The Determinants and Consequences of
School Choice Errors in Kenya[†]

By Adrienne M. Lucas and Isaac M. Mbiti*

School choice programs are a relatively new phenomenon in the United States, but such systems have been in place in some developing countries since the 1960s. While the allocation mechanism by which students are assigned to a school varies, most programs combine students' expressed preferences with either a merit or lottery system that determines student prioritization in the allocation process. Because of the high-stakes nature of a student's stated school preferences, information asymmetries at the time of application can result in a student making ill-advised school choices, which could ultimately reduce the quality of the school he or she attends. In this paper, we examine the relationship between student characteristics and school choice errors among high-ability students applying to the most elite secondary schools in Kenya. We further examine the association between these errors and the probability of admission to these elite schools.

Recent research by Hastings and Weinstein (2008) in Charlotte-Mecklenburg, Lai, Sadoulet, and de Janvry (2009) in Beijing, and Ajayi (2011) in Ghana has shown the potential for school choice systems to result in the stratification of the education system by socioeconomic status. These papers highlighted the importance of information constraints in school choice programs where parents with low levels of education or income and students in low-quality primary schools were

more likely to make ex post judgment errors in the school choice process and less likely to apply to the highest-quality schools, resulting in these children attending lower-quality schools. Moreover, Hastings and Weinstein (2008) and Ajayi (2011) found that providing students and parents with more information about the schools and/or the school choice process eliminated some of this disparity, thereby encouraging economic mobility.

We use data from the nationwide secondary-school admissions process in Kenya to explore the relationship between a student's observable characteristics and errors in the secondary-school choice process. We further examine the association between these errors and the student's probability of admission into the highest-quality schools in Kenya. The Kenyan system admits students to secondary school based on achievement on a national examination, students' stated preferences, and predetermined district quotas for each school. Because of the high degree of heterogeneity in school quality, students must be extremely judicious in stating their school preferences, especially since the system limits the number of schools to which they can apply. Thus, ill-advised application decisions could ultimately lead to two students with the same test score from the same district attending schools of vastly different quality. We focus our analysis on the admission process for the relatively small set of the most prestigious secondary schools in Kenya and restrict our sample to the highest-ability students in Kenya, those with realistic expectations of gaining admission to these schools. Even among high-ability students with a small set of schools from which to choose, we still find a significant prevalence rate of school choice errors where girls, students with lower exam scores, and students from public and lower-quality primary schools are more likely to make selection errors. These errors are associated with a reduction in the likelihood that a student is admitted to

*Lucas: Department of Economics, University of Delaware, 416C Purnell Hall, Newark, DE 19716 (e-mail: alucas@udel.edu); Mbiti: Department of Economics, Southern Methodist University, 3300 Dyer St., Suite 301, Dallas, TX 75275 and J-PAL (e-mail: imbiti@smu.edu). We are especially thankful to the Kenya National Examination Council for providing us with the data. The contents of this paper do not reflect the opinions or policies of this agency. The authors bear sole responsibility for the content of this paper. For useful comments and suggestions we thank Erica Groshen.

an elite school. This paper highlights the potential for school choice programs to have unintended consequences on members of the disadvantaged groups they are designed to help.

I. Background and Setting

The Kenyan educational system consists of eight years of primary and four years of secondary education. To earn a primary school diploma, students take the national Kenya Certificate of Primary Education (KCPE) examination after grade eight. During the registration for the KCPE, prior to learning their scores or taking the exam, students submit their secondary school choices.

Government secondary schools are tiered (or grouped) into three categories: national schools, provincial schools, and district schools. The 18 national schools are the most prestigious schools in the country and admit the best-performing students according to district-based quotas that ensure students from across the country access to these schools. They offer the best infrastructure, teachers with high education levels, high-ability peers, a broad curriculum, and name cachet. Provincial schools are the second tier and admit the best remaining students from within their provinces. Relative to national schools, provincial schools are generally inferior across all observable dimensions of quality. District schools admit students from their own districts who were not admitted to national or provincial schools. These schools are generally the lowest-quality secondary schools in Kenya.

As a part of the KCPE registration, students list the school codes that correspond to their top two choices from each tier. Students are admitted to secondary school based on their scores on the KCPE, preferences, and district quotas. Each national school has a district quota—a set number of students that will be admitted from each district. Once the exams have been scored, students are ranked within a district. The highest-ranked student receives his first choice of schools and the selection continues down the ranked students, filling the district quota for each national school. If a student's preferred national schools are full, then he is admitted to a preferred provincial school (if those schools still have space) even if other national schools have space remaining. As all national schools are single-sex boarding schools, the ranking

and allocation effectively occur within district-gender cells.

Each year, newspapers publish the top secondary schools based on their students' scores and the scores, names, and schools of the top students in the country. The public availability of this information could ameliorate some of the information asymmetries present in other contexts.

II. Conceptual Framework and Empirical Strategy

We examine the relationship between various school choice errors and student characteristics. We are unable, however, to isolate a single causal mechanism due to the possibility of a correlation between unobservable student and parental attributes and both observable characteristics and choice errors. We employ the following linear probability model:

$$(1) \quad y_i = \beta_0 + \beta_1 \, female_i + \beta_2 \, KCPE_i$$
$$+ \beta_3 \, public_i + \beta_4 \, KCPEschool_i$$
$$+ \delta_j + \varepsilon_i,$$

where y_i is a binary variable equal to one if student i made a particular type of error (discussed below) in his or her application, *female* is a dummy equal to one for female students, *KCPE* is the student's standardized KCPE score, *public* is a dummy variable equal to one if the student graduated from a public primary school, *KCPEschool* is the average standardized KCPE score at the student's primary school, δ_j are district fixed effects, and ε is an error term assumed to be independent between districts but allowed to be correlated between students in the same district. The *public* dummy variable includes all attributes of a student that are correlated with public school attendance including lower socioeconomic status and being subject to any potential public school shortcomings (e.g., less information or lack of guidance).[1] *KCPEschool* combines both peer and, potentially, school quality and any related shortcomings.

---

[1] Lucas and Mbiti (forthcoming) found that the Kenyan Free Primary Education program, implemented in 2003, increased the socioeconomic stratification between public and private primary schools and led to a deterioration in the student-teacher ratio in public primary schools.

We examine three types of school choice errors. First, students who filled out their preferences forms incorrectly by leaving blanks, listing a nonnational school, or listing a school that does not admit students of their gender clearly made a very costly mistake. We use this binary definition of a mistake as a dependent variable in equation (1). Second, a student could choose a school that would not have been expected to take anyone from their district, i.e., a quota error. Since some national schools do not admit students from particular districts and the quotas change very little from year to year, we would expect that well-informed students would know not to select such a school. Third, students could incorrectly rank their schools based on the schools' relative selectivity.[2] By district we determine the relative selectivity of schools by the score that was necessary to gain admission to each school from that district, with more-selective schools having higher score thresholds.[3] If a student's score was not sufficient for a less-selective school, it would not have been sufficient for the more-selective school. Therefore, a student who lists a school that is more selective as choice two would never be admitted to his choice two school, effectively wasting a school choice and making a selectivity error.[4]

All of the above errors combine, along with a student's score, stated choices, and district quotas, to determine whether a student is admitted to a national school. We further estimate the relationship between these errors and the probability of being admitted to a national school by including these errors as additional covariates in equation (1) with a dummy variable for being admitted to a national school as the dependent variable.

III. Data

Our student data are administrative testing records from the Kenya National Examination Council for the 2004 KCPE cohort. For each student these data include primary school name and type (public or private), gender, and the six stated school preferences, two for each school tier. We use the national school district-gender quotas from the Kenyan Ministry of Education along with the allocation rule to solve for the school to which a student should have been admitted.[5] We limit our sample to those students whose KCPE scores place them in the top five percent for their gender-district; i.e., those students with a realistic opportunity to be admitted to a national school.[6] Within our sample, 3.7 percent of students listed at least 1 invalid school code, 1.8 percent listed 2 such codes, 3.7 percent listed a school that did not admit anyone from their district, and 36 percent of students listed as second choice a school that was more selective than their first choice. The incidence of the first two types of mistakes is substantially higher for the full score distribution of students with students 2.7 times as likely to list one invalid school code and 4.0 times as likely to list two such codes, but the other errors occur at approximately the same rate in our high test score samples as in the whole population of test takers.

IV. Results

Table 1 contains the estimation of equation (1) with dummy variables that indicate different types of mistakes. Column 1 shows that female students are 1.8 percentage points more likely to enter an incorrect school code for at least one of their choices. Moreover, girls are 0.9 percentage points more likely than boys to enter two incorrect national school codes (column 2), eliminating any chance of being admitted to a national school through the centralized admission process. Further, based on these first two columns, public school students are more likely to make at least one mistake, while students with

<hr>

[2] Lai, Sadoulet, and de Janvry (2009) consider similar types of mistakes "ex post judgment" errors.

[3] See Lucas and Mbiti (2012) for further discussion of the determination of the score thresholds.

[4] We use a contemporaneous measure of selectivity. The relative selectivity of schools is mostly constant over time, but can vary from year to year. Our results are substantively similar if we use a measure of selectivity based on the 2001 KCPE cohort.

[5] Some students who are not initially admitted to a national school appeal directly to a school principal and gain admission through an unofficial second round. These admissions can be based on parental income and influence and are not considered here.

[6] The lowest KCPE score in our sample is one-half a standard deviation above the countrywide mean, and most students scored two standard deviations above the countrywide mean. Our results are substantively similar if we define the sample as the top two percent, those within one standard deviation of the top score for their district-gender, or those within one-half of a standard deviation of a national school effective-score threshold. See Lucas and Mbiti (2012) for a discussion of the effective score threshold.

TABLE 1—DETERMINANTS OF SCHOOL CHOICE ERRORS

	Stated choice			
	Is not a valid national school code		Did not admit students from district	Two is more selective than choice one
	At least one (1)	Both (2)	(3)	(4)
Female	0.018*** (0.005)	0.009** (0.004)	−0.033*** (0.009)	0.039** (0.015)
Standardized KCPE score	−0.021*** (0.005)	−0.014*** (0.005)	−0.026*** (0.006)	−0.082*** (0.014)
Public primary school	0.011* (0.006)	0.009 (0.006)	0.000 (0.006)	0.003 (0.010)
Primary school average KCPE score	−0.016** (0.008)	−0.012* (0.007)	0.016*** (0.004)	0.012 (0.007)
Observations	32,935	32,935	32,935	32,935
R^2	0.055	0.067	0.076	0.026

Notes: Linear probability models; standard errors clustered at the district level appear in parentheses; additional controls: district fixed effects.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

higher KCPE scores and students from schools with higher average KCPE scores are less likely to make mistakes. Each standard deviation improvement in a student's own score or school score reduces the likelihood of making a mistake by about two percentage points.[7] Column 3 shows that girls and students with higher scores are less likely to apply to schools that do not regularly admit students from their district, and students from schools with higher KCPE scores are more likely to list such schools. The coefficient for girls could reflect that fewer quotas for girls are 0, leaving less opportunity for girls to make such an error. The relationship between school KCPE and this type of error is somewhat surprising, but could signify that students at the highest-quality schools are trying to list a lesser-known national school as a backup, especially since no students listed two such schools. Finally, in column 4 we test for selectivity errors where a student's second-choice school was more selective than the first-choice school. This mistake is more common for girls and decreasing in own KCPE score.[8]

Overall, these results suggest that girls and students with lower KCPE scores, even conditional on school quality, are more likely to make various errors in their secondary school choices. Moreover, we find only limited evidence of systematic errors by public school students, conditional on our measure of school quality. While our results do not enable us to distinguish informational constraints from other plausible hypotheses, they do highlight that females and the lowest-scoring students could potentially benefit the most from programs or interventions that increase the information available to students.

Table 2 assesses the relationship between errors and admission outcomes, controlling for the same demographic characteristics as in

---

[7] We find substantially similar results for gender and own KCPE score when we replace primary school characteristics with primary school fixed effects, indicating that the results in Table 1 do not reflect an unobserved attribute of the primary schools that girls and students with lower scores attended.

[8] The standardized test scores for students and schools measure a student's and school's position relative to the countrywide distribution. When we included measures of a student's percentile within their school and a school's percentile within its district as additional covariates in regressions that replicated all four columns of Table 1, their coefficients were not statistically different from 0.

TABLE 2— CONSEQUENCES OF SCHOOL CHOICE ERRORS

	Probability of being admitted to a national school (1)
At least one invalid school code	−0.020***
	(0.008)
Two invalid school codes	0.022*
	(0.012)
District quota of zero	−0.044***
	(0.008)
Choice two more selective than choice one	−0.022***
	(0.003)
Observations	32,935
R^2	0.317

Notes: Linear probability model; additional controls: student KCPE score; average school KCPE score; and female, public primary school, and district dummy variables. Standard errors clustered at the district level appear in parentheses.

  ***Significant at the 1 percent level.
  **Significant at the 5 percent level.
  *Significant at the 10 percent level.

Table 1. We find that net of these characteristics, students who list an invalid school code are two percentage points less likely to be admitted to a national school. Surprisingly, the coefficient on two invalid school codes is positive and marginally statistically significant. Students with such choices listed at least one invalid school code as well, and we cannot reject that the sum of those coefficients is negative. Students who listed a school that did not admit anyone from their district or listed a more-selective school as choice two were also less likely to be admitted to a national school. In our sample, 7.2 percent of students were admitted to a national school. Therefore, making school choice errors is associated with a substantial reduction in the probability of admission to a national school net of available demographics.[9]

V. Discussion and Conclusions

We find that even among the top-scoring students in Kenya who select from a relatively small set of national schools, girls and students with lower average KCPE scores are more likely to make errors in the secondary school selection process. Because national school places are allocated separately for boys and girls, while some girls are penalized by their selections, the total number of girls in national schools is not affected, but these errors could deteriorate the match quality between students and schools. Additionally, we find that net of observable characteristics, these errors are associated with a reduced likelihood that a student is admitted to a national school, generating an inefficient outcome with potentially persistent effects. Therefore, while the Kenyan admissions process, like many other systems, allows all students the opportunity to access the highest-quality secondary schools, school choice errors in the admissions process could undermine offering the best opportunities to the highest-ability students and cause inequalities to persist.

One important consideration is that this sample included only the highest-achieving students. Therefore, the results might be different for students lower in the score distribution, but as those students are far from being admitted to a national school, their national school choices are less relevant. Students are admitted to provincial schools using the same admissions algorithm. The students who are on the margin of being admitted to provincial schools have lower scores than the margin currently under study. Future research will explore the choices selected by these students and the role of information provision on improving the school allocation outcomes for the most ill-informed students.

REFERENCES

Ajayi, Kehinde F. 2011. "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana." Unpublished.

Hastings, Justine S., and Jeffrey M. Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics* 123 (4): 1373–1414.

Lai, Fang, Elisabeth Sadoulet, and Alain de Janvry. 2009. "The Adverse Effects of Parents' School Selection Errors on Academic Achievement: Evidence from the Beijing Open Enrollment Program." *Economics of Education Review* 28 (4): 485–96.

---

[9] The ability to list choices that maximize the admissions probability could be related to the likelihood of errors and other unobserved heterogeneity. Therefore, these results are not necessarily causal.

Lucas, Adrienne M., and Isaac M. Mbiti. Forthcoming. "Access, Sorting, and Achievement: the Short-Run Effects of Free Primary Education in Kenya." *American Economic Journal: Applied Economics*.

Lucas, Adrienne M., and Isaac M. Mbiti. 2012. "Effects of Attending Selective Secondary Schools on Student Achievement: Discontinuity Evidence from Kenya." Unpublished.