Human Development Research Paper 2009/45

Individual Ability and Selection into Migration in Kenya

Edward Miguel
and Joan Hamory

United Nations Development Programme
Human Development Reports
Research Paper

September 2009

Human Development Research Paper 2009/45

Individual Ability and Selection into Migration in Kenya

Edward Miguel
and Joan Hamory

United Nations Development Programme
Human Development Reports
Research Paper 2009/45
September 2009

Individual Ability and Selection into Migration in Kenya

Edward Miguel
and Joan Hamory

Joan Hamory is Researcher at the Center of Evaluation for Global Action, University of California, Berkeley.
E-mail: jrhamory@gmail.com.

Edward Miguel is Professor of economics and Director of the Centre of Evaluation for Global Action, University of California, Berkeley. E-mail: emiguel@econ.berkeley.edu.

Comments should be addressed by email to the author(s).

Abstract

This study exploits a new longitudinal dataset to examine selective migration among 1,500 Kenyan youth originally living in rural areas. We examine whether migration rates are related to individual "ability", broadly defined to include cognitive aptitude as well as health, and then use these estimates to determine how much of the urban-rural wage gap in Kenya is due to selection versus actual productivity differences. Whereas previous empirical work has focused on schooling attainment as a proxy for cognitive ability, we employ an arguably preferable measure, a pre-migration primary school academic test score. Pre-migration randomized assignment to a deworming treatment program provides variation in health status. We find a positive relationship between both measures of human capital (cognitive ability and deworming) and subsequent migration, though only the former is robust at standard statistical significance levels. Specifically, an increase of two standard deviations in academic test score increases the likelihood of rural-urban migration by 17%. Accounting for migration selection due to both cognitive ability and schooling attainment does not explain more than a small fraction of the sizeable urban-rural wage gap in Kenya, suggesting that productivity differences across sectors remain large.

Keywords: Migration, selection, human capital, ability, urban-rural wage gap, productivity.

1. Introduction[1]

Migration is a central issue in the study of labor markets in less developed countries. While the issue of selection into migration has been widely studied in the context of Mexico-U.S. migration (Chiquiar and Hanson 2005), there is little rigorous evidence on patterns of selective rural-urban migration in less developed countries, in large part due to the scarcity of panel datasets that track individuals over time as they make migration decisions (Rosenzweig 1988). Understanding the nature of selection into urban migration as a function of individual ability can help shed light on urban-rural wage gaps, in particular, how much of the gap is due to real productivity differences across sectors versus unobserved differences in average worker ability. Characterizing rural-urban migration is also fundamental to understanding the "structural transformation" out of agriculture that is central to the process of economic development.

We explore selection into rural-urban migration and estimation of the urban-rural wage gap using a new panel data set of Kenyan youth. The Kenyan Life Panel Survey (KLPS) is a unique database, tracking over time 7,500 children who attended primary school in Busia, a rural district of western Kenya, in 1998. In Round 1 of this survey, enumerated during 2003-2005 and referred to hereafter as KLPS-1, longitudinal information was collected for more than 5,200 of these individuals on a wide range of outcomes, including all past residential locations. Round 2 of the KLPS (abbreviated hereafter as KLPS-2), a follow-up survey administered to these same individuals, is currently in the field. Prior to the launch of KLPS-2 enumeration, individuals to be interviewed were randomly divided into two groups (waves), the first to be tracked during 2007/2008, and the second to be tracked during 2008/2009. At the close of Wave 1 in November 2008 nearly 2,500 individuals had been surveyed. This study employs information from these survey respondents, a fully representative subsample of the KLPS population. A main strength of our analysis is the use of this exceptional data source.

The individuals in our analysis were surveyed in 1998, 2003/2005 and 2007/2008, and the latter two surveys collected retrospective migration histories over the intervening periods. As a result, we are able to both measure migration intensity as a series of events (employing the panel aspect of our data), as well as a transition (between survey enumeration rounds). Following

[1] **Acknowledgements:** We thank Sarah Baird, David Evans, Matthew Jukes, and Michael Kremer, our collaborators on the broader KLPS project. Francisco Rodriguez, Duncan Thomas, Chris Woodruff and seminar audiences at UCLA, U.C. Berkeley, and the 2008 ASSA Meetings provided useful comments. We are grateful for financial support from the UNDP. All errors are our own.

Bell and Muhidin (2008), we construct transition measures as descriptive tables early in our paper, measuring migration as a change in "usual home" from the residence in Busia District during the 1998 baseline survey to residence at the time of KLPS-2 survey enumeration in 2007/2008. In the main econometric analysis of selection into urban migration, we then employ the retrospective panel data on all residential moves to capture the full extent of urban migration among rural Kenyan youth.

We focus our analysis on a restricted sample of KLPS-2 respondents with a rich set of pre-migration data on academic test scores, child and household characteristics. Individuals in this age group, primarily 18-26 years old at the time of KLPS-2 tracking, are extremely mobile. During 1998-2008, more than two-thirds of adolescents report migrating from their 1998 residence for a period of at least four months, and 41% report having lived outside of western Kenya and the neighboring parts of Uganda. The vast majority of relocation outside of these local areas is to urban centers elsewhere in Kenya. According to self-reports, schooling, employment search, and lengthy family "visits" are the three most popular reasons given for these moves.

Given this high level of mobility, sample attrition in the KLPS-2 is a natural concern. One of the unique aspects of this survey project is its commitment to locate individuals regardless of where they might have moved: survey enumerators traveled all over Kenya and neighboring Uganda in multiple rounds of long-distance tracking. As a result, 82% of target respondents were interviewed, a remarkably high tracking rate for young adults in a less developed country context. We provide a detailed analysis of tracking patterns to alleviate attrition bias concerns, and fortunately find little evidence that key explanatory variables are systematically related to attrition.

Our main empirical emphasis is two-fold. First, we examine the relationship between individual ability and subsequent migration. Such a relationship can be thought of in the context of a Roy (1951) selection model, as formulated in Borjas (1987). Previous empirical work has used schooling attainment as a proxy for ability (see appendix table A1 for a summary of main results). Resulting evidence is mixed, with most studies finding a positive association between attainment and later migration (Chiquiar and Hanson 2005, McKenzie et al 2006, Grogger and Hanson 2007), but some finding no relationship or even a negative relationship (Ibarraran and Lubotsky 2007). Hunt (2004) finds that long-distance migrants within Germany tend to be high-skilled. The evidence on the relationship between ability and migration in Africa and other low-income regions generally suggests that urban migrants are positively selected. Hoddinott (1994) examines one rural sub-location in western Kenya, and finds a positive relationship between years of schooling and urban migration. Lanzona (1998) similarly finds a positive relationship between years of schooling and migration out of rural areas in Philippines. Zhao (1999) examines migration among inhabitants of China's rural Sichuan province in 1994-5, but finds a small and only weakly positive relationship between years of schooling and migration.

Most of the empirical work on selective migration focuses on a single measure of ability, schooling attainment. We explore a broader definition, including cognitive ability as well as health status. We employ a pre-migration primary school academic test score as a proxy for cognitive aptitude, which to our knowledge is the first use of measure of this kind in a migration selection study. We also exploit pre-migration randomized assignment to a primary school deworming treatment program as a source of exogenous variation in health status, another component of human capital, and thus can more credibly identify the impact of improved health on later migration decisions.

We find only one of these ability measures to be significantly and robustly related to subsequent rural-urban migration, cognitive test scores. This suggests that cognitive aptitude is valued in the urban labor market and physical robustness perhaps less so on average. Specifically, we find that an increase of two standard deviations in 1998 academic test score increases the likelihood of subsequent migration to a city by 17%. Results are robust to several different specifications, including conditioning on measures of parent education and household asset ownership. We conclude young adults with higher cognitive ability are more likely to migrate to urban areas in Kenya. In an interesting contrast with the existing literature, schooling attainment is not associated with urban migration once cognitive ability is accounted for.

Given the high level of migration into Uganda among individuals in our sample, we extend this analysis further to explore selection into international migration. We find no relationship between our multiple measures of individual ability and subsequent international migration, likely because most adolescents moving from Busia, Kenya settle just across the border in similarly rural areas of Uganda, where cognitive and other skills are apparently not as highly valued as they are in urban labor markets.

In the second part of the analysis, we use these improved ability measures to provide more credible estimates of the urban-rural wage gap in Kenya. Specifically, we estimate how much of the massive observed Kenyan urban wage premium – urban wages in our sample are nearly twice as large as rural wages – falls when cognitive and other ability terms are included as controls in the analysis. Cognitive ability and schooling attainment are both meaningful predictors of higher wages, particularly for men. However, accounting for both individual cognitive ability and schooling attainment can explain only a small fraction of the urban-rural wage gap in our sample of Kenyan youth. This suggests that the large urban-rural wage gap in Kenya is driven by large productivity differences, or perhaps by some measures of individual ability not well captured in the variables we employ in our analysis (e.g., personality traits).

The paper proceeds as follows: section 2 describes the data, section 3 lays out a Roy selection framework, section 4 provides the main empirical evidence on selective migration, section 5 estimates the selection-corrected urban-rural wage gap in Kenya, and the final section concludes.

2. Data

In 1998, the Primary School Deworming Project (PSDP), an intestinal helminth treatment program, was launched in Busia, a rural district in western Kenya. Under this program, a local non-governmental organization (NGO) provided deworming treatment to over 30,000 primary school children aged 6-18. In order to evaluate the effects of this health intervention, baseline data was collected on individual school participation, academic performance, health and household characteristics.[2] Five years later a follow-up survey known as the Kenyan Life Panel Survey Round 1 (KLPS-1) was launched. Between 2003 and 2005, this survey tracked a representative sample of 7,500 of these adolescents who were confirmed enrolled in primary school grades 2-7 in Busia District in 1998.[3] Survey data on a wide range of outcomes was successfully collected for over 5,200 of these young adults, including panel information on all residences inhabited for a period of at least four months between 1998 and 2005. In mid-2007, a second round of the Kenyan Life Panel Survey (KLPS-2) went to the field. All sample individuals were randomly divided into two groups, to be tracked in two separate waves of data

[2] Miguel and Kremer (2004) provide more background information on the PSDP.
[3] Note that this population is still broadly representative of the adolescent population in western Kenya: according to the 1998 Kenya DHS, 85% of children in Western Province aged 6-15 are enrolled in school.

collection, both of which are fully representative of the main sample. Wave 1 of the KLPS-2 was completed in November 2008, and contains survey information for nearly 2,500 individuals that form the core of the analysis in this paper. (Wave 2 of KLPS-2 data collection is currently underway, and will be completed in late 2009 and included in future analyses.)

In the current analysis, we employ both the baseline PSDP and the follow-up Wave 1 KLPS-2 data. We focus on a restricted sample of 1,518 individuals with detailed baseline academic test score, school participation and survey data in addition to the KLPS panel residential location information. Baseline academic test score and survey data exist for individuals who were present in school on the pre-announced day the test or survey was administered, and includes only students in grades 3 through 7 in 1998.

A key strength of the KLPS is its respondent tracking methodology. In addition to interviewing individuals still living in Busia District, survey enumerators scoured Kenya and Uganda to interview those who had moved out of local areas. Information was collected on each location inhabited since 1998 for a period of four months or more, as well as reasons for the move and any known contacts in the new location. This endeavor results in a dataset well-suited to the study of migration. Furthermore, the KLPS-2 collects detailed information on the employment and wage history of respondents, providing a rare opportunity to explore labor market outcomes among a group of highly mobile African youth.

In addition to the panel information on residential location, employment and wages, we focus on two unique variables contained in the baseline PSDP data: a pre-migration academic test score and an exogenously assigned proxy for pre-migration individual health. The baseline academic test score data comes from an exam administered to primary school students in grades 3-8 as part of the initial PSDP evaluation. The test was based on standard Kenya Ministry of Education exams, and covered three subjects – English, Math, and Science/Agriculture. Each grade level was administered a separate exam.[4] Students present in school on the day the test was administered are included in the sample. In addition, a small sample of students who had dropped out of school during 1998 were tracked to their homes and also asked to complete the exam, and we use this latter group for robustness checks in our analysis.

[4] We implicitly assume that normalized test score at different ages captures ability to the same extent. This is plausible given our data – each year only 2-8% of students stop attending school between the grades of 3 and 7, suggesting only a second-order ability bias in higher grade levels.

Our measure of pre-migration health is based on the randomized deworming treatment provided to primary school children in Busia District under the PSDP. A parasitological survey conducted by the Kenya Ministry of Health, Division of Vector Borne Diseases in early 1998 suggested that this district is characterized by an extremely high intestinal worm infection rate, on the order of 92% among sampled children in grades 3 through 8 (Miguel and Kremer 2004). Intestinal helminth infections, especially more severe cases, lead to a broad range of negative health outcomes, including abdominal pain, anemia, malnutrition, stunting, wasting, and lethargy. Since intestinal worms have life spans of just one to three years and do not replicate in the human host, periodic deworming treatment can greatly reduce infection.

Under the PSDP, a local NGO provided deworming treatment to individuals in seventy-five schools in Busia District. Due to administrative and financial constraints, the program was phased in over a four-year period. Schools were randomly divided into three groups, with Group 1 schools receiving treatment starting in 1998, Group 2 schools receiving treatment starting in 1999, and Group 3 schools receiving treatment starting in 2001. Thus, Group 3 children received three fewer years of treatment than Group 1, and Group 3 children initially in grades 6 or 7 received no treatment at all.[5]

Below, we examine the relationship between the randomized deworming treatment and subsequent migration. Evidence on the link between the intervention and individual health status has been established elsewhere. Miguel and Kremer (2004) evaluate the short-run impacts of the PSDP, and find significant self-reported health and height-for-age gains after just one year of treatment. Such improvements could be associated with greater strength and labor productivity. The authors also found a drop in school absenteeism by one quarter in treatment schools, although no early academic or cognitive test impacts were found; they suggest this lack of an academic performance effect could be due in part to increased classroom congestion.

Miguel, Baird and Kremer (2007) examine the longer-run impacts of the program, using the KLPS-1 follow-up survey. The authors find long-term height and weight gains for those in lower grades in 1998, females, and for those that live in particularly high infection areas. Recognizing the difficulty in disentangling particular health impacts from each other, a mean effects approach is also used to determine the overall impact of the deworming intervention, and

[5] Although only children who were in school on the day of the drug administration received treatment, compliance rates were high, on the order of 70% (Miguel and Kremer 2004).

the authors report a positive impact of the treatment on height, weight and general health. Together, these studies suggest that deworming treatment has significant positive impacts on individual health. Such effects could continue to work through later life health, strength, and cognitive ability. We will not attempt to disentangle these effects here, but instead we focus on the randomized deworming intervention as a proxy for pre-migration individual health status.[6]

3. A Model of Selective Migration

The Roy (1951) selection model provides a useful framework for considering rural-urban migration in less developed countries, as further developed in Borjas' (1987) work. Consider an economy with two sectors, one urban and one rural. Wages in both sectors, denoted w_U and w_R, depend on individual ability, h_i. Further, there is some individual cost to migration, c_i. The Roy model suggests that individuals move to exploit wage differences across different sectors or regions. The migration decision can be characterized as:

$$\text{Migrate if } w_U(h_i) - w_R(h_i) - c_i \geq 0 \qquad (1)$$

It is natural to consider positive returns to ability in both sectors, $w_U'(h) > 0, w_R'(h) > 0$. There are many ways to think about individual ability. Traditionally, this trait has been modeled in terms of school attainment. However, ability can be thought of as a multidimensional variable, also including cognitive aptitude and health.

Migration costs can be modeled more explicitly as a function of observed (X_i) and unobserved (e_i) individual and household characteristics. For instance, if credit constraints matter, then costs could be related to household income or wealth. This leads to a natural specification for a cost function:

$$c_i = -X_i'b - e_i \qquad (2)$$

Allowing the urban-rural wage gap to be defined as

$$g(h_i) = w_U(h_i) - w_R(h_i) \qquad (3)$$

then it follows that the migration decision can be rewritten in a standard discrete choice framework:

---

[6] We recognize that the measure we use would be more easily interpretable if it were linked more concretely to a particular health outcome. However, as shown by these previous studies on the wide-ranging effects of the deworming treatment, choosing a single health outcome such as height or weight is restrictive.

$$Migrate_i = 1\{g(h_u) + X_i'b + e_i \geq 0\} \quad (4)$$

Such a formulation leads to a probit specification in which individuals choose to migrate as long as the return from doing so is greater than the cost. Here, higher ability people are more likely to migrate if there are greater returns to ability in the urban sector, $w_U'(h_i) > w_R'(h_i)$, conditional on any migration costs. This is quite plausible, for instance if cognitive ability matters more in factory or office work than it does on the farm.

4. Empirical results

4.1 Attrition

Searching for individuals in rural Kenya is an onerous task, and migration of target respondents is particularly problematic in the absence of information such as forwarding addresses or phone numbers. This difficulty is especially salient for the KLPS, which follows young adults in their teens and early twenties. This age group is likely to be extremely mobile due to marriage, schooling, and labor market opportunities. Thus, it is essential to carefully examine survey attrition. If our key explanatory variables are related to attrition, then any resulting estimation will likely be biased.

Table 1, Panel A provides a summary of tracking outcomes for the individuals we study. Nearly 86% of adolescents were located by the field team, such that 82% were surveyed and 4% refused participation, were found but unable to survey, or were found to be deceased.[7] Tables 1 and 2 break out these statistics by PSDP deworming treatment group, gender and 1998 age group. These figures suggest that tracking rates are fairly similar across treatment groups, though they are somewhat higher for males than females, and decrease monotonically with age.

We have detailed information on where all surveyed respondents were living at the time of KLPS-2 tracking. Table 1, Panel B and Table 2, Panel B summarize this information. These statistics suggest a great deal of migration in the cross-section: the crude migration intensity capturing moves outside of Busia District from 1998 until the KLPS-2 survey is 28%. Since

[7] The 7,500 individuals sampled for KLPS-2 participation were divided in half, to be tracked in two separate waves. KLPS-2 Wave 1 tracking launched in Fall 2007 and ended in November 2008. During the first several months of Wave 1, all sampled individuals were tracked. In August 2008, a random subsample containing approximately one-quarter of the remaining unfound focus respondents was drawn. Those sampled were tracked "intensively" for the remaining months, while those not sampled were no longer tracked. We re-weight those chosen for the "intensive" sample by their added importance. As a result, all figures reported here are "effective" rates – calculated as a fraction of those found, or not found but searched for during intensive tracking, with weights adjusted properly. For a detailed explanation of this methodology, see Orr et. al (2003).

individuals we did not find, and did not obtain residential information for, are even more likely to have moved away, these figures almost certainly understate true migration rates.[8]

More than 7% of individuals had moved to neighboring districts, including just across the border into the districts of Busia and Bugiri, Uganda. Over 20% of those with location information were living further afield, with nearly 80% of these individuals inhabiting the five major urban areas in Kenya – Nairobi, Mombasa, Kisumu, Nakuru and Eldoret.[9] Five percent of individuals had moved outside of Kenya, nearly all into the neighboring country of Uganda.

Migration rates are fairly similar across deworming treatment groups, with a slightly higher proportion of Group 3 individuals located outside of Busia and nearby districts. Females appear to have somewhat higher migration rates than males, primarily to regions neighboring Busia. This can likely be explained by high female mobility due to marriage. We also see strong evidence of migration rates increasing with age, particularly with regard to migration outside of Busia and its environs, as well as outside of Kenya as a whole.

Table 3 provides a more formal analysis of survey attrition, with focus on two key measures of individual ability, the 1998 academic test score and years assigned deworming treatment during 1998-2003, in probit specifications. The first column contains the deworming measure by itself, along with a set of controls for gender and 1998 grade, as well as baseline individual and household characteristics (whose descriptive statistics are presented in Table 4). Column (2) adds individual test score to this base specification, and column (3) further includes a control for average baseline school participation. Column (4) includes interactions of both ability measures with each other, gender and age, and columns (5) and (6) repeat earlier specifications using a linear probability model including school fixed effects. We find no evidence that years assigned deworming is systematically related to whether or not an individual was surveyed, and only weak evidence that higher pupil test scores contribute to survey attrition. This latter result is consistent with our findings below, namely that individuals with higher test scores are also more likely to migrate, and thus are generally more difficult to find. Together, this indicates that biases related to differential sample attrition in our main analysis are unlikely

[8] This figure is roughly comparable to Bell and Muhidin's (2008) estimate of lifetime migration intensity, 20% using IPUMS data, though we study migration from 1998 origin rather than birthplace. Our rate is higher, likely in part due to the younger age and rural origin of our focus population.

[9] We define urban areas as those with populations of greater than 150,000. Our measure of location is imperfect in that we observe districts of residence rather than cities. However, the 1999 Kenyan Census indicates that 100% of Nairobi and Mombasa districts – our respondents' main destinations – are urban, with lesser fractions for Kisumu, Nakuru and Uasin Gishu.

to be severe, but indeed likely work against our finding a selection effect: we may actually be slightly understating the relationship between cognitive ability and migration if more high ability migrants are lost from the analysis, as seems plausible given the results in Table 3.

4.2 Migration in the KLPS-2

Over 28% of young adults were no longer living in Busia District at the time of KLPS-2 enumeration. This cross-sectional figure understates total migration among this age group, however. Panel residential location information for the period 1998-2008 among surveyed individuals suggests that 55% of adolescents migrated outside of Busia District at some point for a period of at least four months. This is perhaps not surprising: most individuals in the study group are in their early twenties at the time of KLPS-2 tracking, a period in their lives of tremendous flux as they embark on marriage, job searches or higher education.

Figure 1 displays locations of residence for individuals in our data during 1998-2008.[10] Nearly all adolescents report living in Busia district at some point or in the neighboring areas of Kenya's Western Province and the bordering districts of Uganda. The most popular residential destination by far outside of these local areas is the capital city of Nairobi. Comparatively large fractions of individuals also lived in Rift Valley Province (which houses the major urban areas of Nakuru and Eldoret, and is also an important tea-growing region with large plantations providing relatively well-paid employment), Coastal Province (home to Mombasa), and Nyanza Province (home to Kisumu). In this analysis, we characterize urban migration as residence in cities in Kenya with populations of over 150,000, as well as foreign cities (e.g, Kampala). More than one-third of individuals report living in such locations at some point during the study period. Finally, migration outside of Kenya is substantial: nearly 13% of individuals lived in Uganda at some point. More than 80% of this international migration, however, entailed a move just across the heavily trafficked and porous border between the two countries into neighboring rural districts. Migration to the Ugandan cities of Kampala or Jinja remains comparatively rare.[11]

---

[10] Note that since many individuals lived in more than one location over the eleven-year period, these figures sum to greater than 100%. Further, these figures are not re-weighted to maintain initial population proportions.

[11] Indeed, the authors of this study themselves once unwittingly found out just how porous the Kenya-Uganda border can be. They crossed into Uganda while walking around what they thought was the outskirts of Busia Town in Kenya, and actually strolled for some time in Busia, Uganda before being stopped (and sent back to Kenya with a warning) by a plainclothes Ugandan policeman who noticed the two apparently suspicious-looking economists.

Table 5 provides a simple comparison between individuals who have migrated to a city and those who have not, over a range of individual and household characteristics. Females and older individuals are much more likely to have lived in an urban area. Children who received more years of deworming treatment are actually less likely to live in urban areas, a result which may in part reflect that these individuals tend to be younger (and hence were able to participate in the primary school treatment program longer), and that younger individuals are less likely to have migrated. Individuals with higher baseline body weight are more likely to have migrated, a finding that again may reflect the positive association between urban migration and age instead of a nutrition effect per se. These patterns call for a more rigorous multivariate regression analysis, which we provide below. Interestingly, in the cross-section urban migration is associated with both higher baseline test scores and more years of education attained. This finding goes to the heart of our interest in the measurement of cognitive ability, and we disentangle these two measures in later regressions. Mother's educational attainment is higher for the sample of migrants, though father's attainment does not seem to matter. Finally, urban migrants have more elder siblings on average, a finding perhaps related to family social networks that ease the information and financial costs of migration.

Table 6 displays this same set of simple comparisons, this time for individuals who migrated outside of Kenya (to Uganda) at some point during the survey period versus others. These results differ greatly from the rural-urban migration patterns. First, there is no significant difference in gender between international migrants and non-migrants, although in general older individuals are still more likely to have moved. There also does not appear to be any association between baseline test scores and later migration, and those with higher educational attainment are actually somewhat less likely to have moved outside of Kenya. Finally, migrants are more likely to come from households without a latrine, and thus perhaps come from homes of lower socio-economic status, and have fewer siblings. Together, Tables 5 and 6 demonstrate that rural-urban migration and international migration patterns differ sharply in the Kenyan context, consistent with the finding that an overwhelming proportion of migrants to Uganda settle in rural districts near the Kenyan border. We explore the differences between these migration patterns with further descriptive statistics and a more detailed regression analysis below.

Table 7 provides descriptive information on the frequency of moves and length of stay among these rural Kenyan youth, for both urban and international migrants. Panel A focuses on the former group. As previously noted, over one-third of adolescents report living in a city at some point during the 1998-2008 period, and rates are slightly higher for females and older individuals. Individuals who report rural-urban migration moved on average 2.38 times during 1998-2008, and the average length of stay in a city among these movers is 2.25 years.[12] Though older females are more likely to have ever lived in an urban area, it is older males who tend to stay longer. This may be due to the activities undertaken in the new location – as shown below, women who move to the city tend to work in domestic service jobs as temporary or casual laborers, while men are more likely to obtain permanent positions in an industrial sector.

Panel B of Table 7 explores these same figures for international migrants, and again patterns are quite different. Individuals who have lived outside of Kenya tend to be older and male, while it is the older females who stay abroad longer. Again, this appears to be related to the migrants' activities: a large share of female migration into Uganda is due to marriage, which is typically a long-term proposition.

Table 8 breaks down the stated reasons for migration. The three most popular motivations for urban migration are visiting friends or relatives, schooling/training and employment search, although marriage is also a leading factor in female migration. The former reasons fit well with the temporal pattern of moves. As Figure 2 suggests, most urban migration occurs in December and January, at the close of the calendar school year and when one might move to begin a new course of schooling, to look for a new job, or for an extended holiday with friends or relatives.

Panel B of Table 8 suggests a similar set of broad motivations for international migration. One key difference here is that few women migrate abroad to look for work, and instead most move for marriage. However, the temporal pattern of international migration remains quite similar to that of urban migration, with most moves occurring in January (not shown).

Thus far we have discussed when and why young Kenyan adults move out of their rural homes into urban areas, or to international locations (which are almost entirely rural districts of Uganda). Table 9 presents individual characteristics at the time of survey enumeration for those living in rural versus urban locations. Compared to their counterparts, young adults living in a city are slightly less likely to ever have been married or pregnant, and this effect is largely driven by younger males. While over 25% of young adults living in rural areas are still attending school, this is true for only 14% of individuals who have migrated. In contrast, urban migrants

[12] Many of these stays were censored, i.e., were still ongoing at the time of enumeration, so this is an underestimate.

are much more likely to be in a vocational training program, both men and women alike. Inhabitants of rural areas are apt to run their own business (almost entirely in the informal sector), while those in urban areas are more likely to be employed in formal sector jobs.[13] Unemployment rates are high in both the rural and urban samples, and are similar across age and gender among those living in a city.

4.3 The Kenyan Demographic and Economic Climate

Our study focuses on young adults in Kenya. This age group, composing nearly a quarter of the Kenyan population, is extremely important in shaping both current and future economic outcomes. In order to better understand the migration decisions and labor market activities of these individuals, a brief discussion of the Kenyan demographic and economic setting is useful.

The Kenyan population has increased rapidly since independence, with urban areas experiencing the fastest growth (Republic of Kenya 2002a). Nairobi in particular has grown much faster than any other province, with population increasing by more than 60% each decade. In fact, Nairobi and the Rift Valley province have shown consistent increases in their share of the national population over this period, while shares in other provinces have stagnated or decreased (Republic of Kenya 2001).

This urban population expansion has been fueled in large part by internal migration. Tabulations from the 1999 Kenyan Census suggest that nearly 70% of individuals living in Nairobi at the time of enumeration were born elsewhere, and similarly 57% in Mombasa, 48% in Nakuru, 39% in the district containing Eldoret and 34% in Kisumu. In contrast, only 13% of inhabitants of Busia District (our baseline study district) had migrated there. Further, net migration figures show large influxes of migrants to four of the five main urban areas (Kisumu being the exception), with the numbers of migrants increasing each decade since 1979. Statistics describe a net increase in migrants aged 10-29 for females and males in these four urban centers, the age group we study here (Republic of Kenya 2002b).

The Kenyan economy has also undergone dramatic changes in the post-independence period. Average annual GDP growth was highest in the 1970s, and has slowed since. Indeed, the second half of the 1990s saw shrinking per capita income. Annual GDP growth rates more

[13] Employment in the KLPS-2 is defined as working for pay, volunteering, or interning, and does not include most home agricultural activities.

recently have been extremely volatile, ranging during 1998-2007 from 0.5% to nearly 7% depending on the year (World Bank 2007). In addition, the sectoral composition of national income has shifted considerably. National accounts data (presented in Figure 3) demonstrate a growing importance of the services sector since the late 1970s, now accounting for over half of value-added, while agriculture has waned and industry stagnated. Focusing more specifically on 1998-2007, the share of agriculture in value-added fell from 32% to 23%, and industry's share increased slightly from 18% to 19%, while value-added in services increased sharply from 50% to 58% (World Bank 2007).[14]

One recent survey finds that nearly half of all Kenyans are unable to meet daily minimum food and non-food requirements (World Bank 2008). Consumption growth is quite uneven across Kenyan provinces, and poverty is especially salient in rural areas. Indeed, the Kenya Poverty and Inequality Assessment (2008) finds that mean household consumption grew 24% in urban areas during 1997-2006, while only growing 1.5% in rural areas over the same period. However, it is interesting to note that this same study suggests that poverty rates are lower in households with a migrant. This is perhaps because better-off or more able people migrate, or that migration opens up more opportunities for income creation. We seek to partly disentangle these possibilities in the main analysis that follows (in sections 4.4 and 5).

Despite macroeconomic volatility, the labor force has continued to grow. Census data reveals a nearly seven percentage point increase in the labor force participation rate between 1989 and 1999, with faster growth for females than males.[15] Unsurprising for a country with a high fertility rate, the majority of the labor force remains young, with the largest proportion of individuals between the ages of 20 and 29. Educational attainment among the economically active has also improved dramatically in recent years: the proportion of Kenyan workers with no formal education declined from one-third to one sixth during 1989-1999, though the majority of workers have still attained no more than a primary school education (Republic of Kenya 2002c).

A snapshot of the labor force in a 1998/99 national survey finds more than three-quarters of Kenyans economically active: 66% working and 11% unemployed. Just over half of individuals in the 15-24 age group are labor force participants—38% are employed and 14%

[14] It should be noted that these figures may not fully account for growth in the increasingly important informal sector in Kenya, as enterprises in this sector are generally not officially registered. For a discussion of national accounts source data and the poor quality of data on the informal sector, see IMF (2005).

[15] The following discussion of the labor force focuses on individuals aged 15-64 unless otherwise noted.

unemployed—while many of the inactive individuals are still undoubtedly pursuing their education. More recent figures suggest youth unemployment is now over 20% (World Bank 2008). Labor force participation rates are higher for women than men in this group, but higher for men in older cohorts. Labor market participation rates among individuals aged 15-64 are substantially higher in urban areas than rural ones, as are unemployment rates, with young women having the most severe unemployment. The national data further suggests that unemployed men generally seek paid work in both rural and urban areas, while unemployed women focus their search in urban areas (Republic of Kenya 2003).

Figures from the KLPS-2 provide a similar snapshot for 2007/2008. Among the KLPS population nearly 60% of adolescents are active in the labor force, with more than one-third employed or self-employed, and approximately one-quarter unemployed. Labor force participation is higher in urban areas (76%) than in rural ones (55%), and unemployment is also higher in cities. One key divergence from the national figures is that young adults in the KLPS-2 sample show higher participation rates among men (67%) than women (48%).

According to nationally representative data, small-scale agriculture is the dominant sector of employment in rural areas, while urban workers tend to be employed in the modern (formal) and informal sectors. The 1998/99 Integrated Labour Force Survey (ILFS) reports that 51% of urban employees work in the modern sector, while 39% work in the informal sector. Employment in both sectors has increased in recent years (Republic of Kenya 2005; Republic of Kenya, various years).

Table 10 utilizes the KLPS-2 data to outline the industrial breakdown of working adolescents in urban versus rural locations. Note that agriculture for own use, which is the primary activity for rural individuals, is not included in our definition of employment and hence is left out. Among those working for pay or family gain, or self-employed, most rural inhabitants work in retail or other unclassified industries. In contrast, urban migrants primarily work in manufacturing, domestic service, retail and other service industries. The first and last of these are dominated by male migrants, while female migrants are much more likely to work in retail and domestic service.

Employment questions in the KLPS-2 survey attempt to, but cannot always, distinguish perfectly between formal and informal sector employment. However, it is likely that most of our respondents work in informal sector jobs. Table 10 shows that urban female migrants are most often employed as "house girls" (domestic servants), the quintessential female informal sector job. Furthermore, individuals' employment status presented in Table 11 suggests that most positions are temporary or casual, for rural and urban workers alike, again implying largely informal sector employment. Finally, the types of industries in which most KLPS respondents work (restaurants, domestic service and other service industries) line up closely with employment large in the informal sector (World Bank 2008).

Modern sector real average wages per employee in Kenya have generally increased over the past two decades, with notable exceptions in the early-to-mid 1990s. Between 2000 and 2005, wage growth was fastest in the private sector industries of transport and communications; finance, insurance, real estate and business services; and community, social and personal services. The fastest growing wages in the public sector were in transport and communications, as well as in trade, restaurants and hotels. Although wage growth was slow in some private sector industries over this period—especially in commercial agriculture—public sector wage growth was actually negative in mining and quarrying, and in manufacturing (Republic of Kenya, various years).[16]

The last panel of Table 11 presents figures on average monthly wages from paid employment, generated using the KLPS-2 sample. Cash salaries and in-kind payments taken together are twice as high in urban areas than rural areas. Among those living in a city, remuneration is nearly twice as high for men than for women. Recall that large large shares of KLPS-2 urban women work in generally low-paying domestic service jobs.

This description of the Kenyan demographic and economic climate has highlighted several key differences between urban and rural regions. Migration rates are largest to urban areas, where average wages are much higher and jobs in manufacturing and service sectors are concentrated. There is also evidence that families with migrants tend to have lower poverty rates. We now proceed into our main analysis, examining which individuals migrate and whether such selection can explain the large observed urban-rural wage gap in Kenya.

---

[16] The Kenyan government has outlined a minimum wage policy since Kenyan independence in 1963, and guidelines are adjusted on nearly an annual basis. However, this policy does not apply to formal public sector employment (in which wages are determined by service and periodic performance reviews) or to informal sector employment (due to legal weak enforcement), and thus does not constrain wages for most employees.

4.4 Empirical Evidence on Selection into Urban Migration in Kenya

Table 12 presents the main empirical results on the migration selection analysis. Column (1) displays results using a linear probability model, including one of the two key variables of interest, years of assigned deworming treatment, as well as individual and household control variables. Although the point estimate on deworming is positive and of moderate magnitude, it is not statistically significant at traditional confidence levels. This is true across all specifications the table. It may be that health status is not valued more highly in urban sector jobs than it is on the farm. (We will reevaluate this relationship in future analysis featuring both the Wave 1 and Wave 2 KLPS-2 subsamples.)

The 1998 academic test score is positively and significantly related to subsequent urban migration (column 2), and this holds robustly across specifications in this and ensuing tables. Note that none of the other individual characteristics or proxies for household socioeconomic status are robustly related to migration, with the exception of mother's educational attainment, which is also positively correlated with urban migration. The finding that household assets and other socio-economic characteristics do not predict migration argues weakly against the hypothesis that credit constraints are a major impediment to rural-urban migration in this context. A probit model produces similar results (column 3), and suggests that a two standard deviation increase in academic test score results in 17% increase in the likelihood of rural-urban migration. Disaggregating the 1998 test score measure by subject (English, mathematics, science/ agriculture) did not reveal that a single subject was driving the results (not shown).

Results are robust to the inclusion of additional regression controls. Column (4) includes a measure of individual school attendance in 1998. The size and significance of the main results are unchanged, suggesting that, above and beyond how frequently an individual attended school, cognitive ability has a positive relationship with later migration. Column (5) includes an interaction between the two main ability variables of interest, as well as their interactions with gender and age, but these interaction results are not large in magnitude nor significant.

Figure 4 displays the relationship between the individual test score and migration using a cubic polynomial fit for the full sample (a variety of polynomial controls or nonparametric methods produce visually similar relationships). The strong positive association between test score and migration at higher scores is apparent especially for those with scores greater than one standard deviation above the mean, although we cannot reject a linear relationship. Splitting the sample by gender produces similarly positive relationships (not shown).

Columns 6 and 7 of Table 12 include school fixed effects, and produce similar results, although standard errors increase somewhat, not surprisingly. The school fixed effects might better capture local socio-economic status measures or transport costs not adequately picked up in the earlier regressions, hence this is an important robustness check. Here we focus on the academic test score results; the deworming treatment was randomized at the school level, and so there is not sufficient within-school variation to estimate impacts (any variation comes from differences across initial grade level).[17] The test results support the earlier findings, with an almost identical positive relationship between pupil test score and subsequent rural-urban migration (column 6) and weak interaction effects (column 7).[18]

The results in Table 13 examine the role of schooling attainment in urban migration, and provides an interesting contrast to existing studies of selective migration. We consider the relationship between urban-rural migration and schooling attainment – the almost universal measure of individual ability in the literature – in column 1, and find it to be positive, of moderate magnitude, and highly statistically significant. A three year increase in schooling increases the likelihood of migration by more than 5 percentage points, or roughly 16%. However, the magnitude of this coefficient is cut nearly in half, and loses statistical significance at traditional confidence levels, when controls for parent education are added to the specification (column 2). Mother's education is particularly influential, as in Table 12. When the test control for individual cognitive ability is also included (column 3), we continue to find a strong positive relationship between pre-migration test score and subsequent urban migration, nearly unchanged from Table 12, while the coefficient on schooling attainment falls close to zero. These results provide evidence that cognitive ability is a more decisive determinant of urban migration in the Kenyan context than schooling attainment, and that results of existing studies might be revised if authors had had access to detailed test score data, such as that in the current study.

---

[17] Years assigned deworming treatment is still included as a control in Table 12, columns 6 and 7, nonetheless.

[18] The test score information utilized in the forgoing analysis was only available for individuals present on the day the test was administered. To provide robustness checks on these results, we include additional test score information obtained from a sample of students who had dropped out of school during 1998, but were tracked to their homes and asked to complete the exam. This increases the sample size slightly to 1531 individuals. As before, there is a strong relationship between pupil test scores and subsequent urban-rural migration (not shown).

Table 14 provides results on international migration into Uganda. These results reinforce the earlier descriptive findings of a sharp contrast with the urban migration results. Neither cognitive test scores, nor deworming, nor educational attainment significantly predict international migration in our sample, nor does mother's education (though the latter actually has a small and weak negative relationship). There do appear to be some socioeconomic correlates of international migration but these are inconsistent in sign and difficult to interpret: years of father's schooling is positively linked to migration to Uganda, but those from households with latrines (who tend to be better off households in our setting) are less likely to move.

Overall, there is no evidence that any dimension of ability is related to international migration in our sample. This stands in sharp contrast to the large literature on Mexico-U.S. international migration discussed in the introduction, but of course an important difference between the two settings are the relative living standards in each pair of countries: the U.S. is much wealthier than Mexico, while Kenya and Uganda are at broadly similar levels of economic development. Further the vast majority of international migrants in our sample move just across the border into in the rural districts of eastern Uganda, settings where ex ante few would expect migration to be strongly selected on individual ability.

5. Estimating the Urban Wage Premium in the presence of selective migration

In this section, we use cognitive test score data as an improved measure of individual ability in order to provide more credible estimates of the urban-rural wage gap in Kenya. There is a massive urban wage premium in this setting: conditioning on all of the household and school controls in the previous tables, except for the cognitive test score and schooling attainment, average urban wages in our sample remain twice as large as rural wages (Table 15, column 1). This premium is much larger for men than women in magnitude, at 2648 Kenya shillings per month for men and only 1113 shillings for women (not shown in the table), although the proportional urban wage premium is more similar given men's much higher average earnings.

As expected, both schooling attainment and higher cognitive test score performance are associated with much higher wages, although the test score effect is only marginally statistically significant. A three year increase in schooling is associated with 22% higher wages in the whole sample, while an increase of two standard deviations in the 1998 cognitive test is associated with a roughly 17% wage gain in our sample conditional on other covariates (column 2), and both effects are almost entirely driven by male workers (not shown).

The question we ask is whether the observed urban wage premium continues to hold when these observed ability measures are taken into account, given the strong link between cognitive tests and urban migration documented in Tables 12 and 13. We estimate this in column 3, including controls for the test score, schooling attainment and interactions of each with urban location, to assess whether there are differential returns to skill in urban areas.[19]

We find in column 3 that the large Kenyan urban wage premium is largely robust to including these controls, and running these regressions separately with the two ability measures yields largely similar results (not shown). Both the test score and schooling attainment measures in this table are demeaned, and thus the urban wage premium is 1933.3 Shillings per month. The overall average urban wage premium (in column 1) is 2111.1, which implies that considering observed schooling attainment reduces the urban wage premium by only 8.4%.[20]

Figures 5 and 6 show this graphically. The urban versus rural returns to cognitive test scores and schooling attainment (both conditional on other household and school characteristics) are presented in these two figures, respectively. The relationships are strongly upward sloping, indicating that higher skilled individuals earn higher wages, and there remains a large urban rural wage gap in both cases. Together, Table 15 and Figures 5 and 6 provide evidence that the urban-rural wage gap in our sample of Kenyan youths is largely robust to observed schooling attainment and cognitive test score differences between urban and rural residents, due to large inherent productivity differences across sectors, or perhaps due to some measures of individual ability not well captured in the variables we employ in our analysis (e.g., personality traits), rather than due to migration selection along individual ability.

6. Conclusions and Future Work

We conclude from this analysis that high ability young adults are more likely to migrate out of rural Kenya and into cities, and the magnitude of these effects is quite large. While perhaps not surprising in and of itself, given the number of recent studies that also find positive selection into

[19] This is conceptually related to the Blinder-Oaxaca decomposition.

[20] These are all nominal wage differences. In future work, we will consider urban-rural prices differences and thus real wage differences across sectors. Nonetheless, the main conclusion that ability measures cannot explain the urban wage premium will remain largely unchanged.

migration, our use of a true panel dataset of young adults over a decade and novel measures of ability – including both pre-migration cognitive aptitude and health status – sets this work apart from previous studies. Our ability to exploit exogenous variation in health status induced by randomized assignment to deworming treatment is also a strength. Future work will extend the analysis by considering the KLPS-2 Wave 2 sample, which will roughly double the sample size.

In addition to building on the results of previous selective migration studies, the novel cognitive test score data allows us to make further progress on the classic issue of determining how much of the urban wage premium is due to actual productivity differences rather than selection on unobserved ability. We find that including controls for both schooling attainment and cognitive performance does not appreciably diminish the very large observed urban-rural wage gap observed in our sample of Kenyan youths, in which urban jobs appear to pay roughly twice as much as rural employment. At least in this population, there appear to be very large productivity differences across sectors – perhaps due to agglomeration externalities or other characteristics of the urban environment – beyond what can be explained by selective urban migration.

Our analysis focuses on a population of young adults born in rural areas, and as such not all findings will likely generalize to older workers or those born in urban areas of Kenya. In particular, a study by the World Bank (2008) notes that in general youth unemployment rates are twice those for adults and their wages are much lower. Despite these caveats regarding generalizeability, rural youths remain a key and arguably understudied population, and one which composes a large fraction of the population of many African societies.

Another important issue is whether these findings generalize beyond Kenya. If migration depends on relative returns to skill across sectors, then the extent of technological sophistication in agriculture and the types of urban sector jobs will be critical in determining relative returns to skill. Kenya has relatively unsophisticated agriculture and plentiful formal and informal sector jobs in Nairobi – East Africa's largest city – and such opportunities continue to improve in Kenyan cities given the country's recent economic growth. This is exactly the type of setting in which we would expect to see a great deal of selective urban migration for skilled young adults. It is possible, however, that different patterns would prevail in other countries where cities are smaller and skilled employment opportunities less abundant. We leave this for future research.

7. References

Baird, Sarah, Joan Hamory and Edward Miguel. (2008). "Tracking, Attrition and Data Quality in the Kenyan Life Panel Survey Round 1 (KLPS-1)." U.C. Berkeley Center for International and Development Economics Working Paper C08-151.

Bell, Martin, and Salut Muhidin. (2008). "Cross-National Comparison of Internal Migration". Report prepared for the United Nations Human Development Report Office for the *2009 Human Development Report*.

Borjas, George J. (1987). "Self-Selection and the Earnings of Immigrants." *The American Economic Review,* 77(4), 531-553.

Chiquiar, Daniel, and Gordon Hanson. (2005). "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *Journal of Political Economy*, 113(2), 239-281.

Grogger, Jeffrey and Gordon H. Hanson. (2008). "Income Maximization and the Selection and Sorting of International Migrants", NBER Working Paper #13821.

Hoddinott, John. (1994). "A Model of Migration and Remittances Applied to Western Kenya." *Oxford Economic Papers,* 46(3), 459-476.

Hunt, Jennifer. (2004). "Are migrants more skilled than non-migrants? Repeat, return, and same-employer migrants." *Canadian Journal of Economics*, 37(4), 830-849.

Ibarraran, Pablo, and Darren Lubotsky. (2007). "Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census." In George Borjas (ed.), *Mexican Immigration*. Cambridge: National Bureau of Economic Research, Inc.

International Monetary Fund (2005). "Kenya: Report on the Observance of Standards and Codes (ROSC) – Data Module." IMF Country Report 05/388.

Lanzona, Leonardo A. (1998). "Migration, Self-Selection and Earnings in Philippine Rural Communities." *Journal of Development Economics,* 56, 27-50.

Lucas, Robert E.B. and Oded Stark. (1985). "Motivations to Remit: Evidence from Botswana." *The Journal of Political Economy,* 93(5), 901-918.

McKenzie, David, John Gibson, and Steven Stillman. (2006). "How Important is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration." *IZA Discussion Paper #2087.*

MEASURE DHS STATcompiler. (2007). http://www.measuredhs.com.

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

Miguel, Edward, Sarah Baird and Michael Kremer (2007). "The Long Run Impacts of a School Health Intervention in Kenya", unpublished working paper, U.C. Berkeley.

Orr, Larry, Judith D. Feins, Robin Jacob, Erik Beecroft, Lisa Sanbonmatsu, Lawrence F. Katz, Jeffrey B. Liebman, and Jeffrey R. Kling (2003). "Moving to Opportunity Interim Impacts Evaluation." Washington, DC: U.S. Department of Housing and Urban Development.

Republic of Kenya: Central Bureau of Statistics. (2001). "1999 Population and Housing Census, Volume I: Population Distribution by Administrative Areas and Urban Centers." Government Printer: Nairobi.

Republic of Kenya: Central Bureau of Statistics. (2002a). "1999 Population and Housing Census, Volume III: Analytical Report on Population Dynamics." Government Printer: Nairobi.

Republic of Kenya: Central Bureau of Statistics. (2002b). “1999 Population and Housing Census, Volume VI: Analytical Report on Migration and Urbanization.” Government Printer: Nairobi.

Republic of Kenya: Central Bureau of Statistics. (2002c). “1999 Population and Housing Census, Volume IX: Analytical Report on Labor Force.” Government Printer: Nairobi.

Republic of Kenya: Central Bureau of Statistics. (2003). “Report of the 1998/99 Labor Force Survey.” Government Printer: Nairobi.

Republic of Kenya: Central Bureau of Statistics. (2005). *Statistical Abstract: 2005*. Government Printer: Nairobi.

Republic of Kenya: National Bureau of Statistics. (various years). *Economic Survey*. Nairobi.

Rosenzweig, Mark R. (1988). “Labor Markets in Low-Income Countries.” In H. Chenery and T.N. Srinivasan (eds.), *Handbook of Development Economics, Volume I* (pp. 713-762). New York: Elsevier Science Publishers B.V.

Roy, A.D. (1951). “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers,* 3(2), 135-146.

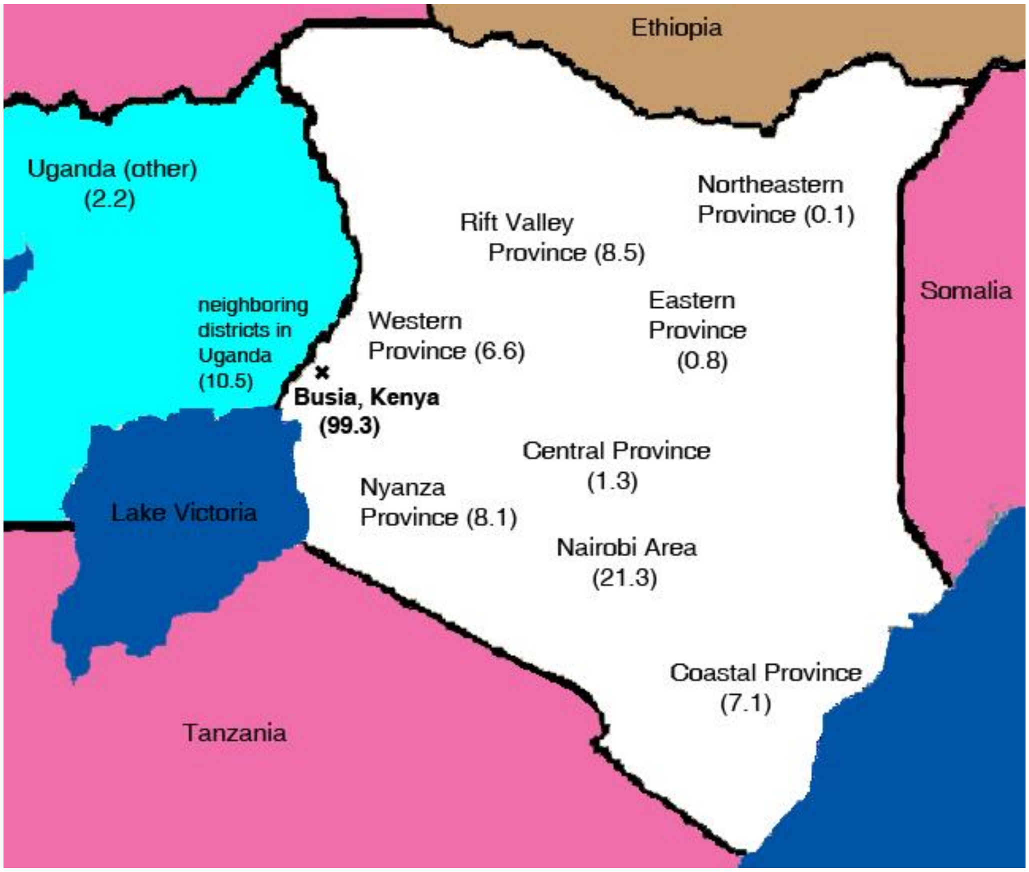
United Nations Statistics Division. (2007). “Detailed Structure and Explanatory Notes: ISIC Revision 2”.

World Bank. (2007). *World Development Indicators*. Washington, DC.

World Bank. (2008). “Kenya Poverty and Inequality Assessment: Executive Summary and Synthesis Report.” Draft Report No. 44190-KE.

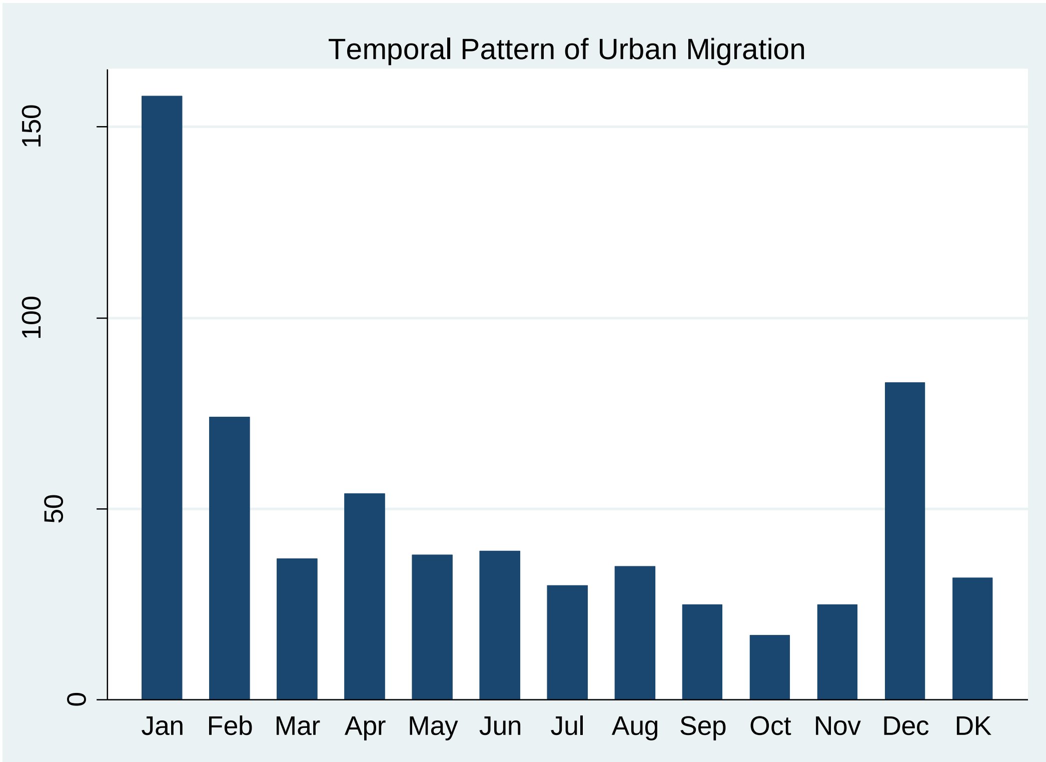
Zhao, Yaohui. (1999). “Labor Migration and Earnings Differences: The Case of Rural China.” *Economic Development and Cultural Change,* 47(4), 767-782.

Figure 1: Locations of residence during 1998-2008



Notes: The sample here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. Values signify percentage of sample that inhabited a given location at some point during 1998-2008. Values will sum to greater than 100, as individuals lived in multiple locations during the survey period. These figures are not weighted to maintain initial population proportions.

Figure 2: Temporal pattern of migration, among urban migrants



Notes: The sample used here includes all individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, who were surveyed and report migration to a city during 1998-2008. Date of migration information is missing for 48 individuals. In addition, one observation with an extreme 1998 ICS test score was dropped from the sample, as well as one observation missing date of survey and five observations missing age information. Figures are not weighted to maintain initial population proportions.

Figure 3: Share of Value-Added in GDP by Sector

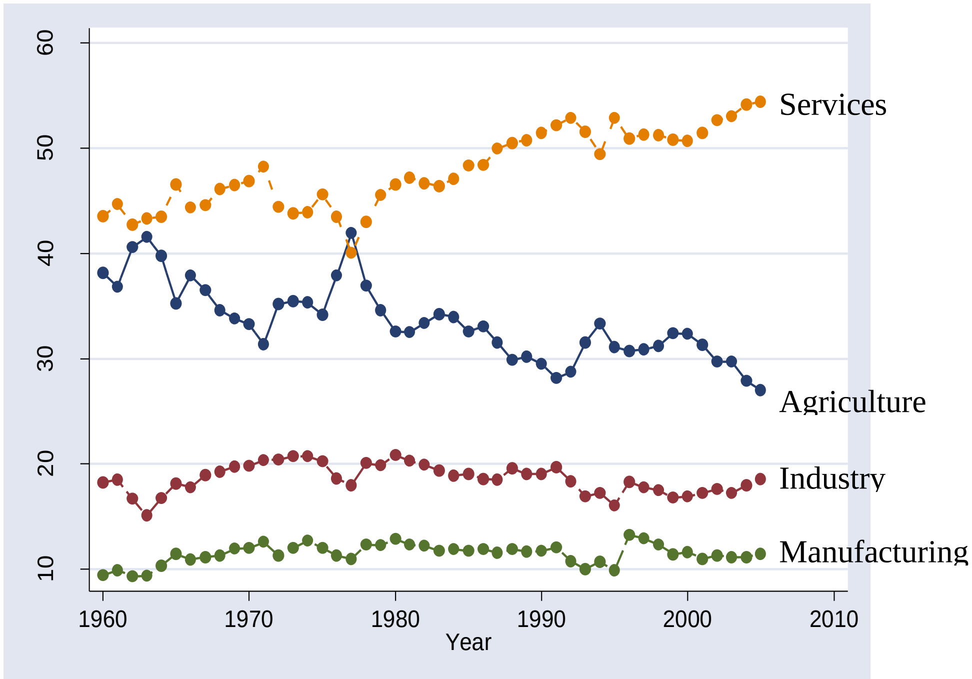
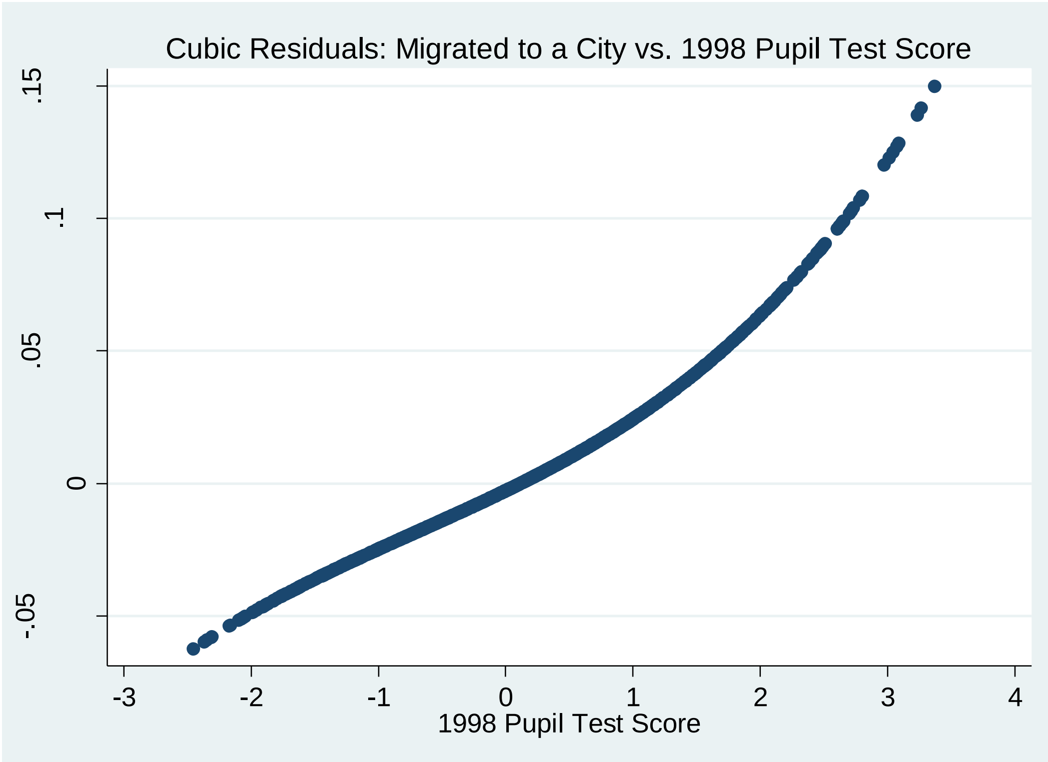
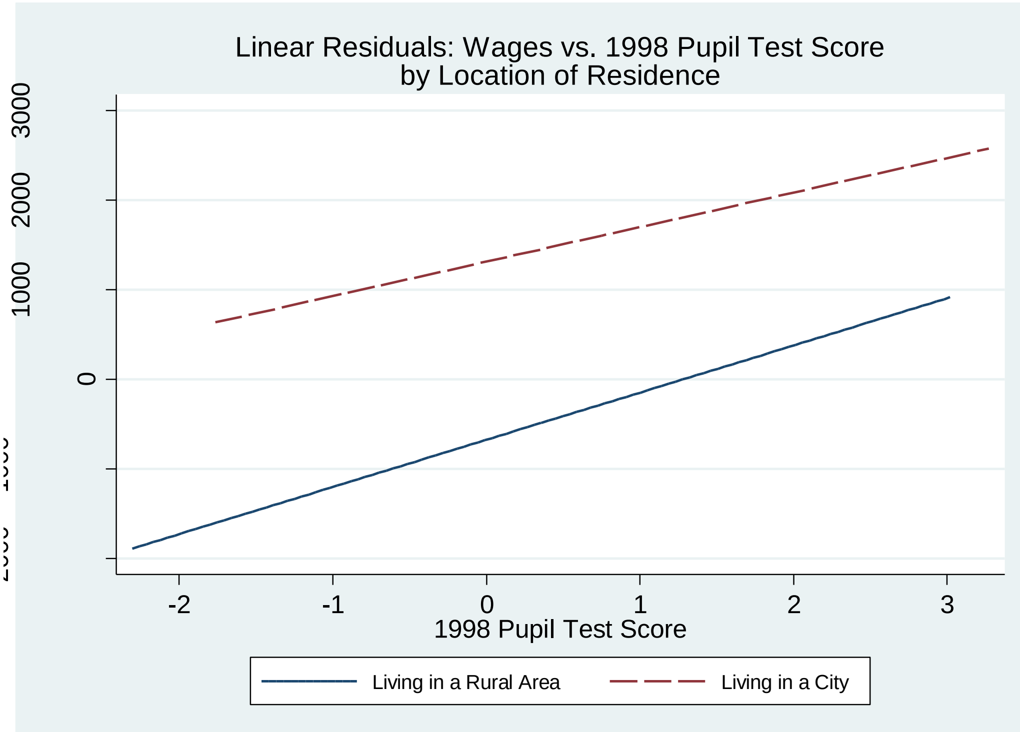


Figure 4: Cubic plot of urban migration on test score



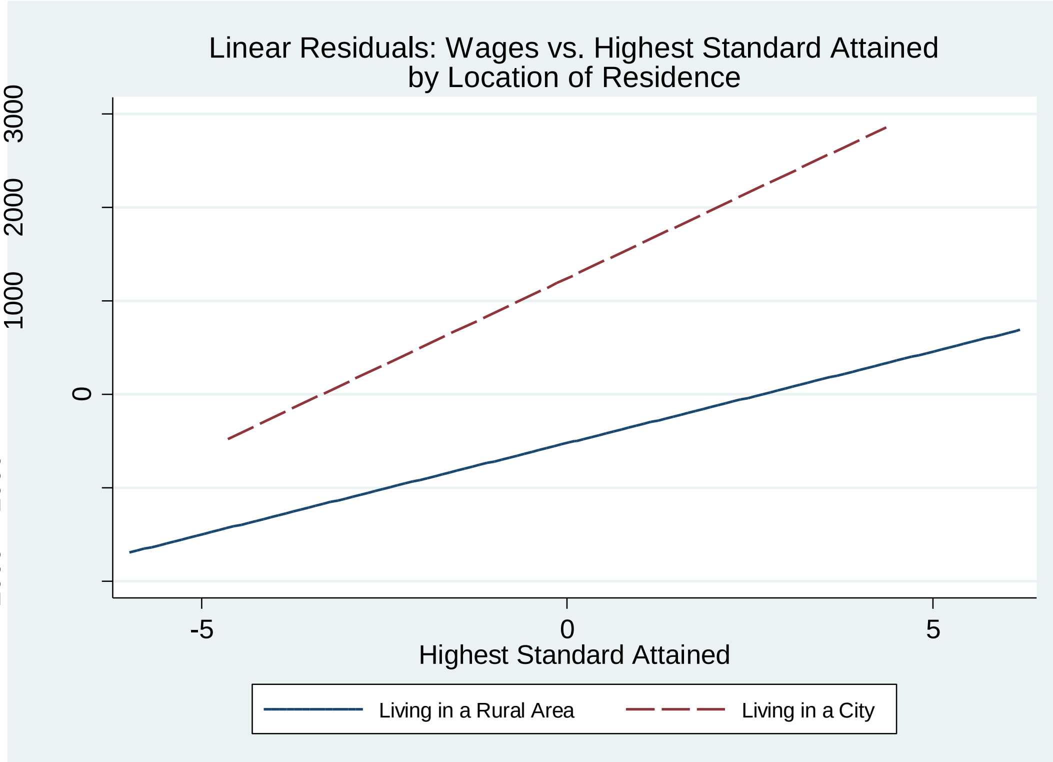
Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as one observation missing date of survey and five observations missing age information. Residuals result from regressions of migration to a city and test score using the specification reported in Table 12, column (2).

Figure 5: Linear residuals fit of wages on test score, by location of residence



<u>Notes:</u> The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, as well as information on wages. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as one observation missing date of survey and five observations missing age information. Wages are measured as cash salary in the last month. Both wages and test score are presented here as residuals from a regression of each on a set of individual and household-level controls.

Figure 6: Linear residuals fit of wages on schooling attainment, by location of residence



Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, ICS test score and schooling attainment data, as well as information on wages. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as one observation missing date of survey and five observations missing age information. Wages are measured as cash salary in the last month. Both wages and schooling attainment are presented here as residuals from a regression of each on a set of individual and household-level controls.

Table 1: Summary statistics on sample attrition and residential location

Means		Treatment Group			Gender	
	All	1	2	3	Female	Male
Panel A: Sample attrition, KLPS-2 I-Module						
Found (effective tracking rate)[a]	0.855	0.853	0.837	0.876	0.844	0.865
Surveyed (effective response rate)	0.820	0.810	0.814	0.838	0.810	0.829
Not surveyed, dead	0.014	0.021	0.010	0.010	0.010	0.017
Not surveyed, refused	0.018	0.019	0.009	0.026	0.023	0.014
Panel B: Residential location information[b]						
Residence in Busia District	0.718	0.725	0.725	0.704	0.709	0.726
Residence in districts neighboring Busia District[c]	0.074	0.082	0.074	0.063	0.094	0.056
Residence outside of Busia and neighboring districts[d]	0.208	0.193	0.201	0.234	0.197	0.218
In Nairobi	0.101	0.090	0.069	0.146	0.099	0.104
In Mombasa	0.037	0.042	0.046	0.021	0.033	0.040
In Nakuru	0.008	0.008	0.010	0.005	0.011	0.005
In Kisumu	0.017	0.020	0.012	0.017	0.014	0.019
Residence outside of Kenya	0.053	0.054	0.053	0.051	0.057	0.049
Number of Observations	1665	588	526	551	826	839

Notes: The sample used here includes all individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, who were surveyed, found deceased, refused participation, found but unable to survey, or not found but searched for during intensive tracking. All figures are weighted in order to maintain initial population proportions.

[a] The proportion found is the combined rates of pupils surveyed, found deceased, refused and found but unable to survey.

[b] Residential location information is available for surveyed individuals only.

[c] Districts neighboring Busia include Siaya, Busia (Uganda), Bugiri (Uganda) and other districts in Kenya's Western Province.

[d] The categories of "Residence outside of Busia and neighboring districts" and "Residence outside of Kenya" are not mutually exclusive.

Table 2: Summary statistics on sample attrition and residential location, by age group

Means	1998 Age				
	6-11	12-13	14-15	16-20	Missing Age
Panel A: Sample attrition, KLPS-2 I-Module					
Found[a]	0.921	0.859	0.849	0.838	0.707
Surveyed	0.878	0.835	0.816	0.771	0.668
Not surveyed, dead	0.009	0.013	0.019	0.020	0.008
Not surveyed, refused	0.032	0.009	0.011	0.034	0.031
Panel B: Residential location information[b]					
Residence in Busia District	0.791	0.763	0.643	0.645	0.680
Residence in districts neighboring Busia District[c]	0.048	0.066	0.093	0.065	0.120
Residence outside of Busia and neighboring districts[d]	0.161	0.172	0.264	0.290	0.200
In Nairobi	0.054	0.100	0.140	0.105	0.083
In Mombasa	0.042	0.023	0.043	0.052	0.039
In Nakuru	0.006	0.002	0.010	0.035	0.000
In Kisumu	0.018	0.002	0.035	0.007	0.010
Residence outside of Kenya	0.035	0.044	0.058	0.083	0.106
Number of Observations	366	530	509	149	111

Notes: The sample used here includes all individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, who were surveyed, found deceased, refused participation, found but unable to survey, or not found but searched for during intensive tracking. All figures are weighted in order to maintain initial population proportions.

[a] The proportion found is the combined rates of pupils surveyed, found deceased, refused and found but unable to survey.

[b] Residential location information is available for surveyed individuals only.

[c] Districts neighboring Busia include Siaya, Busia (Uganda), Bugiri (Uganda) and other districts in Kenya's Western Province.

[d] The categories of "Residence outside of Busia and neighboring districts" and "Residence outside of Kenya" are not mutually exclusive.

Table 3: Impact of deworming and test score on being surveyed

	Dependent Variable: Indicator for Individual Surveyed					
	(1)	(2)	(3)	(4)	(5)	(6)
Pupil test score (1998)		-0.025	-0.026	-0.060	-0.032	-0.068
		[0.015]*	[0.015]*	[0.034]*	[0.020]	[0.036]*
Years assigned deworming	-0.005	-0.007	-0.008	0.000	0.016	0.005
	[0.013]	[0.013]	[0.013]	[0.014]	[0.071]	[0.076]
Pupil test score * Female				0.030		0.024
				[0.030]		[0.031]
Pupil test score * Age				0.021		0.015
				[0.010]**		[0.011]
Pupil test score * Deworming				0.009		0.010
				[0.010]		[0.011]
Deworming * Female				-0.002		-0.007
				[0.023]		[0.023]
Deworming * Age				-0.012		-0.008
				[0.004]***		[0.005]
Age, demeaned (1998)	-0.014	-0.016	-0.014	0.029	-0.013	0.020
	[0.012]	[0.012]	[0.012]	[0.016]*	[0.012]	[0.018]
Falls sick often, self-report (1998)	-0.003	-0.008	-0.011	-0.008	-0.006	-0.007
	[0.029]	[0.029]	[0.029]	[0.027]	[0.031]	[0.030]
Household owns cattle (1998)	0.023	0.023	0.017	0.024	0.039	0.036
	[0.030]	[0.030]	[0.030]	[0.030]	[0.033]	[0.032]
Household has a latrine (1998)	-0.045	-0.044	-0.039	-0.056	-0.063	-0.063
	[0.041]	[0.040]	[0.041]	[0.034]*	[0.046]	[0.040]
Weight, kg (1998)	-0.002	-0.002	-0.002	-0.004	-0.002	-0.003
	[0.003]	[0.003]	[0.003]	[0.003]*	[0.003]	[0.003]
Average school participation, 1998			0.149			
			[0.091]			
Controls for gender and 1998 grade	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1665	1665	1665	1665	1665	1665
R-squared					0.153	0.174
Mean [std dev] of dependent variable	0.820	0.820	0.820	0.820	0.820	0.820
	[0.384]	[0.384]	[0.384]	[0.384]	[0.384]	[0.384]

Notes: Columns (1)-(4) contain probit specifications, with marginal effects evaluated at mean values. Columns (5) and (6) contain linear probability specifications, including school fixed effects. The sample used for all regressions includes individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, who were surveyed, found deceased, refused participation, found but unable to survey, or not found but searched for during intensive tracking. Regressions are weighted in order to maintain initial population proportions, and standard errors are corrected for clustering at the 1998 school level. Robust standard errors in brackets. Test scores are standardized within grade. Years assigned deworming is calculated using treatment group of school and individual's grade in 1998, and is not adjusted for females over the age of 13. Missing age data was replaced with mean values. All specifications include a control for missing age data, and (4) and (6) include interactions between this indicator, deworming and test score. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level. Tests of joint significance for the test score terms in columns (4) and (6) fail to reject the hypothesis that the coefficients are jointly equal to zero. A test of joint significance for the deworming terms in column (4) rejects the hypothesis that the coefficients are jointly equal to zero at the 1% level, and in column (6) at the 5% level.

Table 4: Summary statistics for other variables, subsample with KLPS-2 data

Variable	Mean	Std	# Obs
Female	0.470	0.499	1518
Grade (1998)	4.85	1.41	1518
Age (1998)	12.59	2.28	1518
Years of assigned deworming treatment during 1998-2003	2.97	1.80	1518
Falls sick often, self-report (1998)[a]	1.93	0.51	1518
Weight (kg, 1998)	35.08	8.41	1518
Test score (1998)[b]	0.000	0.987	1518
Highest grade attended	8.85	2.46	1485
Average school participation (1998)	0.924	0.163	1518
Years of mother's education	6.06	4.19	777
Years of father's education	9.70	5.21	727
Household owns cattle (1998)	0.522	0.500	1518
Household has a latrine (1998)	0.794	0.404	1518
Group 1 school	0.388	0.487	1518
Group 2 school	0.297	0.457	1518
Budalangi division school	0.330	0.470	1518

Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. All figures are weighted in order to maintain initial population proportions.

[a] Child falls sick often takes on values of 1 (never/rarely), 2 (sometimes), and 3 (often).

[b] Test score is standardized by 1998 grade.

Table 5: Summary statistics, urban migrants versus non-migrants

	Individuals Who Have Lived in a City	Individuals Who Have Not Lived in a City	Difference
Female	0.523	0.442	0.081** [0.035]
Age (1998)	13.01	12.37	0.64*** [0.15]
Years of assigned deworming treatment during 1998-2003	2.65	3.13	-0.48*** [0.13]
Test score (1998)[a]	0.077	-0.040	0.117* [0.067]
Average school participation (1998)	0.932	0.920	0.012 [0.010]
Highest grade attained at time of survey	9.03	8.75	0.28* [0.16]
Falls sick often, self-report (1998)[b]	1.94	1.92	0.03 [0.04]
Weight (kg, 1998)	37.03	34.06	2.97*** [0.58]
Years of mother's education	6.50	5.75	0.76** [0.39]
Years of father's education	9.99	9.49	0.50 [0.54]
Household owns cattle (1998)	0.498	0.535	-0.037 [0.035]
Household has a latrine (1998)	0.826	0.778	0.048 [0.032]
Number of living siblings[c]	4.53	4.37	0.16 [0.17]
Number of older living siblings[c]	2.28	1.95	0.33** [0.14]
Number of observations	525	993	1518

Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. All figures are weighted in order to maintain initial population proportions.

[a] Test score is standardized by 1998 grade.

[b] Child falls sick often takes on values of 1 (never/rarely), 2 (sometimes), and 3 (often).
[c] Information on siblings is only available at the time of survey.

Table 6: Summary statistics, international migrants versus non-migrants

	Individuals Who Migrated Outside of Kenya	Individuals Who Did Not Migrate Outside of Kenya	Difference
Female	0.422	0.495	-0.056 [0.054]
Age (1998)	13.07	12.51	0.56*** [0.21]
Years of assigned deworming treatment during 1998-2003	2.83	2.99	-0.15 [0.21]
Test score (1998)[a]	-0.047	0.008	-0.056 [0.091]
Average school participation (1998)	0.938	0.921	0.017 [0.013]
Highest grade attained	8.32	8.93	-0.60** [0.30]
Falls sick often, self-report (1998)[b]	2.04	1.91	0.13** [0.05]
Weight (kg, 1998)	35.88	34.95	0.94 [0.75]
Years of mother's education	5.68	6.13	-0.46 [0.48]
Years of father's education	10.53	9.53	1.00 [0.81]
Household owns cattle (1998)	0.543	0.519	0.024 [0.057]
Household has a latrine (1998)	0.682	0.813	-0.131** [0.059]
Number of living siblings[c]	4.01	4.49	-0.48* [0.29]
Number of older living siblings[c]	2.08	2.06	0.03 [0.22]
Number of observations	183	1335	1518

Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. All figures are weighted in order to maintain initial population proportions.

[a] Test score is standardized by 1998 grade.

[b] Child falls sick often takes on values of 1 (never/rarely), 2 (sometimes), and 3 (often).

[c] Information on siblings is only available at the time of survey.

Table 7: Summary statistics on migration history

Means	All	Gender: Female	Male	1998 Age: At/Below Median	Above Median
Panel A: Urban Migration					
Individuals who lived in a city during 1998-2008	0.343	0.381	0.308	0.292	0.410
Among those with information on date of move[a]:					
Number of total moves	2.38	2.67	2.06	2.28	2.47
Number of urban moves	1.30	1.38	1.23	1.21	1.40
Length of stay in urban area (yr)[b]	2.25	2.10	2.41	1.85	2.62
Panel B: International Migration					
Individuals who lived outside Kenya during 1998-2008	0.142	0.128	0.155	0.121	0.172
Among those with information on date of move[c]:					
Number of total moves	2.46	2.42	2.49	2.40	2.51
Number of international moves	1.20	1.15	1.24	1.18	1.22
Length of stay in foreign location (yr)[b]	2.50	3.25	1.92	2.28	2.70
Number of Observations	1518	770	748	864	654

Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. All figures are weighted in order to maintain initial population proportions.

[a] This data exists for 91% of those who report living in an urban area during 1998-2008.

[b] Note that this is an underestimate, as many of these stays are still ongoing.

[c] This data exists for 80% of those who report living outside of Kenya during 1998-2008.

Table 8: Reasons for migration

Means[a]		Gender		1998 Age	
	All	Male	Female	At/Below Median	Above Median
Panel A: Among Urban Migrants					
Schooling/training	0.328	0.435	0.222	0.323	0.332
To look for work	0.279	0.388	0.171	0.195	0.348
To start a new job	0.115	0.100	0.130	0.078	0.145
Marriage	0.062	0.009	0.115	0.048	0.074
Parent/guardian moved	0.018	0.018	0.017	0.014	0.020
Return to permanent home	0.005	0.000	0.011	0.000	0.010
Just visiting	0.355	0.231	0.478	0.408	0.312
Other	0.093	0.093	0.092	0.039	0.137
Number of observations[b]	434	218	216	192	242
Panel B: Among International Migrants					
Schooling/training	0.372	0.353	0.396	0.421	0.335
To look for work	0.249	0.419	0.029	0.103	0.359
To start a new job	0.086	0.153	0.000	0.091	0.083
Marriage	0.177	0.000	0.405	0.199	0.160
Parent/guardian moved	0.008	0.000	0.019	0.012	0.006
Return to permanent home	0.026	0.024	0.029	0.041	0.015
Just visiting	0.132	0.104	0.167	0.214	0.069
Other	0.066	0.015	0.131	0.069	0.063
Number of observations[c]	140	68	72	67	73

Notes: The sample used here includes all individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, who were surveyed and report migration to a city during 1998-2008 (panel A) or to a foreign country (panel B). One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. All figures are weighted in order to maintain initial population proportions.

[a] It is possible for respondents to move multiple times, and to have multiple reasons for each move. An indicator was thus generated to take on a value of 1 if the person migrated for a

given reason, and a zero if they did not migrate for that reason. Thus, proportions likely sum to greater than one.

[b] Information on reasons for migration is missing for 91 of the 525 individuals reporting living in a city since 1998. Statistics presented here are fractions of the non-missing information.

[c] Information on reasons for migration is missing for 43 of the 183 individuals reporting living outside of Kenya since 1998. Statistics presented here are fractions of the non-missing information.

Table 9: Activities of individuals at time of enumeration, by urban migration status

Means			Among those living in a city:			
			Gender		1998 Age	
	Living in a Rural Area	Living in a City	Male	Female	At/Below Median	Above Median
Ever been married	0.395	0.301	0.274	0.335	0.162	0.418
Ever been pregnant[a]	0.499	0.438	0.381	0.503	0.238	0.605
In school[b]	0.257	0.141	0.163	0.117	0.195	0.097
In vocational training	0.232	0.346	0.34	0.353	0.321	0.367
Working, self-employed[c]	0.153	0.076	0.082	0.068	0.043	0.103
Working, not self-employed	0.174	0.389	0.488	0.276	0.302	0.463
Unemployed[d]	0.234	0.300	0.292	0.309	0.306	0.294
Number of Observations	1274	244	127	117	106	138

<u>Notes:</u> The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. Figures are weighted to maintain initial population proportions.

[a] For males, indicates "a partner has ever been pregnant with your child".

[b] Respondent attended school at some time during year of survey enumeration.

[c] Individuals who farm for themselves are not included among those who are self-employed.

[d] Note that the working and unemployment categories do not add up to one, as the remainder of individuals are out of the labor force (which in our definition includes those engaged in agricultural activities for the home).

Table 10: Distribution of working persons by industry, by urban migration status

Means	Living in a Rural Area	Living in a City	Among those living in a city: Gender Male	Female	1998 Age At/Below Median	Above Median
Manufacturing	0.046	0.117	0.139	0.073	0.083	0.134
Trade contractors	0.078	0.043	0.065	0.000	0.034	0.048
Wholesale trade	0.052	0.042	0.044	0.037	0.088	0.018
Retail	0.281	0.131	0.112	0.167	0.218	0.087
Restaurants, cafes, etc.	0.028	0.096	0.103	0.082	0.034	0.127
Domestic Service	0.022	0.143	0.008	0.409	0.205	0.112
Government Services	0.036	0.075	0.070	0.083	0.121	0.051
Passenger transport	0.035	0.014	0.021	0.000	0.041	0.000
Medical, dental and health services	0.015	0.000	0.000	0.000	0.000	0.000
Other services	0.070	0.182	0.233	0.082	0.091	0.229
Other	0.337	0.158	0.204	0.067	0.085	0.195
Number of observations	379	104	66	38	34	70

Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, who were either self-employed or employed by someone else at the time of survey. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. Figures are weighted to maintain initial population proportions.

Table 11: Summary of employment characteristics, by urban migration status

Means	Living in a Rural Area	Living in a City	Among those living in a city: Gender: Male	Female	1998 Age: At/Below Median	Above Median
Employment Status						
Permanent	0.078	0.089	0.133	0.000	0.135	0.063
Temporary	0.227	0.386	0.384	0.391	0.059	0.568
Casual	0.597	0.465	0.406	0.584	0.680	0.345
Unpaid	0.091	0.052	0.078	0.000	0.103	0.024
Working Pattern						
Full time	0.601	0.815	0.747	0.949	0.913	0.759
Part time	0.297	0.147	0.196	0.051	0.087	0.181
Seasonal	0.102	0.038	0.057	0.000	0.000	0.060
Earnings (Ksh)[a]						
Cash salary[b]	2601	5005	5995	3058	3893	5636
In kind	161	326	284	413	227	172
Benefits/allowances	154	179	213	108	182	178
Number of observations	172	88	59	29	29	58

Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, who were either self-employed or employed by someone else at the time of survey. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. Figures are weighted to maintain initial population proportions. Between August 2007 and October 2008, the average exchange rate was 0.0154.

[a] Earnings data is only available for individuals employed by a person or business.

[b] Defined as cash salary in the previous month.

Table 12: Impact of deworming treatment and test score on urban migration

	Dependent Variable: Indicator for Ever Moved to a City						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pupil test score (1998)		0.028	0.029	0.029	0.057	0.029	0.056
		[0.011]**	[0.011]**	[0.011]**	[0.035]	[0.014]**	[0.037]
Years assigned deworming	0.015	0.017	0.018	0.019	0.007	0.046	0.08
	[0.014]	[0.014]	[0.015]	[0.015]	[0.020]	[0.083]	[0.084]
Pupil test score * Female					0.003		0.011
					[0.034]		[0.033]
Pupil test score * Age at tracking					-0.001		-0.003
					[0.007]		[0.007]
Pupil test score * Deworming					-0.011		-0.011
					[0.009]		[0.009]
Deworming * Female					0.025		0.012
					[0.027]		[0.028]
Deworming * Age at tracking					0.008		0.009
					[0.005]		[0.005]*
Age at tracking, demeaned	0.005	0.006	0.007	0.007	-0.016	0.004	-0.022
	[0.009]	[0.009]	[0.010]	[0.010]	[0.017]	[0.009]	[0.016]
Years of mother's education	0.011	0.011	0.011	0.011	0.011	0.011	0.011
	[0.005]**	[0.005]**	[0.005]**	[0.005]**	[0.005]**	[0.005]**	[0.005]**
Years of father's education	0.000	-0.001	-0.001	-0.001	-0.002	-0.001	-0.002
	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]
Falls sick often, self-report (1998)	0.037	0.039	0.04	0.04	0.039	0.041	0.037
	[0.030]	[0.030]	[0.032]	[0.032]	[0.032]	[0.032]	[0.033]
Household owns cattle (1998)	-0.014	-0.013	-0.014	-0.013	-0.016	-0.008	-0.01
	[0.034]	[0.034]	[0.036]	[0.036]	[0.036]	[0.037]	[0.037]
Household has a latrine (1998)	0.056	0.056	0.065	0.064	0.073	0.027	0.034
	[0.037]	[0.036]	[0.039]*	[0.039]*	[0.038]*	[0.038]	[0.038]
Weight, kg (1998)	0.002	0.002	0.002	0.002	0.002	0.003	0.003
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Average school participation, 1998				-0.027			
				[0.099]			
Controls for gender and 1998 grade	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1518	1518	1518	1518	1518	1518	1518
R-squared	0.099	0.103				0.174	0.179
Mean [std dev] of dependent variable	0.343	0.343	0.343	0.343	0.343	0.343	0.343
	[0.475]	[0.475]	[0.475]	[0.475]	[0.475]	[0.475]	[0.475]

Notes: Columns (1), (2), (6) and (7) contain linear probability model specifications, with (6) and (7) also including school fixed effects. Columns (3)-(5) contain probit specifications, with marginal effects evaluated at mean values. The sample employed in all regressions includes surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. Regressions are weighted in order to maintain initial population proportions, and standard errors are corrected for clustering at the 1998 school level. Robust standard errors in brackets. Test scores are standardized within grade. Years assigned deworming is calculated using treatment group of school and individual's standard in 1998, and is not adjusted for females over the age of 13. Missing parent education data is replaced with the mean, and all specifications include a control for missing parent education data. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level. Tests of joint significance for years assigned deworming and its interactions in column (4) reject the hypothesis that the coefficients are jointly equal to zero at the 5% level. These same tests in column (7), as well as tests of joint significance for pupil test score and its interactions, cannot reject this hypothesis.

Table 13: Impact of test score and educational attainment on urban migration

	Dependent Variable: Ever Moved to a City			
	(1)	(2)	(3)	(4)
Pupil test score (1998)			0.027	0.027
			[0.012]**	[0.012]**
Highest grade attended	0.018	0.010	0.006	0.006
	[0.006]***	[0.006]	[0.007]	[0.007]
Years of mother's education		0.011	0.011	0.011
		[0.005]**	[0.005]**	[0.005]**
Years of father's education		0.000	0.000	0.000
		[0.004]	[0.004]	[0.004]
Age at tracking, demeaned	0.022	0.010	0.010	0.010
	[0.010]**	[0.010]	[0.010]	[0.010]
Falls sick often, self-report (1998)	0.028	0.035	0.036	0.036
	[0.032]	[0.033]	[0.032]	[0.032]
Household owns cattle (1998)	-0.039	-0.026	-0.022	-0.022
	[0.037]	[0.038]	[0.038]	[0.038]
Household has a latrine (1998)	0.06	0.063	0.066	0.066
	[0.042]	[0.040]	[0.040]*	[0.040]*
Weight, kg (1998)	0.001	0.002	0.002	0.002
	[0.002]	[0.002]	[0.002]	[0.002]
Average school participation, 1998				-0.003
				[0.093]
Controls for gender and 1998 grade	Yes	Yes	Yes	Yes
Controls for years assigned deworming	No	Yes	Yes	Yes
Number of observations	1485	1485	1485	1485
Mean [std dev] of dependent variable	0.346	0.346	0.346	0.346
	[0.476]	[0.476]	[0.476]	[0.476]

Notes: This table displays probit specifications, with marginal effects evaluated at mean values. The sample employed in all regressions includes surveyed individuals with 1998 Pupil Questionnaire, school participation, ICS test score, and school attainment information. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. Regressions are weighted in order to maintain initial population proportions, and standard errors are corrected for clustering at the 1998 school level. Robust standard errors in brackets. Test scores are standardized within grade. Missing parent education data is replaced with the mean. All specifications include a control for missing parent education data. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 14: Impact of deworming treatment and test score on international migration

	Dependent Variable: Indicator for Ever Lived Outside of Kenya						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pupil test score (1998)		-0.006	-0.007	-0.007	-0.015	-0.002	-0.022
		[0.013]	[0.013]	[0.013]	[0.024]	[0.016]	[0.025]
Years assigned deworming	0.012	0.012	0.011	0.011	0.014	-0.024	-0.013
	[0.010]	[0.010]	[0.009]	[0.009]	[0.011]	[0.059]	[0.062]
Pupil test score * Female					0.011		0.028
					[0.022]		[0.024]
Pupil test score * Age at tracking					0.001		0.003
					[0.005]		[0.005]
Pupil test score * Deworming					0.001		0.003
					[0.007]		[0.008]
Deworming * Female					-0.008		-0.015
					[0.018]		[0.020]
Deworming * Age at tracking					0.004		0.002
					[0.003]		[0.003]
Age at tracking, demeaned	0.010	0.010	0.010	0.010	0.000	0.008	0.004
	[0.008]	[0.008]	[0.007]	[0.007]	[0.010]	[0.008]	[0.011]
Highest grade attended	-0.009	-0.008	-0.008	-0.008	-0.007	-0.008	-0.008
	[0.007]	[0.007]	[0.006]	[0.006]	[0.006]	[0.008]	[0.008]
Years of mother's education	-0.006	-0.006	-0.006	-0.006	-0.006	-0.008	-0.008
	[0.004]	[0.004]	[0.003]*	[0.003]*	[0.003]*	[0.004]*	[0.004]*
Years of father's education	0.009	0.009	0.008	0.008	0.007	0.009	0.009
	[0.005]*	[0.005]*	[0.004]**	[0.004]**	[0.004]**	[0.005]*	[0.005]*
Falls sick often, self-report (1998)	0.057	0.057	0.058	0.058	0.057	0.039	0.039
	[0.019]***	[0.019]***	[0.018]***	[0.018]***	[0.018]***	[0.018]**	[0.019]**
Household owns cattle (1998)	0.034	0.033	0.035	0.035	0.035	0.024	0.023
	[0.027]	[0.026]	[0.024]	[0.024]	[0.024]	[0.026]	[0.027]
Household has a latrine (1998)	-0.079	-0.079	-0.073	-0.073	-0.068	-0.018	-0.018
	[0.048]	[0.048]	[0.042]*	[0.042]*	[0.040]*	[0.048]	[0.047]
Weight, kg (1998)	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Average school participation, 1998				-0.015			
				[0.077]			

Controls for gender and 1998 grade	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1485	1485	1485	1485	1485	1485	1485
R-squared	0.057	0.057				0.156	0.159
Mean [std dev] of	0.129	0.129	0.129	0.129	0.129	0.129	0.129
dependent variable	[0.335]	[0.335]	[0.335]	[0.335]	[0.335]	[0.335]	[0.335]

Notes: Columns (1), (2), (6) and (7) contain linear probability model specifications, with (6) and (7) also including school fixed effects. Columns (3)-(5) contain probit specifications, with marginal effects evaluated at mean values. The sample employed in all regressions includes surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. Regressions are weighted in order to maintain initial population proportions, and standard errors are corrected for clustering at the 1998 school level. Robust standard errors in brackets. Test scores are standardized within grade. Years assigned deworming is calculated using treatment group of school and individual's standard in 1998, and is not adjusted for females over the age of 13. Missing parent education data is replaced with the mean, and all specifications include a control for missing parent education data. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level. Tests of joint significance for years assigned deworming, pupil test score and their interactions in columns (4) and (7) fail to reject the hypothesis that the coefficients are jointly equal to zero.

Table 15: Estimation of the selection-corrected urban-rural wage gap in Kenya

	Dependent Variable: Wages (residual)		
	(1)	(2)	(3)
Indicator for residence in a city (residual)	2111.1***		1933.0***
	(441.4)		(410.7)
1998 Test Score (residual)		289.8	381.7
		(188.4)	(259.8)
Highest grade attended (residual)		253.0**	138.0
		(113.3)	(124.6)
City (residual) * 1998 Test Score (residual)			-285.4
			(409.7)
City (residual) * Highest grade attended (residual)			233.6
			(231.5)
Number of observations	254	248	248
R-squared	0.094	0.052	0.136
Mean [std dev] of dependent variable	3381.98 [3481.83]	3381.98 [3481.83]	3381.98 [3481.83]

Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, as well as information on wages. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as one observation missing date of survey and five observations missing age information. Wages are measured as cash salary in the last month. All variables presented here are residuals from a regression of each on the set of individual and household-level controls in Table 12.

Appendix Table 1: Existing literature on selective migration, and comparison to current study

Study	Country(ies)	Empirical ability measure	Relationship with migration
Chiquiar and Hanson (2005)	Mexico to U.S.	Schooling attainment	Positive
Grogger and Hanson (2007)	Cross country analysis	Schooling attainment	Positive
Hoddinott (1994)	Kenya (urban)	Schooling attainment	Positive
Hunt (2004)	Germany (urban)	Schooling attainment	Positive
Ibarraran and Lubotsky (2007)	Mexico to U.S.	Schooling attainment	Negative
Lanzona (1998)	Philippines (urban)	Schooling attainment	Positive
McKenzie et al (2006)	Tonga to New Zealand	Schooling attainment	Positive
Zhao (1999)	China (urban)	Schooling attainment	None
Current study:			
Hamory and Miguel (2009)	Kenya (urban)	Schooling attainment	Positive; but none conditional on a cognitive test score
		Cognitive tests	Positive
		Health status	None / weakly positive
	Kenya to Uganda	Schooling attainment	None
		Cognitive tests	None
		Health status	None