# The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis*

Owen Ozier†
Development Research Group
The World Bank

June 1, 2016

**Abstract**

I estimate the impacts of secondary school on human capital, occupational choice, and fertility for young adults in Kenya. Probability of admission to government secondary school rises sharply at a score close to the national mean on a standardized 8th grade examination, permitting me to estimate causal effects of schooling in a regression discontinuity framework. I combine administrative test score data with a survey of young adults to estimate these impacts. My results show that secondary schooling increases human capital. For men, I find a drop in low-skill self-employment; for women, I find a reduction in teen pregnancy.

JEL Codes: I21,I28,J00,O12,O15

---

*I am indebted to Lori Beaman, Blastus Bwire, Pascaline Dupas, Esther Duflo, Steve Fazzari, Frederico Finan, Sebastian Galiani, Justin Gallagher, Francois Gerard, Erick Gong, Joan Hamory Hicks, Guido Imbens, Gerald Ipapa, Pamela Jakiela, Seema Jayachandran, Pat Kline, Michael Kremer, Ashley Langer, Karen Levy, Isaac Mbiti, Jamie McCasland, Justin McCrary, Edward Miguel, Salvador Navarro, Carol Nekesa, Rohini Pande, Bruce Petersen, Robert Pollak, Jon Robinson, Alex Rothenberg, and Kevin Stange, as well as conference and seminar participants, for their helpful comments. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author, and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments of the countries they represent. All errors are my own.

†Please direct correspondence to oozier@worldbank.org.

# 1 Introduction

The expansion of schooling in sub-Saharan Africa over the last fifty years has made basic education more accessible to many of the world's poorest: between 1970 and 2005, average schooling attained by young Africans rose from 2.6 years to 6.1 years, and it continues to grow.[1] Increases in schooling have coincided with rising literacy and formal sector employment across the continent, though the direction of causality is not clear. Wage returns to education have been shown in other developing country contexts (Duflo 2001), but similar patterns have not been demonstrated as convincingly in Africa. One complicating factor is that rates of employment are quite low. In 2008, for example, only 38 percent of Kenyan men were employed by someone outside their family.[2] In this context, the effects of education on human capital accumulation and occupational choice may be more relevant measures of the returns to schooling than wage effects.

Yet recent empirical studies have not provided convincing evidence that African schools even have effects on learning outcomes. Two recent papers find no positive academic effects at all: Lucas and Mbiti (2014) show that admission to higher quality secondary schools in Kenya neither raises the probability of completing secondary school nor increases 12th grade test scores; de Hoop (2010) shows that admission to a higher quality secondary school in Malawi increases the probability of remaining enrolled in an assigned school, but has no effect on test scores. These studies, however, only measure the change in academic performance brought about by increases in secondary

[1]Source: Barro and Lee (2010), tabulation based on 33 countries in sub-Saharan Africa.

[2]Source: KNBS and ICF Macro 2009. This includes all age groups and covers both rural areas and urban centers.

school quality. One might reasonably expect the effect of attending *any* secondary school to be much larger. The rise in primary school completion associated with the achievement of the Millenium Development Goals means that a large cohort is about to reach the age of secondary schooling, which until now has been rationed in much of sub-Saharan Africa. Though the impact of secondary schooling on the marginal admitted student would be the policy-relevant effect if one were considering relaxing this constraint, few studies to date have been able to estimate this effect in Africa.[3]

In this paper, I use a regression discontinuity approach to estimate the impacts of secondary schooling in Kenya. The discontinuity I use is based on a standardized 8th grade test, the Kenya Certificate of Primary Education (KCPE). Probability of admission to government secondary school rises sharply at a cutoff score close to the national mean on the examination. I collect an administrative KCPE dataset, and combine it with a recent, detailed survey of young adults in Kenya that includes educational attainment, along with a number of other outcomes. With these two datasets, I use a technique from time series econometrics to identify the structural breaks in patterns of secondary school completion, thereby locating the test score cutoffs in Kenya's secondary school admission policy. I am able to confirm that the KCPE score popularly perceived to constitute "passing" the examination is empirically the most important for boys, while a slightly lower cutoff appears more relevant for girls. This is consistent with a recent survey of local secondary school administrators, who report lower admissions criteria

[3]Outside Africa, one of the rare examples of studies in a developing country is provided by Filmer and Schady (2014), who found that scholarship-driven attendance of lower secondary school in Cambodia had no effect on either test scores or fertility choices.

for girls.

At the admissions cutoff, I find a 15 percent jump in the probability of completing high school. This is a large effect compared to many commonly used instruments for education. I perform relevant specification tests, and find that, for men, this effect is significant and stable across a range of specifications, bandwidths, controls, and sample restrictions. For women, the first stage is significant in some specifications, but not in others, suggesting that results pertaining to women (i.e., fertility) should be interpreted with caution.

Students on either side of the admissions cutoff are very similar demographically, and in a neighborhood of the test score cutoff, admission to secondary school is "as good as randomized" (Lee 2008). This allows me to treat the rise in schooling at the admissions cutoff as a source of exogenous variation for estimating the impact of secondary school. I find that completing secondary school has a substantial impact on human capital accumulation, as measured by performance on vocabulary and reasoning tests in adulthood. I estimate a performance improvement of 0.6 standard deviations attributable to the completion of secondary school. This is the first paper to show such positive effects of secondary schooling in Africa.

For labor market outcomes, I consider rates of employment and low-skill self-employment. I find clear causal effects: for men in their mid-twenties, completing secondary school decreases the probability of low-skill self-employment by roughly 50 percent. There is also suggestive evidence of a 30 percentage point increase in the probability of formal employment, though this is not significant in all specifications. It is important to note that most self-employment in this context is not innovative entrepreneurship. Instead, it is what Lewis (1954) refers to as "casual labour" or "petty trade," and Schoar (2009) describes as "subsistence entrepreneurship." It is the transition away from this sector that marks economic development (Lewis 1954, p.189).

I also find qualified evidence that secondary schooling causes a sharp drop in the probability of teen pregnancy. Studies of the correlation between education and fertility have emphasized a number of possible ramifications: human capital accumulation in the next generation, rates of population growth, and household bargaining, for example (Strauss and Thomas 1995). My results suggest a causal effect of secondary schooling on early fertility, opening an avenue for further study as this population grows older. While my estimates of the effect are relatively large, this sort of reduction is in accord with the findings of Ferré (2009) and Duflo, Dupas, and Kremer (2015) in Kenya, as well as Baird, Chirwa, McIntosh, and Özler (2010) in Malawi.[4]

Thus, I show large effects of secondary schooling on a number of important outcomes. A feature of this work, as compared to other recent studies on secondary schooling in Africa, is that I estimate impacts on the marginal student who attends secondary school. As a result, these estimates are directly interpretable as consequences of potential policy changes that would make secondary school rationing less restrictive. The magnitude of these effects in a population of this age suggests that permanent differences may be revealed as this cohort grows older, opening a clear avenue for further study.

[4]This contrasts with the work of McCrary and Royer (2011), who find that increases in educational attainment induced by age-at-school-entry rules have no such impact in the US; their instrument acts through a different channel on a different subpopulation, however, partially explaining the difference in findings.

This estimation, of effects on the marginal admittees to secondary school, also has clear limitations. It does not necessarily generalize to students with very different levels of preparation or skill. It does not take into account any congestion effects that might come into force, for example, under a policy of universal secondary schooling. It does, however, demonstrate that secondary schooling plays more than simply a signaling role in this context.

An additional contribution of this paper is to show whether cross-sectional analysis, controlling for available covariates, delivers estimates comparable to the causal effects estimated in the regression discontinuity design. Following the general approach of Altonji, Elder, and Taber (2005), I specifically explore the stability of cross-sectional results in the framework suggested by Oster (2015), with implications for other settings in which quasi-experimental variation may not be available. I find that an academic test control variable is valuable in the context of learning outcomes, but is not nearly as useful for labor market outcomes.

The remainder of the paper is organized as follows: Section 2 provides a description of relevant facets of the Kenyan educational system, and the data I use for estimation.[5] Section 3 explains the estimation strategies employed for different types of analysis. Section 4 presents detailed specification checks and results, Section 5 discusses the robustness of cross-sectional analysis on the same outcomes, and Section 6 concludes.

[5] The data appendix provides additional details on the assembly of these datasets.

# 2 Context and Data

Since 1985, the Kenyan education system has included eight years of primary schooling and four years of secondary (Eshiwani 1990, Ferré 2009). At the end of primary school, students take a national leaving examination, the Kenya Certificate for Primary Education (KCPE). A score of 50% or higher—currently 250 points out of 500—is considered to be a passing grade. This examination is the chief determinant of admission to secondary schools (Glewwe, Kremer, and Moulin 2009).

Those who are not admitted to any government school may choose to re-take the examination the following year, or may consider private schools with different standards, vocational education, or schooling outside Kenya. Though an official letter of admission to a government secondary school is rare below this cutoff, admission is still not guaranteed for those above it because the number of candidates passing the KCPE may exceed the number of spaces available in public schools (Aduda 2008, Akolo 2008). Among those who are admitted to secondary school, however, many are still unable to afford tuition and assorted fees: while primary school has been inexpensive for many years, and was made nominally "free" in 2003, even the lowest-tier district secondary schools cost hundreds of dollars per year during the period (2007-2009) observed in this study.[6]

[6]Policy changes after 2008 made low-tier secondary schools considerably less expensive.

## 2.1 Data: KLPS2 Surveys

This admission rule suggests a fuzzy regression discontinuity design for estimating the impacts of secondary schooling. The primary dataset used in this study is the Kenya Life Panel Survey (KLPS), an ongoing survey of respondents originally from Funyula and Budalangi Divisions of Busia District, Kenya (Baird, Hamory, and Miguel 2008).[7] The respondents were sampled from the population attending grades 2 through 7 at rural primary schools in 1998. The first round of surveying (KLPS1) was carried out from 2003 to 2005, while the second (KLPS2) ran from 2007 to 2009, both times tracking respondents across provincial and even national boundaries.[8] Because the outcomes of interest occur only for adult respondents, I mainly use the more recent round of survey data (KLPS2), treating it as cross-sectional data for 5,084 individuals.

The KLPS2 survey is comprehensive, including questions on education, employment, and fertility, as well as cognitive tests. The education section includes yearly school participation questions, from which secondary school completion, grade repetition, and other measures can be constructed; it also includes self-reported KCPE scores for students who complete primary

[7]While Busia, now one of Kenya's 47 counties, cannot provide a perfect representation of life everywhere in Kenya, it is not atypical. For example, according to the 2014 Demographic and Health Survey report, Busia County literacy rates for women (87.0 percent for women, ) are close to the national average (87.8 percent), while for men Busia's literacy (81.3 percent) is below the national average (92.1 percent). Busia County has lower education levels than does Kenya as a whole (median of 6.5 years for women and 7.1 years for men, compared to national statistics of 7.6 years and 7.9 years, respectively), though it is quite similar to several other counties, including Narok and Lamu, in this regard. (KNBS and ICF Macro 2015)

[8]The effective survey tracking rate in KLPS2 is 82.7 percent. (Baird, Hicks, Kremer, and Miguel forthcoming)

school. The cognitive tests administered as part of the survey assess English vocabulary and non-verbal reasoning; the labor market section includes employment and self-employment history, including the dates and sectors of employment, as well as wages.[9]

In order to use a regression discontinuity design, I restrict analysis to respondents who report completing primary school and taking the KCPE which reduces sample size from N=5,084 to N=3,305. Table 1 shows summary statistics for the restricted KLPS2 sample. The KLPS2 respondents are between 14 and 31 years old (though more than 99 percent are between 16 and 29 years old), with a mean age of roughly 22 years. The 3,305 respondents reporting test scores have higher educational attainment, more educated parents, and lower teen pregnancy rates than the full sample; this is to be expected, since these are the respondents who did not drop out during primary school. The lower panels of Table 1 also describe characteristics conditional on further sample restrictions, including restricting the sample to the males from the oldest two of six primary school grades in 1998; these respondents are 24 years old on average, ranging in age from 19 to 31 years old (though more than 99 percent are between 20 and 29 years old).

[9]Non-verbal reasoning is measured using Raven's Matrices, one of the more reliable measures of general intelligence (Cattell 1971); the vocabulary instrument is based on the Mill Hill test, originally designed by J. C. Raven to complement the Matrices. The employment survey question is, "Are you currently employed, working for pay?" The self-employment survey question is, "Other than in farming, are you currently self-employed or running a business to earn a living?" Neither is required to be exclusive of schooling.

## 2.2 Data: Test Scores

While most KLPS variables are quite stable over survey rounds, self-reported KCPE scores are not. Errors in test scores could pose several problems, since I will use KCPE score as the regression discontinuity running variable. Endogenous re-taking and mis-reporting both contribute to a problematic self-reported test score distribution, shown in the upper left panel of Figure 1, and discussed in greater detail in Appendix Section A.2. An alternative source of test score data is thus essential.

To complement the KLPS data, I gathered an auxiliary dataset of 17,384 official KCPE scores. Records were collected from the relevant District Education Offices and, when the district-level offices did not have the records because of recent political changes, a team visited every primary school from which students had been drawn for KLPS, ultimately providing 88 percent coverage of desired administrative data.[10]

With this set of administrative KCPE data, including names, test years, and schools, I use an algorithm to match records to more than 2,500 of the 3,305 KLPS2 respondents reporting a score.[11] For 2,273 respondents, I find exactly one test score; for 263 more, I find two scores in different (typically consecutive) years. Using the KLPS2 survey to determine whether matched scores are first or second attempts, I am able to clearly identify 2,167 first

[10]Officials were forthcoming in all cases, but because the records had been kept on paper and were, in some cases, more than ten years old, this set of administrative records is missing roughly 12 percent of the data from KLPS schools in the relevant years.

[11]I can cross-check KCPE scores for roughly 77 percent of the KLPS2 respondents who report taking the KCPE. While many self-reported scores are in accordance with the official records, there is substantial misreporting. I discuss the matching process and characterize misreporting in Appendix section A.1.5.

test scores.[12] Their distribution is plotted in the upper right panel of Figure 1 and is tested for a density break in the lower right panel. I find no evidence of manipulation of administratively reported first test scores around the score of 250.[13]

## 2.3 Gender-Specific Discontinuities

The KCPE cutoff for secondary school admission is well-known in Kenya; national media reported that "Out of the over 695,000 candidates who sat the KCPE examination, 350,000 candidates attained over 250 marks, making them eligible to join secondary school."[14] However, a survey of secondary schools in the area suggests that, though 250 is the modal 2009 cutoff score reported by school administrators, many competitive schools use higher cutoffs.[15] Importantly, while no school reported a cutoff below 250 for boys (and many report exactly 250 for boys), seven out of eighteen reported cutoff scores below 250 for girls. As such, 250 may not be the cutoff where the largest fraction of girls are exogenously induced to attend secondary school.[16] To address the problem, I apply a technique from the structural break literature, following Card, Mas, and Rothstein (2008): I first restrict attention

[12] For some cases where I observe only one test score, it either appears to be a second score (because the respondent reports repeating Standard 8 and the score comes from the last year in which the respondent reports enrollment in Standard 8), or it is unclear whether it is a first or second score. I exclude these when using only first test score.

[13] I also plot the density of administrative test scores, including unmatched scores, in Appendix Figure A5, and find no density break.

[14] Excerpted from Akolo (2008).

[15] Edward Miguel and Matthew Jukes, unpublished data (2009).

[16] Because the recent survey only included 2009 cutoffs, I re-visited secondary schools to find out their history of admissions rules, but current school administrations were not able to provide records of admissions rules covering the period of study in this paper.

to a window of scores between 150 and 350 points on the KCPE; I then regress the outcome (completing secondary school) on indicators for hypothetical discontinuities from 200 to 300 points and a piecewise linear control for KCPE score, one potential discontinuity at a time, separately for men and women. For each sex, I consider the discontinuity whose regression produces the highest value of $R^2$ to be the "true" cutoff. Results are shown in Figure 2. For men, the $R^2$-maximizing cutoff is 251 points rather than 250 (a close second place). For women, the best cutoff in this sense is 234 points.[17] Considering these to be the "true" discontinuities, I use these values for the cutoff, $c$, in specification checks for the first stage and in the estimation that follows.[18]

[17]Thus, for men, the the $R^2$-maximizing cutoff is almost exactly the popularly-known and frequently-reported one. It is worth noting that for women, however, the $R^2$-maximizing cutoff is not as close to any value reported by any school surveyed.

[18]Several features of this process are worth noting. Prior to Card, Mas, and Rothstein (2008), this technique was also used in the context of schooling by both Kane (2003) and Chay, McEwan, and Urquiola (2005). Estimation of the location of the discontinuity, in the presence of a discontinuity, is super-consistent (Hansen 2000), and the error is not asymptotically normally distributed; this is also evident in Monte Carlo simulations using a data generating process designed to mimic the one I estimate here. Sampling error in the location of the discontinuity can be ignored in estimation of the magnitude of the discontinuity, so standard errors in subsequent estimation need not be adjusted (Card, Mas, and Rothstein 2008). I use the same data for estimating the location of the discontinuity as for estimating the impact on outcomes; Card, Mas, and Rothstein (2008) have a much larger sample, and are able to use half the data to locate the discontinuity, and the other half to estimate the rest of their model. Since my use of the data could create an endogeneity concern, I carry out robustness checks (selected checks shown in the Appendix) with the highest discontinuity for women below 250 reported by any surveyed secondary school in the region—240 rather than 234—and the *ex ante* cutoff of 250 rather than 251 for men. I obtain similar empirical results, though the first stage loses power substantially for women.

# 3 Empirical Strategy

Consider an equation characterizing the causal relationship between whether an individual completes secondary school, $Sec_i$, and outcome $Y_i$:

$$Y_i = \pi_0 + \pi_1 Sec_i + \pi_2 KCPE_i + \pi_3 X_i + \varepsilon_i \tag{1}$$

Equation 1 controls for academic ability, proxied by KCPE score, $KCPE_i$; other observable individual characteristics, $X_i$; and both a constant term $\pi_0$ and idiosyncratic error $\varepsilon_i$. Estimation of Equation 1 using ordinary least squares (OLS) may lead to biased estimates of $\pi_1$ for the usual reasons: measurement error in educational attainment could bias coefficients downwards, while any positive correlation between $\varepsilon_i$ and $Sec_i$, perhaps due to unobserved ability, could bias estimates upwards (Griliches 1977, Card 2001).

Instead, I use a regression discontinuity approach to identify the effect of secondary school on outcomes. As described in Section 2, Kenyan students who take the primary school leaving examination (KCPE) face an admission rule, possibly by gender ($g$): below a cutoff score, $c_g$, it is more difficult to gain admission to secondary school. The identifying assumptions in my analysis are that all other outcome-determining characteristics except for the probability of secondary school attendance vary smoothly near the cutoff, and that outcomes change at the cutoff only because of the induced change in schooling. Because the probability of attendance does not jump from zero to one, this is a "fuzzy" regression discontinuity (Imbens and Lemieux 2008), so the causal effect of secondary school on outcomes is:

$$\tau_{FRD} = \frac{\lim_{k \downarrow c_g} E[Y|KCPE = k] - \lim_{k \uparrow c_g} E[Y|KCPE = k]}{\lim_{k \downarrow c_g} E[Sec|KCPE = k] - \lim_{k \uparrow c_g} E[Sec|KCPE = k]} \tag{2}$$

As long as the order of polynomial in the running variable and the data window are the same for the first and second stage outcomes, estimation of $\tau_{FRD}$ in equation 2 is equivalent to an instrumental variables approach, where the first and second stages are:

$$Sec_i = \alpha_0 + \alpha_1 Above_i + \alpha_2 K_i + \alpha_3 K_i \cdot Above_i + \alpha_4 X_i + \zeta_i \tag{2a}$$

$$Y_i = \beta_0 + \beta_{FRD} Sec_i + \beta_2 K_i + \beta_3 K_i \cdot Above_i + \beta_4 X_i + \xi_i \tag{2b}$$

In equations 2a and 2b, I use normalized KCPE scores, $K_i = KCPE_i - c_g$, shifted so that the discontinuity occurs at $K_i = 0$; the variable $Above_i$ (the instrument) is equal to 1 if $K_i \geq 0$, and 0 otherwise; the parameter of interest is $\beta_{FRD}$; I allow the relationship between $Y_i$ and $K_i$ to have different slopes on either side of the discontinuity. This is an estimation based on *compliers*, the population who would not complete secondary school if they scored below the cutoff, but who would if they scored above it. The estimated effect is a local average treatment effect at the point in the test score distribution where the cutoff falls. By definition, it is the policy-relevant cutoff for a policy change that would consider moving the cutoff slightly and changing the number of available slots in secondary schools. In this case, however, the cutoff also falls very near the median (and mean) of the test score distribution, which suggests that the effects I measure are relevant for the median Kenyan KCPEtaker, rather than for outliers in the education or skill distribution.[19]

## 3.1 Other Estimation Approaches

In the case of binary outcome variables, such as whether a respondent is pregnant by age 18, a nonlinear instrumental variables approach may be appropriate. In particular, I consider the IV probit, with the same first stage given in equation 2a, but with second stage:

$$\Pr\left[Preg18_i = 1\right] = \Phi\left(\gamma_0 + \gamma_{FRD} Sec_i + \gamma_2 K_i + \gamma_3 K_i \cdot Above_i + \gamma_4 X_i\right) \quad (3)$$

The IV probit estimation procedure is only correctly specified when the first stage residuals are asymptotically normally distributed, and when the first stage is linear.[20] An alternative, when the first stage outcome is binary, is the bivariate probit (Maddala 1983):[21]

$$Sec_i = \mathbb{1}\left(\delta_0 + \delta_1 Above_i + \delta_2 K_i + \delta_3 K_i \cdot Above_i + \delta_4 X_i + \tau_i > 0\right) \quad (4)$$

$$Y_i = \mathbb{1}\left(\phi_0 + \phi_1 Sec_i + \phi_2 K_i + \phi_3 K_i \cdot Above_i + \phi_4 X_i + \omega_i > 0\right) \quad (5)$$

This approach explicitly models endogeneity through the correlation, $\rho$, between $\tau_i$ and $\omega_i$. I follow Greene (2007) and others in imposing a bivariate normal distribution on the error terms. Though in practice, IV probit and

[19] In contrast, many US studies relying on date-of-birth identification strategies are focused on relatively low-achieving students; while studies such as the work of Saavedra (2008) in Colombia estimate the returns only to the highest-quality universities.

[20] A binary endogenous regressor would typically not yield asymptotically normal residuals.

[21] Maddala (1983) presents the model on pp. 122-3; Greene (2007) discusses the model further on pp.823-6; Wooldridge (2002) also discusses it on p.478.

bivariate probit yield marginal effects estimates that are often quite similar to those given by 2SLS, they have the advantage that, when correctly specified, they can provide greater statistical power when the probability of an outcome variable is very close to either zero or one.[22] The cost of this power is additional distributional assumptions, however, so I present results from both linear and nonlinear estimation techniques, when appropriate.

# 4 Results

## 4.1 Specification: Bandwidth and Polynomial Order

For the first stage, I consider a window of data symmetric about the discontinuity, and regress completion of secondary school on an indicator for scoring above the discontinuity and piecewise linear controls in test score. I plot the resulting estimates of the discontinuity magnitude in Figure 3, as a function of the width of the data window; here, I scale down scores by a factor of 100 so that coefficient estimates in subsequent tables are read more easily. The discontinuity estimate fluctuates slightly, but remains significant and of similar magnitude no matter which bandwidth I use.[23] At each bandwidth, I carry out a specification test in which in addition to the discontinuity dummy and the piecewise linear controls, I include indicators for narrow-width bins

[22]This can be shown in Monte Carlo simulations, for example.

[23]Here I use the term *bandwidth* in the sense of Imbens and Lemieux (2008), Lee and Lemieux (2010), and others in the regression discontinuity literature to mean the window of data used for estimation; this is not a non-parametric regression; I use a "rectangular kernel," meaning I do not weight data differently according to distance from the discontinuity. Estimates indicate that the probability of high school graduation increases by between 15 and 24 percentage points depending on choice of bandwidth.

of KCPE scores: 251-260, 261-270, et cetera.[24] I test these indicators for joint significance; if they are significant, I consider the piecewise linear first stage to be mis-specified. This test rejects for widths of 90 points and higher on either side of the discontinuity. The same is true when I include a piecewise quadratic control in test score. Thus, for the rest of this paper, I use a bandwidth of 80 points on either side of the discontinuity.[25] Finally, I use Akaike's information criterion to confirm that the first-order polynomial control is sufficient: piecewise linear (as opposed to constant, quadratic, cubic, or quartic) is the "best" specification according to AIC for both the 80-point bandwidth and nearly all other bandwidths under consideration. I use the same bandwidth and order of polynomial (linear) in both the first and second stage estimation, so that I can simply use 2SLS both for estimation and standard errors.[26]

I carry out validity tests of the smoothness assumption using observables, four of which are depicted graphically in Figure 4. Gender, age, and mother's and father's education vary smoothly at the boundary, with differences that are neither large enough to be important nor statistically significant. This contrasts with Urquiola and Verhoogen (2009), who show that schools' responses to a class-size policy discontinuity in Chile can invalidate a regression discontinuity research design. While they find large and significant differences

[24]For this test, I follow Lee and Lemieux (2010) and Lee and McCrary (2009). The results are similar when I vary bin width, for example using a width selected by a leave-one-out cross-validation procedure.

[25]Alternatively, I can use the procedure suggested by Imbens and Kalyanaraman (2012); this yields similar or smaller "optimal" bandwidths, depending on the outcome. These bandwidths are shown for differing outcomes and subsamples in Table 5, and with the alternative cutoff of 250 for men, in Appendix Table A9.

[26] See, in particular, Lee and Lemieux (2010) Section 4.3.3.

in parents' education levels at the discontinuity (as well as sharp changes in the class size histogram near cutoffs), I find no such patterns here.[27]

Though the fraction female does not change with statistical significance at the gender-specific discontinuity, the appearance of a small change in the lower left panel of Figure 4 suggests that the density of test scores for either men or women could change at the estimated discontinuity location. Tests of this possibility are shown in Appendix Figure A4. While for men, there is no density change, women appear less likely to be found immediately to the right of the estimated discontinuity location. Though this is not likely to be because of misreporting (because of the use of administrative data) or test score manipulation (since women are more likely to appear to just fail the test than to just pass it), I speculate that it may relate to differential difficulty tracking respondents, or perhaps even to sampling variation in response rates across the score distribution interacting with the structural break search procedure. Either way, this density difference at the discontinuity for women should be borne in mind; it implies that additional caution is required in interpreting the results for women in this sample.

## 4.2 First Stage: Discontinuity

The first stage discontinuity is shown in the upper-left panel of Figure A6, and in a regression framework in Table 2.[28] In Table 2, the discontinuity is

[27] See Section A.1.5 and the right panels of Figure 1. In particular, while I cannot rule out all types of cheating on the KCPE, as in the Texas testing context investigated by Martorell (2004), none of the known mechanisms for cheating on the exam would permit endogenous sorting around the discontinuity.

[28] In this case, because the data window constrains predictions to within the unit interval, a logit or probit specification yields marginal effects that are almost identical in magnitude

estimated first with genders pooled (Columns 1-3), then separately among men (Columns 4-6) and women (Columns 7-9).[29] I show the results with and without a piecewise quadratic control and controls for other covariates: age, gender, parents' education levels, and cohort dummies. I cannot reject that the discontinuities for men and women are of the same magnitude, though the smaller point estimate for women is consistent with the lower overall level of secondary schooling for women in this setting. My preferred specifications are given in Columns (2), (5), and (8), in which the discontinuity is measured as a 16-percentage-point change in the probability of completing secondary school for men; a 13 percent change for women, and a 15 percent change when pooled.[30] [31] That controls do not substantially change the point estimate is unsurprising, given that they do not change significantly at the discontinuity.

When the estimation is carried out separately by gender, the discontinuity is significant for both men and women, but the F-statistic is now below the rule of thumb for weak instruments for the subsample of women (Stock

and significance to the discontinuity estimated here in a linear probability model.

[29]Separate first-stage discontinuity estimates by gender are shown in Appendix Figure A7.

[30]Decomposition as suggested by Gelbach (2016) shows that the change in coefficient magnitude from Column (7) to Column (9) is mostly due to the inclusion of the covariate controls; the slightly larger standard error is brought about because of the inclusion of the piecewise quadratic in the running variable. A separate issue is that small fraction of the sample is still in school; this fraction varies slightly at the discontinuity, and as such, the completion of secondary schooling may be viewed as a censored outcome in the first stage, which could be the source of some bias. In practice, restricting the sample to respondents who are surveyed at least five years after they take the KCPE does not substantially alter the results.

[31]Also note that while I find a larger discontinuity for men than women, Uwaifo Oyelere (2010) found that variation in free primary education in Nigeria predicts years of education equally well for men and women. This could be because free primary school induces additional schooling at too young an age for womens' early marriage and fertility decisions to be relevant, and would have been especially true in the period when Nigeria's primary education system was first coming into existence, included in that analysis.

and Yogo 2002). However, because the model is just-identified, the weak-instruments bias towards OLS is not present (Angrist and Pischke 2009), though tests may not be correctly sized.

I also note that, were I to use the popularly-known cutoff of 250, about which there is no evidence of differential attrition in the McCrary density test, rather than the estimated cutoffs, the first stage for men is largely unaffected (because the change is only a single point), while for women, it is no longer statistically significant; this robustness check is shown in Appendix Table A8. This, again, warrants caution in interpreting results for women.

In Figure 5, I show the estimated difference between the cumulative distribution functions (CDFs) for education of the populations on either side of the discontinuity. For each point in the figure, I estimate a separate regression of the probability of attaining more than $x$ years of education (one minus the CDF) on a piecewise linear control and an indicator for the discontinuity. Thus, I estimate the difference in CDFs (probabilities) between those who are just to the right and just to the left of the discontinuity: the impact of passing the cutoff score on the probability of attaining more than each number of years of schooling, $x$; the plot shows the coefficients and confidence intervals on the discontinuity for each of these outcomes. The KCPE discontinuity as an instrument clearly predicts secondary schooling, and moreover, secondary school completion. The estimates, however, drop to insignificance when estimating the probability of attaining more than 12 years of schooling: the KCPE score that induces marginal students to attend and complete secondary school does not induce them to attend a university.

## 4.3 Estimation of Outcomes

### 4.3.1 Human Capital

I begin with analysis of the impact of schooling on human capital. The KLPS2 survey includes a commonly used test of cognitive ability—a subset of Raven's Progressive Matrices—and an English-language vocabulary test based on the Mill Hill synonyms test. Adaptations of both measures have been used internationally for several decades, and each captures different aspects of intelligence.[32] I standardize both outcomes so that they are measured in terms of standard deviations in the KLPS2 population, and show both OLS and 2SLS results for a combined Z-score[33] and separately by test in Panel A of Table 3: completing secondary school improves performance on these tests by 0.6 standard deviations, with very similar estimates given by 2SLS and (potentially biased) OLS.[34] This estimate is robust to the inclusion of controls (Column 4), and when decomposed, is driven by the larger and more precisely estimated effect on vocabulary.[35] The reduced form effect, roughly 0.1 standard deviations at the discontinuity, is shown in the upper

[32]Though standardized to have mean zero and standard deviation one in the population, in Table 1 these two cognitive measures have positive mean and standard deviations slightly less than one, because these summary statistics are only shown for the sample with a restricted range of first KCPE scores. The "Matrices" are often considered to measure something akin to "fluid" intelligence, while the vocabulary test measures something more related to what specialists in the field call "crystallized" intelligence (Cattell 1971). The relationship of the two measures appears similar here to in other settings: in these data, as elsewhere (Raven 1989), their correlation is near 0.5.

[33]Combined Z-score is equivalent to the Kling, Liebman, and Katz (2007) "mean effect."

[34]Note that this OLS specification already includes KCPE score as a control, and restricts the sample substantially; detailed comparisons of OLS and 2SLS are provided in Section 5.

[35]When I use the bandwidth and kernel suggested by Imbens and Kalyanaraman (2012), I find even larger impacts of secondary schooling on human capital, as shown in Column 2 of Table 5.

right panel of Appendix Figure A6. To the extent that subsequent outcomes depend on a mixture of human capital and signaling, this is evidence that secondary schooling in Kenya does not play a purely signaling role: students measurably gain skills from schooling.[36]

These results contrast with the recent work of Filmer and Schady (2014) in Cambodia, who show that increased secondary schooling has no impact on subsequent test scores, as well as the work of Lucas and Mbiti (2014), who show that increased quality of secondary schooling in Kenya (at higher discontinuities in KCPE score) has no impact on academic outcomes.[37]

A clue to reconciling Lucas and Mbiti's findings with mine may lie in the recent work of Pop-Eleches and Urquiola (2013). Using a similar multiple-discontinuity design to estimate the returns to secondary school quality in Romania, they find very modest positive effects, around .04 standard deviations on an academic test. These effects are simply too small to be detectable in the Lucas and Mbiti (2014) study, and when compared with the results I show in Table 3, it is clear that attending *any* secondary school could simply have a much larger effect than increasing the quality of the school.[38]

[36]A pessimistic interpretation might hypothesize that the longer respondents have been out of school, the worse they perform on tests; since secondary schooling delays exit from school, the apparent positive effect is simply a delayed deterioration of human capital. The data do not support such an interpretation: the longer respondents have been out of school (and thus the older they are), the *better* they do on the tests administered during KLPS2; even if it had the opposite sign, the coefficient would still be too small (around 0.02 standard deviations per additional year out of school) to explain an effect more than an order of magnitude larger; and, in any event, the effect remains significant and of the same magnitude in both OLS and 2SLS after controlling for duration out of school.

[37]This appears to be true even when the marginal student admitted into the school is not the worst student in the higher-quality school.

[38]de Hoop (2010) also finds no positive effects of secondary school quality on a standardized test outcome in Malawi, but this is in keeping with the aforementioned studies. On the other hand, the Lucas and Mbiti (2012) finding that increased primary schooling

Reconciling the results here with those of Filmer and Schady is more difficult. A similarity between data from their setting and data in KLPS is that additional completed grades are associated with between 0.2 and 0.3 standard deviations on the cognitive tests. However, their study and this one ultimately examine two (related, but) different sources of variation in schooling, operating in two different contexts, with two different follow-up periods. Any of these differences could be important. For example: the LATE that Filmer and Schady estimate in Cambodia may be for lower-ability students, less able to benefit from additional schooling; alternatively, those who comply with a scholarship treatment in Cambodia (for whom cost drives the schooling decision) may have lower returns to schooling than those who would comply with an admission treatment.

### 4.3.2 Self-Employment and Employment

Next, I examine the impact of education on labor market outcomes. Because many of the younger respondents are still in school, and because men are typically primary earners in Kenya, I consider only the oldest two cohorts of men for this analysis, so that the incapacitation effect of continued schooling does not dominate the patterns of interest.[39] (For reference, I also consider

actually reduced average performance on the KCPE exam is driven by the setting: the universal primary education policy they study couples an increase in years of schooling with increased enrollment. While test scores might have risen for some students who received more schooling, Lucas and Mbiti (2012) note that the class size and compositional changes overwhelm any positive effect on test scores.

[39] As shown in Panel D of Table 1, only 13 percent of the men in the oldest two cohorts are still in school (as compared to 44 percent in the younger four cohorts). Human capital effects of secondary school remain broadly similar when limiting the sample to respondents who were in standards 6 and 7 in 1998, though standard errors widen (predictably) with the lower sample size; results shown in Panel B of Table 3.

these outcomes for women in Appendix Table A11.) According to 2008 Demographic and Health Survey (DHS) data, young men in Kenya without secondary school have a higher employment rate at age 20 than do men who complete secondary schooling, since the latter group has had less time to look for jobs. At roughly age 25 (near the mean age of the older two male KLPS2 cohorts), DHS data show roughly equal employment rates in these two groups; as they grow older still, the better educated are more likely to be employed. I confirm exactly this pattern in KLPS2, shown in Panel A of Table 4, Columns 1 and 2. OLS shows a fairly precise zero effect of secondary schooling on employment at this age.[40] However, the regression discontinuity approach gives very different results: the coefficient on schooling is positive and significant depending on controls, shown in IV probit and bivariate probit specifications in Columns 3-6. While 2SLS is positively signed, it is insignificant; this may be in part because 2SLS is less efficient than estimation via IV probit and bivariate probit when the true model is nonlinear and the mean of the response variable is close to zero or one, as in this case.[41] Depending on the specification, I find a rise in employment of between 24 and 43 percent in response to secondary schooling.

Besides being employed by someone outside their family, many respondents are self-employed. Of these, 88 percent have no employees: common self-employment occupations in KLPS2 include fishing, hawking, and working as a "boda-boda" bicycle taxi driver. On the other hand, among the employed respondents, the degree of skill varies among unskilled (loader of goods onto vehicles), semi-skilled (factory worker, carpenter, mechanic), and

[40] The same is true of probit in the cross-section, shown in Appendix Table A7 Panel A.

[41] As a diagnostic, predicted values from 2SLS clearly lie outside the unit interval.

high-skill professional occupations (electronics repair, teachers, and other government and NGO employees).

As in other labor market studies of relatively young men (Griliches 1977, Zimmerman 1992), I use sector of employment rather than wage to estimate the impact of secondary schooling. Clear patterns emerge when I measure the effect of education on (implicitly low-skill) self-employment, shown as a reduced form graph in the lower left panel of Figure A6, and presented in the second row of Table 4. While secondary education and self-employment are negatively associated in the cross-section (Columns 1 and 2),[42] the causal impact of secondary schooling on low-skill self-employment is much larger; marginal effects from IV probit and bivariate probit estimation are in broad agreement with the 2SLS coefficients: a 40-50 percent lower probability of being self-employed among those who go to secondary school because they pass the KCPE cutoff.[43]

As an additional robustness check in relation to labor market outcomes, I construct an "unemployment" variable, indicating which respondents are neither employed, self-employed, nor in school (though that is relatively uncommon in this age range). I then use 2SLS and IV probit to estimate effects on this outcome, shown in Appendix Table A10. I find no significant effects in any specification, for either of two variants of the outcome (depending on how I handle vocational training). This is consistent with the pattern of increases

[42]The same is true of probit in the cross-section, shown in Appendix Table A7 Panel B.

[43]When using the Imbens and Kalyanaraman (2012) bandwidth, 2SLS and IV Probit coefficients without controls are similar to those discussed here: for the employment outcome, they are similar in magnitude (and still statistically insignificant); for self-employment, they are larger in magnitude (and still statistically significant). These robustness checks are shown in Columns 3-6 of Table 5.

in formal sector employment being offset by decreases in self-employment at this age, as discussed in the preceding paragraphs.

Finally, I show that the main patterns of results are robust to simultaneously excluding women from the sample entirely and using the popularly-known cutoff of 250 (rather than the estimated cutoff), while restricting analysis to the Imbens-Kalyanaraman bandwidth. This check is shown in Appendix Table A9.

#### 4.3.3 Fertility

While labor market outcomes are of primary interest for the men in this sample, fertility and health outcomes are of importance for women. In Panel B of Table 4 (and in a reduced form graph, shown in the lower right panel of Appendix Figure A6), I look at the probability of pregnancy by age 18 among female KLPS2 respondents.[44] The association between secondary schooling and decreased early fertility is strong: in the first two columns, OLS shows a roughly twelve percentage point drop in teen pregnancy among secondary school finishers.[45] While these are only cross-sectional associations, their sign agrees with associations seen elsewhere in the world, summarized by Schultz (1988). Two-stage least squares (shown in the last columns) predicts outside the unit interval, since again, this is a low-probability outcome, so, in Columns 3 through 6, I use IV probit and bivariate probit estimation and find a near elimination of teen pregnancy among compliers at the discontinuity,

[44] Note that the change in KCPE density at the data-driven cutoff for women warrants caution in interpreting these results.

[45] The same is true of probit in the cross-section, shown in Appendix Table A7 Panel C.

robust to the inclusion of the usual controls.[46] [47]

This contrasts with the work of McCrary and Royer (2011), who find no conclusive effect of education on timing of womens' first births. Their study, however, is based on variation in the timing of school entry, rather than school exit, as is the case here. It is school exit timing that is more likely to interact with fertility. Thus, while they find essentially no impact of education on early fertility, it may still be sensible that in contrast to their work, I find large effects.[48] While Filmer and Schady (2014) also find no effect of additional schooling in Cambodia on pregnancy rates, the median respondent in their survey was still only 14 years old.

Other studies in sub-Saharan Africa have found similar (though smaller) effects of schooling on teen pregnancy in relation to those I show here. Ferré (2009) finds that a policy shift reclassifying 8th grade from secondary to primary school increased the fraction of students reaching 8th grade, thereby reducing teen pregnancy by 10 percentage points in Kenya in the 1980s.

[46]Using the Imbens and Kalyanaraman (2012) bandwidth, I find even more statistically significant impacts of secondary schooling on fertility, as shown in Columns 7 and 8 of Table 5.

[47]Since many of the secondary schools are single-sex, one interpretation could be that teens in secondary school simply see members of the opposite sex less frequently than they otherwise would, so lower rates of pregnancy follow. This interpretation is not supported by the data, though: when I categorize secondary schools as single-sex or mixed, I see no significant difference in the pregnancy decline across the two types of schools. In the cross section, the reductions in teen pregnancy associated with going to the two types of schools are also similar and statistically indistinguishable: 9 percentage points for girls at mixed schools, and 10 percentage points for those who attend all-girls' schools.

[48]Their date-of-birth instrument also predicts educational attainment among those who, for the most part, stop schooling almost as soon as possible. In my case, however, the KCPE discontinuity only has an effect on those who would consider continuing beyond primary education, given the opportunity. These may be higher ability students, relative to the Kenyan distribution, than the McCrary and Royer (2011) students in relation to the US distribution.

Duflo, Dupas, and Kremer (2015) observe a 1.5 percentage point reduction in teen childbearing in Kenya in response to a school uniform distribution program that helped girls stay in school; and Baird, Chirwa, McIntosh, and Özler (2010) find that a conditional cash transfer to bring dropouts back into school reduces teen pregnancies by 5 percentage points in Malawi. In Kenya, the practical mutual exclusivity of pregnancy and schooling means that high-ability girls at the KCPE discontinuity may simply face a choice between attending secondary school and starting a family immediately.[49]

# 5 Cross-Sectional Analysis with Key Covariates

Though the present analysis turns on a clear source of quasi-experimental variation, not all contexts offer such opportunities. Without such variation, any analysis—of the impacts of schooling, in this case—is concerned with whether omitted variables may complicate the measurement or interpretation of empirical patterns. The nature of omitted variable bias, in turn, depends on both the importance of unobserved determinants of an outcome, and the correlation of those variables with the independent variable of interest.

[49]In Kenya, dropping out of school is more common among girls than boys, and is most pronounced once girls enter their teens (Kremer, Miguel, and Thornton 2009). This is closely linked to pregnancy: girls in the Kenyan schools are "required to discontinue their studies for at least a year[50]" if they become pregnant. Schooling and childbearing in Kenya are in practice nearly mutually exclusive, as is true in many other contexts (Field and Ambrus 2008). Though I am aware of no *rule* prohibiting teen mothers from returning to school—though rules of that sort exist in other sub-Saharan countries (Ferré 2009)—teen mothers still face stigmatization in Kenyan primary and secondary schools (Omondi 2008), so even after giving birth, they are unlikely to continue their schooling.

This study creates at least two opportunities for assessing the level of omitted variable bias in a cross-sectional analysis. The first approach is to simply compare the regression discontinuity estimates to cross-sectional OLS estimates. If they are in agreement, OLS may not be a bad approach in this setting. While the cross-sectional approach estimates the average treatment effect (ATE) rather than the local average treatment effect (LATE) at the discontinuity, this comparison may still be informative.

The second approach is to measure whether the OLS estimates vary with the inclusion of controls. Following Altonji, Elder, and Taber (2005), the movement of point estimates with the inclusion of controls speaks to selection on observables; if one assumes an econometric similarity between selection on observables and selection on unobservables, the stability of an estimate under the inclusion of controls speaks to its stability more generally. In this framework, Oster (2015) has pointed out that the movements in R-squared are as equally important; intuitively, if the observables explain very little variation, they may simply be the wrong variables for the question at hand, and may not tell us much about variation explained by the unobservables. I follow Oster (2015) and Gonzalez and Miguel (2014) in assessing the stability of OLS estimates. The approach entails setting a maximum $R^2$ that could be attained if unobservables were included, and asking how much the coefficient of interest could change; a conservative step that Oster discusses is to scale up current $R^2$ by a factor of 1.3.

For the human capital and labor market outcomes, I begin by comparing OLS to the regression discontinuity estimates, then I discuss the Oster approach to stability. Because of the issues surrounding the discontinuity for women, I do not carry out this exercise for fertility.

In Tables 6, 7, and 8, I show how the cross-sectional (OLS) coefficient changes over two sample restrictions, and the inclusion of KCPE score as a control. The first sample restriction is a restriction to KCPE test-takers; the second is the 80-point bandwidth of test scores near the cutoff.

The estimated relationship between secondary schooling and human capital (combined vocabulary and Raven's Matrices score), shown in the first column of Table 6, is twice as large as the regression discontinuity estimate (and is statistically distinguishable from it). Restricting the sample in Columns 2 and 3 brings the coefficient much closer to the regression discontinuity estimate, and including KCPE score as a control reduces it even further (though none of Columns 2, 3, and 4 is statistically distinguishable from the estimate in Column 5). This pattern is consistent with OLS being biased by unobserved ability in Column 3 as compared to Column 4. A regularity across these specifications is that secondary schooling always appears to have an effect on subsequent test scores in this context.

Including a measure of earlier ability (KCPE) as a control increases $R^2$ appreciably; another way of saying this is that the KCPE control has a t-statistic of 12.8 in this regression: the 8th-grade test score is strongly predictive of the survey-based test score years later. Using the criterion suggested by Oster (2015), I test whether additional unobservables could drive the true cross-sectional effect to zero, and find that under the assumptions she suggests, unobservables would not change the finding that secondary schooling in Kenya increases subsequent human capital.[51] Thus, having a pre-treatment

[51]Specifically, I use Oster's `psacalc` program in Stata, first testing coefficient stability if unobservables could scale $R^2$ up by 1.3, then with a maximum feasible $R^2$, denoted

measure of academic capabilities could be sufficient for analysis of impacts on learning outcomes with reasonably low-magnitude bias.

When I turn to labor market outcomes in Tables 7 and 8, the patterns are quite different. For both outcomes, the cross-sectional point estimates are of much smaller magnitude than the point estimate from the regression discontinuity; however, for both outcomes, given the wide confidence interval in the regression discontinuity design, it is not statistically distinguishable from the OLS values.

In the case of employment, the cross-sectional relationship is never statistically distinguishable from zero; in the case of self-employment, it is always significantly negative. These regressions never have very high $R^2$ values, however, and the inclusion of the KCPE control does not change this by very much. As such, the bounding approach of Oster (2015) is very sensitive to reasonable choices of the maximum $R^2$ that unobservables would yield. For both outcomes, the bounding approach cannot reject null effects or effects of the opposite sign with $R_{max} = 0.1$. However, with Oster's suggested approach (for example, setting the ratio of maximum $R^2$ to the current $R^2$ equal to 1.3), the self-employment pattern appears robust in the cross-section. It is worth noting that an alternative control, respondent age, is much more predictive of employment outcomes than KCPE score is, but this control does not improve $R^2$ for self-employment at all. In sum, this does not lead to any instructive conclusions for non-experimental estimates of schooling impacts on labor market outcomes in this setting: OLS and 2SLS are not in

$R_{max}$ equal to 0.5. This follows the suggestions in Oster (2015) in light of the discussion in Gonzalez and Miguel (2014); the Vocabulary + Raven's outcome in the KLPS survey is constructed from relatively few items (compared to an hour-long test like the PIAT to which Oster refers).

close accord, and $R^2$ is low, so the potential role of omitted variables, not well-proxied-for by KCPE score, still looms large.

# 6 Conclusion

This paper identifies the effects of secondary schooling using a discontinuous change in secondary school admission resulting from score cutoffs on Kenya's end-of-primary-school examination. Secondary schooling in Kenya has large effects on human capital, reducing low-skill self employment and weakly increasing formal employment for older cohorts of young men by the time of the survey. I also find qualified evidence that teen pregnancy is sharply reduced by secondary schooling.

The discontinuity I exploit occurs near the mean score—and highest density of scores—on the national primary school leaving examination: perhaps as policy-relevant or externally valid as a discontinuity could be. An expansion of secondary schooling that preserved the quality of secondary schools but reduced the minimum required score would be likely to bring about the effects I esimate on roughly 15 percent of the population near the discontinuity: the compliers. As governments (including Kenya's) consider the expansion of secondary schooling against other policy options, this study should provide a useful guidepost for understanding the consequences of such expansion, as long as it does not substantially alter the characteristics of the schools.

The difference between the unambiguously positive human capital findings in this paper and the less cheery conclusions from other studies of education in Kenya suggest that increased school enrollment in sub-Saharan Africa will have varying consequences, depending on how it is undertaken. The findings in this paper, and in other experimental and quasi-experimental papers, are contingent on the nature of the exogenous variation: the secondary school admission instrument I use, at the KCPE discontinuity, induces both a rise in secondary school completion, and a resulting delay in pregnancy among female compliers; a date-of-birth instrument in the US that also induces additional secondary education has no such effect, however, perhaps both because of the timing of the education effects and because of the underlying skills and preferences of compliers with the different instruments.

In terms of robustness, the pattern of findings across outcomes is robust to specifying the instrument in terms of either years of schooling or an indicator for completing secondary school, and is robust to alternative bandwidths as suggested by Imbens and Kalyanaraman (2012). Besides these variations, the results for men are largely robust to changes in control variables, discontinuity definition, and other empirical variations. Results for women are not as robust to such variations, both because the popularly-known KCPE cutoff does not induce a statistically significant change in secondary schooling in this sample, and because at the estimated cutoff, there is evidence of potential differential attrition for women. Thus, this paper's findings for men may be viewed as more definitive than those for women.

OLS and 2SLS do not always produce similarly signed effects in this analysis: cross-sectional analysis does not reveal the impact of secondary schooling on employment on this age, but in a causal framework, the pattern emerges. In the cross-section, controlling for academic ability is clearly important for outcomes that are closely linked to it, such as the human capital measure in this study. Ability, traditionally linked to academic tests, may not be nearly as useful a control for other outcomes, at least for the young Kenyan adults in this study.

This study also highlights a possible avenue for researchers interested in the consequences of education throughout sub-Saharan Africa: many countries have examinations much like the KCPE, with analogous cutoff rules for secondary school admission. An important caveat is that while some questions in survey data show a very high degree of reliability, this cannot be said for KCPE scores. Combining administrative test data with a rich follow-up survey overcomes this obstacle, and may yield novel findings establishing causal links between education, fertility, and labor markets throughout the developing world.

# References

Aduda, D. (2008): "Girls Shine in KCPE," *The Daily Nation*, December 30, Online: http://www.nation.co.ke/News/-/1056/508980/-/u0q1pe/-/index.html, Accessed on May 24, 2016.

Ajayi, K. F. (2013): "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana," Discussion Paper 259, IED.

Akolo, J. (2008): "KCPE results indicate highest gender parity," *Kenya Broadcasting Corporation*, Online: http://www.kbc.co.ke/story.asp?ID=54707, Accessed on May 31, 2009.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113(1), 115–184.

Angrist, J. D., and J.-S. Pischke (2009): *Mostly Harmless Econometrics*. Princeton University Press, Princeton.

Baird, S., E. Chirwa, C. McIntosh, and B. Özler (2010): "The Short-Term Impacts of a Schooling Conditional Cash Transfer Program on the Sexual Behavior of Young Women," *Health Economics*, 19(S1), 55–68.

Baird, S., J. Hamory, and E. Miguel (2008): "Tracking, Attrition and Data Quality in the Kenyan Life Panel Survey Round 1 (KLPS-1)," Working Paper C08-151, Center for International and Development Economics Research, University of California at Berkeley.

Baird, S., J. H. Hicks, M. Kremer, and E. Miguel (forthcoming): "Worms at Work: Long-run Impacts of a Child Health Investment," *Quarterly Journal of Economics*.

Barro, R. J., and J.-W. Lee (2010): "A New Data Set of Educational Attainment in the World, 1950-2010," Working Paper 15902, National Bureau of Economic Research, Online: http://www.nber.org/papers/w15902, Accessed on May 24, 2016.

Card, D. (2001): "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica*, 69(5), 1127–1160.

CARD, D., A. MAS, AND J. ROTHSTEIN (2008): "Tipping and the dynamics of segregation," *Quarterly Journal of Economics*, 123(1), 177–218.

CATTELL, R. B. (1971): *Abilities: Their Structure, Growth, and Action.* Houghton Mifflin Company, Boston.

CHAY, K. Y., P. J. MCEWAN, AND M. URQUIOLA (2005): "The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools," *American Economic Review*, 95(4), 1237–1258.

DE HOOP, J. (2010): "Selective Secondary Education and School Participation in Sub-Saharan Africa: Evidence from Malawi," Discussion Paper TI 2010-041/2, Tinbergen Institute.

DUFLO, E. (2001): "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91(4), 795–813.

DUFLO, E., P. DUPAS, AND M. KREMER (2015): "Education, HIV, and Early Fertility: Experimental Evidence from Kenya," *American Economic Review*, 105(9), 2757–2797.

ESHIWANI, G. S. (1990): "Implementing Educational Policies in Kenya," Africa Technical Department Series Discussion Paper 85, The World Bank.

FAN, J. (1992): "Design-adaptive Nonparametric Regression," *Journal of the American Statistical Association*, 87(420), 998–1004.

FERRÉ, C. (2009): "Age at First Child: Does Education Delay Fertility Timing? The Case of Kenya," Policy Research Working Paper 4833, The World Bank.

FIELD, E., AND A. AMBRUS (2008): "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh," *Journal of Political Economy*, 116(5), 881–930.

FILMER, D., AND N. SCHADY (2014): "The Medium-Term Effects of Scholarships in a Low-Income Country," *Journal of Human Resources*, 49(3), 663–694.

GELBACH, J. (2016): "When Do Covariates Matter? And Which Ones, and How Much?," *Journal of Labor Economics*, 34(2), 509–543.

Glewwe, P., M. Kremer, and S. Moulin (2009): "Many Children Left Behind? Textbooks and Test Scores in Kenya.," *American Economic Journal: Applied Economics*, 1(1), 112–135.

Gonzalez, F., and E. Miguel (2014): "War and Local Collective Action in Sierra Leone: A Comment on the Use of Coefficient Stability Approaches," mimeo, University of California at Berkeley.

Greene, W. H. (2007): *Econometric Analysis*. Prentice Hall, Upper Saddle River, 6th edn.

Griliches, Z. (1977): "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45(1), 1–22.

Hahn, J., P. Todd, and W. Van der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201–209.

Hansen, B. E. (2000): "Sample Splitting and Threshold Estimation," *Econometrica*, 68(3), 575–603.

Imbens, G. W., and K. Kalyanaraman (2012): "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, 79(3), 933–959.

Imbens, G. W., and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142(2), 615–635.

Kane, T. J. (2003): "A Quasi-Experimental Estimate of the Impact of Financial Aid on College-Going," Working Paper 9703, National Bureau of Economic Research.

Kenya National Bureau of Statistics (KNBS), and ICF Macro (2010): *Kenya Demographic and Health Survey 2008-09*. KNBS and ICF Macro, Calverton, Maryland.

——— (2015): *Kenya Demographic and Health Survey 2014*. KNBS and ICF Macro, Calverton, Maryland.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75(1), 83–119.

Kremer, M., E. Miguel, and R. Thornton (2009): "Incentives to Learn," *Review of Economics and Statistics*, 91(3), 437–456.

Lee, D. S. (2008): "Randomized experiments from non-random selection in U.S. House elections," *Journal of Econometrics*, 142(2), 675–697.

Lee, D. S., and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(2), 281–355.

Lee, D. S., and J. McCrary (2009): "The Deterrence Effect of Prison: Dynamic Theory and Evidence," mimeo, University of California, Berkeley.

Lewis, W. A. (1954): "Economic Development with Unlimited Supplies of Labour," *Manchester School of Economic and Social Studies*, 22(2), 139–191.

Lucas, A. M., and I. M. Mbiti (2012): "Access, Sorting, and Achievement: the Short-Run Effects of Free Primary Education in Kenya," *American Economic Journal: Applied Economics*, 4(4), 226–253.

——— (2014): "Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya," *American Economic Journal: Applied Economics*, 6(3), 226–253.

Maddala, G. S. (1983): *Limited-dependent and qualitative variables in econometrics*, Econometric Society Monographs. Cambridge University Press, New York.

Martorell, F. (2004): "Do High School Graduation Exams Matter? A Regression Discontinuity Approach," mimeo, University of California at Berkeley.

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

McCrary, J., and H. Royer (2011): "The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth," *American Economic Review*, 101(1), 158–195.

Miguel, E., and M. Kremer (2004): "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72(1), 159–217.

Nichols, A. (2016): "RD: Stata module for regression discontinuity estimation," Online: http://EconPapers.repec.org/RePEc:boc:bocode:s456888, Accessed on May 24, 2016.

Omondi, G. (2008): "Teen Mothers Face Ridicule in Schools," *The Daily Nation*, May 7, Online: http://allafrica.com/stories/200805070435.html, Accessed on May 24, 2016.

Orlale, O. (2000): "Fewer Exams in 8-4-4 Shake-Up," *The Daily Nation*, September 9, Online: http://allafrica.com/stories/200009260281.html, Accessed on May 24, 2016.

Oster, E. (2015): "Unobservable Selection and Coefficient Stability: Theory and Evidence," mimeo, Brown University.

Pop-Eleches, C., and M. Urquiola (2013): "Going to a Better School: Effects and Behavioral Responses," *American Economic Review*, 103(4), 1289–1324.

Raven, J. (1989): "The Raven Progressive Matrices: A Review of National Norming Studies and Ethnic and Socioeconomic Variation Within the United States," *Journal of Educational Measurement*, 26(1), 1–16.

Saavedra, J. E. (2008): "The Returns to College Quality: A Regression Discontinuity Analysis," mimeo, Harvard University.

Schoar, A. (2009): "The Divide between Subsistence and Transformational Entrepreneurship," in *NBER Innovation Policy and the Economy*, ed. by J. Lerner, and S. Stern, chap. 3, pp. 57–81. University of Chicago Press, Chicago.

Schultz, T. P. (1988): "Education Investments and Returns," in *Handbook of Development Economics*, ed. by H. Chenery, and T. Srinivasan, vol. 1, Chapter 13, pp. 543–630. Elsevier Science Publishers B.V., Amsterdam.

Stock, J. H., and M. Yogo (2002): "Testing for Weak Instruments in Linear IV Regression," Technical Working Paper 284, National Bureau of Economic Research.

Strauss, J., and D. Thomas (1995): "Human Resources: Empirical Modeling of Household and Family Decisions," in *Handbook of Development Economics*, ed. by J. Behrman, and T. N. Srinivasan, vol. 3A, Chapter 34, pp. 1883–2023. Elsevier Science Publishers B.V., Amsterdam.

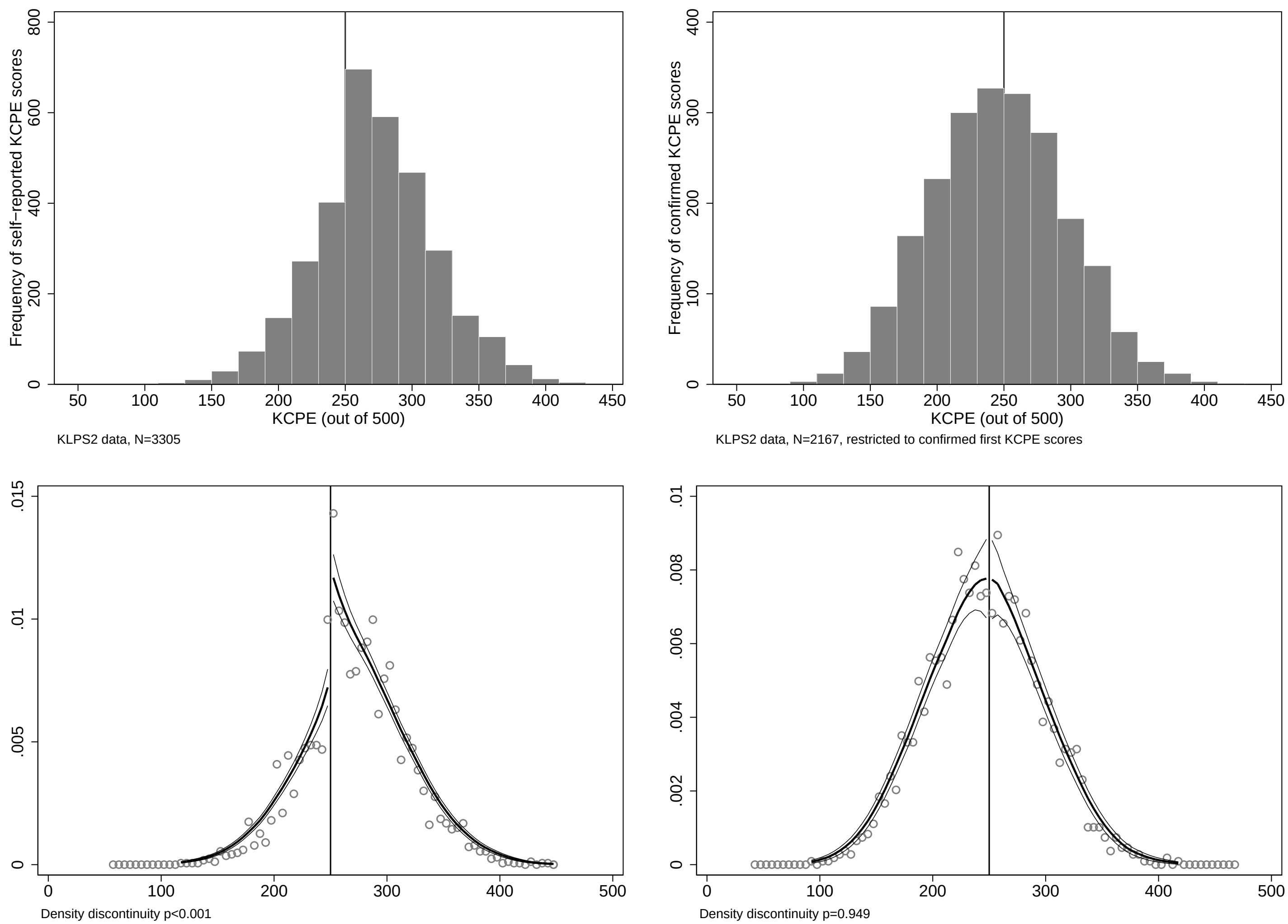
Urquiola, M., and E. Verhoogen (2009): "Class-Size Caps, Sorting, and the Regression-Discontinuity Design," *American Economic Review*, 99(1), 179–215.

Uwaifo Oyelere, R. (2010): "Africa's education enigma? The Nigerian story," *Journal of Development Economics*, 91(1), 128–139.

Wooldridge, J. M. (2002): *Econometric analysis of cross section and panel data*. MIT Press, Cambridge, 6th edn.

Zimmerman, D. J. (1992): "Regression Toward Mediocrity in Economic Stature," *American Economic Review*, 82(3), 409–429.

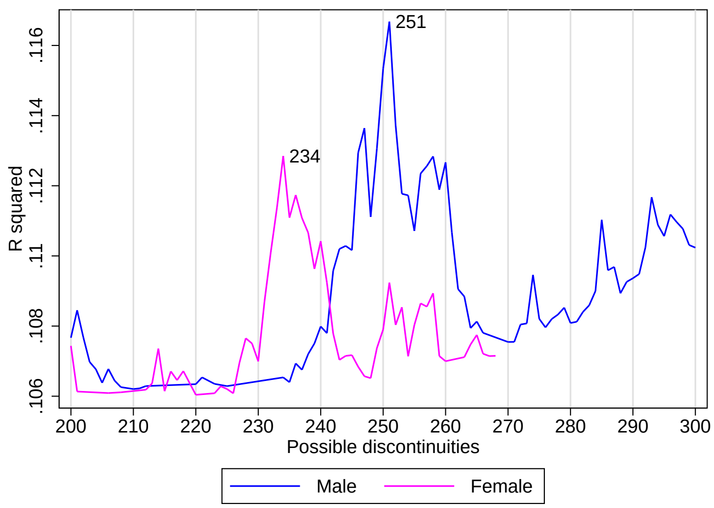
Figure 1: Self-reported and confirmed KCPE scores with density tests.




Left panels: self-reported scores; right panels: confirmed official scores.

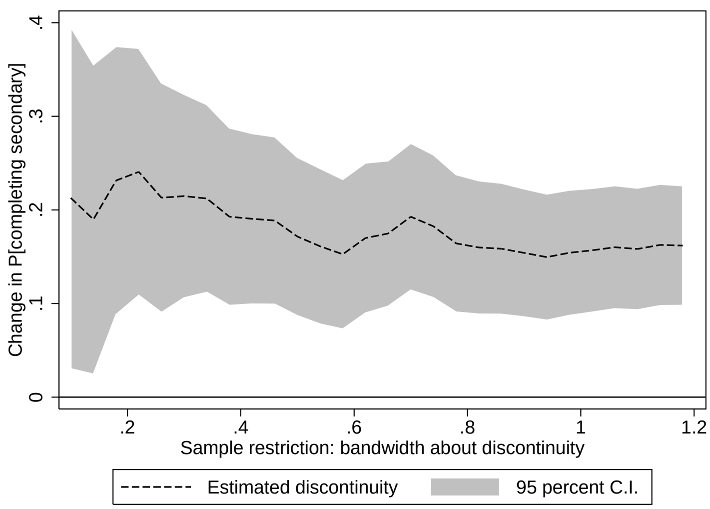
Note: KCPE scores prior to 2001 are converted to the current 500-point scale; density graphs generated by the McCrary (2008) Stata program.

Figure 2: Structural break search




Estimation based on method used in Card, Mas, and Rothstein (2008). X-axis is KCPE score. (Note: KCPE scores prior to 2001 are converted to the current 500-point scale.)

Figure 3: Discontinuity: a function of bandwidth




Estimates and confidence intervals based on piecewise linear specification.; X-axis is bandwidth in terms of 100s of points of KCPE score. (Note: KCPE scores prior to 2001 are converted to the current 500-point scale.)

Figure 4: RD Validity: local quadratic regressions of covariates on KCPE scores.

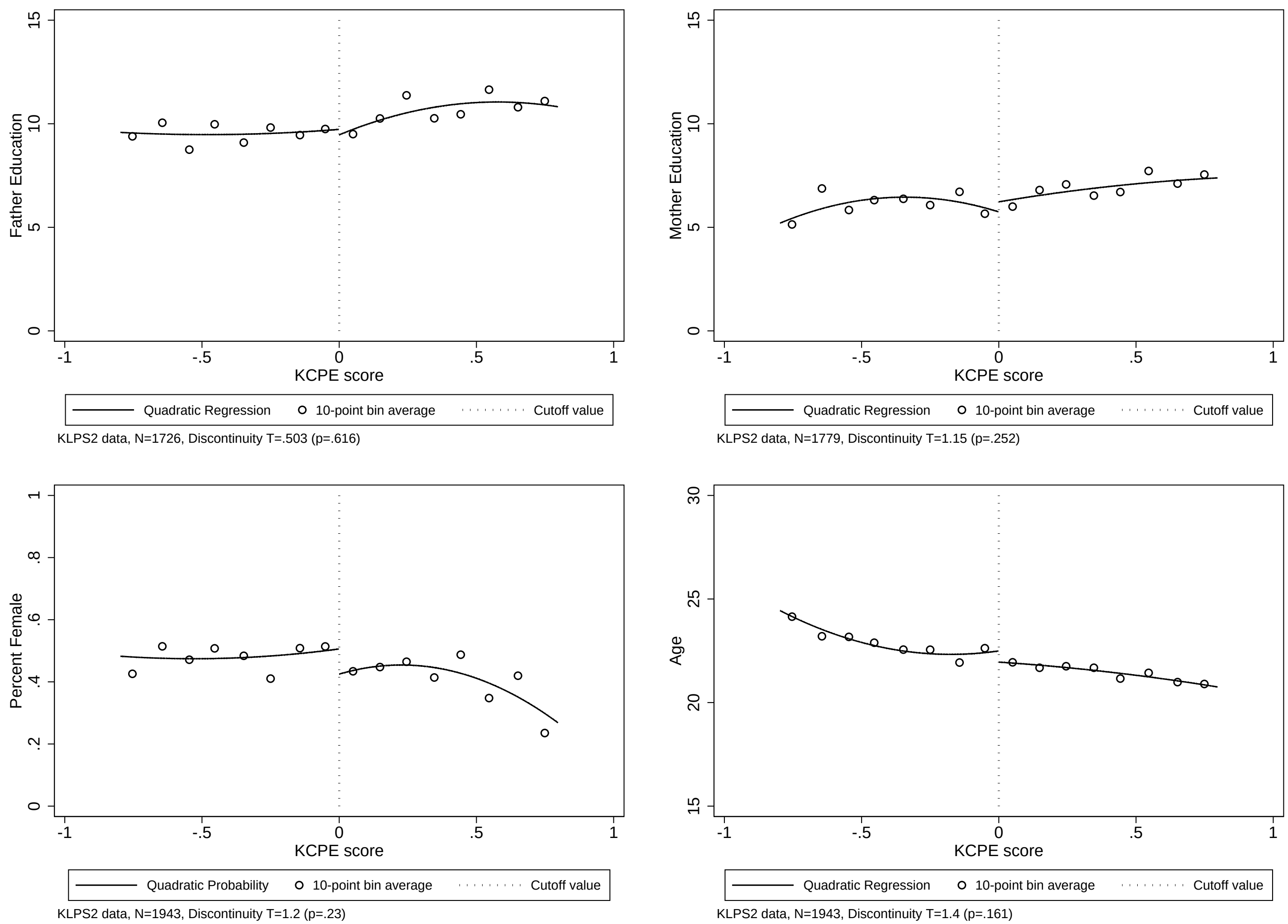
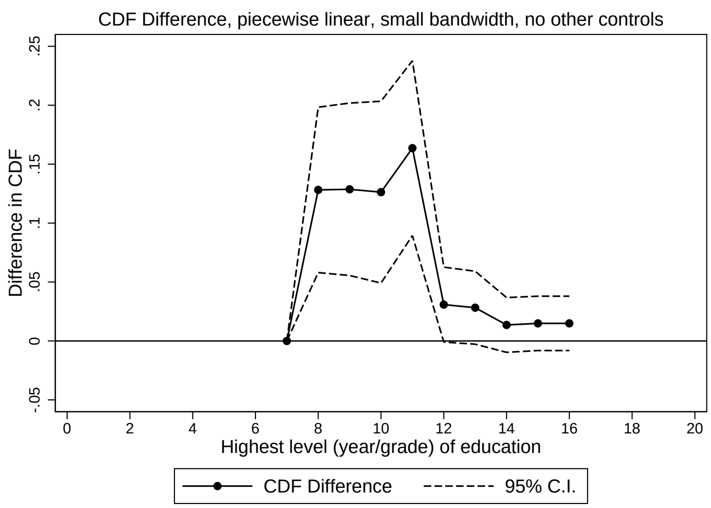


Figure 5: CDF difference




Difference in cumulative distribution functions for years of education at the discontinuity. This is equivalent to the difference in the fraction of individuals whose years of schooling exceed a given value on the x-axis. Thus, the difference in probability of exceeding 7 years is zero, since everyone in the sample, whether they score well or poorly on the KCPE, must exceed 7 years (reach standard 8) in order to take the KCPE in the first place. The difference in probability of exceeding 8-11 years (thus, attaining at least 9-12 years) is clearly statistically distinguishable from zero.

Figure 6: First stage and reduced forms: cognitive performance; self-employment among older men; pregnancy by 18 among women.

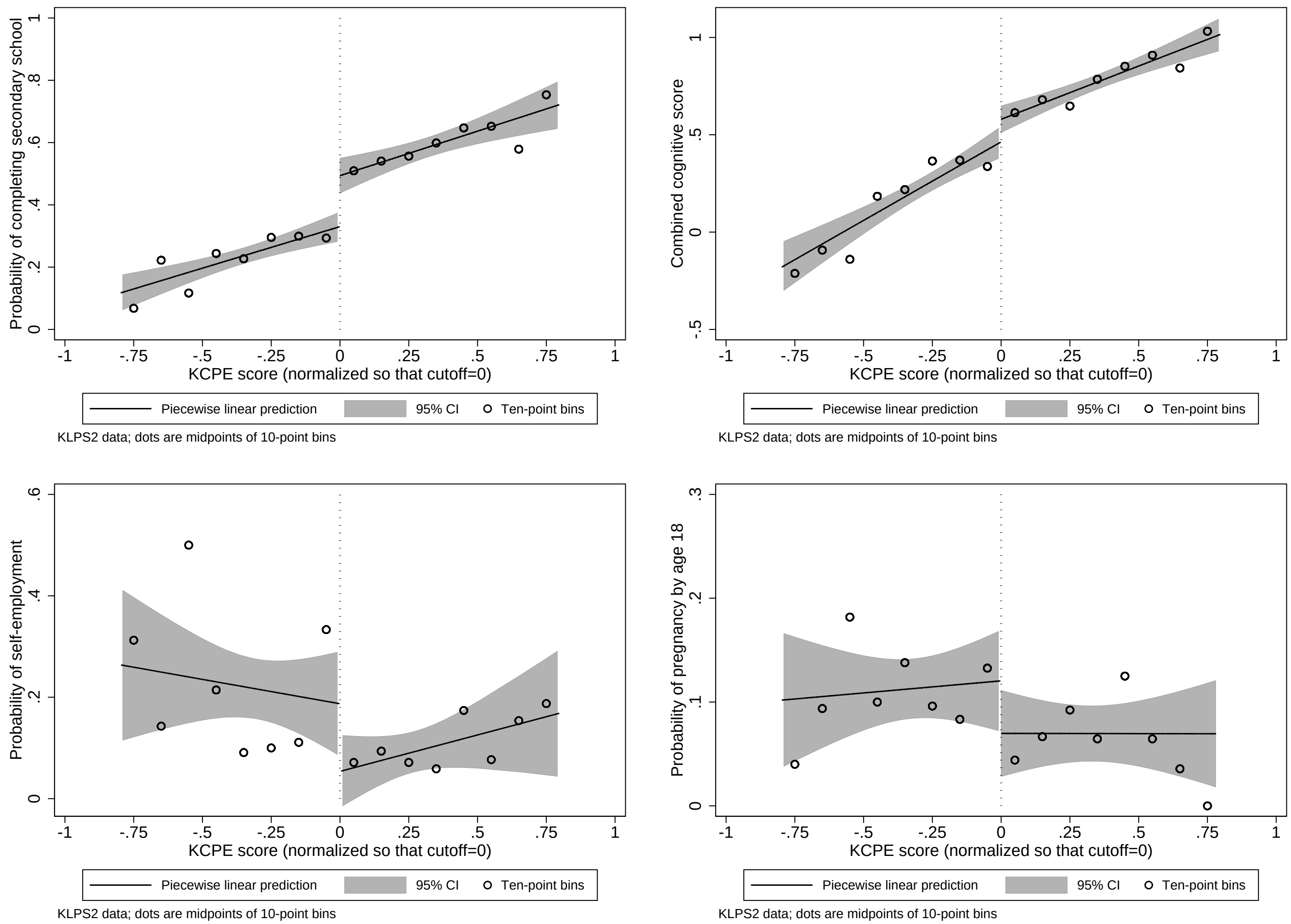


Table 1: KLPS2 Summary Statistics

| Characteristic | Mean | Standard Dev. | N |
|---|---|---|---|
| *Panel A: Respondent characteristics among those with KCPE scores* | | | |
| Age | 22.06 | (2.57) | 3306 |
| Female | 0.45 | (0.50) | 3306 |
| Father's level of education | 10.05 | (4.99) | 2954 |
| Mother's level of education | 6.61 | (4.18) | 3050 |
| *Panel B: Education outcomes among those with KCPE scores* | | | |
| Self-reported KCPE Score (out of 500) | 254.50 | (52.23) | 3306 |
| Years of Education | 10.14 | (2.09) | 3306 |
| Still attending school | 0.30 | (0.46) | 3306 |
| Any secondary schooling | 0.62 | (0.49) | 3306 |
| Complete (4y) secondary schooling | 0.37 | (0.48) | 3306 |
| Post-secondary schooling | 0.04 | (0.18) | 3306 |
| *Panel C: Cognitive outcomes within 80-point bandwidth* | | | |
| Vocabulary test (standardized) | 0.53 | (0.70) | 1942 |
| Raven's matrices (standardized) | 0.34 | (0.91) | 1923 |
| Standardized vocabulary + Raven's | 0.49 | (0.76) | 1923 |
| *Panel D: Labor market outcomes for older men within 80-point bandwidth* | | | |
| Age \| male, oldest cohorts | 24.32 | (1.98) | 378 |
| Attending school \| male, oldest cohorts | 0.13 | (0.34) | 378 |
| Employed \| male, oldest cohorts | 0.34 | (0.47) | 378 |
| Self-employed \| male, oldest cohorts | 0.16 | (0.37) | 378 |
| *Panel E: Fertility within 80-point bandwidth* | | | |
| Age \| female, at least 18 y.o. | 22.06 | (2.28) | 853 |
| Pregnant by 18 \| female, at least 18 y.o. | 0.09 | (0.29) | 853 |

Note that this is a subsample of the KLPS2 data; by conditioning on the presence of a KCPE score, I eliminate all respondents who left school before completing 8th grade (N=3,305 rather than 5,084). Also note that the average grade in 1998 is between 4 and 5 because the KLPS sample has essentially equal numbers of pupils drawn from each grade from 2 through 7. The variable "Still attending school" is measured in 2007, 2008, or 2009, depending on when the survey took place; as one would expect, it declines with age and grade cohorts; likewise, employment rates trend in the opposite direction. KCPE scores prior to 2001 have been converted to the current 500-point scale.

Table 2: Discontinuity (First Stage) Estimation.

| Outcome: Completing secondary school | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample restriction: | Pooled | | | Male | | | Female | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| KCPE≥cutoff | 0.16*** | 0.15*** | 0.17*** | 0.17*** | 0.16*** | 0.21*** | 0.16*** | 0.13** | 0.12 |
| | (0.04) | (0.03) | (0.05) | (0.05) | (0.05) | (0.06) | (0.06) | (0.05) | (0.07) |
| KCPE centered at cutoff | 0.27*** | 0.27*** | 0.07 | 0.3*** | 0.28*** | 0.07 | 0.24*** | 0.25*** | 0.06 |
| | (0.06) | (0.05) | (0.18) | (0.09) | (0.09) | (0.31) | (0.08) | (0.08) | (0.26) |
| (KCPE≥cutoff)×KCPE | 0.02 | 0.006 | 0.2 | -0.02 | -0.01 | -0.03 | -0.006 | 0.05 | 0.5 |
| | (0.09) | (0.08) | (0.3) | (0.11) | (0.11) | (0.41) | (0.14) | (0.13) | (0.48) |
| Constant | 0.33*** | 0.44*** | 0.41*** | 0.39*** | 0.41** | 0.37** | 0.27*** | 0.34* | 0.32 |
| | (0.02) | (0.14) | (0.14) | (0.04) | (0.17) | (0.18) | (0.04) | (0.19) | (0.19) |
| Piecewise Quadratic | No | No | Yes | No | No | Yes | No | No | Yes |
| Controls | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Discontinuity F-stat | 19.46 | 21.55 | 14.87 | 11.13 | 12.42 | 10.94 | 7.50 | 5.81 | 2.71 |
| Observations | 1943 | 1943 | 1943 | 1064 | 1064 | 1064 | 879 | 879 | 879 |
| $R^2$ | 0.14 | 0.23 | 0.23 | 0.14 | 0.24 | 0.24 | 0.12 | 0.2 | 0.2 |

Notes for all regression tables: Standard errors, clustered at the KCPE-score level, are in parentheses. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level. Coefficients on KCPE score and interactions with it have been scaled up by a factor of 100. KCPE score has been re-centered at the discontinuity (251 for men; 234 for women), so that the coefficient on the discontinuity may be interpreted directly. KCPE scores prior to 2001 have been converted to the current 500-point scale. Controls, when indicated, include age, parents' education levels (and an indicator for survey nonresponse), and indicators for all but one of the six KLPS cohorts.

Table 3: Human capital: all cohorts

| **Outcome:** | **Mean effect: Vocabulary and Raven's Matrices** | | | | **Vocabulary** | **Matrices** |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | OLS | 2SLS | 2SLS | 2SLS | 2SLS |
| *Panel A: Full sample* | | | | | | |
| Completing Std 12 | 0.612*** | 0.584*** | 0.67** | 0.596** | 0.645** | 0.399 |
| | (0.032) | (0.033) | (0.282) | (0.3) | (0.275) | (0.432) |
| KCPE centered at cutoff | 0.663*** | 0.607*** | 0.637*** | 0.602*** | 0.607*** | 0.448* |
| | (0.085) | (0.086) | (0.168) | (0.17) | (0.16) | (0.232) |
| (KCPE≥cutoff)×KCPE | -0.311** | -0.302** | -0.311** | -0.302** | -0.468*** | -0.061 |
| | (0.127) | (0.124) | (0.127) | (0.123) | (0.112) | (0.175) |
| Female | -0.19*** | -0.222*** | -0.183*** | -0.22*** | -0.136*** | -0.25*** |
| | (0.029) | (0.03) | (0.042) | (0.051) | (0.047) | (0.073) |
| Constant | 0.361*** | 1.055*** | 0.334** | 1.048*** | 1.580*** | 0.273 |
| | (0.031) | (0.204) | (0.14) | (0.274) | (0.219) | (0.389) |
| Controls | No | Yes | No | Yes | Yes | Yes |
| Discontinuity F-stat | . | . | 20.496 | 23.070 | 23.070 | 23.070 |
| Observations | 1923 | 1923 | 1923 | 1923 | 1923 | 1923 |
| $R^2$ | 0.331 | 0.345 | 0.33 | 0.345 | 0.404 | 0.153 |
| *Panel B: Sample restricted to oldest two cohorts* | | | | | | |
| Completing Std 12 | 0.689*** | 0.648*** | 0.685* | 0.62 | 0.958** | 0.129 |
| | (0.049) | (0.05) | (0.385) | (0.429) | (0.379) | (0.569) |
| Controls | No | Yes | No | Yes | Yes | Yes |
| Discontinuity F-stat | . | . | 10.783 | 9.041 | 9.041 | 9.041 |
| Observations | 693 | 693 | 693 | 693 | 693 | 693 |
| $R^2$ | 0.42 | 0.428 | 0.42 | 0.428 | 0.452 | 0.184 |

(See Notes for all regression tables below Table 2.) In Panel B, though only the coefficient on secondary schooling is shown, the specifications are the same as in Panel A, except that the sample is restricted to the oldest two cohorts.

Table 4: Employment outcomes for men (oldest two cohorts) and fertility outcomes for women

| **Outcome** | **Estimation** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) IVP | (4) IVP | (5) BVP | (6) BVP | (7) 2SLS | (8) 2SLS |
| *Panel A: Employment outcomes, men* | | | | | | | | |
| P[Formally employed] | -0.036 | 0.036 | 0.263 | 0.427** | 0.240 | 0.359** | 0.291 | 0.549 |
| | (0.055) | (0.058) | (0.253) | (0.216) | (0.192) | (0.171) | (0.352) | (0.486) |
| P[Self-employed] | -0.104*** | -0.12** | -0.459*** | -0.516*** | -0.464*** | -0.347** | -0.502* | -0.601* |
| | (0.040) | (0.049) | (0.092) | (0.103) | (0.147) | (0.136) | (0.273) | (0.359) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Discontinuity F-statistic | . | . | 9.031 | 5.986 | 9.031 | 5.986 | 9.031 | 5.986 |
| Observations | 378 | 378 | 378 | 378 | 378 | 378 | 378 | 378 |
| *Panel B: Fertility, women* | | | | | | | | |
| P[Pregnant by 18] | -0.119*** | -0.138*** | -0.454 | -0.583*** | -0.199** | -0.184 | -0.333 | -0.389 |
| | (0.020) | (0.022) | (0.300) | (0.191) | (0.086) | (0.123) | (0.238) | (0.286) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Discontinuity F-statistic | . | . | 6.993 | 5.589 | 6.993 | 5.589 | 6.993 | 5.589 |
| Observations | 853 | 853 | 853 | 853 | 853 | 853 | 853 | 853 |

(See Notes for all regression tables below Table 2.) Only the coefficient on completed secondary schooling is shown; each coefficient comes from a separate regression. Abbreviations: BVP and IVP denote bivariate probit and IV probit, respectively; marginal effects are shown for both. Standard errors for bivariate probit estimates are obtained via bootstrapping with 1,000 draws.

Table 5: Robustness: Imbens-Kalyanaraman Bandwidth.

| | **Outcome:** Completing secondary | **Outcome:** Cognitive score | **Outcome:** Employment | | **Outcome:** Self-employment | | **Outcome:** Pregnancy by 18 | |
|---|---|---|---|---|---|---|---|---|
| | OLS | 2SLS | 2SLS | IVP | 2SLS | IVP | 2SLS | IVP |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | 0.216*** | 1.462*** | 0.238 | 0.284 | −0.813** | −0.474*** | −0.752* | −0.640*** |
| | (0.059) | (0.515) | (0.366) | (0.259) | (0.374) | (0.045) | (0.418) | (0.030) |
| Observations | 1184 | 793 | 260 | 260 | 180 | 180 | 521 | 521 |
| IK Bandwidth: | 37 | 23 | 43 | 43 | 29 | 29 | 36 | 36 |
| Restrictions: | | | Male; Oldest two cohorts | | Male; Oldest two cohorts | | Female; 18+ years old | |

Here, the Stata `rd` command (Nichols 2016) is used to esimate the first stage, followed by the fuzzy RD, for each of the main outcomes in the paper. The default kernel (triangular) and bandwidth (Imbens-Kalyanaraman, IK) are shown, with results from the `rd` command, in columns (1), (2), (3), (5), and (7). The Imbens-Kalyanaraman bandwidth is then used for IV probit estimation, rather than 2SLS, in columns (4), (6), and (8), for comparability with the rest of the analysis in this paper; this estimation is carried out using the Stata `ivprobit` command but restricting to the bandwidth given by the `rd` command; in these columns, marginal effects are presented with standard errors obtained via the delta method (provided by the `margins` command). Bandwidth is given in terms of KCPE points on either side of the discontinuity. Significance is assessed from p-values based on z-statistics.

Table 6: Controls and bias: human capital, all cohorts

| | **Outcome: combined vocab/Raven's score** | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | OLS | OLS | OLS | OLS | 2SLS |
| Completing Std 12 | 1.226*** | 0.821*** | 0.754*** | 0.612*** | 0.67** |
| | (0.022) | (0.033) | (0.034) | (0.032) | (0.282) |
| Constant | -0.196*** | 0.213*** | 0.26*** | 0.309*** | 0.334** |
| | (0.02) | (0.03) | (0.03) | (0.026) | (0.14) |
| Restrict to KCPE takers | No | Yes | Yes | Yes | Yes |
| Restrict KCPE bandwidth | No | No | Yes | Yes | Yes |
| Control for KCPE | No | No | No | Yes | Yes |
| Observations | 4885 | 2149 | 1923 | 1923 | 1923 |
| $R^2$ | 0.326 | 0.292 | 0.271 | 0.329 | 0.33 |

(See NOTES FOR ALL REGRESSION TABLES below Table 2.) In Tables 6 through 8, Column (1) uses only robust standard errors, however, because KCPE data are not included in the first specification. All columns in this table include a control for respondent gender.

Table 7: Controls and bias: employment for older two cohorts of men

| | Outcome: formally employed | | | | | |
|---|---|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) OLS | (4) OLS | (5) OLS | (6) 2SLS |
| Completing Std 12 | -0.014 | -0.017 | -0.039 | -0.035 | 0.015 | 0.291 |
| | (0.036) | (0.047) | (0.051) | (0.055) | (0.058) | (0.352) |
| Constant | 0.327*** | 0.344*** | 0.362*** | 0.36*** | -0.857*** | 0.223 |
| | (0.026) | (0.033) | (0.036) | (0.038) | (0.316) | (0.197) |
| Restrict to KCPE takers | No | Yes | Yes | Yes | Yes | Yes |
| Restrict KCPE bandwidth | No | No | Yes | Yes | Yes | Yes |
| Control for KCPE | No | No | No | Yes | Yes | Yes |
| Control for age | No | No | No | No | Yes | No |
| Observations | 664 | 429 | 378 | 378 | 378 | 378 |
| $R^2$ | 0.0002 | 0.0003 | 0.002 | 0.002 | 0.036 | . |

See Notes for Table 6.

Table 8: Controls and bias: self-employment, older two cohorts of men

| | Outcome: self-employed | | | | | |
|---|---|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) OLS | (4) OLS | (5) OLS | (6) 2SLS |
| Completing Std 12 | -0.145*** | -0.125*** | -0.127*** | -0.105*** | -0.101** | -0.502* |
| | (0.026) | (0.037) | (0.04) | (0.04) | (0.044) | (0.273) |
| Constant | 0.246*** | 0.225*** | 0.23*** | 0.218*** | 0.111 | 0.403** |
| | (0.02) | (0.03) | (0.033) | (0.032) | (0.276) | (0.17) |
| Restrict to KCPE takers | No | Yes | Yes | Yes | Yes | Yes |
| Restrict KCPE bandwidth | No | No | Yes | Yes | Yes | Yes |
| Control for KCPE | No | No | No | Yes | Yes | Yes |
| Control for age | No | No | No | No | Yes | No |
| Observations | 802 | 429 | 378 | 378 | 378 | 378 |
| $R^2$ | 0.034 | 0.029 | 0.03 | 0.034 | 0.034 | . |

See Notes for Table 6.

# A Online Appendix - not for print publication

## A.1 Data

### A.1.1 KCPE Logistics

When students take the exam,[52] they indicate a list of secondary schools they would prefer to attend, including one or two from each of three tiers of secondary schools: national, provincial, and district.[53] National schools are the most competitive and are considered to be of the highest quality; district schools, the least so.

Initial admission cutoffs are determined centrally by the Ministry of Education. The cufoffs may differ for boys and girls, and vary according to characteristics of each secondary school.[54] After the cutoffs are decided and test results are available, heads of secondary schools meet in groups to determine matches between schools and students, and make admission offers. Following this initial selection round, a variety of additional selection mechanisms are employed: students may accept offers; students may contact other schools which they would prefer to attend, to see whether the admissions committee is willing to accept them;[55] and school leaders may meet again to carry out a second round of selection if an insufficient number of students accept offers in the first round. Nevertheless, the odds of secondary school admission jump up sharply at a KCPE score of 250, and continue to rise thereafter, as does the quality of the school to which a student is admitted.

### A.1.2 KCPE Scoring

From 1985 to 2000, the KCPE covered seven subjects, each scored on a 100-point scale: English, Swahili, math, science, geography, arts, and business;

[52]The Kenyan academic year coincides with the calendar year; the KCPE takes place in early November, and results are announced in the last few days of December.

[53]This mirrors the "rule-of-thumb" tiered system of secondary school choice that Ghana adopted in 2008, for example (Ajayi 2013).

[54]Some school cutoffs are below 250, while others are above. The cutoff governing the largest fraction of schools in the region, however, is 250.

[55]Often the school-imposed KCPE cutoff is higher for these cases, but exceptions are made at the discretion of the admissions committee for especially meritorious or needy students.

from 2001 onward, the last two of these—arts and business—were removed (Orlale 2000, Kremer, Miguel, and Thornton 2009). As a consequence, the KLPS data include observations in which the maximum score is 700, and observations where the maximum is 500. Throughout this paper, I normalize all scores to the 500-point scale.

Those who are not admitted to any government school have several options if they wish to continue their education: they may repeat eighth grade and re-take the KCPE; they may still have access to private secondary schools and vocational schools; or they may travel to Uganda to enroll in school there.

### A.1.3 Re-taking the KCPE

One clear pattern both from the survey data and the administrative records is that students sometimes re-take the test. In the 2003-2005 round of surveying (KLPS1), the questionnaire asked not only for respondents' KCPE score, but also how many times they had taken the KCPE. Of KCPE-takers in the older cohorts (who had reached eighth grade before being interviewed in KLPS2), approximately 87 percent said they took it exactly once, 13 percent said that they had taken the exam twice, and around one tenth of one percent said they took it three times. The reason such a small fraction re-take such an important examination, according to my interviews with with both teachers and pupils, is that it is costly: they have to repeat eighth grade in order to do it. The survey data are in agreement: more than 98 percent of respondents who report re-taking the KCPE also report repeating standard 8; conversely, of those who take the KCPE only once, comparatively few respondents (less than 3 percent) repeat standard 8 for any reason. While a pupil's decision to re-take the test is conditioned on the the pupil's unobserved ability as well as the relationship of her first score to the discontinuity—thereby skewing the second score distribution and any conclusions drawn from it, as in the case studied by Martorell (2004)—a pupil's first test score should not show any sign of manipulation around the discontinuity.

### A.1.4 KLPS Data

All 73 schools are rural, and together represent 80% of the schools in those two administrative Divisions. From this population of roughly 22,000 students, a representative 7,530 pupils were randomly sampled for two follow-ups. The survey acts both as a follow-up to the Primary School Deworming

Project (PSDP) (Miguel and Kremer 2004), and as a longitudinal study representative of an entire region.

Of 7,530 sampled pupils, 5,084 were surveyed during KLPS2. Though this is only 67.5 percent of the sample, some of the original sample has been confirmed deceased, and because some pupils were easier to locate than others after ten years, each survey wave was carried out in two phases: regular and intensive. Only a random sample of respondents who were not found during the regular phase were sought during the intensive phase: 63 percent of respondents were located in the regular phase, and of the remaining 37 percent, more than half of an intensive sample was located, bringing the effective tracking rate of KLPS2 to above 80 percent.[56]

### A.1.5 Matching and Correcting KCPE Scores

If pupils taking KCPE could somehow manipulate their test score to place themselves just above the secondary school cutoff, it would invalidate the research design. Administrative data for a recent year across the entire province, however, depicted in Figure A1, does not show this characteristic. Examination papers from any particular school are graded by separate teachers for each subject, and are never graded by teachers from the school where the papers originated, so precise manipulation around the discontinuity would not be straightforward in any event. To resolve the discrepancy between the distributions of self-reported and administratively reported KCPE scores, I gathered an auxilliary dataset of 17,384 official KCPE scores from the Government of Kenya, via district education offices and school visits. These official data do not include all schools in all years for several reasons: recent political upheavals made some records inaccessible, re-districting changed which offices were responsible for maintaining the records in question; and record-keeping over the past eleven years at local primary schools has occasionally been frustated by natural disasters. Nevertheless, the data I gathered include roughly 88 percent of the KLPS schools[57] during the years

[56]Discussion of tracking logistics, intensive sampling, and sample attrition in the earlier 2003-2005 round of this survey, KLPS1, may be found in Baird, Hamory, and Miguel (2008).

[57]In the process of visiting many of the schools myself in order to collect this data, in addition to all the schools in the original Primary School Deworming Project study, I was also able to visit a number of schools in neighboring districts where some KLPS respondents had transferred by the time of their KCPE examinations. These records are included in the 17,384 total.

of interest in this study. I match the KCPE records to the KLPS surveys by pupil name,[58] and by the year(s) and school(s) in which the pupil took the KCPE. After condensing spelling variations of the same name, I am able to match KCPE records to KLPS2 surveys whenever there is no better match in the year and school in which the respondent took KCPE—and at least two names agree across the two datasets.

Comparing the administrative test scores with the survey data, I am able to ask what predicts misreporting; the lower the true score, the higher the chances of misreporting.[59]

Based on confirmed first test scores, I am also able to chart the probability of re-taking the test, shown in Appendix Figure A3. As expected, the probability of re-taking the test is highest for respondents whose first score is below 250 points. The average test score improvement from the first attempt to the second is 54 points, just above one standard deviation on the test.

[58] Names in Kenya are not as fixed as in the United States: they may be spelled differently even within the same document ("Winnstone" for "Winston," for example); order of names is also typically not fixed (so that "Juma Winston" is likely to be the same person as "Winston Juma"); and the subset of names reported ("Juma Winston Wandera") varies from record to record. I should also note that the distribution of names is skewed more towards the most common names in western Kenya than in it is in the United States. "Smith" was the most common surname in the 1990 US census, with just over one percent of the population; no other surname exceeded one percent. In this region, there are five names that occur with frequencies above three percent each. Despite this concentration, unique identification of pupils is made feasible by the typically small exam cohorts from each school.

[59] In a linear regression framework, respondents with low ability as measured by cognitive tests at survey time, those who round their test scores to a multiple of five, and those who took the test further in the past are all more likely to misreport the score. None of these, however, is a very reliable predictor: they do not yield large differences in the probability of misreporting. The only sharp predictor is whether the respondent took the KCPE before the KLPS1 survey was administered, and reports the same score in KLPS1 and KLPS2; in that case, there is an 86 percent chance that the respondent is reporting the truth, though this is without conditioning on any other predictors. Nearly all of these cases, of course, are respondents who scored above 250 on the test. I am able to include these data in robustness checks that expand the sample slightly, from 2,167 to 2,236 first test scores, and my results do not change appreciably.

## A.2 Measurement Error and Re-taking

While most KLPS variables are quite stable over survey rounds, self-reported KCPE scores are not. Grade in school in 1999, for example, has a correlation of 0.95 between responses given in KLPS1 and four years later in KLPS2, while self-reported test score has a correlation closer to 0.7. The noise in test scores could pose several problems, since I use KCPE score as the regression discontinuity running variable.[60]

Noise in the form of classical measurement error for only a random subset of the data would simply reduce the power of the regression discontinuity design. Classical measurement error in all of the data could eliminate the discontinuity entirely. Non-classical error could invalidate the regression discontinuity design, if either mis-reporting or test repetition were driven by unobservables correlated with outcomes.[61] A histogram of the self-reported scores, in the upper left panel of Figure 1, shows that the distribution of scores shows signs of non-classical error, in the form of manipulation of the reported scores around the "passing" point; a test for density smoothness proposed by McCrary (2008), shown in the lower left panel, rejects at this point.

This feature of the distribution could arise simply from repeated test-taking: if many of those who fail the test try again until they pass, the distribution of most recent test scores could include more mass just to the right of the cutoff than to the left (see discussion below for a concrete example.) To see whether this phenomenon is solely responsible for the shape of the distribution, I consider a slice of the data, available in KLPS1, in which respondents provided every test score for as many times as they had taken the KCPE. Even if the most recent test score is endogenous with respect to the respondent's type and the location of the cutoff, the first test score should not be. Appendix Figure A2 shows that although the problem is less severe in this restricted sample, even these first scores do not have a smooth density at the discontinuity. However, administrative data on scores in the region display no such irregularity at the cutoff score, so I conclude that the self-reports are, in many cases, incorrect, and administrative data must be matched to the KLPS2 dataset in order to use a regression discontinuity design I show the distribution of regional 2008 test scores in Appendix A.5.

[60] Some authors, such as Imbens and Lemieux (2008), refer to this as a "forcing variable."
[61] See Martorell (2004) for discussion of multiple potential effects of test repetition.

## A.3 Example

Following the notation and discussion of Hahn, Todd, and Van der Klaauw (2001), consider outcome $y_i$, and a binary indicator for secondary schooling, $x_i$. Let $y_i = \alpha_i + x_i \cdot \beta$. Label KCPE score (centered at the admission cutoff) $z_i$, so that $\Pr[x_i = 1|z_i = z]$ is discontinuous at $z = 0$. Define:

$$\begin{aligned}
x^+ &= \lim_{z \to 0^+} E[x_i|z_i = z] \\
x^- &= \lim_{z \to 0^-} E[x_i|z_i = z] \\
y^+ &= \lim_{z \to 0^+} E[y_i|z_i = z] \\
y^+ &= \lim_{z \to 0^-} E[y_i|z_i = z]
\end{aligned}$$

Assuming that these limits exist, and that $E[\alpha_i|z_i = z]$ is continuous in $z$ at 0:

$$\beta = \frac{y^+ - y^-}{x^+ - x^-}$$

This is true no matter what the correlation between $\alpha_i$ and $x_i$ is, as long as the continuity assumption holds. If $E[x_i|z_i = z] = \Pr[x_i = 1|z_i = z]$ has a discontinuity of $x^+ - x^- = \phi$ at $z = 0$, then $y^+ - y^- = \phi \cdot \beta$, and the result follows; this is the Hahn, Todd, and Van der Klaauw (2001) argument for identification in the regression discontinuity design.

### A.3.1 Continuous Classical Measurement Error in Running Variable

Let $\tilde{z}_i = z_i + \eta_i$, where $\eta_i$ is a continuous random variable independent of $x_i$, $z_i$, and $\alpha_i$ with probability density function $f_\eta(\cdot)$. Suppose we observe $y_i$, $x_i$, and $\tilde{z}_i$, but not $z_i$. Define:

$$\begin{aligned}
\tilde{x}^+ &= \lim_{z \to 0^+} E[x_i|\tilde{z}_i = z] \\
\tilde{x}^- &= \lim_{z \to 0^-} E[x_i|\tilde{z}_i = z]
\end{aligned}$$

By iterated expectations:

$$E[x_i|\tilde{z}_i = z] = \int_{-\infty}^{\infty} E[x_i|z_i = t] f_\eta(z - t) dt$$

If $f_\eta(\cdot)$ is differentiable everywhere, application of Leibniz' rule to the expression for $E[x_i|\tilde{z}_i = z]$ above shows that $E[x_i|\tilde{z}_i = z]$ is differentiable everywhere, even if $E[x_i|z_i = z] = \Pr[x_i = 1|z_i = z]$ is not. Thus, in this setting, $\tilde{x}^+ - \tilde{x}^- = 0$: there is no discontinuity when the running variable is measured with this type of error.

### A.3.2 Alternative Forms of Measurement Error

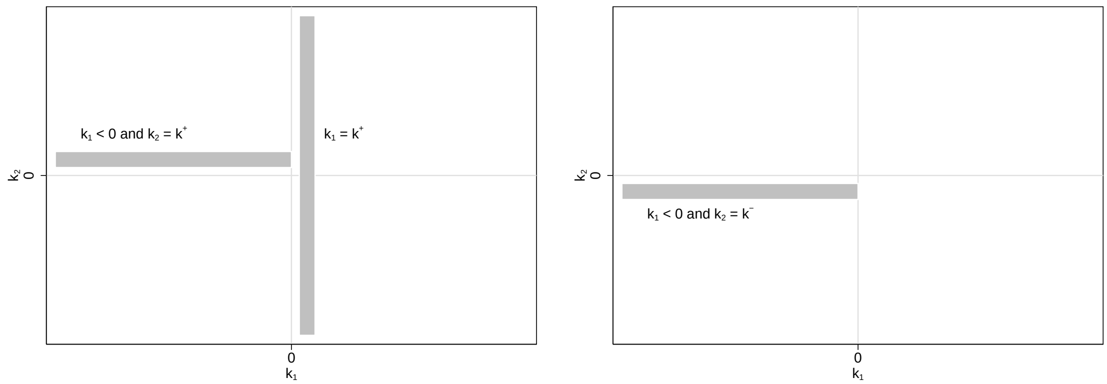
Suppose that instead of $\tilde{z}_i = z_i + \eta_i$, we observe $\tilde{z}'_i = z_i + \eta_i \cdot \zeta_i$, where $\zeta_i$ is binary. Let $\Pr(\zeta_i = 0) = p$, with $\zeta_i$ independent of $\eta_i$, $x_i$, $z_i$, and $\alpha_i$. Using analogously defined limit expressions, $\tilde{x}'^+ - \tilde{x}'^- = p \cdot \phi$, and $\tilde{y}'^+ - \tilde{y}'^- = p \cdot \phi \cdot \beta$, so the regression discontinuity can still be used to consistently estimate $\beta$, though the discontinuity is made smaller. In neither of these cases ($\tilde{z}_i$ or $\tilde{z}'_i$), however, should the density of the observed running variable be discontinuous at $z = 0$ if the underlying density of $z_i$ is smooth at $z = 0$.

### A.3.3 Re-taking

In the presence of test re-taking, a regression discontinuity estimate might or might not yield the desired local average treatment effect, as discussed by Martorell (2004). However, it could easily yield an "artificial" discontinuity in the *density* of reported test scores, as follows: For this discussion, let $k_1$ and $k_2$ be the first and second test scores a student receives on the KCPE. Let $k_1, k_2 \sim iid\, \mathcal{N}(0, \sigma^2)$ with mean zero (at the cutoff). The student only learns the second test score if he does not pass the first time. The student might then report only the most recent score; the first score if he passes the first time, the second if he takes the test a second time.

$$k_{recent} = k_1 \cdot \mathbb{1}[k_1 \geq 0] + k_2 \cdot \mathbb{1}[k_1 < 0]$$

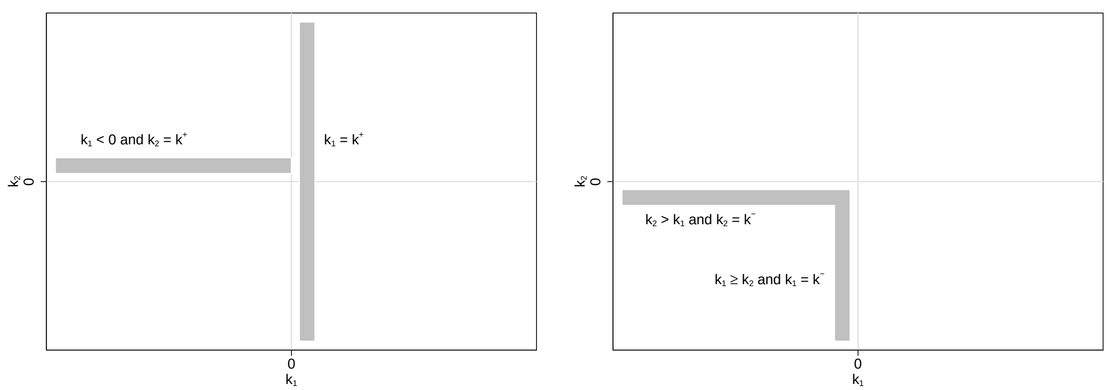
Though the distributions of $k_1$, $k_2$, and even $\max(k_1, k_2)$ are smooth, the density of $k_{recent}$ is discontinuous at the cutoff. This is because for $k^+ \geq 0$, of $k_{recent}$ can take the value $k^+$ either when $k_1 < 0$ and $k_2 = k^+$, or when $k_1 = k^+$. For $k^- < 0$, only the former condition applies. Graphically:




Alternatively, the student might follow the same re-taking rule, then report the best score:

$$k_{best} = k_1 \cdot \mathbb{1}[k_1 \geq 0] + \max(k_1, k_2) \cdot \mathbb{1}[k_1 < 0]$$

Again, the density of $k_{best}$ is discontinuous at the cutoff, because for $k^+ \geq 0$, of $k_{best}$ can still take the value $k^+$ either when $k_1 < 0$ and $k_2 = k^+$, or when $k_1 = k^+$. But now, for $k^- < 0$, a modification of the former condition applies: $k_{best} = k^-$ either when $k_2 = k^-$ and $k_2 > k_1$, or when $k_1 = k^-$ and $k_1 > k_2$. Graphically:




In either case, the density of scores just to the left of the cutoff will be lower than to the right, and the McCrary (2008) test should reject smoothness at the cutoff point.

## A.4 Years of Schooling

Rather than being seen as causing an increase in the probability of completing secondary school, the KCPE discontinuity may also be interpreted as simply causing an increase in years of schooling. This version of the first stage is shown in Appendix Table A2. This first stage is evident in all the same specifications as before, and the coefficient magnitudes are roughly four times larger, since the indicator for completing secondary school represented four years of schooling. Appendix Tables A3, A4, and A5 show the results under this first stage, and for the most part, the coefficients are simply four times smaller. This interpretation is misleading, however: while compliers at the discontinuity do gain approximately 0.16 standard deviations on the cognitive tests for each additional year of schooling (Appendix Table A3, columns 3 and 4), this is true because nearly all the compliers at the discontinuity gain exactly 4 years of schooling (Figure 5), and thus just above 0.6 standard deviations on the tests (Table 3, columns 3 and 4). The relevant policy experiment is not to extend secondary school by an additional year, but to change the cutoff so that a larger fraction of the population attends—and completes—secondary school. Nevertheless, results are largely robust to the alternative specification.

## A.5 Additional Figures and Tables

Figure A1: True administrative distribution from 2008

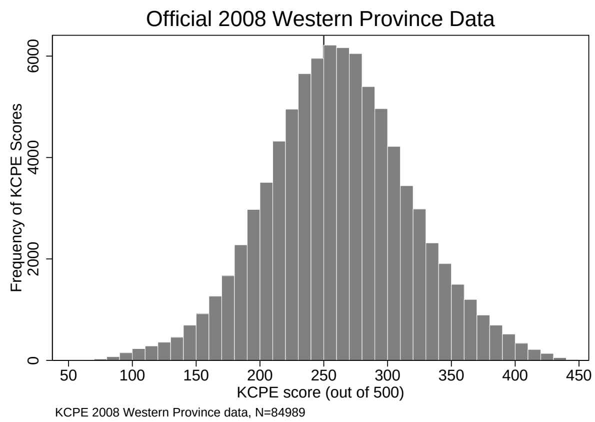
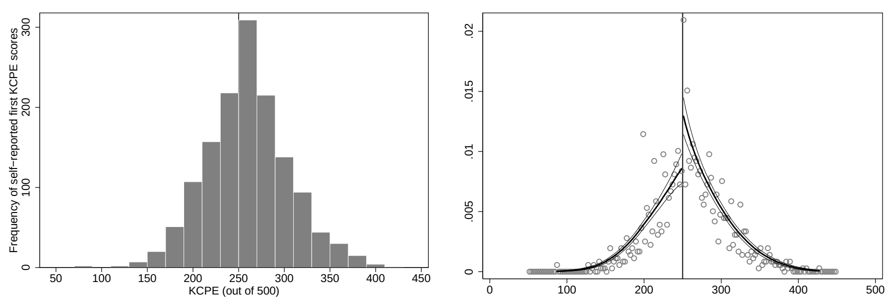
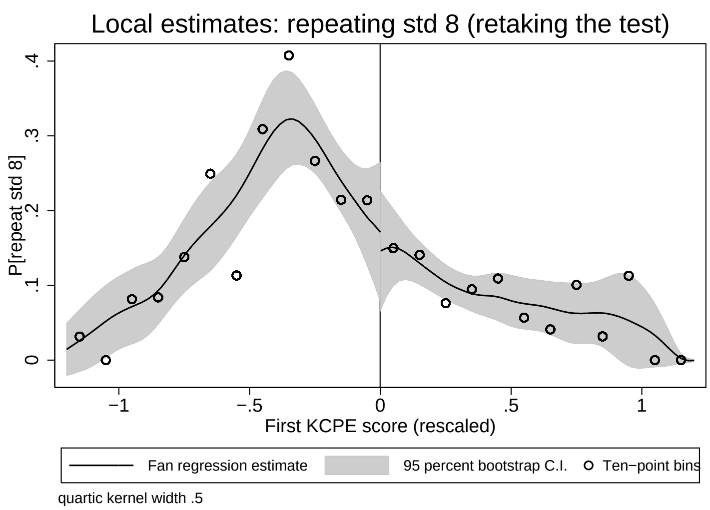



Figure A2: Self-reported KLPS1 first KCPE scores with smoothness test



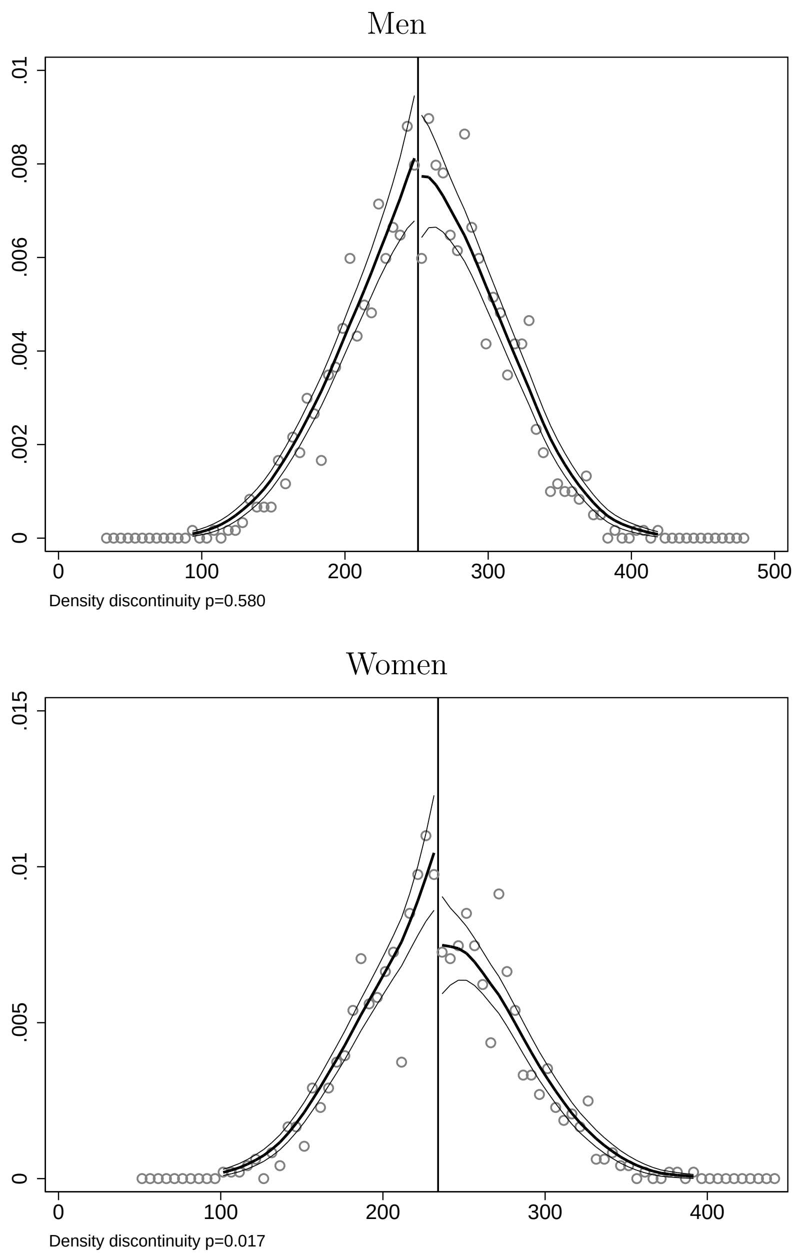

Generated using the Stata program developed by McCrary (2008).

Figure A3: Retaking the test: local linear estimates




Graph generated using the algorithm proposed by Fan (1992).

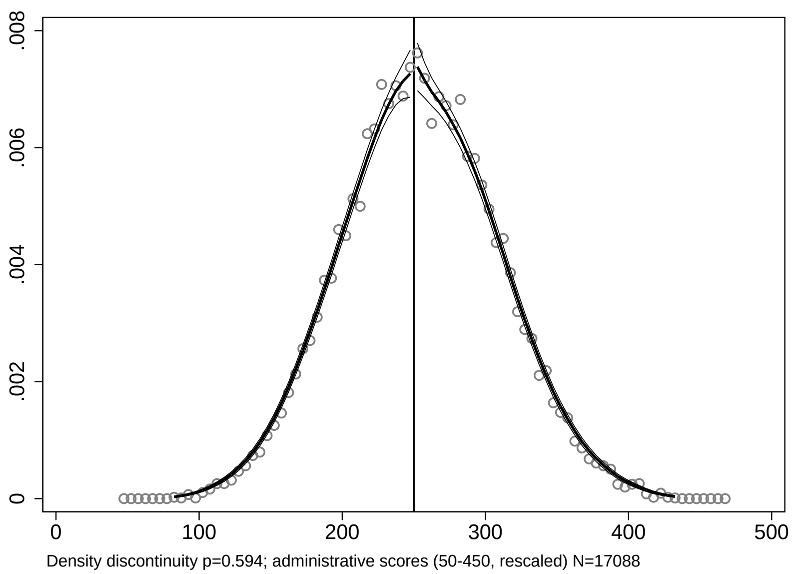
Figure A4: Confirmed KCPE scores with density tests around structural breaks, by gender.




Note: KCPE scores prior to 2001 are converted to the current 500-point scale; density graphs generated by the McCrary (2008) Stata program.

Figure A5: Administrative KCPE scores with density test at 250

All administrative records gathered:




Note: KCPE scores prior to 2001 are converted to the current 500-point scale; density graph generated by the McCrary (2008) Stata program.

Figure A6: First stage and reduced forms: cognitive performance; self-employment among older men; pregnancy by 18 among women.

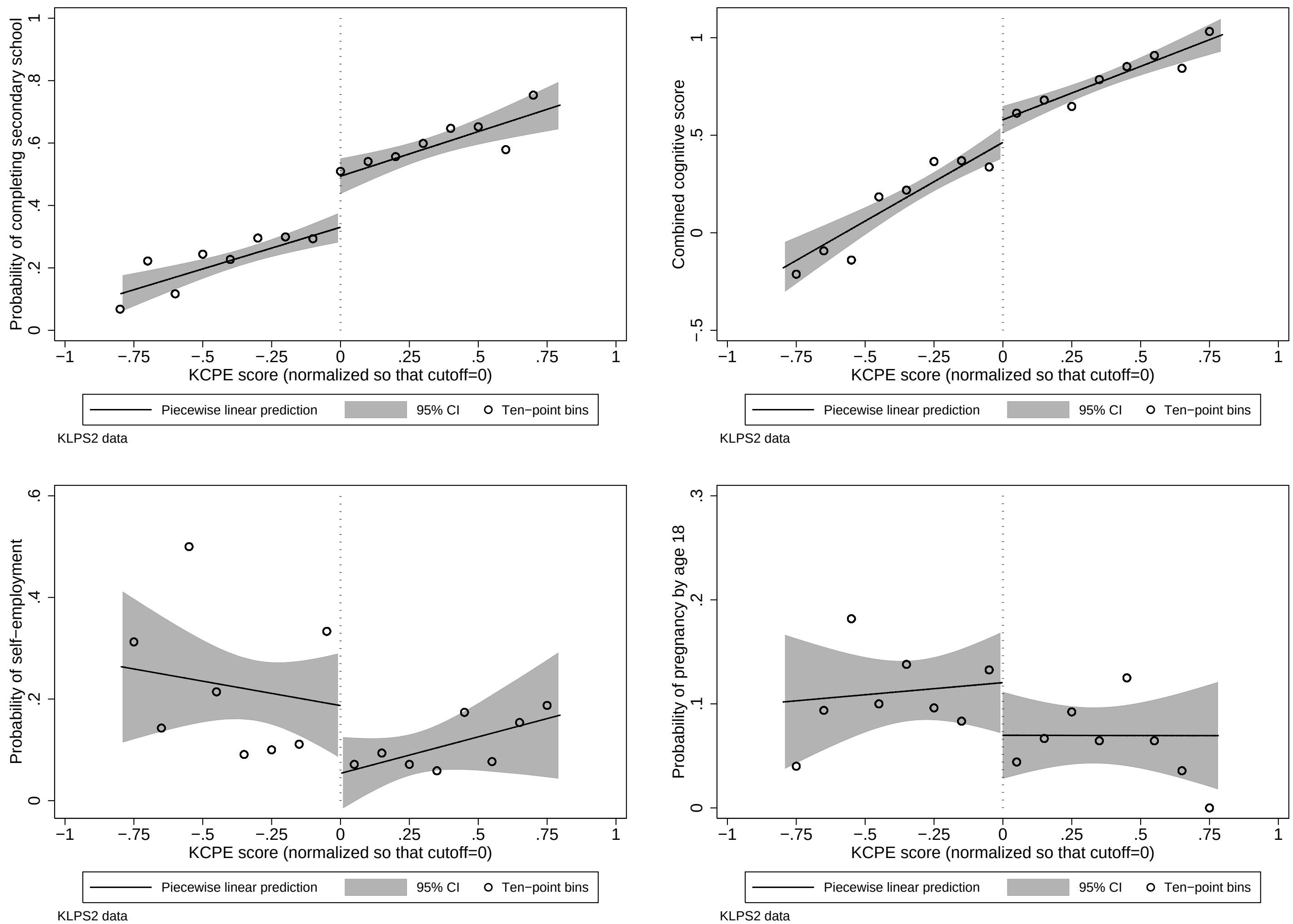
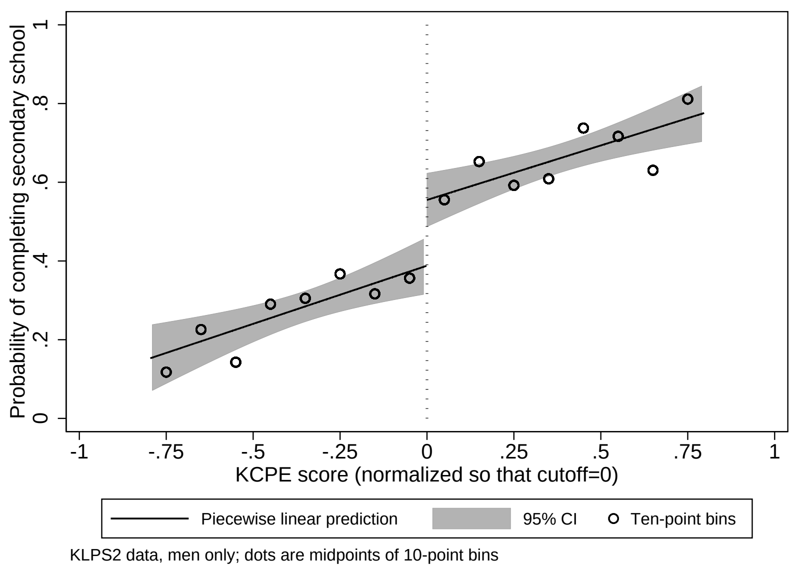


Figure A7: First stage by gender

Men




Women

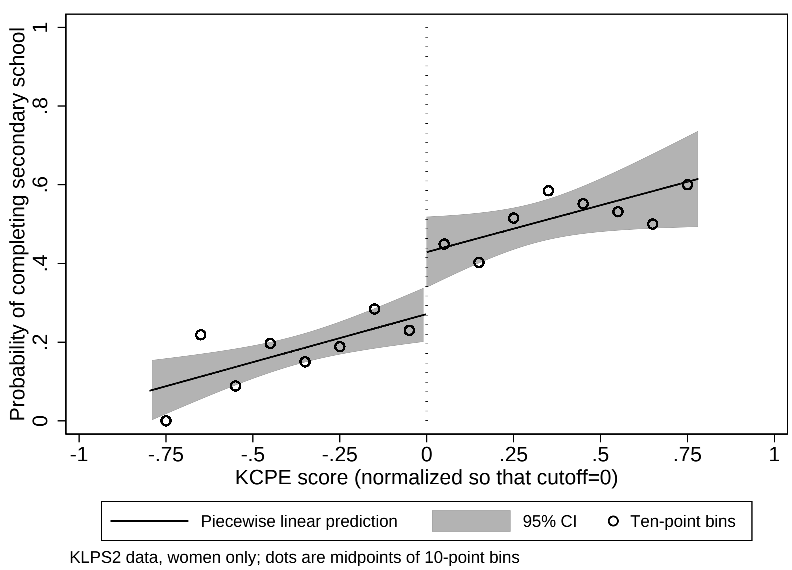


Table A1: Cross-section relationship between cognitive performance and wage

| OUTCOME: LOG(WAGE), CONDITIONAL ON OBSERVATION | (1) | (2) | (3) |
|---|---|---|---|
| Standardized cognitive measure | 0.256*** | 0.199*** | 0.291*** |
| | (0.035) | (0.039) | (0.084) |
| Constant | 7.874*** | 7.940*** | 8.052*** |
| | (0.035) | (0.04) | (0.082) |
| Observations | 772 | 592 | 208 |
| $R^2$ | 0.066 | 0.042 | 0.055 |

In column 1, wages are regressed on the standardized measure of cognitive ability (standardized sum of vocabulary and Raven's Matrices Z-scores). In column 2, the sample is restricted to men; in column 3, it is restricted to men in the oldest two cohorts (Standards 6 and 7 in 1998). Wages are reported in Kenyan Shillings per month.

Table A2: Robustness: alternative discontinuity (first stage) estimation.

| Outcome: Highest grade level of educational attainment | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample restriction: | Pooled | | | Male | | | Female | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| KCPE≥cutoff | 0.65*** | 0.57*** | 0.83*** | 0.68*** | 0.6*** | 0.84*** | 0.64*** | 0.56** | 0.82** |
| | (0.16) | (0.14) | (0.22) | (0.21) | (0.19) | (0.29) | (0.24) | (0.23) | (0.34) |
| KCPE centered at cutoff | 1.80*** | 1.59*** | 0.92 | 1.76*** | 1.53*** | 1.32 | 1.86*** | 1.63*** | 0.4 |
| | (0.25) | (0.24) | (0.78) | (0.33) | (0.33) | (1.14) | (0.37) | (0.37) | (1.22) |
| (KCPE≥cutoff)×KCPE | -0.52 | -0.56 | -1.39 | -0.34 | -0.32 | -1.86 | -1.10** | -0.91 | -0.76 |
| | (0.41) | (0.37) | (1.38) | (0.53) | (0.49) | (1.85) | (0.56) | (0.57) | (2.06) |
| Constant | 10.06*** | 13.04*** | 12.90*** | 10.24*** | 12.70*** | 12.64*** | 9.88*** | 12.76*** | 12.55*** |
| | (0.1) | (0.59) | (0.59) | (0.12) | (0.74) | (0.76) | (0.16) | (0.87) | (0.85) |
| Piecewise Quadratic | No | No | Yes | No | No | Yes | No | No | Yes |
| Controls | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Discontinuity F-stat | 16.63 | 15.73 | 13.98 | 10.22 | 9.88 | 8.46 | 7.25 | 5.68 | 5.73 |
| Observations | 1943 | 1943 | 1943 | 1064 | 1064 | 1064 | 879 | 879 | 879 |
| $R^2$ | 0.18 | 0.28 | 0.28 | 0.19 | 0.29 | 0.29 | 0.16 | 0.25 | 0.25 |

(See Notes for all regression tables below Table 2.)

Table A3: Robustness: human capital, all cohorts

| **Outcome:** | **Mean effect: Vocabulary and Raven's Matrices** | | | | **Vocabulary** | **Matrices** |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | OLS | 2SLS | 2SLS | 2SLS | 2SLS |
| Educational attainment | 0.162*** | 0.152*** | 0.167** | 0.153** | 0.166** | 0.103 |
| | (0.009) | (0.009) | (0.07) | (0.077) | (0.067) | (0.112) |
| KCPE centered at cutoff | 0.53*** | 0.521*** | 0.517** | 0.519*** | 0.518*** | 0.392 |
| | (0.086) | (0.086) | (0.203) | (0.2) | (0.184) | (0.284) |
| (KCPE$\geq$cutoff)$\times$KCPE | -0.21* | -0.21* | -0.207 | -0.21 | -0.368*** | 0.0008 |
| | (0.125) | (0.123) | (0.138) | (0.136) | (0.127) | (0.193) |
| Female | -0.181*** | -0.203*** | -0.178*** | -0.202*** | -0.117** | -0.238*** |
| | (0.029) | (0.03) | (0.044) | (0.059) | (0.051) | (0.086) |
| Constant | -1.070*** | -0.666*** | -1.125 | -0.676 | -0.286 | -0.882 |
| | (0.096) | (0.233) | (0.744) | (1.058) | (0.9) | (1.560) |
| Controls | No | No | Yes | Yes | Yes | Yes |
| Discontinuity F-stat | . | . | 17.290 | 16.965 | 16.965 | 16.965 |
| Observations | 1923 | 1923 | 1923 | 1923 | 1923 | 1923 |
| $R^2$ | 0.352 | 0.358 | 0.352 | 0.358 | 0.443 | 0.15 |

(See Notes for all regression tables below Table 2.) Note that this differs from Table 3 in that the first stage uses highest level of educational attainment rather than an indicator for completing secondary school.

Table A4: Robustness: employment outcomes, older two cohorts of men

| **Outcome** | **Estimation** | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | OLS | IVP | IVP | 2SLS | 2SLS |
| P[Formally employed] | -0.022** | -0.009 | 0.151 | 0.087* | 0.062 | 0.118 |
| | (0.01) | (0.011) | (0.159) | (0.046) | (0.078) | (0.113) |
| P[Self-employed] | -0.019*** | -0.023*** | -0.382*** | -0.12*** | -0.106* | -0.129 |
| | (0.006) | (0.008) | (0.075) | (0.026) | (0.059) | (0.08) |
| Controls | No | Yes | No | Yes | No | Yes |
| Discontinuity F-stat | . | . | 8.118 | 5.346 | 8.118 | 5.346 |
| Observations | 378 | 378 | 378 | 378 | 378 | 378 |

(See Notes for Table 4.) Note that this differs from Panel A of Table 4 in that the first stage uses highest level of educational attainment rather than an indicator for completing secondary school; the bivariate probit is not appropriate for a continuous first stage and is omitted.

Table A5: Robustness: fertility outcome (women)

| **Outcome** | **Estimation** | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | OLS | IVP | IVP | 2SLS | 2SLS |
| P[Pregnant by 18] | -0.038*** | -0.045*** | -0.544*** | -0.134*** | -0.079 | -0.09 |
| | (0.006) | (0.007) | (0.197) | (0.048) | (0.055) | (0.064) |
| Controls | No | Yes | No | Yes | No | Yes |
| Discontinuity F-stat | . | . | 6.993 | 5.624 | 6.993 | 5.624 |
| Observations | 853 | 853 | 853 | 853 | 853 | 853 |

(See Notes for Table 4.) Note that this differs from Panel B of Table 4 in that the first stage uses highest level of educational attainment rather than an indicator for completing secondary school; the bivariate probit is not appropriate for a continuous first stage and is omitted.

Table A6: Robustness: alternative discontinuity location.

| Outcome: Highest grade level of educational attainment | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample restriction: | Pooled | | | Male | | | Female | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| KCPE≥cutoff | 0.14*** | 0.14*** | 0.15*** | 0.14*** | 0.15*** | 0.17*** | 0.13** | 0.13** | 0.14* |
| | (0.04) | (0.04) | (0.05) | (0.05) | (0.05) | (0.07) | (0.06) | (0.05) | (0.08) |
| KCPE centered at cutoff | 0.3*** | 0.28*** | 0.26 | 0.32*** | 0.28*** | 0.11 | 0.28*** | 0.27*** | 0.4 |
| | (0.07) | (0.06) | (0.23) | (0.09) | (0.09) | (0.32) | (0.09) | (0.09) | (0.34) |
| (KCPE≥cutoff)×KCPE | 0.02 | 0.004 | -0.05 | -0.006 | -0.001 | 0.11 | 0.009 | 0.03 | -0.29 |
| | (0.1) | (0.09) | (0.33) | (0.12) | (0.11) | (0.43) | (0.15) | (0.13) | (0.51) |
| Constant | 0.35*** | 0.46*** | 0.46*** | 0.4*** | 0.43** | 0.41** | 0.31*** | 0.38** | 0.4** |
| | (0.03) | (0.13) | (0.14) | (0.04) | (0.17) | (0.18) | (0.04) | (0.19) | (0.2) |
| Piecewise Quadratic | No | No | Yes | No | No | Yes | No | No | Yes |
| Controls | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Discontinuity F-stat | 10.66 | 14.49 | 7.96 | 7.16 | 9.99 | 6.81 | 4.94 | 5.81 | 3.21 |
| Observations | 1935 | 1935 | 1935 | 1059 | 1059 | 1059 | 876 | 876 | 876 |
| $R^2$ | 0.14 | 0.23 | 0.23 | 0.14 | 0.24 | 0.24 | 0.12 | 0.21 | 0.21 |

(See Notes for all regression tables below Table 2.) Note that this table differs from Table 2 in that the discontinuity is located at KCPE=250 for men and KCPE=240 for women, rather than the locations detected automatically

Table A7: Robustness: Probit versus OLS for employment and fertility.

| **Outcome** | **Estimation** | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | OLS | OLS | Probit | Probit |
| *Panel A: Employment, oldest two cohorts of men* | | | | |
| Completing Std 12 | −0.036 | −0.013 | −0.036 | −0.013 |
| | (0.055) | (0.056) | (0.055) | (0.055) |
| Observations | 378 | 378 | 378 | 378 |
| *Panel B: Self-employment, oldest two cohorts of men* | | | | |
| Completing Std 12 | −0.104** | −0.120*** | −0.106*** | −0.119*** |
| | (0.040) | (0.045) | (0.041) | (0.045) |
| Observations | 378 | 378 | 378 | 378 |
| *Panel C: Fertility (pregnancy by age 18), women* | | | | |
| Completing Std 12 | −0.119*** | −0.131*** | −0.117*** | −0.125*** |
| | (0.020) | (0.021) | (0.018) | (0.019) |
| Observations | 853 | 853 | 853 | 853 |

(See Notes for all regression tables below Table 2.) Only the coefficient on completed secondary schooling is shown; each coefficient comes from a separate regression. The first two columns here duplicate the first two columns of Table 4 for easy reference. Marginal effects are shown for probit estimation. Standard errors for probit marginal effects are obtained via the delta method (provided by the `margins` command in Stata 13).

Table A8: Discontinuity (First Stage) Estimation, Robustness to 250 Cutoff.

| Outcome: Completing secondary school | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Sample restriction: | Pooled | | | Male | | | Female | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| KCPE≥250 | 0.1** | 0.12*** | 0.12** | 0.14*** | 0.15*** | 0.17*** | 0.06 | 0.08 | 0.07 |
| | (0.04) | (0.04) | (0.06) | (0.05) | (0.05) | (0.07) | (0.06) | (0.06) | (0.08) |
| KCPE (score) | 0.36*** | 0.33*** | 0.4* | 0.32*** | 0.28*** | 0.11 | 0.38*** | 0.36*** | 0.73** |
| | (0.06) | (0.06) | (0.24) | (0.09) | (0.09) | (0.32) | (0.08) | (0.09) | (0.37) |
| KCPE≥250×KCPE | 0.003 | -0.02 | -0.22 | -0.006 | -0.001 | 0.11 | 0.01 | -0.03 | -0.72 |
| | (0.09) | (0.08) | (0.34) | (0.12) | (0.11) | (0.43) | (0.15) | (0.13) | (0.54) |
| Constant | 0.39*** | 0.48*** | 0.49*** | 0.4*** | 0.43** | 0.41** | 0.39*** | 0.47** | 0.53** |
| | (0.03) | (0.13) | (0.14) | (0.04) | (0.17) | (0.18) | (0.04) | (0.2) | (0.21) |
| Piecewise Quadratic | No | No | Yes | No | No | Yes | No | No | Yes |
| Controls | No | Yes | Yes | No | Yes | Yes | No | Yes | Yes |
| Discontinuity F-stat | 5.73 | 9.21 | 4.52 | 7.16 | 9.99 | 6.81 | 0.92 | 2.18 | 0.64 |
| Observations | 1925 | 1925 | 1925 | 1059 | 1059 | 1059 | 866 | 866 | 866 |
| $R^2$ | 0.14 | 0.23 | 0.23 | 0.14 | 0.24 | 0.24 | 0.13 | 0.21 | 0.22 |

Specifications in this table are identical to those in Table 2, but with KCPE score centered at 250 rather than the empirically located discontinuities.

Table A9: Robustness: Imbens-Kalyanaraman Bandwidth, men only, 250 alternative cutoff.

| | **Outcome:** Completing secondary | **Outcome:** Cognitive score | **Outcome:** Employment | | **Outcome:** Self-employment | |
|---|---|---|---|---|---|---|
| | OLS | 2SLS | 2SLS | IVP | 2SLS | IVP |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | 0.154** | 1.437* | 0.076 | 0.172 | −2.980 | −0.503*** |
| | (0.065) | (0.736) | (0.475) | (0.401) | (3.079) | (0.016) |
| Observations | 929 | 670 | 383 | 383 | 143 | 143 |
| IK Bandwidth: | 63 | 40 | 81 | 81 | 23 | 23 |
| Restrictions: | | | Oldest two cohorts | | Oldest two cohorts | |

This table contains estimation identical to that in Appendix Table 5, but restricting attention to men, and to the cutoff score of 250 rather than the data-driven cutoff of 251.

Table A10: Robustness: Unemployment outcome.

| | (1) 2SLS | (2) 2SLS | (3) 2SLS | (4) 2SLS | (5) IVP | (6) IVP | (7) IVP | (8) IVP |
|---|---|---|---|---|---|---|---|---|
| P[Unemployed] | 0.026 | −0.004 | 0.179 | 0.170 | 0.029 | −0.003 | 0.174 | 0.168 |
| | (0.358) | (0.387) | (0.360) | (0.387) | (0.353) | (0.382) | (0.311) | (0.339) |
| P[Unemployed] | 0.026 | −0.004 | 0.179 | 0.170 | 0.029 | −0.003 | 0.174 | 0.168 |
| | (0.358) | (0.387) | (0.360) | (0.387) | (0.353) | (0.382) | (0.311) | (0.339) |
| Observations | 378 | 378 | 378 | 378 | 378 | 378 | 378 | 378 |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Exclude vocational | No | No | Yes | Yes | No | No | Yes | Yes |

(See NOTES FOR ALL REGRESSION TABLES below Table 2.) Only the coefficient on completed secondary schooling is shown; each coefficient comes from a separate regression. All results presented in the table are 2SLS or IV probit estimates of the effect of secondary schooling on the outcome variable. (For IV probit, marginal effects are shown.) Above, the outcome variable is an indicator for being "unemployed," meaning neither in school, self-employed, nor employed by anyone else. Columns (3), (4), (7), and (8) also exclude enrollment in a vocational training program (which is not necessarily full-time) from the definition of "unemployment" for the purpose of the regression.

Table A11: Additional outcomes: Labor variables, older cohorts of women.

| | (1) OLS | (2) OLS | (3) IVP | (4) IVP | (5) 2SLS | (6) 2SLS |
|---|---|---|---|---|---|---|
| P[Employed] | 0.099** | 0.103** | −0.282 | −0.404 | −0.308 | −0.484 |
| | (0.040) | (0.040) | (0.452) | (0.382) | (0.499) | (0.645) |
| P[Self-employed] | −0.132*** | −0.139*** | −0.468 | −0.498 | −0.480 | −0.534 |
| | (0.044) | (0.045) | (0.301) | (0.327) | (0.468) | (0.581) |
| Observations | 325 | 325 | 325 | 325 | 325 | 325 |
| Controls | No | Yes | No | Yes | No | Yes |

(See NOTES FOR ALL REGRESSION TABLES below Table 2.) Only the coefficient on completed secondary schooling is shown; each coefficient comes from a separate regression. IVP denotes IV probit, for which marginal effects are shown.