomes, fertility, health, migration, and cognitive skills, among other domains. It is one of the most comprehensive panel datasets available in Sub-Saharan Africa, notable for its intensive long-distance tracking protocols that have achieved effective tracking rates of roughly 85–90 percent of the original sample in each major round (Hamory et al., 2021).

The fifth round of KLPS (KLPS-5) is currently ongoing. Importantly for this study, the

KLPS-5 Cognition module (C+ Module) has been completed, introducing a neuropsychological battery of more than 20 assessments covering memory, executive function, language, and visuospatial skills (Gross et al., 2025). These assessments are based on the Harmonized Cognitive Assessment Protocol (HCAP), which was developed to enable cross-national comparability of cognitive measures across the Health and Retirement Study (HRS) and its international partner studies. Now widely regarded as the benchmark tool for international research on cognition and aging, HCAP provides a standardized framework for measuring adult cognition. We use these new data to examine the long-run effects of secondary schooling on cognition, complementing the broader socioeconomic outcomes collected in earlier KLPS rounds.

Ozier (2018) previously used the KCPE cutoff to estimate the effects of secondary school admission with data from KLPS-2 (2007–2009). At that time, respondents were still in their early 20s, limiting the ability to capture long-term impacts. In contrast, we incorporate outcomes from subsequent KLPS rounds, including the KLPS-5 cognition module, providing an additional 10–15 years of follow-up. This allows us to track respondents into their mid-to-late 30s and examine the long-run consequences of secondary schooling not only for educational attainment, but also for cognition, fertility, earnings, and intergenerational outcomes, substantially extending the scope of prior research.

3.2 Administrative Data

KLPS surveys do not include official exam records. We therefore link respondents to an auxiliary archive of 17,384 official KCPE scores originally assembled in earlier work and expanded for this study (Ozier, 2018). The archive is candidate-level and includes each examinee's name, sex, exam year, primary school, and total KCPE score. Records were obtained with permission from District Education Offices and, when necessary, directly from primary schools. The collection targets the years in which the KLPS cohort would have sat for the KCPE (1999-2006). Paper preservation was incomplete in some districts; the

resulting dataset nonetheless covers the large majority (more than 80%) of school–year cells in which sample members could have taken the exam.

This paper also expands the KLPS-KCPE linkage relative to the earlier study. Using a modernized matching approach, we verify identities against name strings, school–year information, and basic demographics, yielding about 1,000 additional verified matches to the KCPE archive and a substantially larger analytic sample.

4 Econometric Approach

We exploit the admission rule governing entry into Kenyan secondary schools, which created a discontinuity in the probability of admission at the Kenya Certificate of Primary Education (KCPE) cutoff. This rule generates quasi-experimental variation that allows us to identify the causal impacts of secondary school admission on long-run outcomes. Our empirical strategy follows the regression discontinuity (RD) framework in Ozier (2018), but we focus exclusively on reduced-form effects.

Formally, let K_i denote the KCPE score for individual i, normalized so that the cutoff is at zero, and let Y_i represent an outcome of interest. The reduced-form parameter of interest is the discontinuity in the conditional expectation of the outcome at the cutoff:

$$\tau_{RF} = \lim_{k \downarrow 0} E[Y_i \mid K_i = k] - \lim_{k \uparrow 0} E[Y_i \mid K_i = k]. \tag{1}$$

We estimate τ_{RF} using the `rdrobust` procedure of Calonico et al. (2014, 2019). This method fits local linear regressions on either side of the cutoff, applies robust bias correction based on higher-order local polynomials, and constructs confidence intervals that remain valid under data-driven bandwidth selection. The local linear regression specification can be written as:

$$Y_i = \beta_0 + \tau_{RF} \cdot 1\{K_i \geq 0\} + \beta_1 K_i + \beta_2 \left(K_i \cdot 1\{K_i \geq 0\}\right) + \varepsilon_i, \tag{2}$$

where τ_{RF} captures the discontinuity in outcomes at the cutoff, and separate slopes in K_i are allowed on either side. Our primary estimates use the MSE-optimal bandwidth chosen by `rdrobust`.

Because admission cutoffs were not uniform across gender, we follow Ozier (2018) in estimating the effective cutoff separately for men and women using a structural break approach based on Card et al. (2008). For boys, the effective cutoff is approximately 250 points; for girls, the effective cutoff is lower, at 234 points.[1] In all specifications, KCPE scores are normalized relative to these gender-specific thresholds.

Identification relies on the assumption that, absent the admission rule, potential outcomes evolve smoothly in KCPE scores around the cutoff. We assess this assumption using the Cattaneo et al. (2020) density test as well as the McCrary (2008) density test. Both tests are for manipulation of the running variable; neither test finds any evidence of such manipulation, with p-values of approximately 0.43 and 0.65. (Results are shown in Appendix Figure A2.) In this setting, manipulation is unlikely to arise in a centrally marked examination, but one could imagine data generating processes that could lead to different probabilities of matching records on either side of the discontinuity; regardless, we see no evidence of this kind of phenomenon.

For the purposes of the present analysis, we do not scale the reduced-form estimates by the discontinuity in secondary school completion in order to estimate a local average treatment effect (LATE) as one would in a conventional fuzzy regression discontinuity design, though we could. The cutoff affects admission to secondary school, and not necessarily completion—or even an additional year—of schooling. Empirically, it does turn out—as in Ozier (2018)—that the cutoff affects 12th grade completion just as much as it affects reaching 9th grade, but one could even imagine that admission itself could change outcomes directly, through

[1]For girls, the R^2-maximizing cutoff is 234, exactly as in Ozier (2018). For boys, the R^2-maximizing cutoff is 258, but we use 250 as our reference point since it was both the cutoff widely reported in national media and the modal value cited by secondary school administrators in 2009. Appendix Figure A1 shows the structural break estimates: the R^2 values at 250 and 258 are nearly identical, so our choice of 250 is substantively inconsequential.

altered expectations, peer groups, or parental investments, regardless of whether students ultimately complete more years of school. As such, the exclusion restriction required for the IV interpretation could be argued not to hold. We show reduced-form results here. If one were comfortable with the exclusion restriction, one could proceed with fuzzy regression discontinuity analysis in the present dataset as in Ozier (2018), and the LATEs would be the ratio of the reduced form and first stage results shown here.

By focusing on reduced-form effects, we capture the overall causal impact of being just above the secondary school admission threshold on a wide range of later-life outcomes—including educational attainment, cognition, fertility, earnings, and intergenerational measures—without requiring that all of these impacts operate solely through additional schooling. This is also analogous to the approach (Duflo et al., 2025) use to study the longer-term effect of secondary scholarships in Ghana. In both cases, the reduced form is a policy-relevant parameter, as it can be interpreted as the effect on the margin of shifting the admission threshold.

5 Results

5.1 Education Impacts

We begin by examining the first-stage impacts on educational attainment. Table 1 reports reduced-form estimates for completed years of schooling and key attainment thresholds. Individuals scoring just above the KCPE cutoff completed, on average, 0.6 more years of schooling than those just below, significant at the 5 percent level. Consistent with this, the probability of attending any secondary school rises by 14 percentage points, while the likelihood of completing Form 4 increases by 12 percentage points. These are substantial effects relative to the control means, which indicate that only 56 percent of respondents in the below-cutoff sample ever attended secondary school, and 30 percent completed Form 4.

Figure 1 visualizes the discontinuity in secondary school completion. The sharp jump at

the cutoff confirms that KCPE scores generated meaningful quasi-experimental variation in secondary attainment. Figure 2 further shows that the cumulative distribution of years of schooling shifts upward for those above the cutoff, with especially pronounced differences at thresholds of 9–12 years corresponding to secondary grades.

Appendix Table A1 reports results separately for men and women. Both groups exhibit significant increases in secondary enrollment and completion at the cutoff, though the magnitudes are somewhat larger for women. This is consistent with the fact that the effective cutoff for women is lower (234 points versus 250 for men), meaning that girls just above the threshold were particularly likely to gain access to secondary school.

Together, these results replicate the first-stage findings in Ozier (2018) but with updated KLPS data through 2021. The magnitudes are comparable to earlier estimates, reinforcing the interpretation that admission policies induced substantial increases in attainment. Importantly, these effects motivate our focus on reduced-form outcomes: while admission strongly increases the probability of starting secondary school, it does not deterministically translate into additional years of completed schooling for all compliers. Admission itself may alter trajectories through expectations, peer groups, or parental investments, even absent full completion of secondary school. Our reduced-form strategy therefore captures the overall causal impact of crossing the admission threshold.

5.2 Cognitive Impacts

We next examine the effect on adult cognitive performance. Table 2 reports reduced-form estimates for domain-specific factor scores derived from the KLPS-5 cognitive battery of assessments (Gross et al., 2025). Individuals just above the KCPE cutoff score higher on both general cognition and memory, with gains of 0.19 and 0.25 standard deviations, respectively. The effect on memory is significant at the 5 percent level, while the general factor is significant at the 10 percent level. For the other domains—including language, executive function, orientation, and visuospatial ability—the point estimates are positive but relatively small

and imprecisely estimated.

Appendix Table A2 presents results separately by gender. Among men, we find a large and significant improvement in memory (0.34 SD), with smaller and less precise gains in general cognition. For women, the estimates are consistently positive across domains, though none are statistically significant at conventional levels. Taken together, these findings suggest that access to secondary schooling yields durable improvements in adult cognitive performance.

5.3 Fertility Impacts

We then examine the effects on fertility. Table 4 reports reduced-form estimates for realized fertility (Panel A) and fertility intentions (Panel B). Adults just above the KCPE cutoff have substantially fewer children two decades later: total births fall by about one-half child (-0.477, s.e. 0.193), and births conditional on having any children also decline (-0.446, s.e. 0.174), both significant at the 5 percent level. In contrast, extensive-margin fertility shows little change: the probability of having any children is -0.032 (s.e. 0.042), and age at first birth rises by 0.49 years (s.e. 0.48), neither statistically significant. Child mortality outcomes are negative but imprecise (under-5: -0.017, s.e. 0.025; under-1: -0.012, s.e. 0.018).

Fertility intentions are directionally consistent with these realized outcomes. Desired family size falls by 0.42 children in KLPS2 and by 0.32 children in KLPS3 (both significant at the 5 percent level), indicating that admission to secondary school shifts preferences toward smaller families well before many births occur. Corresponding estimates by gender appear in Appendix Tables A3 (males) and A4 (females).

Overall, the evidence points to a clear change in fertility trajectories—roughly a half-child reduction over the long run—accompanied by earlier-reported declines in desired family size. These patterns are broadly consistent with prior experimental and quasi-experimental evidence on the downstream demographic effects of secondary schooling in the region.

Figure 3 contrasts average cumulative fertility by secondary completion status for men

and women. The divergence opens early in adulthood and continues to widen through the late 20s and early 30s. Consistent with this timing, Figure 4 plots reduced-form effects on cumulative births by year with 95% confidence intervals; the pooled effect becomes increasingly negative over adulthood.

5.4 Other Economic Outcomes

Table A5 reports reduced-form impacts at the KCPE cutoff on economic outcomes over the prior 12 months. Crossing the cutoff has no detectable effect on total individual earnings or on per-capita household consumption. By component, farming profits decline (-$16.52, s.e. $7.78; 5% level)[2], while self-employment profits are negative but imprecise; wage earnings are positive yet imprecise. Total household earnings are also close to zero. These patterns suggest modest reallocation across income sources rather than large changes in total resources.

Disaggregating by gender yields similar conclusions. Among men, farming profits fall (-$20.68, s.e. 10.967; 10% level), with other components near zero or imprecisely estimated; per-capita consumption is essentially unchanged. Among women, self-employment profits decline (-$128.36, s.e. $74.38; 10% level), while wage earnings, total earnings, and per-capita consumption remain close to zero and imprecise. Overall, crossing the KCPE cutoff appears to shift the composition of earnings slightly without clear effects on aggregate earnings or consumption. This pattern is consistent with the initial reduction in (largely low-skill) self-employment shown by Ozier (2018).

6 Conclusion

The lasting changes in cognition that we document are novel, but in agreement with the work of Duflo et al. (2025) in Ghana. That study documents lasting increases in PISA-inspired test performance in math and reading comprehension. By contrast, the measures we include

[2]Measured in 2017 USD PPP.

in this study are not based on these specific skills, but are instead based on general abilities; the largest cognitive effects we document are lasting benefits for memory.

We see very little evidence of increased earnings following secondary schooling. Our overall income estimates are relatively imprecise, but we can see definitive evidence of a decrease in the farm earnings of the individual who attended secondary school, consistent with earlier work by Ozier (2018). The lack of definitive changes in overall income is consistent with work in Ghana by Duflo et al. (2025), suggesting that in these lower middle income countries in Sub-Saharan Africa, secondary school alone may not be transformative in terms of income.

We find impacts of secondary schooling on desired fertility lasting many years after completing secondary school, though that effect appears to dissipate in the most recent survey round. The reduction in desired fertility contrasts with the findings of Brudevold-Newman (2021), who sees no impact on fertility intentions. One distinction may arise from the two study designs. The compliers in Brudevold-Newman's design experience a policy change that everyone around them also experiences, whereas in our regression discontinuity design, compliers gain an opportunity that their peers do not. Whether one sees oneself as following the same patterns as one's peers may influence outcomes including fertility intentions.

Consistent with Duflo et al. (2021) and Ozier (2018), we see marked and lasting reductions in fertility. Drawing on fertility history data and years of longitudinal surveys, we can see that the pattern could be consistent with a delay in fertility rather than a reduction in total fertility, but either could prove to be true in the years to come.

Though our study may not have the power to detect small changes in income, secondary school has changed the trajectory of students' lives in Kenya. Their human capital, measured by memory and other cognitive skills later in life, is raised two decades later. There is a shift away from farm income, and a demographic change in the timing and number of children, an effect evident for both men and women. In sum, we document a collection of longer-lasting changes than had ever been quantified before with quasi-experimental designs in Kenya, and

we show that the human capital benefits of education can last a lifetime.

Table 1: Education Impacts

Outcome	Mean	Reduced Form CC&T RD		Sample sizes
Max years of schooling attained				
	10.035	0.610**	(0.273)	N: 2825; 1388
Any Secondary Schooling				
	0.564	0.142***	(0.047)	N: 2825; 1467
Completed Form 4				
	0.301	0.121**	(0.052)	N: 2825; 1575
13+ Years of Schooling				
	0.156	0.033	(0.043)	N: 2825; 1838

Notes: Education attainment measures are constructed from all available KLPS survey rounds (1–4), providing annual schooling histories from 1998–2021. Vocational training codes are counted as additional years of education only for respondents who had already completed Form 4; otherwise they are recoded as lower secondary attainment. Respondents never re-interviewed after the baseline roster are coded as missing. *Any Secondary Schooling* is an indicator equal to 1 if the individual completed at least 9 years of education. *Completed Form 4* is an indicator equal to 1 if the individual completed at least 12 years. *13+ Years of Schooling* is an indicator equal to 1 if the individual completed 13 or more years. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively.

Table 2: Cognitive Impacts

Outcome	Mean	Reduced Form CC&T RD		Sample sizes
General Cognitive Factor Score				
	0.043	0.191*	(0.102)	N: 2226; 994
Memory Factor Score				
	-0.012	0.248**	(0.117)	N: 2226; 1075
Language Factor Score				
	0.029	0.001	(0.083)	N: 2226; 1024
Executive Function Factor Score				
	0.063	0.070	(0.091)	N: 2226; 963
Orientation Factor Score				
	-0.029	0.097	(0.064)	N: 2226; 1079
Visuospatial Factor Score				
	0.065	0.080	(0.081)	N: 2226; 962

Notes: Cognitive outcomes are drawn from the KLPS-5 C+ Module, which implemented the Harmonized Cognitive Assessment Protocol (HCAP) in the Kenyan context (Gross et al., 2025). The battery includes more than 20 neuropsychological assessments spanning orientation (orientation to person, place, and time), memory (CERAD Word List Learning and Recall, East Boston Memory Test, Logical Memory), executive function and attention (Raven's Progressive Matrices, Token Test, Go–No–Go, Trail Making, Symbol Cancellation), language and verbal fluency (animal naming, object naming, similarities/differences), and visuospatial ability (CERAD Constructional Praxis, interlocking pentagons, clock drawing). We construct domain-specific factor scores using confirmatory factor analysis and normalize them to mean zero and unit variance across the full KLPS-5 sample. Because standardization is based on the full sample, the means in our regression sample are not necessarily zero. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and , respectively.

Table 3: Fertility Impacts

Outcome	Mean	Reduced Form CC&T RD		Sample sizes
Panel A: Realized Fertility				
Indicator for any children				
	0.870	-0.032	(0.042)	N: 2828; 1318
Total number of children				
	2.648	-0.477**	(0.193)	N: 2828; 1249
Age at first birth				
	23.412	0.486	(0.477)	N: 2301; 1253
Number of live births (cond> 0)				
	3.045	-0.446**	(0.174)	N: 2302; 1150
Under five mortality				
	0.056	-0.017	(0.025)	N: 4381; 1881
Under one mortality				
	0.034	-0.012	(0.018)	N: 6112; 3153
Panel B: Fertility Intentions				
Desired No. Children (KLPS2)				
	3.379	-0.417**	(0.165)	N: 2301; 885
Desired No. Children (KLPS3)				
	3.347	-0.324**	(0.131)	N: 2315; 1348
Desired No. Children (KLPS4)				
	5.380	0.107	(0.365)	N: 2279; 1175

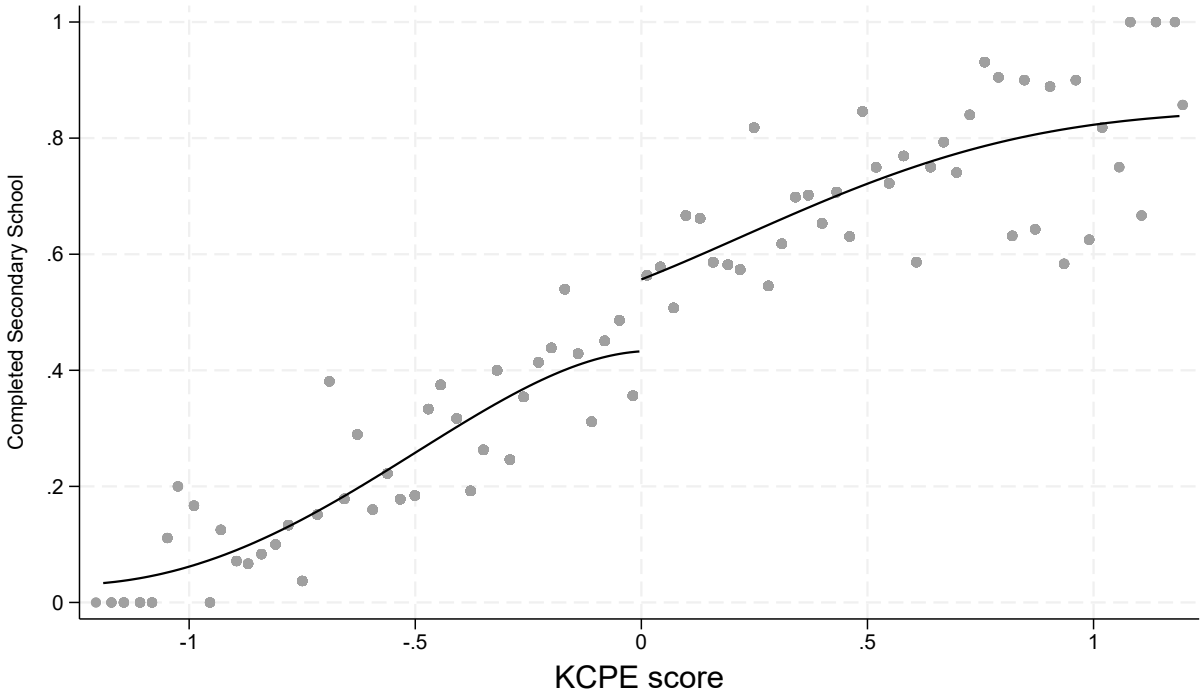
Notes: Panel A reports realized fertility outcomes, constructed from retrospective fertility histories in KLPS rounds 1–4. Outcomes include indicators for having any children, the total number of children ever born, age at first birth, number of live births (conditional on having at least one), and child mortality rates (under-five and under-one). Panel B reports fertility intentions, capturing respondents' desired total number of children (including those already born) as reported in successive KLPS rounds. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and * respectively.

Table 4: Fertility Impacts - 2SLS

Outcome	Mean	2SLS CC&T RD		Sample sizes
Panel A: Realized Fertility				
Indicator for any children				
	0.870	-0.222	(0.318)	N: 2828; 1287
Total number of children				
	2.648	-3.372**	(1.718)	N: 2828; 1381
Age at first birth				
	23.412	2.768	(2.411)	N: 2301; 1057
Number of live births (cond> 0)				
	3.045	-2.284**	(0.973)	N: 2302; 994
Under five mortality				
	0.056	-0.116	(0.125)	N: 4381; 2493
Under one mortality				
	0.034	-0.054	(0.103)	N: 6112; 3448
Panel B: Fertility Intentions				
Desired No. Children (KLPS2)				
	3.379	-2.160*	(1.138)	N: 2301; 885
Desired No. Children (KLPS3)				
	3.347	-2.134*	(1.150)	N: 2315; 1026
Desired No. Children (KLPS4)				
	5.380	0.629	(2.671)	N: 2279; 1333

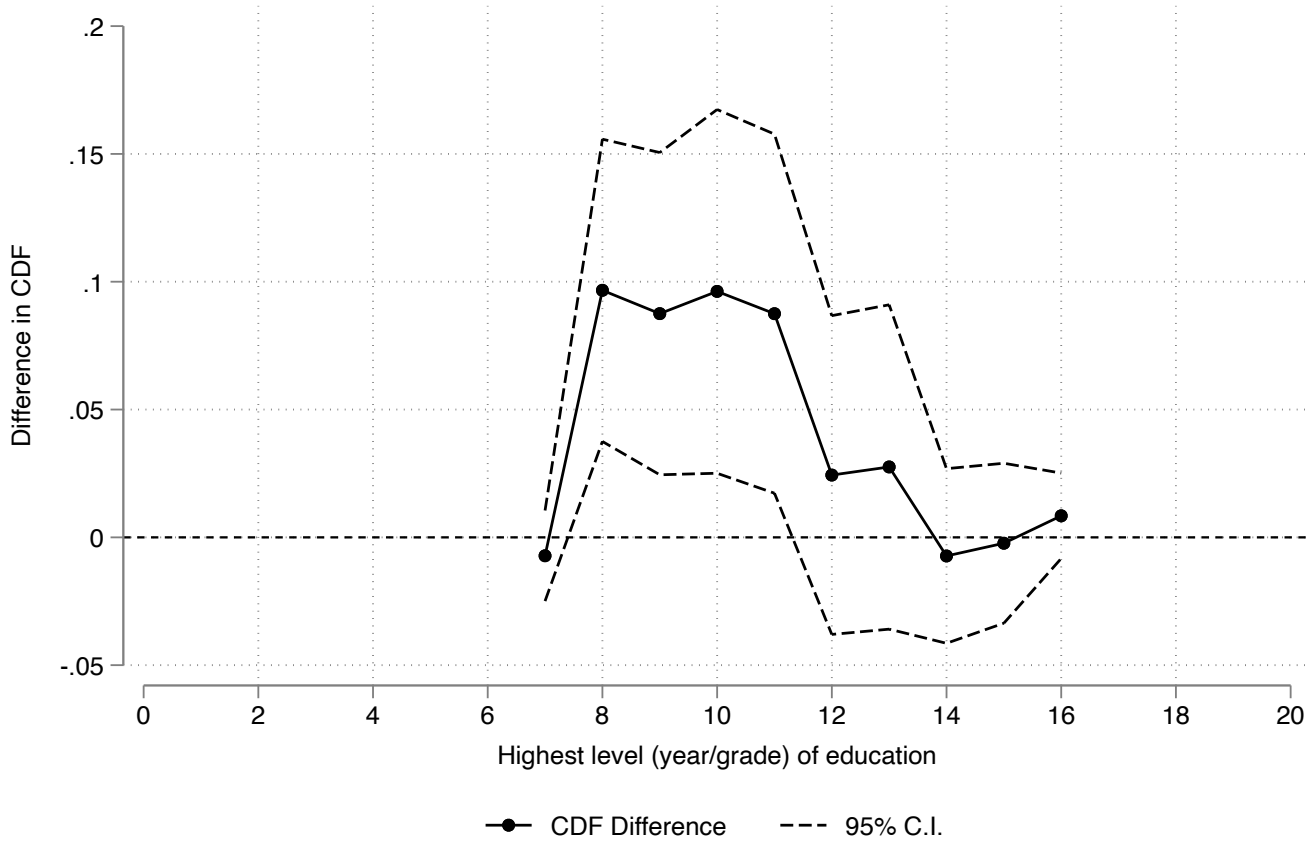
Notes: Panel A reports realized fertility outcomes, constructed from retrospective fertility histories in KLPS rounds 1–4. Outcomes include indicators for having any children, the total number of children ever born, age at first birth, number of live births (conditional on having at least one), and child mortality rates (under-five and under-one). Panel B reports fertility intentions, capturing respondents' desired total number of children (including those already born) as reported in successive KLPS rounds. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and * respectively.

Figure 1: Secondary School Completion



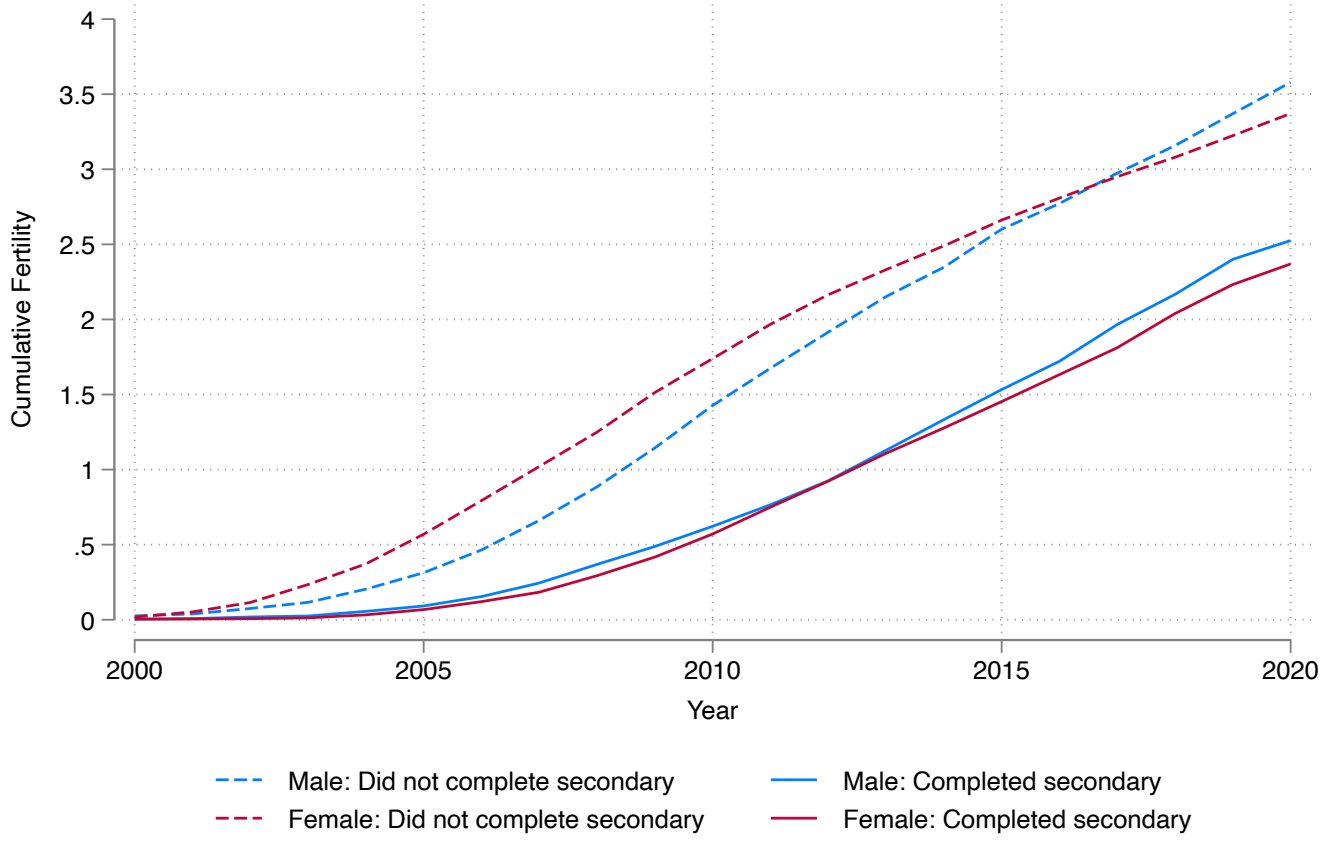
Notes: This figure plots the probability of completing Form 4 (final year of secondary school) as a function of KCPE exam scores. Completion outcomes are constructed from KLPS survey rounds 1–4, using retrospective schooling histories (see Table 1 for details on education variable construction). Because admission cutoffs differed by gender, KCPE scores are normalized so that zero corresponds to the effective cutoff separately for men and women. Mean completion rates are shown in bins of the running variable using the `rdplot` command, which partitions the support of KCPE scores into evenly spaced bins and overlays local linear fits on either side of the cutoff. The vertical axis reports the share completing Form 4; the horizontal axis is standardized KCPE score relative to the cutoff.

Figure 2: CDF Difference



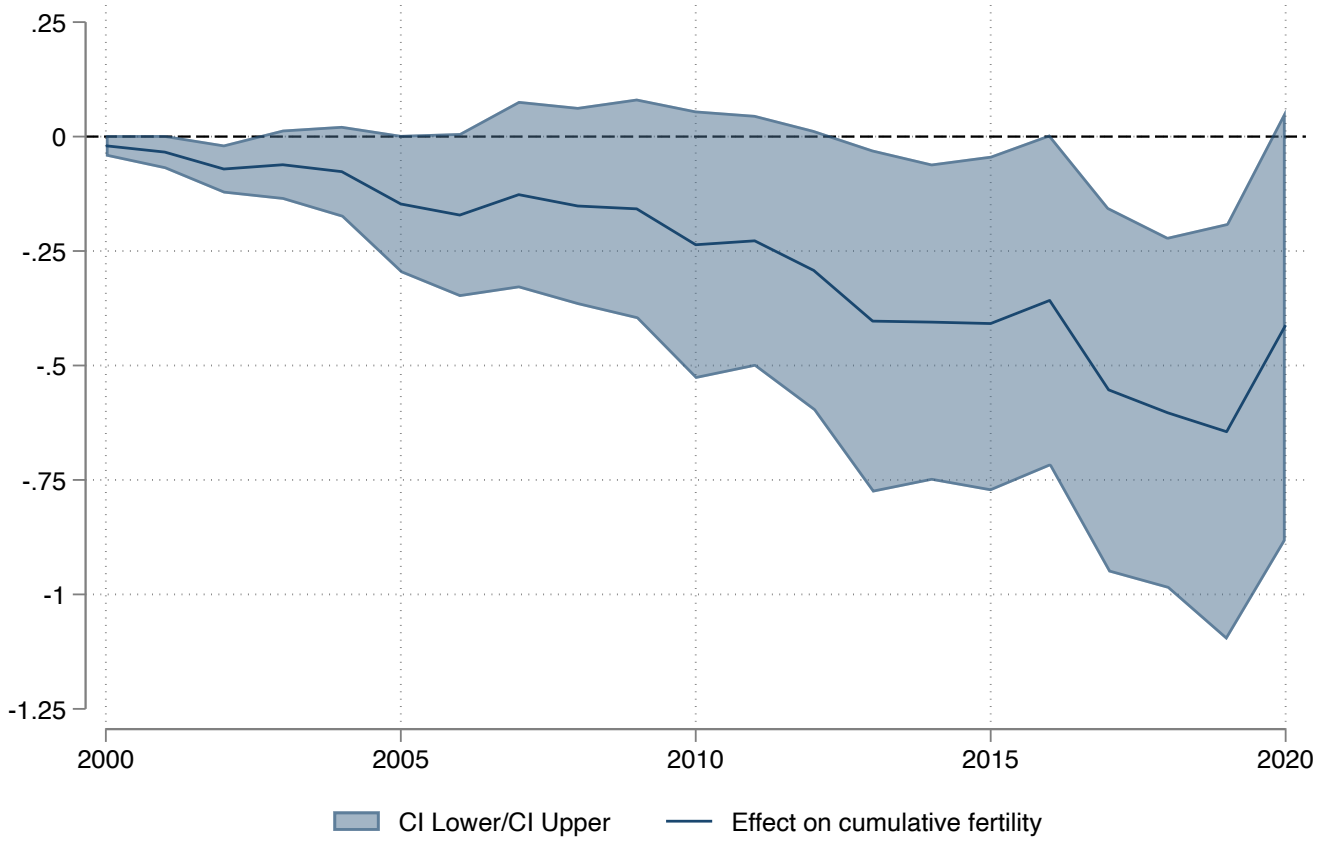
Notes: This figure shows differences in the cumulative distribution of completed years of schooling around the KCPE cutoff. The horizontal axis marks thresholds of educational attainment, while the vertical axis reports the discontinuity in the probability of reaching at least that level. By design, all respondents in the sample must have completed eight years to sit for the KCPE, so the difference at that point is mechanically zero. At higher thresholds—particularly 9 through 12 years, corresponding to secondary grades—the estimated gaps are large and statistically significant, indicating that admission to secondary school increased the likelihood of advancing further in the educational system. Years of schooling are derived from harmonized KLPS attainment measures (see Table 1).

Figure 3: Average Fertility by Secondary School Completion



Notes: This figure plots differences in realized fertility by age between respondents who completed secondary school and those who did not, separately for men and women. Fertility outcomes are measured from retrospective birth histories collected in KLPS rounds 1–4. The vertical axis reports mean differences in the number of children ever born at each age, while the horizontal axis shows respondent age in years.

Figure 4: Effect on Cumulative Fertility - Pooled



Notes: This figure plots reduced-form regression-discontinuity estimates of the effect of scoring above the KCPE threshold on cumulative fertility by calendar year (2000–2020). Cumulative fertility is the number of children ever born through that year, constructed from retrospective birth histories in KLPS Rounds 1–4. KCPE scores are normalized so that zero corresponds to the empirically relevant cutoff separately for men and women. The solid line shows local-linear estimates; shaded bands are 95% confidence intervals based on `rdrobust` with robust bias correction and MSE-optimal bandwidths. Negative values indicate fewer births among those just above the cutoff. Estimates are pooled across men and women.

References

Banks, J. and F. Mazzonna (2012). The effect of education on old age cognitive abilities: Evidence from a regression discontinuity design. *The Economic Journal 122*(560), 418–448.

Breierova, L. and E. Duflo (2004, May). The impact of education on fertility and child mortality: Do fathers really matter less than mothers? NBER Working Paper 10513, National Bureau of Economic Research, Cambridge, MA.

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.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics 101*(3), 442–451.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Card, D., A. Mas, and J. Rothstein (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics 123*(1), 177–218.

Cattaneo, M. D., M. Jansson, and X. Ma (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association 115*(531), 1449–1455.

Duflo, E., P. Dupas, and M. Kremer (2021, June). The impact of free secondary education: Experimental evidence from ghana. NBER Working Paper 28937, National Bureau of Economic Research, Cambridge, MA.

Duflo, E., P. Dupas, and M. Kremer (2025). The impact of secondary school subsidies on career trajectories in a dual labor market: Experimental evidence from Ghana. Technical report, Technical report.

Duflo, E., P. Dupas, E. Spelke, and M. P. Walsh (2024, July). Intergenerational impacts of secondary education: Experimental evidence from Ghana. Working Paper 32742, National Bureau of Economic Research.

Ferreira, J. R. and W. A. Sandholtz (2024). Sibling spillovers and free schooling. Technical report, CESifo Working Paper.

Garlick, R. (2019). Quality-quantity tradeoffs in pricing public secondary education. *Economic Research Initiatives at Duke (ERID) Working Paper* (290).

Glymour, M. M., I. Kawachi, and C. S. Jencks (2008). Does childhood schooling affect old age memory or mental status? using state schooling laws as natural experiments. *Journal of Epidemiology & Community Health 62*(6), 532–537.

Gross, A. L., M. Duhon, E. Ochieng, J. N. Ikanga, W. H. Dow, J. Lee, M. W. Walker, M. Layvant, A. Ngugi, J. R. Ehrlich, et al. (2025). Midlife cognitive testing in Africa: validity of the harmonised cognitive assessment protocol in the Kenya Life Panel Survey. *BMJ open 15*(8), e096619.

Hamory, J., E. Miguel, M. W. Walker, M. Kremer, and S. J. Baird (2021). Twenty year economic impacts of deworming. *Proceedings of the National Academy of Sciences*.

Kremer, M., E. Miguel, and R. Thornton (2009). Incentives to learn. *Review of Economics and Statistics 91*(1), 437– 456.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics 142*(2), 698–714.

Miguel, E. and M. Kremer (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72*(1), 159–217.

Ohba, A. (2011). The abolition of secondary school fees in kenya: Responses by the poor. *International Journal of Educational Development 31*(4), 402–408.

Osili, U. O. and B. T. Long (2008, August). Does female schooling reduce fertility? evidence from nigeria. *Journal of Development Economics 87*(1), 57–75.

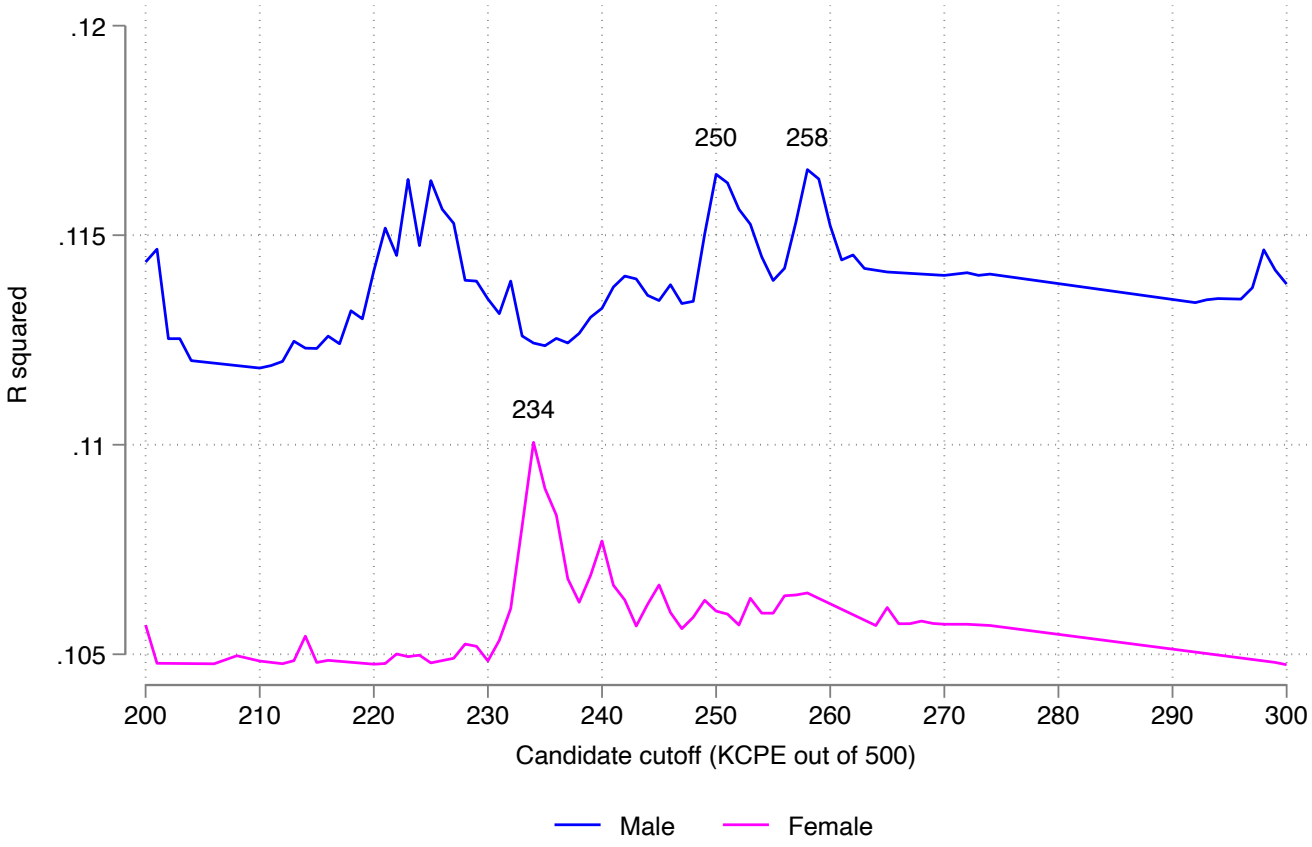
Ozier, O. (2018). The impact of secondary schooling in Kenya: A regression discontinuity analysis. *Journal of Human Resources 53*(1), 157–188.

Schneeweis, N., V. Skirbekk, and R. Winter-Ebmer (2014). Does education improve cognitive performance four decades after school completion? *American Economic Journal: Applied Economics 6*(1), 1–19.

Vilhuber, L. (2023). Reproducibility and transparency versus privacy and confidentiality: Reflections from a data editor. *Journal of Econometrics 235*(2), 2285–2294.

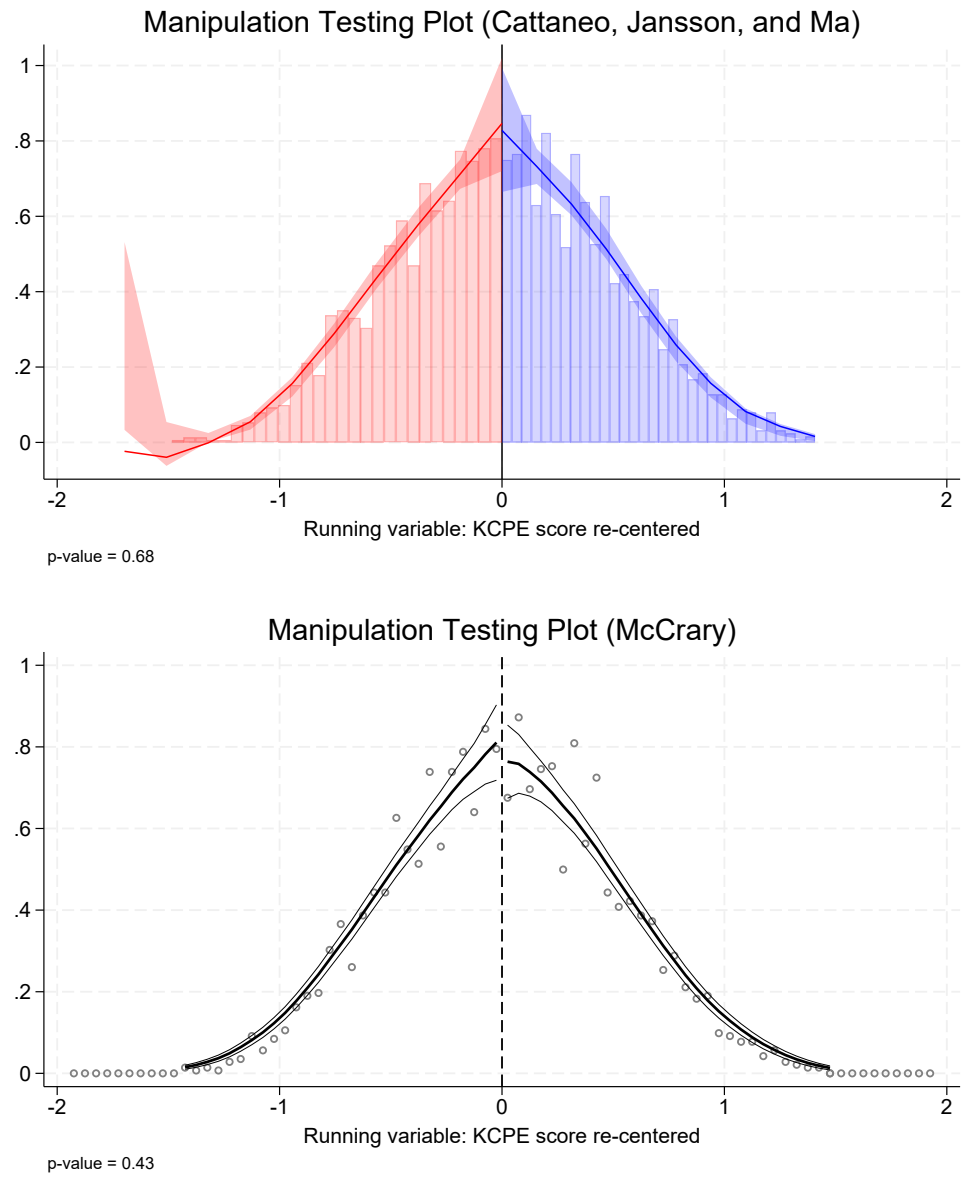
A Additional Tables and Figures

Figure A1: Structural Break Search



Notes: This figure implements a structural-break search (Card–Mas–Rothstein, 2008–style) to locate the KCPE score at which the probability of completing Form 4 rises discontinuously. For each candidate cutoff on the 0–500 scale, we estimate a piecewise linear model for completion within a symmetric ±70-point window and plot model fit (R-squared) separately for men and women. The KCPE scale changed over time: scores from 1999–2000 were reported on a 700-point scale and are converted to 500-point equivalents; the exam also shifted in the early 2000s, so we repeat the search by subperiod (1999–2000, 2001–2002, 2003–2005). In the pooled series the best fit occurs near 258, but we adopt 250 for men—both for institutional comparability and because it is more stable across subperiods—and 234 for women, as indicated by the female-specific search.

Figure A2: Density test



Notes: This figure reports manipulation tests for the running variable at the KCPE cutoff. The running variable is the KCPE score re-centered at the gender-specific threshold. The top panel shows the local-polynomial density test of Cattaneo et al. (2020) (`rddensity`); the bottom panel shows the McCrary (2008) discontinuity test (`DCdensity`). Both tests fail to reject continuity at the cutoff ($p = 0.68$ and $p = 0.43$, respectively), consistent with no sorting in the test score around the threshold.

Table A1: Education Impacts by Gender

Outcome	Mean	Reduced Form CC&T RD		Sample sizes
Among men:				
Max years of schooling attained				
	10.338	0.363	(0.355)	N: 1581; 798
Any Secondary Schooling				
	0.635	0.094*	(0.054)	N: 1581; 879
Completed Form 4				
	0.352	0.084	(0.071)	N: 1581; 854
13+ Years of Schooling				
	0.178	0.059	(0.067)	N: 1581; 792
Among women:				
Max years of schooling attained				
	9.688	0.811*	(0.427)	N: 1244; 582
Any Secondary Schooling				
	0.482	0.199**	(0.087)	N: 1244; 520
Completed Form 4				
	0.243	0.169**	(0.084)	N: 1244; 593
13+ Years of Schooling				
	0.131	0.050	(0.073)	N: 1244; 612

Notes: This table replicates the Table 1 specification separately for men and for women; pooled estimates appear in Table 1. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and *, respectively.

Table A2: Cognitive Impacts by Gender

Outcome	Mean	Reduced Form CC&T RD		Sample sizes
Among men:				
General Cognitive Factor Score				
	0.104	0.211	(0.138)	N: 1223; 576
Memory Factor Score				
	-0.025	0.340**	(0.165)	N: 1223; 597
Language Factor Score				
	0.050	0.112	(0.110)	N: 1223; 580
Executive Function Factor Score				
	0.174	-0.048	(0.118)	N: 1223; 502
Orientation Factor Score				
	-0.003	0.092	(0.080)	N: 1223; 576
Visuospatial Factor Score				
	0.159	0.013	(0.106)	N: 1223; 531
Among women:				
General Cognitive Factor Score				
	-0.023	0.157	(0.136)	N: 1003; 565
Memory Factor Score				
	0.002	0.162	(0.161)	N: 1003; 535
Language Factor Score				
	0.006	-0.162	(0.134)	N: 1003; 419
Executive Function Factor Score				
	-0.057	0.137	(0.132)	N: 1003; 512
Orientation Factor Score				
	-0.057	0.097	(0.092)	N: 1003; 595
Visuospatial Factor Score				
	-0.038	0.141	(0.120)	N: 1003; 472

		Reduced Form	
Outcome	Mean	CC&T RD	Sample sizes

Notes: This table replicates the Table 2 specification separately for men and for women; pooled estimates appear in Table 2. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and , respectively.

Table A3: Fertility Impacts: Males

		Reduced Form		
Outcome	Mean	CC&T RD		Sample sizes
Realized Fertility:				
Indicator for any children				
	0.850	-0.039	(0.064)	N: 1583; 722
Total number of children				
	2.564	-0.584**	(0.276)	N: 1583; 681
Age at first birth				
	24.528	0.217	(0.663)	N: 1245; 575
Number of live births (cond> 0)				
	3.017	-0.652**	(0.272)	N: 1245; 514
Under five mortality				
	0.050	-0.034	(0.028)	N: 2278; 1184
Under one mortality				
	0.028	-0.016	(0.026)	N: 3295; 1406
Fertility Intentions:				
Desired No. Children (KLPS2)				
	3.481	-0.540**	(0.216)	N: 1287; 601
Desired No. Children (KLPS3)				
	3.439	-0.303	(0.186)	N: 1328; 662
Desired No. Children (KLPS4)				
	5.555	0.350	(0.491)	N: 1265; 634

Notes: This table replicates the Table 3 specification separately for men; pooled estimates appear in Table 3. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and * respectively.

Table A4: Fertility Impacts: Females

Outcome	Mean	Reduced Form CC&T RD		Sample sizes
Realized Fertility:				
Indicator for any children				
	0.892	-0.029	(0.054)	N: 1245; 500
Total number of children				
	2.744	-0.407	(0.291)	N: 1245; 497
Age at first birth				
	22.197	0.873	(0.703)	N: 1056; 613
Number of live births (cond> 0)				
	3.075	-0.330	(0.268)	N: 1057; 451
Under five mortality				
	0.062	-0.003	(0.038)	N: 2103; 913
Under one mortality				
	0.040	-0.009	(0.023)	N: 2817; 1974
Fertility Intentions:				
Desired No. Children (KLPS2)				
	3.265	-0.207	(0.208)	N: 1014; 456
Desired No. Children (KLPS3)				
	3.236	-0.370*	(0.193)	N: 987; 634
Desired No. Children (KLPS4)				
	5.183	-0.098	(0.504)	N: 1014; 636

Notes: This table replicates the Table 3 specification separately for women; pooled estimates appear in Table 3. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and * respectively.

Table A5: Earnings and Consumption Impacts

		Reduced Form		
Outcome	Mean	CC&T RD		Sample sizes
Total earnings, 12m				
	1287.504	-73.686	(203.528)	N: 7414; 4154
Farming profits, 12m				
	13.232	-16.515**	(7.780)	N: 7485; 3104
Self-employment profits, 12m				
	261.090	-87.559	(53.416)	N: 7441; 3889
Wage earnings, 12m				
	906.171	160.911	(194.578)	N: 7406; 3647
Total per capita household consumption, 12m				
	2151.033	28.908	(260.258)	N: 2842; 1367
Total household earnings, 12m				
	1232.128	131.932	(310.000)	N: 2533; 1419
Among men:				
Total earnings, 12m				
	1829.610	-182.290	(287.041)	N: 4168; 2483
Farming profits, 12m				
	15.166	-20.680*	(10.965)	N: 4238; 1944
Self-employment profits, 12m				
	342.443	-60.668	(76.466)	N: 4194; 2066
Wage earnings, 12m				
	1313.018	199.075	(270.856)	N: 4159; 2234
Total per capita household consumption, 12m				
	2474.574	0.952	(412.923)	N: 1583; 746
Total household earnings, 12m				

1466.519	-132.429	(536.669)	N: 1417; 746
Among women:			
Total earnings, 12m			
664.436	-95.852	(213.400)	N: 3246; 1652
Farming profits, 12m			
10.973	-7.931	(6.530)	N: 3247; 2098
Self-employment profits, 12m			
167.453	-128.359*	(74.384)	N: 3247; 1605
Wage earnings, 12m			
438.366	49.390	(176.196)	N: 3247; 2026
Total per capita household consumption, 12m			
1790.864	11.589	(243.622)	N: 1259; 515
Total household earnings, 12m			
966.409	302.747	(238.014)	N: 1116; 577

Notes: Monetary outcomes are expressed in 2017 USD PPP. *Total earnings* equal the sum over the last 12 months of wage earnings across all jobs, non-agricultural self-employment profits across all businesses, and farming profits. *Farming profit* is the net revenue from non-crop and crop farming activities for which the respondent provided all reported household labor hours and was the main decision-maker, defined as output value less cash and in-kind input costs (including hired labor, land rental, storage, and inputs). *Self-employment profit* is net profit from non-agricultural businesses in the last 12 months. *Wage earnings* is the sum of wage earnings across all jobs within the last 12 months. *Total per-capita household consumption* is the total value of food and non-food consumption (purchases, gifts, and home production) in the last 12 months divided by household size, with an urban–rural price adjustment for Nairobi and Mombasa. *Total household earnings* equal the sum of wage earnings, self-employment profits, and agricultural profits across household members in the last 12 months, divided by household size. Reported sample sizes reflect the number of nonmissing observations and those within the RD bandwidth. Standard errors are from `rdrobust` with robust bias correction; significance at the 1, 5, and 10 percent levels is denoted by ***, **, and * respectively.

B Matching Approach

This appendix describes the procedure used to link individual records from the Kenya Life Panel Survey (KLPS) to the Kenya Certificate of Primary Education (KCPE) administrative data. The KLPS dataset contains detailed survey information on pupils, including multiple reported names across survey rounds and records of the primary schools they attended. The KCPE data contain the official examination records with candidate names, schools, years, and exam scores. We follow a strict matching procedure with the goal of linking each KCPE record to a single KLPS respondent, based on consistency between reported names and schools.

Step 1: Assembling identifying information. We began by compiling all available name and school variables from the KLPS surveys. In each KLPS round, respondents are asked to report their name again, and across rounds they often provide slightly different versions or spellings. We take advantage of all names ever reported in KLPS to maximize the chances of identifying a correct match in the KCPE records. In addition, KLPS collects information on schooling history, including the primary schools attended, which we draw on alongside names. On the KCPE side, we use the candidate's full name, school in which they sat for the exam, and the year of examination. By working with the full history of reported names and schools, we aimed to capture as much identifying information as possible, recognizing that spelling or reporting could vary over time.

Step 2: Cleaning and standardizing names. Beyond these generic edits, we created specific replacement rules for common variants, such as standardizing "Maureen/Morine/Maurine" to MAUREEN, "Bernard/Benard" to BERNARD, and "Jackline/Jacqueline" to JACKLINE. We also harmonized common regional variants of surnames, for example treating ODUOR, ODUAR, and ODOUR as the same name standardized to ODUOR. Each name was then broken into tokens (for example, "John Paul Otieno" became {John, Paul, Otieno}), allowing us to match across different name orders and combinations.

Step 3: Cleaning and standardizing schools. A similar cleaning process was applied to school names. We removed suffixes like "Primary School," "PS," or "PRI SCH," and corrected frequent transcription mistakes. In cases where the same school was referred to with multiple spellings or abbreviations, these were collapsed into a single standardized form.

Step 4: Restricting matches by school. The first filter required that survey data

and KCPE records agree on school. Using the cleaned school names, we matched each survey school to the KCPE schools either exactly or through a fuzzy string comparison that allowed for small spelling differences (Jaro-Winkler xx). This step limited the pool of potential matches for each child to those who sat for the KCPE exam at the same school.

Step 5: Matching names within schools. Within the restricted set of schools, we compared the cleaned and tokenized names from KLPS and KCPE. We measured overlap between the sets of tokens, while allowing for small differences in spelling using fuzzy string distance (Jaro-Winkler xx). Candidate matches were then scored based on two factors: (1) the share of KCPE tokens that appeared in the survey names, and (2) whether there was an exact or close school match. The scoring was designed so that a perfect coverage of the KCPE name tokens combined with a school agreement would yield the maximum possible score.

Step 6: Identifying perfect matches. A match was considered "perfect" if the survey names captured all tokens in the KCPE record and the school matched. These are high-confidence matches. If a respondent had exactly one such perfect match, it was accepted automatically. When multiple perfect matches occurred for the same respondent, additional filters were needed to resolve them.

Step 7: Applying additional filters for multiple matches. Two further checks were introduced. First, we compared gender: matches where the KCPE record listed a sex that directly contradicted the survey information were dropped. Second, we checked whether the exam year was consistent with the respondent's reported grade progression in KLPS. Specifically, we estimated the expected KCPE year based on the grade the child reported being in at baseline and compared this with the official exam year. Matches that were implausibly early were removed, and if more than one plausible record remained, we selected the earliest KCPE year consistent with the child's schooling trajectory.

Step 8: Resolving collisions across respondents In some cases, multiple KLPS respondents matched to the same KCPE record. These "collisions" required additional adjudication. We compared how closely each respondent's expected exam year aligned with the actual KCPE year. If only one respondent fell within a small acceptable margin (within zero to two years), that respondent was selected.