Title
Essays in Development Economics

Author
Hicks, Joan Hamory

Publication Date
2009

Peer reviewed|Thesis/dissertation

Essays in Development Economics

by

Joan Hamory Hicks

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Edward Miguel, Chair
Professor Enrico Moretti
Professor David I. Levine

Fall 2009

Table of Contents

Chapter 1
Income Shocks, School Attendance and Child Ability

Chapter 2
Individual Ability and Selection into Migration in Kenya

Chapter 3
Civil Unrest, Social Cohesion and Informal Credit in Rural Kenya

Dedication

This dissertation is dedicated to my parents, Bruce and Ann Hamory, who instilled in me a deep appreciation for learning, encouraged me to take advantage of every available opportunity, and, through their own accomplishments, inspired me to pursue this degree. Without their unwavering love and support, this work would not have been possible.

Acknowledgments

I am forever indebted to a number of individuals for their roles in the creation of this dissertation. Edward Miguel, my committee chair, provided me with gainful employment for much of my time in graduate school, sent me to Kenya to acquire invaluable field experience, granted me access to his data for numerous solo and collaborative research projects, and offered helpful feedback and advice at every stage of my graduate career. Margaret McMillan, my undergraduate thesis advisor, guided me through my first efforts in rigorous economic research, offered me two incredible research assistant opportunities straight out of college, and helped me to navigate the graduate school application process (with great success!). Without the mentorship and guidance of both Ted and Maggie, I would not be where I am today.

I am eternally grateful to my parents, Bruce and Ann Hamory, who have worked hard throughout my life to afford me every possible opportunity, and set no limits on my pursuits – academic or otherwise. My parents – as well as my brother, Matthew, and sister-in-law, Beth – have graciously tolerated many "working holidays" and sponsored numerous trips back to the East Coast for family visits. I also thank my dear husband, Daniel Hicks, who read and re-read drafts of papers, spent countless hours discussing research ideas, and provided me with constant encouragement throughout graduate school.

The research presented in various chapters of this dissertation improved significantly through conversations with Daniel Hicks, David Levine, Edward Miguel, Enrico Moretti, Alex Rothenberg, and Elisabeth Sadoulet. I am further grateful for feedback from participants of the UC Berkeley Development Lunch.

Finally, I would like to acknowledge the incredible financial support I have received along the way. Specifically, I am grateful for the National Science Foundation Graduate Research Fellowship, which funded the first three years of my graduate education, and to my parents, who have financed my studies for more than two decades.

Abstract

Essays in Development Economics

by

Joan Hamory Hicks

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Edward Miguel, Chair

According to the World Bank's *World Development Report 2007*, there are 1.3 billion young adults aged 12-24 living in less-developed countries today. Individuals in this age group are going through a period of tremendous flux in their lives as they embark on marriage, job searches or higher education, and their experiences during this time will shape the next generation of decision makers. Research focusing on the choices of these individuals, as well as the circumstances under which they are made, is urgently needed. The present collection of essays seeks to advance such research by utilizing a recent longitudinal survey to examine the decisions of young adults in rural Kenya as they relate to education, migration and behavior in the wake of violent civil conflict.

Chapter 1 explores the extent to which individual academic and cognitive ability is factored into household decisions concerning education. Panel information on schooling for nearly 1,900 rural Kenyan youth over the period 1998-2008 is combined with satellite precipitation data in order to examine the effects of agricultural income variability on school attendance. A unique early-age academic test score proxies for child ability. Regression analysis indicates that during times of plenty, there is an 11 percent increase in attendance of high ability relative to low ability individuals, suggesting households recognize and value ability when making schooling decisions. This finding is framed using a model of human capital accumulation in which schooling decisions are a function of individual ability. Surprisingly, although youth on the whole are less likely to attend school during negative income shocks, there are no differential attendance changes across individuals of different ability levels. Instead, credit constraints and income shocks may work together in this setting to limit desirable human capital investments. Such consumption smoothing behavior could imply negative long-term effects on household welfare.

Chapter 2 studies selective migration among 1,500 Kenyan youth originally living in rural areas. In particular, this essay examines whether migration rates are related to individual "ability", broadly defined to include cognitive aptitude as well as health, and then uses 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, the present research employs 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. Results suggest 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%. In an interesting contrast with the existing literature, schooling attainment is not significantly associated with urban migration once cognitive ability is accounted for. 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.

Finally, Chapter 3 examines the socioeconomic impacts of two months of protests and violent, primarily ethnic-based clashes that erupted across central and western Kenya in late 2007 following the controversial conclusion of a heavily-contested presidential election. Although not an epicenter of the conflict, Busia District experienced sporadic unrest, an influx of refugees from other parts of Kenya, inflation, supply shortages, and local market closures. Unique and timely survey data collected from young adults living primarily in this district of rural western Kenya in the months surrounding the election permits the use of both differences-in-differences and propensity score matching methodologies to estimate the short- to medium-run impacts of this conflict, and both approaches yield broadly similar findings. Despite little support for lasting effects on labor, migration and nutritional outcomes within weeks of cessation of the violence, there do appear to be persistent consequences for social cohesion and informal financial activities. While there is little indication of change in survey respondents' self-reported attitudes regarding trust of others, analysis confirms large declines in attendance at religious services, participation in community and bible groups, and utilization of non-family members as points of contact for future survey enumeration efforts. These findings highlight a disconnect between *reported* attitudes and *observable* behavior. Furthermore, respondents are between 29 and 53 percent less likely to engage in informal lending and transfers post-conflict. Given the key role played by social networks in informal financial markets in less-developed countries, these results indicate that even brief civil unrest may have lasting negative consequences.

Chapter 1

Income Shocks, School Attendance and Child Ability

Chapter 1 Summary

This chapter explores the extent to which individual academic and cognitive ability is factored into household decisions concerning education in rural Kenya. Panel information on schooling for nearly 1,900 youth over the period 1998-2008 is combined with satellite precipitation data in order to examine the effects of agricultural income variability on school attendance. A unique early-age academic test score proxies for child ability. Regression analysis indicates that during times of plenty, there is an 11 percent increase in attendance of high ability relative to low ability individuals, suggesting households recognize and value ability when making schooling decisions. This finding is framed using a model of human capital accumulation in which schooling decisions are a function of individual ability. Surprisingly, although youth on the whole are less likely to attend school during negative income shocks, there are no differential attendance changes across individuals of different ability levels. Instead, credit constraints and income shocks may work together in this setting to limit desirable human capital investments. Such consumption smoothing behavior could imply negative long-term effects on household welfare.

1. Introduction

Small-holder agriculture continues to be the primary occupation and means of survival for rural populations in less-developed countries today. Limited access to credit and insurance in these regions tightens the linkage between agricultural output and livelihood, and as a consequence crop loss can have a devastating effect on household income. A growing literature seeks to examine the various strategies employed by households to smooth consumption over income shocks in this setting, and to evaluate the effects of these practices on household welfare.

One branch of this literature focuses on household coping strategies involving children, such as intensification of child labor or reduction in investments in child human capital. Findings concentrating on the latter largely conclude that negative income shocks increase the likelihood of school dropout among youth. However, little is known regarding the characteristics of children affected most, or how this practice impacts longer-term individual and household outcomes. A recent study suggests that there may be persistence in school withdrawal (de Janvry *et al.*, 2006). This finding raises many concerns, as children who receive less schooling are likely to have lower earnings potential later in life, thereby perpetuating household poverty. At an aggregate level, such a cycle could sustain the high level of societal inequality that plagues many less-developed countries (Behrman and Knowles, 1999).

The present analysis expands on the existing literature by exploring the extent to which individual academic and cognitive ability is factored into schooling decisions among young adults in rural Kenya. Panel information on schooling for nearly 1,900 youth over the period 1998-2008 is combined with satellite precipitation data in order to examine the effects of agricultural income variability on school attendance, and a unique early-age academic test score is utilized as a proxy for child ability.

Among Kenyan youth, delayed primary school enrollment, grade repetition and dropout upon completion of the primary degree are common. Advancement to secondary school requires both satisfactory scores on a standardized examination and competition with peers for limited placements. As a result, individuals of higher academic and cognitive ability are more likely to be able to proceed further through school. If returns to education are increasing in ability, then decisions made to compromise the schooling of the most promising students during childhood may be especially detrimental to later individual and household earnings potential. This suggests that the long-term welfare effects of smoothing techniques involving school withdrawal in response to income shocks may be large.

Results from regression analysis indicate that during periods of positive income shocks (i.e., higher than average, but not flood-level rainfall) attendance increases more for high ability individuals than their low ability counterparts. This finding is consistent with a model of human capital accumulation in which schooling decisions are a function of individual ability, and both schooling attainment and ability affect human capital accumulation. As household income increases, households are able to send more children to school, and choose those of highest academic ability first in anticipation of later-life earnings of these individuals.

Results of this analysis further suggest that while during periods of negative income shocks (i.e., drought) rural Kenyan youth on the whole are less likely to attend school, surprisingly there are no differences in school attendance according to level of individual ability. Lack of access to credit provides one possible explanation for this finding. Because households in this region are severely credit constrained, it could be that children are being pulled out of school because households lack the resources to send them, and in the absence of sufficient credit, there is no room for discrimination based on ability level. This explanation is supported

by self-reported reasons for dropout among individuals in the sample. Given the issues presented above, that there may be persistence in school withdrawal and that higher ability children advance more easily through school than their lower ability counterparts, such consumption smoothing behavior could imply negative long-term effects on household welfare.

This paper proceeds as follows. Section 2 summarizes previous work exploring income shocks and schooling decisions. Section 3 presents a model of human capital investment, outlining how child ability may play a role in the household decision-making process. Section 4 describes the data employed, and Section 5 describes rainfall, agricultural practices and schooling in the study area. Section 6 presents empirical results, and the final section concludes.

2. Related Literature on Income Shocks and Schooling Decisions

A large literature has advanced over time detailing the determinants of schooling in less-developed countries, and household income is generally thought to play an important role. However, until recently little work has explored the dynamics of this relationship. In regions where credit and insurance markets are thin, as they typically are in rural areas of less-developed countries, households may face a tradeoff between current consumption and human capital investment (and thus future income prospects). This is likely to be especially true for transitory, aggregate (non-idiosyncratic) shocks, which affect many households in a given area simultaneously.[1]

Jacoby and Skoufias (1997) explore this tradeoff between smooth consumption and child human capital investment in India.[2] The authors employ panel data from the late 1970s to examine the relationship between locally aggregate agricultural income shocks, instrumented using deviations in rainfall from historical averages, and school attendance for nearly 260 children in six rural Indian villages. Even though human capital investment has positive returns in both agricultural and nonagricultural labor in this setting, the authors find that school attendance among youth falls in response to shocks.

Jensen (2000) finds a similar result for weather-related income shocks and school enrollment among over 1,500 households in Côte d'Ivoire in the late 1980s – regions experiencing rainfall greater than one absolute standard deviation from historical average see school enrollment decrease by more than one-third, with no difference across gender. Likewise, using panel data for nearly 800 households in the Kagera region of Tanzania, Beegle *et al.* (2006) conclude that transitory agricultural income shocks, measured using an indicator for crop loss due to pests or other calamity, decrease school enrollment.[3] In contrast, Guarcello *et al.* (forthcoming) explore responses of both child labor and schooling to shocks using a recent

[1] A related (and somewhat larger) literature explores the use of child labor in response to transitory income shocks. It has not been established, though, that changes in the intensity of child labor affect schooling outcomes. Ravallion and Wodon (2000) point out that it is common for children in developing countries to both attend school and take part in household chores or farm work, and there is not necessarily a one-to-one tradeoff in time spent between the two. In fact, it could be that hours devoted to both are so low that an increase in labor does not affect schooling at all (Beegle *et al.,* 2006). Indeed, Guarcello *et al.* (forthcoming) find that shock-induced increases in child labor do not necessarily imply corresponding reductions in school attendance. For this reason, the literature reviewed here focuses solely on studies of child schooling responses to shocks.

[2] The setting of the present study is rural western Kenya, a region in which livelihood is dominated by small-holder subsistence agriculture. For that reason, this literature review explores work examining similar circumstances, and ignores research related to other types of income shocks in more developed economies.

[3] The authors further find that this effect is almost entirely mediated by household asset holdings.

survey of Guatemalan households, and observe that non-idiosyncratic shocks (such as earthquakes, floods and fires) substantially increase the likelihood of performing both activities.[4]

The research described thus far has focused on the impact of income shocks on children's school enrollment or attendance in general, but has not explored whether certain types of children are more affected than others. Also, this literature has not addressed the longer-term effects (if any) of this behavior on children's human capital attainment or earnings, or the household's welfare. de Janvry *et al.* (2006) make some progress on these issues. Though the authors find no effect of drought on school attendance among children living in rural Mexico, they do find a strong negative relationship between school attendance and other natural disasters (earthquake, flood, hurricane or plague), particularly for girls, children in primary school, and children of agricultural workers. The authors further show that there is state dependence in schooling (meaning children who are not in school in one period are less likely to be in school the next period) and that because of this transitory shocks can have potentially long-lasting effects. Specifically, the probability of enrollment drops by 3.4 percentage points in the period of a natural disaster and by another 0.4 percentage points the subsequent year, from a base enrollment rate of 81.4 percent. These findings are particularly pronounced for secondary school pupils.

The present research explores in more detail the characteristics of children affected by this form of consumption smoothing. Among Kenyan youth, delayed enrollment in primary school, grade repetition and dropout upon primary school completion (or well before) are common. Individuals must obtain satisfactory scores on a national, standardized examination to even be considered for acceptance into secondary school, and due to an extremely limited number of places available at this level, competition is fierce. As a result, individuals of higher academic and cognitive ability are more likely to successfully proceed through and beyond primary education. If these individuals also receive higher returns later on in life, as research indicates that they do, then decisions made to compromise their schooling in youth may have dramatic effects on longer-term individual and household welfare, implying that the lasting effects of transitory income shocks may be large. With this in mind, this paper investigates the relationship between child ability and school dropout due to weather-related local income variability.

3. A Model of Human Capital Investment and Child Ability

A two-period model of human capital investment is outlined to investigate how child ability factors into household decision making concerning school attendance. This framework is related to a more complex, multi-period model presented in Jacoby and Skoufias (1997) and more recently in Sawada (2003), but simplified to the two-period case and extended to include child ability.

Consider the case in which households gain utility over their (two-period) lifetime from two sources: consumption in the current period (C_t), and the discounted value of human capital that they impart to their child upon adulthood (H_{t+1}).[5] Households expect a positive return to

[4] There is no evidence presented on the time spent in each activity, however, so it may be that while enrollment is unaffected, actual attendance or time spent in school is reduced.

[5] As it is not possible to identify households, only children, in the data used in this analysis, one child per household is assumed here. More information is provided on this in Section 4.1 below.

schooling in later-life income, and final human capital determines the adult income of the child.[6] Thus, in choosing the optimal period t consumption and schooling (S_t), households maximize expected lifetime utility:

$$\max_{\{C_t, S_t\}} U(C_t) + \pi \mathrm{E}_t[V(H_{t+1})] \quad (1)$$

where π is the discount factor ($\pi \in (0,1)$), and U and V are both concave functions.

Child human capital is accumulated in period t. It evolves as a function of schooling, individual academic and cognitive ability ($\alpha \in (0,1)$), other individual characteristics (X_t), and school quality (q_t), such that:

$$H_{t+1} = H_t + h(S_t, \alpha) + f(X_t, q_t) \quad (2)$$

where h is a increasing linear function.[7] It is the inclusion of ability here that serves as a key addition to previous empirical literature examining the household schooling decision making process.

Households maximize lifetime utility subject to a budget constraint, in which they are assumed to consume all of their income and there is no credit:[8]

$$C_t = Y_t^P + Y_t^T - pS_t \quad (3)$$

where Y_t^P is the (permanent) income of the household, Y_t^T is a temporary income shock (positive, negative or zero), and p is the price of schooling. The child's forgone earnings are ignored for the moment, as it is not clear whether most children who are withdrawn from school in this setting do in fact work for an income.

Households choose consumption and schooling after the temporary income shock is realized. Thus the focus of this analysis is on after-the-fact consumption smoothing and risk-coping rather than on income smoothing. Assuming that $U = \ln(C_t)$ and $h(S_t, \alpha) = \alpha S_t$, the first order conditions of this maximization problem imply that:

$$S_t = \left[Y_t^P + Y_t^T - \frac{1}{\alpha\beta \mathrm{E}_t\left[\frac{\partial V}{\partial H_{t+1}}\right]} \right] * \frac{1}{p} \quad (4)$$

This relationship indicates that schooling increases with positive shocks to income ($Y_t^T > 0$), and decreases with negative shocks to income ($Y_t^T < 0$). Furthermore higher child ability is associated with higher rates of schooling. Finally, equation (4) suggests that an increase in both temporary income and ability raises S by more on average for higher ability individuals than for lower ability individuals. To see this, consider Figure 1.

---

[6] This assumption is justified using information specific to the study area in Section 5 below. Even in the context of agricultural employment, individuals may believe that higher levels of human capital lead to use of improved farming practices (Jacoby and Skoufias, 1997).

[7] Linearity will be assumed in this model, although any function more convex than U will suffice in this context. Furthermore, human capital does not depreciate in this model.

[8] The no credit assumption is plausible in the context studied here, where access to formal credit is extremely limited and informal credit is difficult to obtain during periods of bad weather (as multiple households in a given area are all hit by the same rainfall shock).

Figure 1 illustrates how shocks to household income, in conjunction with variation in child academic and cognitive ability, might affect the schooling decision under the assumptions previously specified. For households with income $Y^P < Y_1$, the marginal utility of consumption in the current period strictly exceeds that of the discounted marginal utility derived in the next period from investments in child schooling, so that none of these households will send children to school. In the event of a positive income shock that pushes total income above Y_1, households will send children to school as long as the child's ability level falls above some threshold level, α_0 that corresponds to the new total income Y'. Thus, as income rises, more children will attend school, starting with those of highest ability. Furthermore, this implies that at some sufficiently high income Y_2 all children (high and low ability alike) will attend school. Similarly, this model predicts that in the event of negative income shocks fewer children will attend school, particular those with lower ability on average.

In the analysis in Section 6, schooling will be measured as an indicator variable for attendance during the year. Equation (4) indicates that the following relationship holds:

$$S_t^* = g\left(Y_t^P, Y_t^T, \alpha, \beta, X_t, q_t\right) \quad (5)$$

The empirical specification used in Section 6 is drawn directly from this relationship.

4. Data Description

The data used in the present analysis is drawn from two sources, a longitudinal dataset of rural Kenyan youth and satellite precipitation data. In this section, each are described in turn.

4.1 The Kenyan Life Panel Survey

In 1998, an intestinal helminth treatment program known as the Primary School Deworming Project (PSDP) was introduced in Busia, a rural district of western Kenya. Under this program, a local non-governmental organization provided free deworming treatment to over 30,000 children, all of whom were attending one of 75 primary schools in the district at the time. In order to evaluate the effects of this health intervention, data was collected prior to the start of treatment on various outcomes including individual school participation, academic performance, health and household characteristics.[9]

Five years later a follow-up survey entitled the Kenyan Life Panel Survey Round 1 (KLPS-1) was launched. Between 2003 and 2005, this survey tracked a representative sample of 7,500 adolescents who were confirmed enrolled in primary school grades 2-7 in Busia District in 1998. This sample is fairly representative of the adolescent population in western Kenya: according to a Kenya Demographic and Health Survey, 85 percent of children in Western Province aged 6-15 were enrolled in school in 1998. However it should be noted that eighteen schools in Busia District were excluded from the sample, mainly because they were either economically and geographically unrepresentative of schools in the district or they had already received health and worm treatments under prior programs prior to the PSDP.[10] Survey data on a wide range of outcomes was successfully collected for over 5,200 of these young adults, including retrospective panel information on school attendance for the period 1998-2005.

In mid-2007 a second round of the Kenyan Life Panel Survey (KLPS-2) went to the field, and the entire sample of 7,500 adolescents was sought once again. Individuals were randomly divided into two groups, to be tracked in two separate waves of data collection, both of which are

[9] Miguel and Kremer (2004) provide more background information on the PSDP.

[10] Note that these eighteen schools were not included in the original PSDP study either, and do not count in the 75 schools from which survey respondents were drawn.

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, including data on school attendance extended through 2008.[11]

The present research utilizes the original PSDP information as well as the follow-up KLPS-1 and KLPS-2 Wave 1 data. The analysis is focused on a restricted sample of 1,865 individuals who have, in addition to KLPS panel school participation information, 1998 PSDP academic test score and survey data.

The key variable on which this study is built is this academic test score, which comes from an exam administered to students 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. The exam was given only to students in grades 3-8, thereby restricting the analysis to students enrolled in grades 3-7 in 1998 (recall that the KLPS only follows individuals in 1998 grades 2-7). Each grade level was administered a separate test.[12] Given the high rate of enrollment in Kenya during early primary school, exam scores – especially for students in lower grades in 1998 – are more likely to reflect raw academic and cognitive ability rather than household factors which could influence school enrollment and attendance (such as wealth). Thus, it is asserted that this test score data provides a unique measure of early individual academic and cognitive ability. Further evidence in support of this claim is outlined in Section 5.4 below.

It should be noted that the PSDP and KLPS datasets were sampled and collected at the individual, rather than household, level. It is not possible to definitively identify pupils who reside in the same household. However, given that the KLPS sample was chosen as 7,500 individuals out of a population of approximately 30,000, and only 1,865 of those individuals are included in this analysis, the chances of a large number of siblings occurring in the dataset are quite low. For the purposes of this analysis, therefore, it is assumed that there is only one child per household.

Panels A and B of Table 1 provide basic individual- and household-level information for the sample of interest. Just over half of individuals are male. Pupils were on average 12.6 years old in 1998 (with a range of 6-21) and had attained 4.8 grades of schooling. The vast majority (96 percent) are members of the Luhya tribe. In 1998 these individuals were typically of well below average weight-for-age and reported regular episodes of ill-health, an important reason this district was chosen for the original deworming treatment study. At the time of KLPS-2 survey enumeration, nearly 40 percent report ever having been married and nearly 50 percent report having ever been (or had a partner who was) pregnant. Father's educational attainment averages not quite nine years, and mother's less than six. According to the system of education in place prior to 1985, this means that in general fathers have completed almost two years of lower secondary schooling while mothers generally have not graduated from primary school. Just over half of pupil's households owned cattle in 1998, a strong indicator of wealth in this part of Kenya.

4.2 Rainfall Data

[11] Wave 2 of KLPS-2 data collection is currently underway, and will be completed in late 2009 and included in future analysis.

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

The Climate Prediction Center's Famine Early Warning System (FEWS) publishes daily precipitation estimates for Africa through the Africa Rainfall Climatology (ARC) dataset. These estimates are generated using an algorithm that combines infrared satellite data on cloud top temperature with information from ground-level rain gauges, and are available at a resolution of 0.1 by 0.1 degrees over the period 1995 through the present for the whole of the continent. Evaluation suggests that this data is relatively accurate, although it may underestimate particularly heavy rainfall (Love, Kumar, Xie and Thiaw 2004).

Using GPS information on the locations of the 75 primary schools attended in 1998 by individuals in the sample, the PSDP/KLPS survey information was merged with yearly average ARC estimates in order to obtain a panel of school attendance and annual average rainfall for the period 1998-2008.[13] All schools fall into one of nine different 0.1 by 0.1 degree grids (approximately equal to 10 kilometers by 10 kilometers), which are hereafter referred to as rainfall zones. Household or village level rainfall data would be preferable, but unfortunately is not available for this area. There are very few weather stations located in Busia District, and those that do exist do not report a consistent time series of precipitation information. Section 5 uses the ARC data to describe rainfall in the study district in more detail, and its relationship with agricultural output in particular.

5. Busia District, Kenya

5.1 Agricultural Practices and Rainfall

The setting for this analysis is Busia, Kenya. Busia is a primarily rural district located in western Kenya, abutting Uganda and Lake Victoria. Though the district is rather small in total land area, it contains a wide range of topography – altitude spans from 1,130 meters above sea level at the shores of the lake to 1,375 meters above in the hills that extend diagonally across the district from northeast to southwest. Nearly three quarters of the land is arable, and the main economic activity is subsistence agriculture (and/or fishing for households living near the lake).

As described previously, the present analysis focuses on a subset of youth who were enrolled in primary school in Busia in 1998. At that time, all sample individuals went to school in one of the two largest divisions of the district, Funyula and Budulang'i. Together, these divisions account for approximately half of the total land area of the district and one third of the population. They also encompass its varied geography, extending from Lake Victoria inland through marshland and seasonally flooded riverbeds up into the hills.

As within the rest of the district, the primary occupation in these two divisions is agriculture. Seventy-two percent of individuals in the sample lived in farming households at the time of KLPS-2 survey enumeration (2007/2008) and 67 percent performed farming activities personally.[14] Plot size is just 1.6 acres on average, and farming is primarily for subsistence –

[13] By assigning rainfall to an individual according to the school they attended in 1998, I assume little migration from this location. In fact, 64 percent of individuals in my sample report migration at some point. However, this migration is often very local, reportedly to continue schooling or visit a relative. Furthermore, most individuals live with a spouse or sibling upon moving, suggesting that their parents may remain at the original location. If this is the case, then household income is likely to be supported in large part by agricultural production at the 1998 location. To the extent that rainfall at the 1998 school does not capture income shocks appropriate to the individual, my estimates of the effect of transitory, weather-related shocks on school attendance will be attenuated.

[14] This figure represents a decrease in comparison to reports obtained during the KLPS-1 survey round conducted five years earlier (2003-2005) on the same individuals, at which time 89 percent of households reported farming. It is unlikely that this represents a decrease in farming participation of households, however, as almost 90 percent of

only one-fifth of farmers report crop sales, with receipts averaging Ksh 3,800 annually (approximately US $60, or one-tenth of Kenyan GDP per capita). Forty percent of these households employ some type of improved practice, which generally takes the form of planting improved or hybrid seeds. Just 7.5 percent use irrigation, indicating that agricultural output is extremely dependent on rainfall.

Rainfall in Busia is bimodal, with a "long" rainy season running from approximately March through July and a "short" rainy season from September to November. Planting occurs at the start of each rainy season. The primary crops grown in Budalang'i and Funyula are maize (reported by 77 percent of farmers) and cassava (14 percent). Maize cultivation is both highly labor-intensive and highly rain-dependent, whereas cassava is a relatively drought-tolerant crop that can be left in the ground (field storage) to be harvested as needed by the household.[15] Small fractions of farmers also report cultivation of beans, sorghum, kale and millet, but little land is devoted to cash crops. Thus, farmer survival is largely dependent on maize yields, and the bulk of production occurs during the long rains.

Total annual precipitation varies on average from 1,270 millimeters to 1,790 millimeters depending on location within the district (African Development and Emergency Organization, 2003), and generally increases in quality and reliability with elevation (Conelly and Chaiken, 2000). According to the United States Agency for International Development (USAID, 2006), 80 percent of inhabitants of the seven Kenyan districts nearest to Lake Victoria (of which Busia is one) live in the lowlands rather than the highlands, and are therefore extremely vulnerable to both drought and floods. Indeed, 44 percent of farming households interviewed in KLPS-1 reported crop destruction due to weather in the 12 months preceding the survey, and nearly 60 percent of these were drought-related.[16]

Figure 2 plots total annual rainfall during the long rains season by zone over 1998-2008, for each of the nine ARC rainfall zones in the study area. Rainfall varies over the period of study from a low of approximately 500 millimeters to a high of nearly 1,000 millimeters. According to monthly USAID reports on weather and food security in Kenya over this period, this distribution represents rainfall that fluctuates both above and below average without going too far to a single extreme. Only 2003 listed as a flood year of note, and that flooding appears to have been confined to Budalang'i Division (parts of which compose a catchment area for the River Nzoia). The USAID reports further suggest that crop yields in the region during the time period in question track rainfall fluctuations. Thus, for the purposes of this analysis, a fairly linear relationship between rainfall and crop yields will be assumed, such that higher than average rainfall will be considered "good" (translating into better than average crop yields) while lower than average rainfall will be considered "bad".

Because of the relatively compact size of the study area, there is a somewhat limited amount of variation in rainfall in the dataset. Figure 2 shows that precipitation displays a high degree of co-movement across zones, although there is still some variation in precipitation both across and especially within zones over time.[17] However, although zone fixed effects explain only 7 percent of variation in rainfall, zone and year fixed effects together explain approximately

---

individuals who report no longer living in a farming household at the most recent survey had migrated at some point during 1998-2008.

[15] Flour from maize (and sometimes cassava) is used in the staple food of the area, *ugali.*

[16] These statistics are calculated only among the study sample of households interviewed both in KLPS-2 Wave 1 and in KLPS-1.

[17] Variation in precipitation is not confined to particular zones, a fact illustrated by the kernel densities of rainfall (standardized within zone over time) in Appendix Figure 1.

94 percent of rainfall variation. It is this remaining village-year variation that is exploited here to proxy for local income shocks.

5.2 Rainfall as a Proxy for Income Shocks

Rainfall variability has been utilized as a proxy for income shocks in many studies of rural areas of less-developed countries, to examine a wide range of questions. Paxson (1992) is one of the earliest studies to take this approach, using deviations in rainfall from long-term mean as a measure of transitory income shocks in Thailand to study the savings behavior of agricultural households. The author finds that rainfall shocks have large effects on income, especially in the planting portion of the agricultural cycle. Rose (1999) employs this same measure in a study of the effect of income shocks at birth on the probability of female infant mortality in rural India. Hoddinott and Kinsey (2001) utilize extreme rainfall to study the effect of income variability on infant height in Zimbabwe. Miguel (2005) employs various measures of rainfall variation to examine income shocks as a determinant of violent crime in rural Tanzania.

As described previously, variations in rainfall are especially salient in rural Western Kenya, where livelihoods depend on crop production through rainfed agriculture. Most agricultural output occurs during the long rainy season, and thus this analysis will focus on measures of precipitation including these months only. Income shocks will be captured through a measure of annual deviation of rainfall within each zone from the fourteen-year zonal mean. In this way, only the transitory component of precipitation is recorded.

Due to a lack of time-series information on income or expenditures for KLPS individuals, the analysis presented here will take the form of a reduced form, rather than instrumental variables, regression, in which local rainfall will proxy for household income. Evidence has been provided elsewhere on the size of the effect of rainfall shocks on rural income and consumption in similar environments. Using data from six villages in rural Ethiopia, Dercon (2003) finds that a 10 percent decline in rainfall decreases total consumption by 3 percent and calorie intake by 6 percent. In later work examining a larger set of communities, Dercon *et al.* (2007) find that 10 percent higher rainfall increases consumption by 1.7 percent. Miguel (2005) studies nearby Meatu, Tanzania, and finds that an indicator for extreme rainfall (measured by village reports of drought or flood) is associated with a 25 percent drop in average village income.

For the rainfall shocks measured here to be truly transitory, there cannot be serial correlation in rainfall over time. To test this, a series of regressions by zone of average annual rainy season rainfall have been run on its lag and a constant term. Lagged rainfall does not significantly predict current any of the nine zones, suggesting that the hypothesis that rainfall follows a white-noise process cannot be rejected. As a result, the main analysis will be performed assuming deviations from long-term average rainfall are stochastic and exogenous.

5.3 Schooling

Since 1985, schooling in Kenya has been based on the "8-4-4" system: eight years of primary, four years of secondary, and up to four years of post-secondary education. Children are permitted to enter primary school at the age of six, although many do not immediately enroll. Furthermore, students who have not shown satisfactory mastery of the material are held back, and grade repetition is common. As a result of this delayed enrollment and repetition, it is not unusual for primary school students to range in age from 8-18.

In 2003 the national government introduced Free Primary Education (FPE) and enrollment rates increased dramatically. According to Demographic and Health Survey data for 1998 and 2003, school attendance among children aged 6-15 increased from 85 percent to 95 percent.[18] Despite this program, primary education continues to place a cumbersome financial burden on households. Although schools are not officially allowed to request money from parents, in practice families pay fees for extra tuition, textbooks, standardized exams, activities/sports, uniforms and transport. In fact, a recent study quotes primary school related costs at between Ksh 3,100 and 7,500 per child annually in rural areas (approximately US$40-96, or up to nearly 20 percent of Kenyan GDP per capita), while the government capitation grant rests at 1,020 per child (Elimu Yetu Coalition, 2004).[19]

Secondary school is even more costly to households, and due to limited spaces entry is competitive. Acceptance is based on scores on a national standardized exam known as the Kenya Certificate of Primary Education (KCPE). As a result of these costs and competition, enrollment and transition rates are low. In 2004 the gross and net secondary school enrollment ratios were 30 percent and 17 percent, respectively, and among those completing primary school the transition rate was only 50 percent. However, it does appear that those who are able to attend secondary school go on to complete it, and with much lower rates of dropout and repetition – nearly 85 percent pupils who enrolled in secondary school graduated within four years (Onsomu *et al.*, 2006).

Post-secondary options include college, university and vocational schools. Continuation to higher levels of academic pursuit is quite rare, and most who attend school after primary or secondary opt for a vocational institution. Indeed, data from the 2003 Kenya DHS indicates that average schooling attainment in Kenya's Western Province is no more than primary school, where 69 percent of individuals aged 20-29 have attained some level of primary schooling and 11 percent have not attended school at all.

Panel A of Table 2 summarizes information on schooling from the KLPS-2 respondents that are the focus of the present analysis. As the KLPS sample was chosen from the population of children attending primary school in the Funyula and Budalang'i Divisions of Busia District in 1998, it must be the case that all individuals report attending school at some point during 1998-2008. School participation in 1998, measured by a series of enumerator visits to check attendance, was high, on the order of 90 percent. Furthermore, almost a quarter of youth were still attending school at the time of KLPS-2 enumeration.

It should be noted that attainment is slightly higher in this sample than for the youth population at large[20] – the highest grade attained averaged 9 years (one year of secondary school) among the entire sample, and 8.5 years among those not attending school at the time of survey. Three quarters of individuals have either completed or were still attending primary school, and nearly 40 percent have either completed or were attending secondary school at the time of

[18] This increase was similar across gender.

[19] Information on household expenditures collected in 2004-2005 from 155 respondents of the KLPS-1 survey is revealing. Fifty percent of households pay school fees (though it is not possible to distinguish between fees paid for primary, secondary, or other schooling), and the average annual payment across all households is Ksh 9,700, nearly 20 percent of non-food consumption. The KLPS-1 did not collect information on total number of household members attending school at the time of survey. However, annual fees paid are much higher for households in which the respondent himself or herself is attending school at the time of survey – nearly Ksh 20,000 on average. Note that these figures are not weighted to maintain initial population proportions, but are instead presented as anecdotal evidence on schooling costs in rural Kenya.

[20] At least, in comparison to the DHS estimates of school attainment.

survey. Fewer than 5 percent of respondents had attended an academic post-secondary school, although over one-quarter had attended a vocational institution.

Panel B of Table 2 provides summary statistics on grade repetition and dropout. Over 70 percent of individuals have repeated a grade at some point during 1998-2008. The KLPS-2 collects detailed information on reasons for repetition, and the two most popular include "scores not good enough because didn't try hard" (reported by 53 percent of those who have ever repeated a grade) and "no money for uniform/fees" (reported by 18 percent).

Dropout rates are also high. Measuring dropout including both instances of failure to enroll in school prior to completion of secondary, as well as partial-year attendance, the rate is approximately 75 percent. Including in this definition only individuals who fail to enroll, dropout remains high at 63 percent. Excluding those who conclude their education after attending grades 8 or 12, which are both popular points of termination in Kenya, dropout rates are still 46 percent and 21 percent respectively. It appears that most individuals only drop out of school once, and that this dropout is permanent. By far the most cited reason for dropout prior to grade 12 is "no money for uniform/fees" (reported by 78 percent). Pregnancy is also cited by a many female dropouts.

So what are these youth doing when not attending school? In rural Kenya, child labor in the household or on agricultural land owned by the household is common, for both those attending school and those who have dropped out. At the time of KLPS-1 interview in 2003-2005, 42 percent of individuals in the study sample who were not attending school reported working on the household farm, and similarly 40 percent of those who were attending school performed such labor. In contrast, labor outside of the household appears to be restricted to dropouts. While fewer than 19 percent of surveyed youth attending school report cash labor, nearly 46 percent of those not attending school do. The primary jobs for these individuals are hawking items like clothes and food, or fishing.

5.4 Ability, Schooling and the Kenyan Labor Market

In this analysis, the central variable of interest is a 1998 academic test score recorded from an exam administered as part of the initial PSDP evaluation and based on standard Kenya Ministry of Education exams. Given the high rate of enrollment in Kenya during early primary school, test scores – especially for students in lower grades in 1998 – are more likely to reflect raw academic and cognitive ability than other factors which could influence school enrollment and attendance, such as household wealth or individual health. Table 3 provides some evidence in support of this claim, presenting results from regressions of 1998 individual test score on various observable individual and household characteristics available in 1998.

Column (1) of Table 3 includes the entire sample of individuals in 1998 grades 3-7, and column (2) is limited to individuals in grades 3-5 only. The first thing to notice from both sets of results is that the wide range of controls included is unable to explain much of the variation in academic test score, as indicated by the low R-squared value. Furthermore, in both samples, the coefficients on household ownership of cattle (a robust indicator of wealth in this region), weight-for-age z-score, and self-reported health status indicate no statistically significant relationships with test score, indicating that exam performance is not an indicator of household wealth or individual physical robustness.[21] Age-for-grade is significant and negative in both columns, though to a much lesser degree for youth in lower grades. Such a relationship likely holds for individuals who were held back from grade progression, which may be related to low

[21] The regression coefficient on self-reported health is only weakly significant in the first column.

academic ability. Although females do appear to score lower on the test, this result is not robust for the sample of younger individuals in column (2). Father's educational attainment is positively related to ability for the full sample. Finally, average school participation in 1998 is significantly and positively related to test score within the full sample, although this relationship disappears when examining only individuals in lower grades. The fact that this relationship is significant only for students in higher grades strongly suggests that causality runs from test score to attendance rather than the other way around.

Combined with the fact that enrollment rates for children in early primary school are extremely high, these findings imply that 1998 test scores are likely a measure of ability rather than socioeconomic or health factors. Figure 3 displays the kernel density of the PSDP test score separately by gender.

As suggested previously, due to the costs of schooling and the competitive nature of advancement, those of higher academic and cognitive ability are generally more successful proceeding further through school in Kenya. Panels A and B of Table 4 provide summary statistics on various educational outcomes collected at the time of KLPS-2 survey (2007/2008) for individuals in the sample, broken out by gender and by a binary measure of ability.[22] Individuals classified as high ability according to their 1998 test score have completed nearly 1.5 additional years of schooling on average by the time of KLPS-2 survey, and are significantly more likely to have attained primary and secondary degrees, as well as to have attended some post-secondary schooling. Interestingly, high and low ability individuals are equally likely to have attended a vocational education program. In addition to higher attainment, repetition and dropout rates are much lower for high ability pupils, although dropout for all individuals is very likely to be permanent.

All of these relationships hold across gender, though levels of female attainment are much lower, suggesting that there is strong gender bias in school progression. It should be noted, however, that the differences between high and low ability females are much larger for many outcomes than the corresponding differences across ability levels for males. Further, outcomes for high ability females are substantially better than those for low ability males, indicating that progression bias toward males is not so extreme as to ignore academic and cognitive skill.

The figures presented in Panels A and B of Table 4 indicate that students of higher ability do tend to proceed further through school, but what type of later-life returns does this advanced education provide? Panel C summarizes employment outcomes the study sample. Rates of unpaid agricultural labor for the household, self-employment, work for pay and unemployment are similar across ability levels. At the time of KLPS-2 survey, 83 percent of individuals were still living in rural areas, so these figures largely reflect the rural labor market. Rates of self- and wage-employment are slightly higher among those living in urban areas, and interestingly, in related research Hamory Hicks and Miguel (2009) find that high ability individuals are more likely to migrate to cities (likely in pursuit of higher paying informal and formal sector jobs).

Most farming households in the study area do not employ outside labor. Approximately two-thirds of KLPS individuals report performing agricultural tasks for their households, and among them average time spent is 13 hours per week. The primary industries of self- and wage-employment in rural areas are retail of food/drink/tobacco, retail of clothing/shoes/textiles, and

[22] Although the main regression analysis will make use of all the variation available in this test score measure, for the purposes of summary statistics I find it useful to consider just two levels – "high" ability individuals are defined as those with above-median scores, and "low" ability individuals are those with at- or below-median scores.

fishing. Service industries are an important source of employment in urban areas, including mainly domestic services, engineering/architectural/technical services, restaurant/café/bar services. Although low rates of employment in rural areas could be evidence of a lack of employment opportunities in the rural labor market (and indeed one-quarter of individuals report searching for work), some of this is also likely related to the substantial fraction of high ability individuals still attending school. It deserves some note that wages among those working for pay are significantly higher for higher ability individuals (Ksh 4833 versus Ksh 3509 per month), suggesting that once employed there are significant returns to ability and/or educational attainment.

Because individuals in the KLPS sample are still young and have not yet completed their schooling, the present research will not attempt to use this survey to distinguish between later-life employment and wage outcomes related to individual academic and cognitive ability, and those related to ultimate educational attainment. Instead, the research of Knight and Sabot (1990) sheds some light on this issue. In their study of returns to education among urban wage earners in the Kenyan capital of Nairobi, the authors attempt to disentangle the effects on wages of cognitive skill (which they define as human capital acquired through schooling, as measured by a set of literacy and numeracy tests administered to respondents), reasoning ability (individual ability independent of schooling, as measured by a Raven's Coloured Progressive Matrices test), and years of schooling attained.[23] Regressions of log wages on all three measures, including a measure of labor market experience, suggest a large and positive effect of schooling attainment and cognitive skill, but little to no premium for reasoning ability. However, the authors do find that ability matters through indirect pathways – higher ability individuals complete more years of schooling than their counterparts, and cognitive skill is found to be monotonically higher in ability (with an elasticity approximately one-third). Using the difference in average ability between the top and bottom third of individuals, the authors estimate the total (direct and indirect) effects of ability to increase average wages by 32 percent. This study indicates that returns to individual ability in Kenya, as well as those to school attainment and the knowledge picked up in school, are high. The present research attempts to capture some aspect of this raw ability through the PSDP test score measure.

6. Empirics

6.1 Sample Attrition

Before moving to the main empirical analysis, one must first address the issue of sample 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, an age group likely to be extremely mobile due to marriage, schooling, and labor market opportunities. Furthermore, the analysis presented in this paper utilizes only the subsample of KLPS individuals who were present on the days of PSDP

[23] There are limitations with this study. In order to ensure a large sample of secondary school completers in their data, the authors restricted their survey to urban wage earners. This of course ignores returns to education among rural wage earners, and to other forms of employment such as self-employment. Furthermore, if there is selectivity among secondary school completers in wage-earning positions, then estimates of returns to secondary school will be biased up. The authors deal with this problem by comparing the relative wages of primary to secondary school completers. Since there is likely to be even more selectivity among the former who manage to obtain wage earning positions, any estimates of the relative returns will be biased down.

examination administration in 1998, or among a group of individuals who were absent but tracked to their homes to complete a makeup exam. Because one could reasonably believe that only individuals with particular characteristics are included in this subsample, it is essential to carefully examine survey attrition. If the key test score measure is systematically related to attrition, then any resulting estimation will likely be biased.

Column (1) of Table 5 provides results from the regression of an indicator for inclusion in the main analysis on various individual and school-level characteristics. The sample used in this column includes all individuals sought for KLPS-2 Wave 1 survey – including those surveyed, found but not surveyed, and unfound – and only the few variables that are available for this entire group are utilized in this regression.[24] Approximately 71 percent of these individuals are included in the analysis presented in this paper. Regression results suggest no differential attrition across gender, age, 1998 grade, or a 1996 school average test score (a measure of school quality).

Column (2) of Table 6 is restricted to individuals who additionally have survey data from the 1998 PSDP, and use this additional data to further control for individual health and household wealth in 1998. Once again, none of these measures are significantly related to inclusion in the main analysis. Together these results suggest that there is little evidence of differential attrition within the sample of interest.

Column (3) further extends the attrition analysis to include 1998 individual test score. Here, it does appear that there is a significant and negative relationship between the likelihood of being included in the main analysis and test score. Specifically, high ability individuals (those with test scores one standard deviation above the mean) are nine percent less likely to be included in the analysis than low ability individuals (those with test scores one standard deviation below the mean). This is not too surprising, as Hamory Hicks and Miguel (2009) find that higher ability individuals (using this same test score measure) are significantly more likely to migrate than their lower ability counterparts. However, this finding does have important implications for the present analysis. High ability individuals are underrepresented in the data, and hence results are likely biased, though the direction of this bias is unclear.

6.3 Main Analysis

Equation (5) indicates a relationship between the school attendance decision and a variety of other individual and household characteristics. More specifically, attendance can be expressed as a function of household income, income shocks, child ability, other child attributes, household socio-economic characteristics, and school quality. Because the data used in this study do not include longitudinal information on household income, expenditures or crop yields, the main analysis will employ a reduced-form regression in which rainfall shocks proxy for income changes. Furthermore, the focus here is on how attendance changes in response to temporary changes in income, and whether individual- and household-level factors influence this relationship. Thus, the primary empirical specification is:

[24] Not all of the 3,745 individuals originally sampled to be including in KLPS-2 Wave 1 tracking are included in this sample. During the first several months of Wave 1 tracking, all 3,745 sampled individuals were sought. 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. Those chosen for the "intensive" sample are reweighted in the present analysis by their added importance. For a clear explanation of this methodology, see Orr *et al.* (2003). For a more detailed explanation of the KLPS sample, tracking and weighting methodology, see Baird *et al.* (2008).

$$S_{ijt} = \beta_0 + \beta_1 Y_{jt}^T + \beta_2 \left(Y_{jt}^T * \alpha_{ij} \right) + \beta_3 \left(Y_{jt}^T * X_{ij} \right) + \beta_4 \phi_{dt} + \gamma_{ij} + \delta_t + \varepsilon_{ijt} \qquad (6)$$

where i indexes individuals, j indexes ARC rainfall zones and t indexes time (1998-2008). S denotes attendance, defined using an indicator for primary or secondary school attendance during the year. Y^T represents the rainfall shock, measured as standardized precipitation within each zone (using data by zone over the period 1995-2008). α measures individual ability using the previously described 1998 PSDP test score (standardized within grade). X is a vector of individual- and household-level controls, including gender, age, 1998 grade, parent education (which provides a rough proxy for household socio-economic status), and a 1996 school-level mock test score (a measure of school quality). For a subsample of respondents, additional information exists on 1998 weight-for-age and self-reported health, household ownership of cattle (a proxy for household wealth in the area) and number of small children in the compound. These are included in a subset of regression results.

ϕ_{dt} denotes an indicator for female and 1998 grade level interacted with year, and is included to soak up the natural trend of dropout that occurs over time as cohorts age. γ is an individual-level fixed effect, included to absorb the influence of any time-invariant individual (or household) factors that may be related to attendance. As noted above, because the KLPS data was collected at the individual (rather than household) level, and it is unlikely that numerous pairs of siblings exist in the sample, one child per household is assumed in this analysis. Finally, δ denotes a linear year trend, and ε is a white-noise disturbance.[25]

Table 6 presents results from regressions employing this empirical specification. Column (1) illustrates that higher levels of rainfall are indeed associated with higher attendance rates. For a one standard deviation increase in zonal rainfall (which translates as approximately 116 millimeters during the long rains season, or 16 percent of average long rains rainfall), attendance increases by over 5 percent. Including higher powers of rainfall does not suggest a nonlinear relationship (results not shown), and the USAID food security reports described in Section 5.1 indicate that during the time period studied here more rainfall is indeed associated with higher yields while less rainfall is associated with lower yields. Figure 4 provides a quintic residuals plot of school attendance versus rainfall, and confirms this relatively linear relationship everywhere but the very extremes of precipitation.

Column (2) additionally controls for an interaction between standardized rainfall and 1998 individual test score, and suggests a significantly positive relationship. During periods of better rainfall, higher ability individuals are even more likely than their lower ability counterparts to attend school. Column (3) breaks this interaction up into quintiles of test score, and the results show that this relationship is driven by individuals with scores in the highest quintile. Figure (4) employs a quintic residuals plot to illustrate this graphically. Although the relationship is not quite as clear here, one can see that attendance rates for high ability individuals rises more rapidly with rainfall than rates for lower ability individuals.

Columns (4) and (5) of Table 6 add additional controls to this base regression, and although the test score interaction reduces in magnitude, the significance of its relationship with school attendance remains robust. Results further suggest a weakly negative relationship with

[25] Figure 1 suggests that most variation in rainfall is contained within zones over time, rather than across zones. For this reason year fixed effects are not used in the main analysis, as they would remove much of this variation. Rather, δ is a linear year trend included to soak up the general upward trend in rainfall across all zones over the time period. Year fixed effects are included in robustness checks on the main results, and reveal substantively the same conclusions.

school attendance for females during periods of better rainfall, perhaps due to changes in household labor requirements (female youth in the study area are frequently tasked with the care of younger siblings so that adults may work out in the fields).[26] Older children are less likely to attend school, but those initially in higher grades are weakly more likely to attend, indicating perhaps that older children within a given grade are less academically able, while children in higher grades may already be selected on academic ability.

Father's educational attainment does appear to be an important determinant of school attendance during periods of rainfall shocks, perhaps because this term is acting as a proxy for household income, or perhaps as fathers who received more education themselves value the same for their children. Another possible explanation is that this term is picking up some of the ability effect (recall that Table 3 suggested a significant and positive relationship between paternal schooling and child test scores). A proxy for household wealth in the region, ownership of cattle, has only a weak relationship with attendance, suggesting that it is not the wealthier households (who may be better able to smooth over small changes to income) that see attendance changes during periods of rainfall variation. Indicators of individual health and physical robustness in 1998 (weight-for-age z-score, self-reported health) also do not appear to be related to the schooling decision during extreme rainfall.

Together, the results in Table 6 suggest that there is indeed a modest positive and robust relationship between child ability and school attendance during periods of extreme rainfall. In particular, regression analysis indicates that when moving from a test score of one standard deviation below the mean to one standard deviation above the mean, an increase in rainfall by one standard deviation increases school attendance by an additional 3 percent. These findings are robust to controlling for the rainfall measure interacted with 1998 school participation, and do not appear to be driven by a single portion of the 1998 exam (results not shown).

In order to see whether these results are being driven by "good" (higher than average, but not flood-level) or "bad" (lower than average) rainfall, Tables 7 and 8 display results from a subset of these regressions using indicators for high and low rainfall, respectively. Table 7 provides results in which the rainfall shock is defined as an indicator for rainfall that is one standard deviation or greater above the long-run zonal average, and Table 8 provides results using an indicator for rainfall that is one standard deviation or greater below the long-run zonal average. These results suggest that high rainfall is associated with higher school attendance while low rainfall is associated with lower school attendance, a finding that fits well with the model presented in Section 3.

The results from these two tables further indicate that the test score findings are driven by periods of higher than average rainfall. Moving from a test score of one standard deviation below the mean to one standard deviation above, periods of high rainfall are associated with 11 percent higher attendance. Column (3) suggests that this relationship holds for schooling among individuals of average and above average ability, a finding which makes intuitive sense. The poorest households (which are likely to be most heavily affected by the increase in income) are now able to send their children to school, and would send the highest ability first. The comparatively wealthier households, who were already sending some children to school, would also enroll additional children, but ones of lower comparative ability. However, the analysis appears to be driven more by these "marginal" households; indeed, there is no statistical significance on a control for the triple interaction of rainfall variability, child ability and

[26] Regression analysis indicates that females in households with higher numbers of small children are even less likely to attend school in times of good rainfall.

household cattle ownership (results not shown). These results suggest that in order to lead any households to put the very lowest ability children in school there may need to be a much larger income change (larger than we observe during the study period).

Interestingly, there is no statistically significant relationship between test score and school attendance during periods of low rainfall – negative income shocks lead to school dropout irrespective of ability level. One possible explanation for this is the limited availability of formal credit in the primary study region, and the lack of informal credit during these times of locally aggregate income shocks. Households hit by a negative income shock may lack of funds to support attendance of any child, meaning when income drops all children are pulled out of school. This explanation is supported by self-reported explanations for dropout among individuals in the study sample. A second possible explanation is that negative rainfall shocks have a different effect schooling choices than positive shocks, perhaps because they reduce outside work opportunities.

These results are likely to encompass a large but not complete representation of the youth population of rural Kenya at large. As previously described, there is some selection built into the KLPS sample. Individuals were chosen for participation in the survey based on their attendance in school in 1998. Furthermore, the sample employed in this analysis under-represents individuals of high academic ability.

7. Conclusion

This analysis explores human capital accumulation decisions in rural Kenya, a setting where households attempt to smooth consumption in the face of income shocks by pulling children out of school. Because high ability students accumulate human capital from schooling more rapidly, theory suggests that they should be more likely to attend school during times of plenty, and less likely to be removed from school when the household is hit with a negative shock. Using a unique test score measure to capture individual ability, I show that households do factor child academic and cognitive ability into schooling decisions. Specifically, during periods of higher than average rainfall, when income is high, households exhibit a strong preference in favor of the education of children with higher intellect or academic ability.

A second major finding is that during periods of drought, when households have less money, they show no evidence of discriminating among children based on ability. One possibility for this discrepancy in household behavior when confronted with positive and negative shocks is the existence of credit constraints in the region. Because households are severely credit constrained, and because sending a child to school in the region requires a large share of the typical household's income, during negative shocks households may lack the resources to discriminate based on ability. Faced with a negative shock, households are simply forced to remove all children from school.

Several results from this analysis are worth highlighting. Households appear to exhibit ability preference when making schooling decisions, likely because returns to education are increasing in child ability. Nonetheless, the failure of households to protect high ability students from dropout during rainfall shocks suggests a large and potentially long-term cost to credit constraints in the region in the form of reductions in human capital investment. Importantly, the poorest households are likely the most vulnerable to income shocks, and thus the most liable to employ this mechanism to smooth consumption. Because low income households are apt to be disproportionately affected, this phenomenon would both drive and perpetuate overall economic and social inequality within the region. Policy providing wider availability of credit, lowering

fees for schooling, or facilitating access to sources of income with reduced volatility, would all serve to alleviate the impact of negative shocks on child well-being and help to break the cycle of poverty and under-education for future generations of Kenyans. However, limited rainfall variation in the primary study area limits the generalizability of these results to other settings. Future work is needed to examine this question on a larger scale.

Chapter 1 References

African Development & Emergency Organization. (2003). "Kenya Country Programme: Annual Report."

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 Research Working Paper C08-151.

Beegle, Kathleen, Rajeev H. Dehejia, and Roberta Gatti (2006). "Child Labor and Agricultural Shocks." *Journal of Development Economics*, 81, 80-96.

Behrman, Jere and James Knowles (1999). "Household Income and Child Schooling in Vietnam." *World Bank Economic Review,* 17(2), 229-254.

Conelly, W. Thomas and Miriam S. Chaiken (2000). "Intensive Farming, Agro-Diversity, and Food Security Under Conditions of Extreme Population Pressure in Western Kenya." *Human Ecology,* 28(1), 19-51.

Elima Yetu Coalition (2004). "Monitoring of the Free Primary Education and Establishing the Unit Cost of Primary Education in Kenya." Nairobi.

de Janvry, Alain, Frederico Finan, Elisabeth Sadoulet, and Renos Vakis (2006). "Can Conditional Cash Transfer Programs Serve as Safety Nets in Keeping Children at School and From Working When Exposed to Shocks?" *Journal of Development Economics,* 79, 349-373.

Dercon, Stefan (2003). "Growth and Shocks: Evidence from Rural Ethiopia." Centre for the Study of African Economies Working Paper WPS/2003-12.

Dercon, Stefan, John Hoddinott and Tassew Woldehanna (2007). "Growth and Poverty in Rural Ethiopia: Evidence from 15 Communities1994-2004." Background Paper for the Chronic Poverty Report 2009-09.

Guarcello, L., F. Mealli, and F. Rosati (forthcoming). "Household Vulnerability and Child Labor: the Effect of Shocks, Credit Rationing and Insurance." *Journal of Population Economics.*

Hamory, Joan and Edward Miguel (2009). "Individual Ability and Selection into Migration in Kenya." Mimeo.

Hoddinott, John and Bill Kinsey (2001). "Child Growth in the Time of Drought." *Oxford Bulletin of Economics and Statistics,* 63(4), 409-436.

Jacoby, Hanan G. and Emmanuel Skoufias (1997) "Risk, Financial Markets, and Human Capital in a Developing Country." *Review of Economic Studies,* 64 (3), 311-335.

Jenson, Robert (2000). "Agricultural Volatility and Investments in Children." *The American Economic Review,* 90(2), 399-404.

Knight, John B., and Richard H. Sabot (1990). *Education, Productivity, and Inequality: The East African Natural Experiment.* Oxford University Press, New York.

Love, T.B., V. Kumar, P. Xie, and W. Thiaw (2004). "A 20-Year Daily Africa Precipitation Climatology Using Satellite and Gauge Data." AMS Conference on Applied Climatology, AMS.

MEASURE DHS STATcompiler (2009). Retrieved from http://www.measuredhs.com.

Miguel, Edward (2005). "Poverty and Witch Killing." *Review of Economic Studies,* 72, 1153-1172.

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

Onsomu, Eldah N., David I. Muthaka, Moses W. Ngware and Damiano K. Manda (2006). "Determinants and Strategies for Expanding Access to Secondary Education in Kenya." Kenya Institute for Public Policy Research and Analysis Discussion Paper No. 63.

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.

Paxson, Christina H. (1992). "Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand." *The American Economic Review,* 82(1), 15-33.

Ravallion, Martin and Quentin Wodon (2000). "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy." *The Economic Journal,* 110, C158-C175.

Rose, Elaina (1999). "Consumption Smoothing and Excess Female Mortality in India." *The Review of Economics and Statistics,* 81(1), 41-49.

Sawada, Yasuyuki (2003). "Income Risks, Gender, and Human Capital Investment in a Developing Country." Discussion Paper CIRJE-F-198.

USAID (2006). "Special Report: Kenya's Lake Region Short Rains Rapid Food Security Assessment." Retrieved from www.fews.net/kenya. February.

USAID (various months/years). "Kenya: Food Security Update." Retrieved from www.fews.net/kenya.

Figure 1: Illustration of model of human capital accumulation with child ability

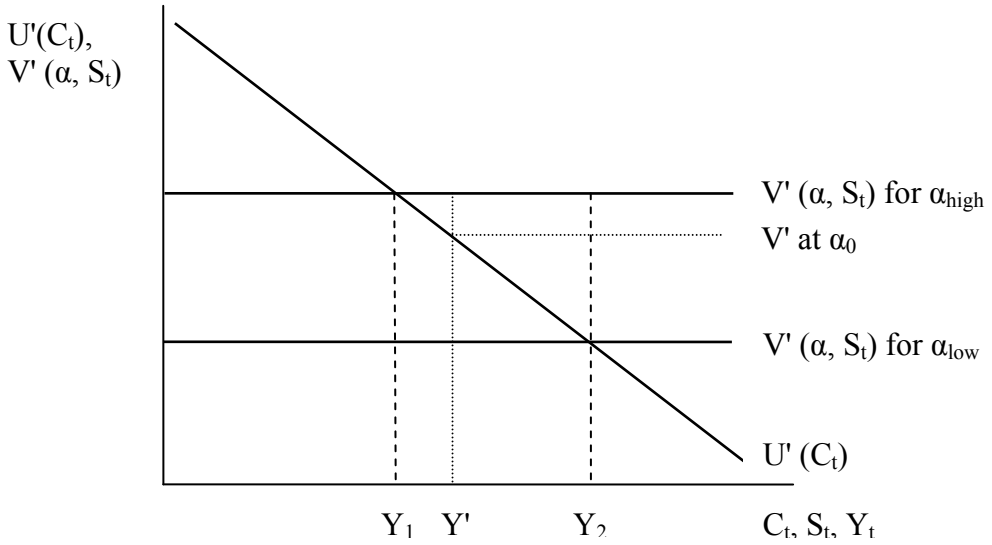
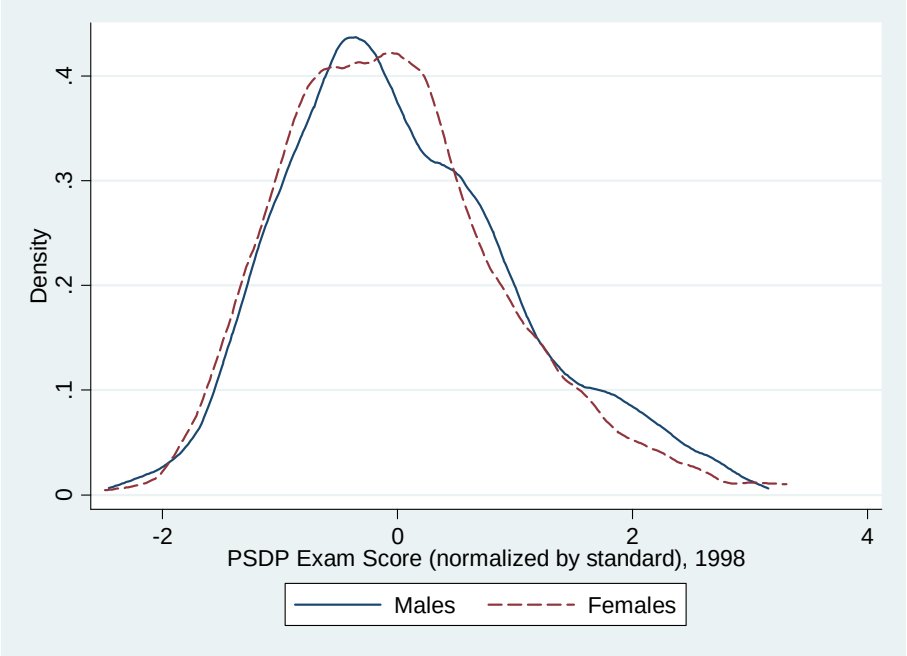


Figure 2: Total rainy season rainfall over time, by zone

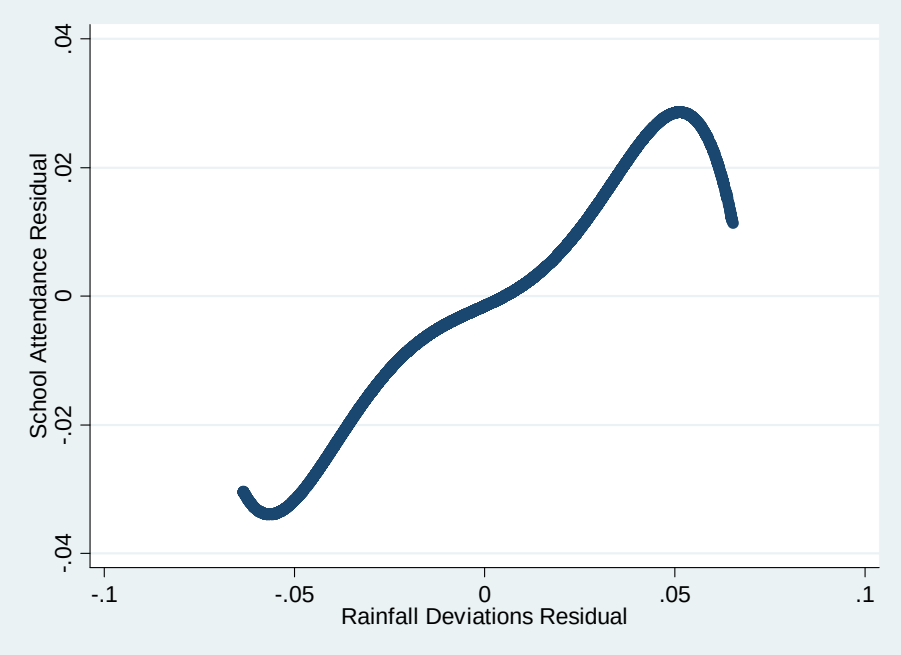
Notes: Rainy season is defined to include the long rains only.

Figure 3: Kernel density of standardized test score, by gender



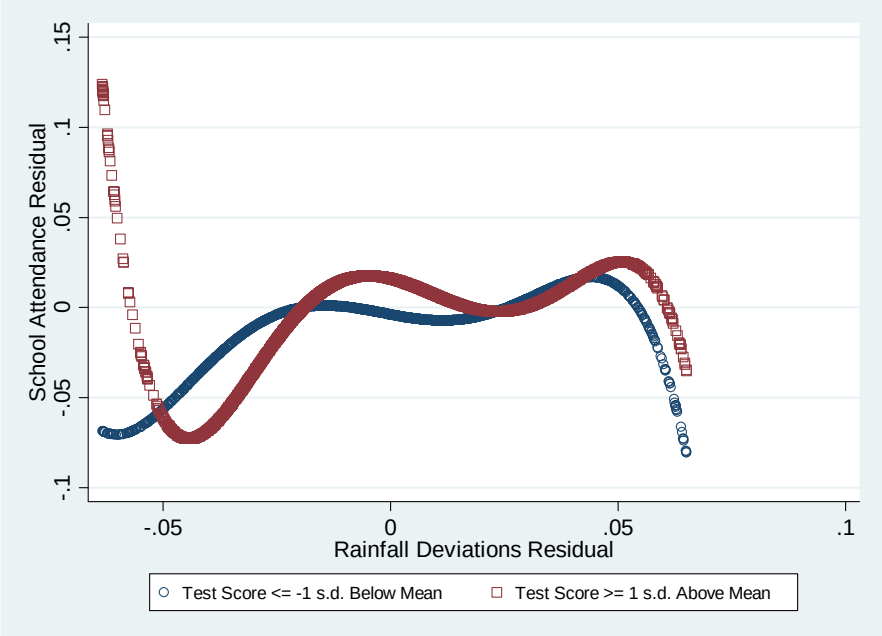
Notes: The sample used here includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score were dropped. Figures are weighted in order to maintain initial population proportions.

Figure 4: Quintic residuals plot of school attendance vs. rainfall shock



Notes: The sample used here includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score, as well as the upper and lower 5 percent tails of rainfall residuals, were dropped. Figures are weighted in order to maintain initial population proportions.

Figure 5: Quintic residuals plot of school attendance vs. rainfall shock, by 1998 test score



Notes: The sample used here includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score, as well as the upper and lower 5 percent tails of rainfall residuals, were dropped. Figures are weighted in order to maintain initial population proportions.

Table 1: Summary statistics for sample, basic

	Mean	Std. Dev.	No. Obs.
Panel A: Individual information			
Female	0.463	0.499	1865
Age (1998)	12.6	2.3	1860
Grade (1998)	4.80	1.41	1865
Member of the Luhya tribe	0.960	0.197	1859
Weight-for-age z-score (1998)	-1.49	0.83	1566
Falls sick often (1998)[a]	1.93	0.51	1566
Years assigned deworming treatment (1998)[b]	2.94	1.80	1865
Ever been married	0.386	0.490	1865
Ever been pregnant[c]	0.490	0.500	1863
Panel B: Household information			
Father's educational attainment[d]	8.71	4.24	1502
Mother's educational attainment[d]	5.58	3.92	1545
Household owns cattle (1998)	0.523	0.500	1565

Notes: The sample used here includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score were dropped. All figures are weighted in order to maintain initial population proportions.

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

[b] 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.

[c] For males, ever been pregnant relates to partners.

[d] Parent educational attainment data was collected in the KLPS-2 survey. It is coded according to the pre-1985 school system, which included seven years of primary, four years of secondary, two years of an advanced sequence and three years of university. Attainment is censored at 15 years, and coded such that 14 years includes those who attended some post-secondary education and 15 years includes those who completed post-secondary.

Table 2: Summary statistics for sample, school-related

	Mean	Std. Dev.	No. Obs.
Panel A: Ability/school achievement information			
ICS test score (1998)[a]	0.007	0.976	1865
School average of mock test score (1996)[b]	0.414	0.416	1865
Attended school at some point during 1998-2008	1.00	0.02	1865
Average school participation (1998)	0.904	0.197	1826
Attending school at time of survey	0.230	0.421	1852
Highest grade attained at time of survey[c,d]	9.08	2.31	1865
Among those not attending school[c,d]	8.52	2.18	1416
Completed or is attending primary school at time of survey	0.760	0.427	1863
Attended secondary school at some point during 1998-2008	0.450	0.498	1865
Completed or is attending secondary school at time of survey	0.389	0.488	1854
Attended post-secondary school at some point during 1998-2008[d]	0.048	0.214	1865
Attended a vocational education program	0.258	0.438	1865
Panel B: Grade repetition and dropout information[e]			
Repeated a grade at some point during 1998-2008	0.705	0.456	1796
Dropped out of school at some point during 1998-2008 (defn A)[f]	0.848	0.359	1865
Not including those in grades 8 or 12 prior to dropout	0.457	0.498	1865
Among this subgroup, number of times	1.17	0.52	802
Among this subgroup, percent that are permanent	0.721	--	802
Dropped out of school at some point during 1998-2008 (defn B)[f]	0.630	0.483	1865
Not including those in grades 8 or 12 prior to dropout	0.208	0.406	1836
Among this subgroup, number of times	1.04	0.21	382
Among this subgroup, percent that are permanent	0.767	--	382

Notes: The sample used here includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score were dropped. All figures are weighted in order to maintain initial population proportions.

[a] PSDP test score is standardized by 1998 grade.

[b] School average test score data has been converted into units of individual standard deviations.

[c] Child educational attainment is coded according to the post-1985 school system, which includes eight years of primary, four years of secondary, and four years of university. Attainment is censored at 13 years, such that 12 years encompasses those who attended through the last year of secondary school and any Ugandan A-level schooling, and 13 years encompasses those who attended any amount of post-secondary school (including college and university).

[d] Post-secondary education figures should not include vocational education programs, but may include vocational education in a few cases due to changes in how attainment questions were framed between the KLPS Round 1 and Round 2 surveys.

[e] Grade repetition and dropout information is only calculated during years of primary and secondary schooling.

[f] Dropout definition A counts as dropouts both those who were attending school and then failed to enroll the following year, as well as those who enrolled in school but only attended for a partial year – this measure thus takes into account both enrollment and actual attendance. Dropout definition B includes only the former group, thereby only taking into account enrollment.

Table 3: Regressions of test score on characteristics

	Dependent Variable: Test Score (1998)	
	(1)	(2)
	All	Grades 3-5
Female	-0.1591**	-0.1297
	(0.0637)	(0.0793)
Age (1998)	-0.0570***	-0.0414*
	(0.0199)	(0.0216)
Grade (1998)	0.0330	
	(0.0348)	
Weight-for-age z-score (1998)	0.0112	-0.0400
	(0.0454)	(0.0637)
Falls sick often (1998)[a]	-0.1025*	-0.1136
	(0.0612)	(0.0825)
Household owns cattle (1998)	-0.0442	-0.0666
	(0.0624)	(0.0814)
Father's education	0.0226**	0.0144
	(0.0092)	(0.0123)
Mother's education	-0.0064	-0.0034
	(0.0098)	(0.0136)
Average school participation (1998)	0.4025***	0.1803
	(0.1500)	(0.1798)
Observations	1531	919
R-squared	0.04	0.02

Notes: The sample used here includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score and 1998 Pupil Questionnaire information. Three observations with an extremely high test score were dropped. Robust standard errors are in parenthesis. All figures are weighted in order to maintain initial population proportions. Age is demeaned. Additional controls include indicators for missing age and parent education data.

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

Table 4: Outcomes by ability level

				Males			Females		
	Total	Low	High	Total	Low	High	Total	Low	High
Panel A: Schooling achievement outcomes									
Attending school	0.230	0.197	0.265	0.270	0.241	0.301	0.183	0.144	0.225
Highest grade attained[a]	9.08	8.42	9.80	9.43	8.81	10.12	8.68	7.96	9.44
Among those not attending school[a]	8.52	7.94	9.21	8.82	8.28	9.47	8.21	7.58	8.94
Completed or is attending primary school	0.760	0.656	0.871	0.811	0.725	0.906	0.700	0.575	0.832
Attended secondary school during 1998-2008	0.450	0.331	0.578	0.510	0.406	0.626	0.379	0.241	0.523
Completed or is attending secondary school	0.389	0.292	0.493	0.461	0.366	0.565	0.305	0.203	0.412
Attended post-secondary school during 1998-2008[b]	0.048	0.019	0.079	0.059	0.028	0.094	0.035	0.009	0.062
Attended a vocational education program	0.258	0.253	0.264	0.238	0.239	0.238	0.281	0.270	0.293
Panel B: Grade repetition and dropout outcomes[c]									
Repeated a grade during 1998-2008	0.705	0.776	0.631	0.708	0.771	0.640	0.702	0.782	0.621
Dropped out of school during 1998-2008 (defn A)[d]	0.457	0.549	0.357	0.441	0.544	0.327	0.476	0.555	0.391
Among dropouts, percent that are permanent	0.721	0.750	0.671	0.597	0.635	0.526	0.854	0.884	0.809
Dropped out of school during 1998-2008 (defn B)[d]	0.208	0.240	0.174	0.180	0.207	0.150	0.242	0.280	0.202
Among dropouts, percent that are permanent	0.767	0.778	0.752	0.640	0.658	0.612	0.878	0.883	0.869
Panel C: Employment outcomes[e]									
Performs agricultural activities for household	0.665	0.671	0.660	0.634	0.627	0.643	0.700	0.720	0.678
Self-employed	0.132	0.143	0.120	0.139	0.150	0.127	0.124	0.135	0.112
Working for pay	0.170	0.167	0.173	0.249	0.252	0.245	0.078	0.064	0.093
Among employed, monthly salary (Ksh)	4159	3509	4833	4435	3484	5522	3132	3623	2778
Unemployed	0.247	0.259	0.234	0.252	0.267	0.235	0.242	0.250	0.234
Panel D: Other outcomes									
Member of a community group	0.581	0.547	0.617	0.617	0.583	0.656	0.538	0.503	0.574
Ever been married	0.386	0.420	0.349	0.293	0.318	0.265	0.494	0.542	0.444
Ever been pregnant[f]	0.490	0.544	0.432	0.378	0.421	0.330	0.620	0.691	0.546
Number of observations	1865	933	932	954	466	488	911	467	444

Notes: The sample used here includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score were dropped. All figures are weighted in order to maintain initial population proportions. Low ability individuals are defined as those with test scores at or below the median, while high ability individuals are those with above median test scores. All outcomes are at time of KLPS-2 survey unless otherwise specified.

[a] Educational attainment is coded according to the post-1985 school system, which includes eight years of primary, four years of secondary, and four years of university. Attainment is censored at 13 years, such that 12 years of schooling encompasses those who attended through the last year of secondary school and any Ugandan A-level schooling, and 13 years encompasses those who attended any amount of post-secondary school (including college and university).

[b] Post-secondary education figures should not include vocational education programs, but may include vocational education in a few cases due to changes in how attainment questions were framed between the KLPS Round 1 and Round 2 surveys.

[c] Grade repetition and dropout information is only calculated during years of primary and secondary schooling.

[d] Definition A counts as dropouts both those who were attending school and then failed to enroll the following year, as well as those who enrolled in school but only attended for a partial year – this measure thus takes into account both enrollment and actual attendance. Definition B includes only the former group, thereby only taking into account enrollment. Neither school dropout measures includes individuals who were in grades 8 or 12, which represent the end of primary and secondary schooling, respectively.

[e] Note that these categories are not mutually exclusive (although neither self-employed nor working for pay overlap with unemployed).

[f] For males, ever been pregnant relates to partners.

Table 5: Analysis of sub-sample attrition

	Dependent Variable: Indicator for Employed in Analysis		
	(1)	(2)	(3)
Female	-0.0369	-0.0124	-0.0209
	(0.0292)	(0.0343)	(0.0314)
Age (1998)	-0.018	-0.0166	-0.0183
	(0.0114)	(0.0128)	(0.0120)
Grade (1998)	-0.0067	-0.0079	-0.0063
	(0.0197)	(0.0255)	(0.0235)
School average mock score (1996)	0.0479	0.0506	0.0462
	(0.0507)	(0.0513)	(0.0476)
Weight-for-age z-score (1998)		-0.0098	-0.0162
		(0.0190)	(0.0169)
Falls sick often (1998)		0.012	-0.0047
		(0.0320)	(0.0299)
Household owns cattle (1998)		0.0361	0.0221
		(0.0342)	(0.0290)
No. of babies in compound (1998)		-0.0062	-0.0024
		(0.0187)	(0.0168)
ICS Test Score (1998)			-0.0371**
			(0.0158)
Observations	2296	1865	1711
Mean (std dev) of dependent variable	0.709 [0.454]	0.734 [0.442]	.822 [0.382]

Notes: The sample used here includes all KLPS-2, Wave 1 individuals who were found or searched for during intensive tracking. Robust standard errors, clustered within 1998 school, are in parenthesis. All figures are weighted in order to maintain initial population proportions. Additional controls include an indicator for missing age data and assigned years of treatment in the original deworming project.

[a] School average test score data has been converted into units of individual standard deviations.

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

Table 6: Effect of rainfall shocks and test score on attendance

	Dep Var: Indicator for Attended School During Year				
	(1)	(2)	(3)	(4)	(5)
Rainfall	0.0332**	0.0352**	0.0244	-0.0022	-0.0027
	(0.0146)	(0.0149)	(0.0181)	(0.0519)	(0.0564)
Rainfall * Test score (1998)		0.0127***		0.0078**	0.0089**
		(0.0039)		(0.0036)	(0.0041)
Rainfall * Female				-0.0332	-0.0530*
				(0.0222)	(0.0266)
Rainfall * Age (1998)				-0.0263***	-0.0262***
				(0.0022)	(0.0021)
Rainfall * Grade (1998)				0.0152*	0.0184*
				(0.0089)	(0.0093)
Rainfall * Father's education					0.0028***
					(0.0010)
Rainfall * Mother's education					-0.0012
					(0.0013)
Rainfall * School avg mock score (1996)					0.0071
					(0.0085)
Rainfall * Weight-for-age z-score (1998)					0.0042
					(0.0050)
Rainfall * Falls sick often (1998)					-0.0031
					(0.0069)
Rainfall * Household owns cattle (1998)					0.0154*
					(0.0088)
Rainfall * Test score (2nd quintile)			-0.0017		
			(0.0121)		
Rainfall * Test score (3rd quintile)			0.0165		
			(0.0126)		
Rainfall * Test score (4th quintile)			0.0097		
			(0.0115)		
Rainfall * Test score (5th quintile)			0.0281**		
			(0.0111)		
Includes child FE and year trend	Yes	Yes	Yes	Yes	Yes
Observations	19191	19191	19191	19191	16056
Number of Respondents	1865	1865	1865	1865	1565
R-squared	0.46	0.47	0.47	0.48	0.49
Mean (std dev) of dep var	0.61	0.61	0.61	0.61	0.61
	(0.49)	(0.49)	(0.49)	(0.49)	(0.49)

Notes: Results are from linear regressions, with robust standard errors (clustered by 1998 school) in parentheses. The sample used in columns (1)-(4) includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score were dropped. Column (5) additionally limit the sample to those with 1998 Pupil Questionnaire information. Rainfall is defined to include the long rains only, and is measured in standard deviations from 14-year zonal mean. Individuals drop out of the sample upon completion of secondary school. All columns control for gender*1998 grade*year. Columns (4)-(5) include a control for rainfall interacted with years assigned deworming treatment in 1998, and column (5) includes a control for rainfall interacted with number of small children in compound in 1998. All figures are weighted in order to maintain initial population proportions.

Table 7: Effect of positive rainfall shocks and test score on attendance

	Dep Var: Indicator for Attended School During Year			
	(1)	(2)	(3)	(4)
Rainfall	0.0503**	0.0557**	0.0139	0.0663
	(0.0236)	(0.0252)	(0.0292)	(0.1095)
Rainfall * Test score (1998)		0.0347***		0.0322***
		(0.0082)		(0.0099)
Rainfall * Female				-0.1020**
				(0.0408)
Rainfall * Age (1998)				-0.0304***
				(0.0049)
Rainfall * Grade (1998)				0.0137
				(0.0190)
Rainfall * Father's education				0.0031
				(0.0021)
Rainfall * Mother's education				0.0019
				(0.0030)
Rainfall * School avg mock score (1996)				0.0016
				(0.0222)
Rainfall * Weight-for-age z-score (1998)				-0.0006
				(0.0114)
Rainfall * Falls sick often (1998)				-0.0082
				(0.0156)
Rainfall * Household owns cattle (1998)				0.0297
				(0.0211)
Rainfall * Test score (2nd quintile)			0.0279	
			(0.0223)	
Rainfall * Test score (3rd quintile)			0.0455**	
			(0.0220)	
Rainfall * Test score (4th quintile)			0.0450*	
			(0.0231)	
Rainfall * Test score (5th quintile)			0.0893***	
			(0.0243)	
Includes child FE and year trend	Yes	Yes	Yes	Yes
Observations	19191	19191	19191	16056
Number of Respondents	1865	1865	1865	1565
R-squared	0.46	0.47	0.47	0.48
Mean (std dev) of dep var	0.61	0.61	0.61	0.61
	(0.49)	(0.49)	(0.49)	(0.49)

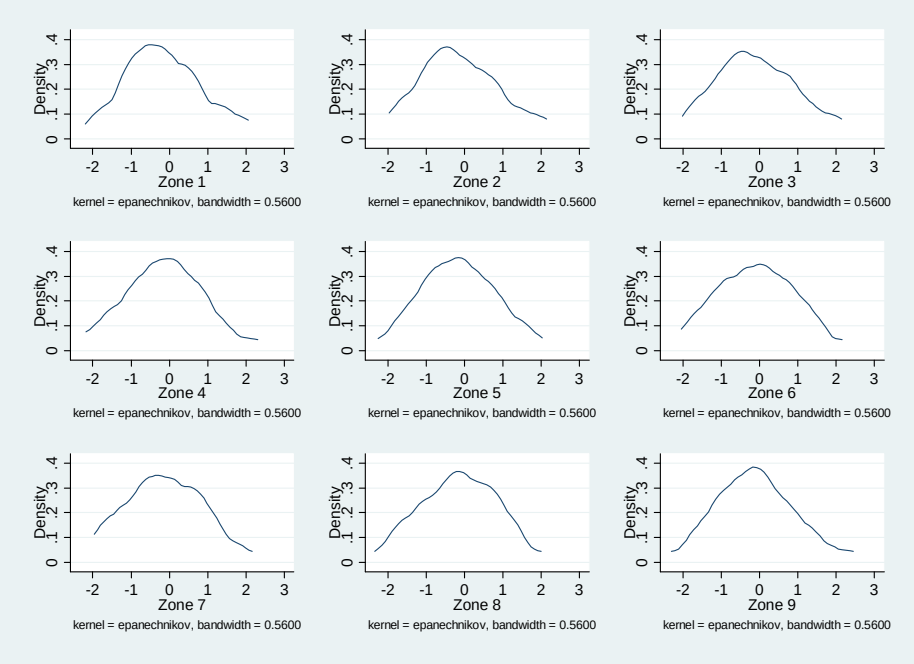
Notes: Results are from linear regressions, with robust standard errors (clustered by 1998 school) in parentheses. The sample used in columns (1)-(3) includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score were dropped. Column (4) additionally limits the sample to those with 1998 Pupil Questionnaire information. Rainfall is defined to include the long rains only, and is measured using an indicator for greater than or equal to one standard deviation above the long-run zonal average. Individuals drop out of the sample upon completion of secondary school. All columns control for gender*1998 grade*year. Column (4) includes a control for rainfall interacted with years assigned deworming treatment in 1998, and rainfall interacted with number of small children in compound in 1998. All figures are weighted in order to maintain initial population proportions.

Table 8: Effect of negative rainfall shocks and test score on attendance

	Dep Var: Indicator for Attended School During Year			
	(1)	(2)	(3)	(4)
Rainfall	-0.0641***	-0.0646***	-0.0708**	0.3975***
	(0.0202)	(0.0200)	(0.0292)	(0.1386)
Rainfall * Test score (1998)		-0.0156		-0.0029
		(0.0099)		(0.0106)
Rainfall * Female				-0.0365
				(0.0378)
Rainfall * Age (1998)				0.0566***
				(0.0058)
Rainfall * Grade (1998)				-0.0992***
				(0.0225)
Rainfall * Father's education				-0.0088***
				(0.0027)
Rainfall * Mother's education				0.0047*
				(0.0028)
Rainfall * School avg mock score (1996)				-0.0302
				(0.0205)
Rainfall * Weight-for-age z-score (1998)				-0.0104
				(0.0125)
Rainfall * Falls sick often (1998)				0.0019
				(0.0165)
Rainfall * Household owns cattle (1998)				-0.0264
				(0.0199)
Rainfall * Test score (2nd quintile)			0.0388	
			(0.0303)	
Rainfall * Test score (3rd quintile)			-0.0083	
			(0.0288)	
Rainfall * Test score (4th quintile)			0.0068	
			(0.0276)	
Rainfall * Test score (5th quintile)			-0.0125	
			(0.0270)	
Includes child FE and year trend	Yes	Yes	Yes	Yes
Observations	19191	19191	19191	16056
Number of Respondents	1865	1865	1865	1565
R-squared	0.46	0.46	0.47	0.49
Mean (std dev) of dep var	0.61	0.61	0.61	0.61
	(0.49)	(0.49)	(0.49)	(0.49)

Notes: Results are from linear regressions, with robust standard errors (clustered by 1998 school) in parentheses. The sample used in columns (1)-(3) includes all KLPS-2, Wave 1 surveyed individuals with 1998 PSDP test score information. Three observations with an extremely high test score were dropped. Column (4) additionally limits the sample to those with 1998 Pupil Questionnaire information. Rainfall is defined to include the long rains only, and is an indicator for greater than or equal to one standard deviation below the long-run zonal average. Individuals drop out of the sample upon completion of secondary school. All columns control for gender*1998 grade*year. Column (4) includes a control for rainfall interacted with years assigned deworming treatment in 1998, and rainfall interacted with number of small children in compound in 1998. All figures are weighted in order to maintain initial population proportions.

Appendix Figure 1: Kernel densities of standardized rainy season rainfall, by zone



Notes: Rainy season is defined to include the long rains only. Rainfall is standardized within zones over time. Densities include rainfall for 1998-2008.

Chapter 2

Individual Ability and Selection into Migration in Kenya

Joan Hamory[a] and Edward Miguel[b]

[a] University of California, Berkeley
[b] University of California, Berkeley and NBER

Chapter 2 Summary

This study exploits a new longitudinal dataset to examine selective migration among 1,500 Kenyan youth originally living in rural areas. More than one-third of individuals report moving to an urban area during the study period. Understanding how this migration differs for people with different ability levels is important for correctly estimating urban-rural wage gaps, and for characterizing the process of "structural transformation" out of agriculture. 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%. Results are robust to conditioning on household demographic and socioeconomic measures that might capture some aspect of credit constraints or household bargaining. In an interesting contrast with the existing literature, schooling attainment is not significantly associated with urban migration once cognitive ability is accounted for. In contrast, academic test score performance is not correlated with international migration to neighboring Uganda. 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.

1. Introduction

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 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.[27] 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.[28] 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 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.[29]

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

[27] Miguel and Kremer (2004) provide more background information on the PSDP.

[28] Note that this population is still fairly representative of the adolescent population in western Kenya: according to a Kenya Demographic and Health Survey, 85 percent of children in Western Province aged 6-15 were enrolled in school in 1998. However it should be noted that eighteen schools in Busia District were excluded from the sample, mainly because they were either economically and geographically unrepresentative of schools in the district or they had already received health and worm treatments under prior programs prior to the PSDP.

[29] Wave 2 of KLPS-2 data collection is currently underway, and will be completed in late 2009 and included in future analyses.

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.[30] 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.

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.[31]

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.

[30] 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.

[31] 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).

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 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.[32]

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:

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

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

[32] 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.

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.[33] 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 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.[34]

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.[35] 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

[33] 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).

[34] 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.

[35] 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.

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 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.[36] 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.[37]

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

[36] 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.

[37] 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.

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.[38] 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

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

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 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.[39] 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 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

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

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).[40]

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 (World Bank, 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.[41] 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% 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)

[40] 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).

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

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 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).[42]

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.[43] Among those living in a city, remuneration is nearly twice as high for men than for women. Recall that 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

[42] 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.

[43] Even replacing the cash salaries of those who report being unemployed with zero, the gap is similarly large, at Ksh 1,061 in rural versus Ksh 2,799 in urban areas.

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.

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) does not reveal that a single subject drives 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).[44] The test results support the earlier findings, with an

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

almost identical positive relationship between pupil test score and subsequent rural-urban migration (column 6) and weak interaction effects (column 7).[45]

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 an arguably preferable measure in the study of selection into urban migration in the Kenyan context, 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.[46]

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,

[45] 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).

[46] Given the high rate of enrollment in Kenya during early primary school, exam scores – especially for students in lower grades in 1998 – are more likely to reflect raw academic and cognitive ability rather than household factors which could influence school enrollment and attendance (such as wealth, results not shown).

average urban wages in our sample remain twice as large as rural wages (Table 15, column 1).[47] 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.[48]

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%.[49,50]

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 migration, our use of a true panel dataset of young adults over a decade and novel measures of

[47] Wages are measured here as the cash salary from primary employment among individuals who are employed. The analysis was repeated includes zero salary for individuals who are unemployed (no current job but are looking for work), and the results are substantively the same (see below).

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

[49] 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.

[50] Replacing wages with zero for unemployed individuals, the urban wage premium in column (3) is Ksh 1336.6, compared to Ksh 1471.7 in column (1). Thus, the main result is the same – considering observed schooling attainment reduces the urban wage premium by only 9.1%.

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.

Chapter 2 References

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.

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.

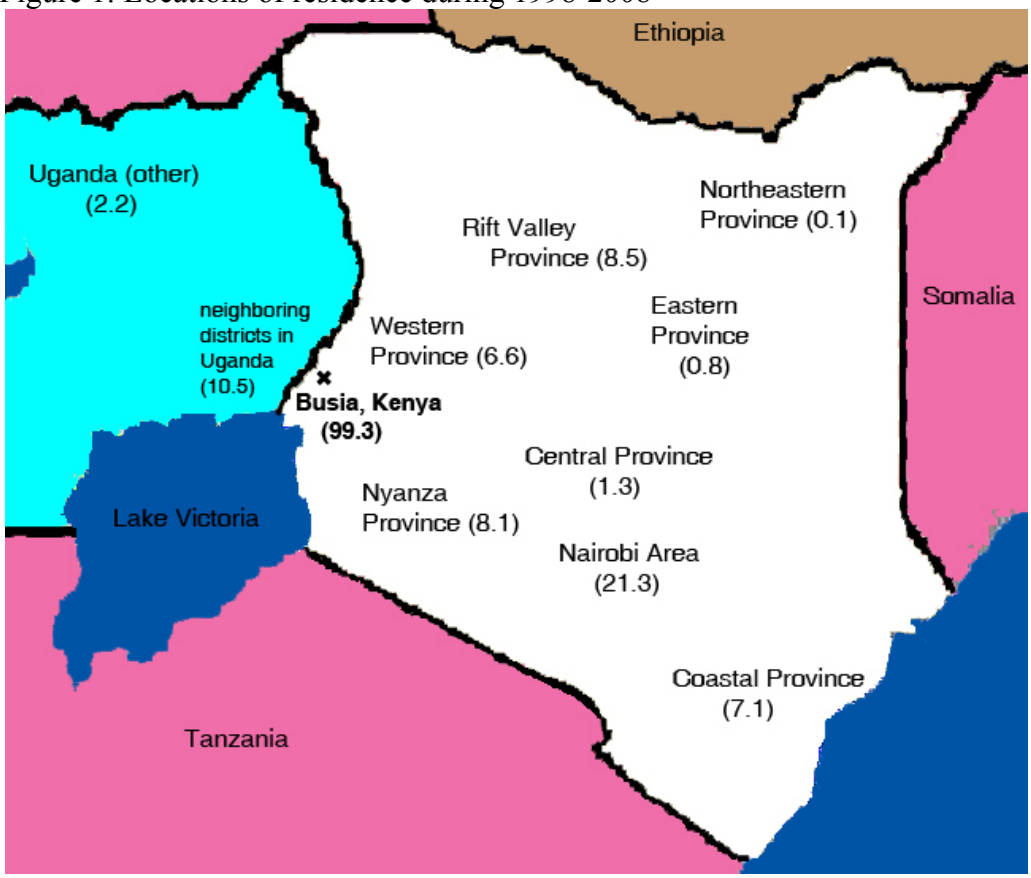
Republic of Kenya: Central Bureau of Statistics. (2002b). "1999 Population and Housing Census, Volume VI: Analytical Report on Migration and Urbanization." Government Printer: Nairobi.

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.

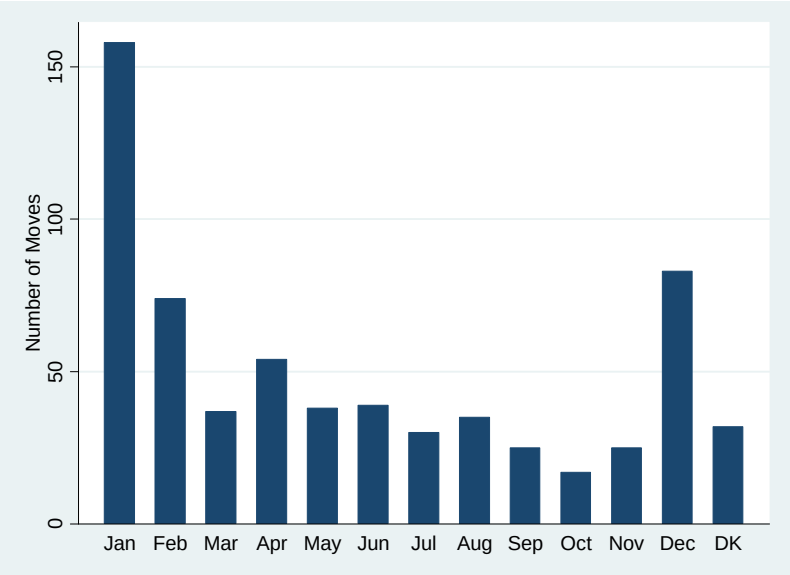
Republic of Kenya: Central Bureau of Statistics. (2005). *Statistical Abstract: 2005*. Government Printer: Nairobi.
Republic of Kenya: National Bureau of Statistics. (various years). *Economic Survey.* Nairobi.
Rosenzweig, Mark R. (1988). "Labor Markets in Low-Income Countries." In H. Chenery and T.N. 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.
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



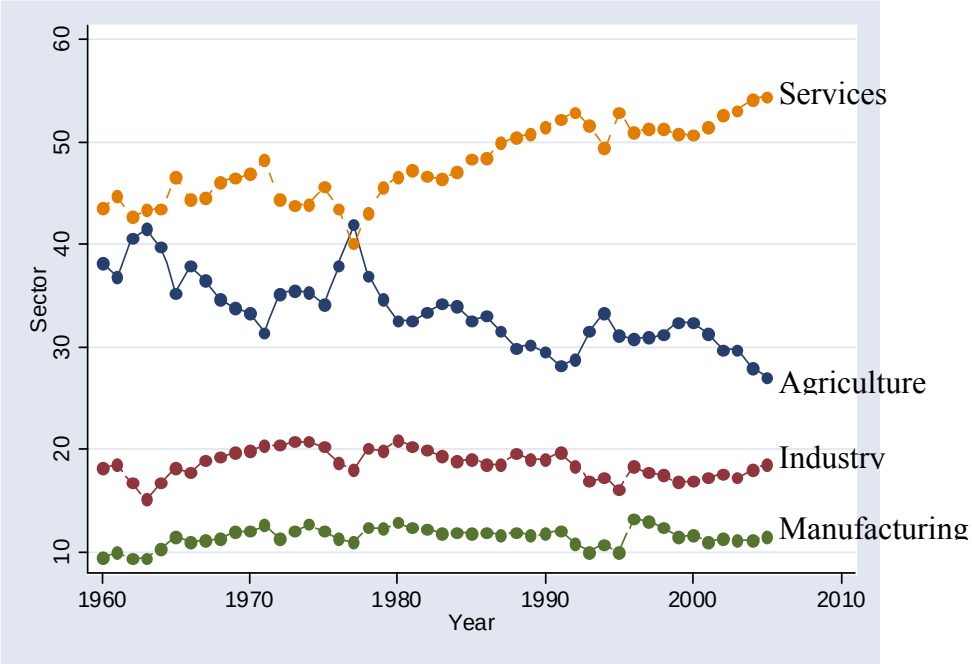
Notes: The sample here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as six observations missing date of survey or age information. Values signify 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



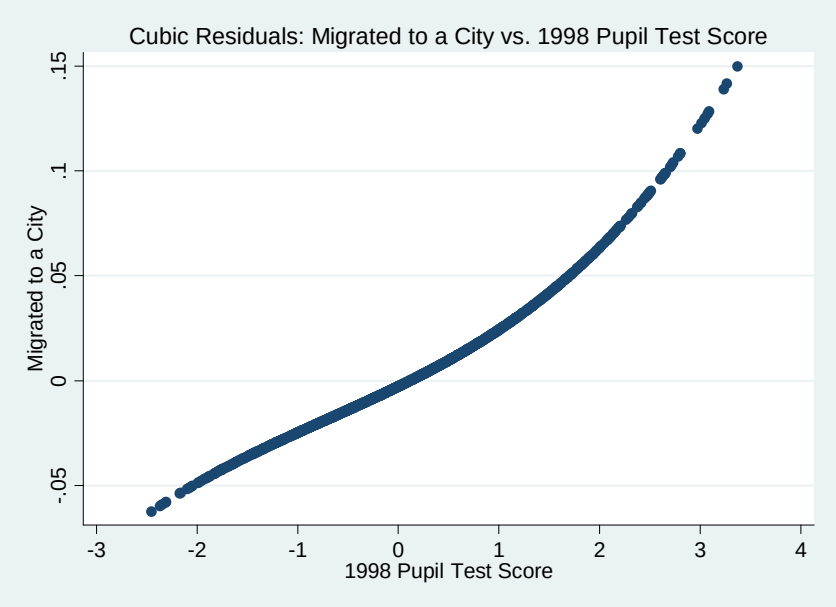
Notes: The sample used here includes all individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data, who were surveyed and report migration to a city during 1998-2008. Date of migration information is missing for 48 individuals. In addition, one observation with an extreme 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



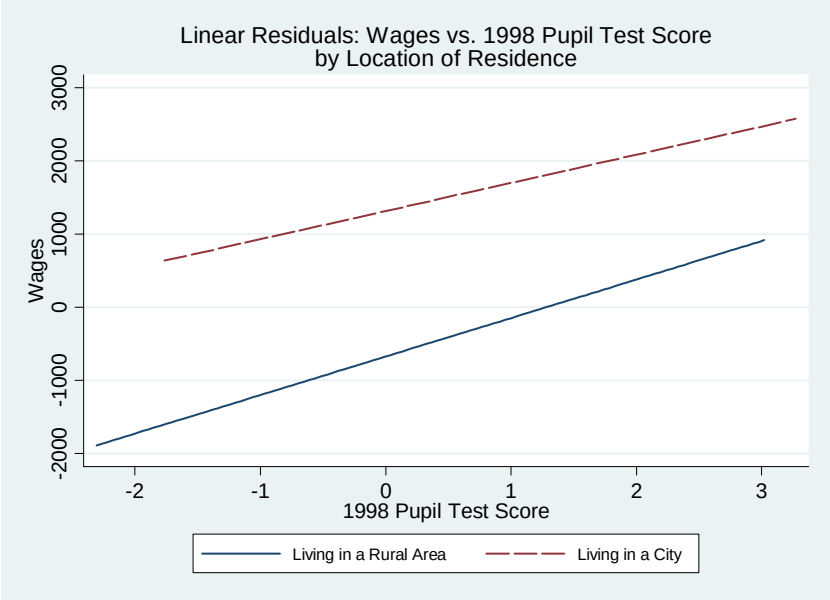
Source: World Bank (2007)

Figure 4: Cubic plot of urban migration on test score



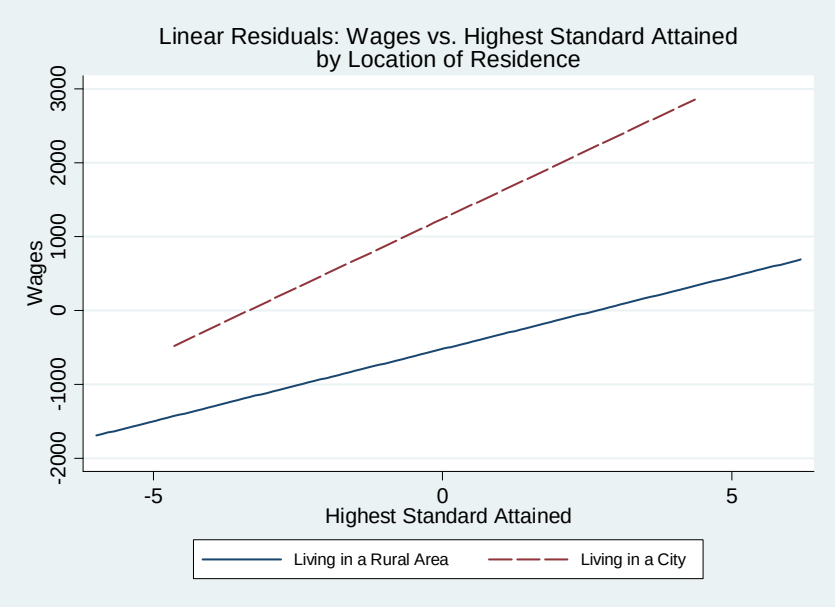
Notes: The sample used here includes all surveyed individuals with 1998 Pupil Questionnaire, school participation, and ICS test score data. One observation with an extreme 1998 ICS test score was dropped from the sample, as well as one observation missing date of survey and five observations missing 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



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. 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 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 near 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 (in Uganda)	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 (in Uganda)	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 to Uganda	Individuals Who Did Not Migrate to Uganda	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		Gender		1998 Age[a]	
	All	Female	Male	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[b]:					
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)[c]	2.25	2.10	2.41	1.85	2.62
Panel B: International Migration					
Individuals who live in Uganda during 1998-2008	0.142	0.128	0.155	0.121	0.172
Among those with information on date of move[d]:					
Number of total moves	2.46	2.42	2.49	2.40	2.51
Number of moves to Uganda	1.20	1.15	1.24	1.18	1.22
Length of stay in Uganda (yr)[c]	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] Median age in 1998 is 13.

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

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

[d] 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[b]	
	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[c]	434	218	216	192	242
Panel B: Among International Migrants (to Uganda)					
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[d]	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] Median age in 1998 is 13.

[c] 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.

[d] 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[a]	
	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[b]	0.499	0.438	0.381	0.503	0.238	0.605
In school[c]	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[d]	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[e]	0.234	0.300	0.292	0.309	0.306	0.294
Number of Observations	1274	244	127	117	106	138

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. Figures are weighted to maintain initial population proportions.

[a] Median age in 1998 is 13.

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

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

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

[e] 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			Among those living in a city:			
			Gender		1998 Age[a]	
	Living in a Rural Area	Living in a City	Male	Female	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.

[a] Median age in 1998 is 13.

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

Means			Among those living in a city:			
			Gender		1998 Age[a]	
	Living in a Rural Area	Living in a City	Male	Female	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)[b]						
Cash salary[c]	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] Median age in 1998 is 13.

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

[c] 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 in Uganda						
	(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 dependent variable	0.129	0.129	0.129	0.129	0.129	0.129	0.129
	[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 among employed individuals in the last month (results are nearly identical replacing cash salary with zero for unemployed individuals, see notes in text). 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

Chapter 3

Civil Unrest, Social Cohesion and Informal Credit in Rural Kenya

Chapter 3 Summary

This study examines the socioeconomic impacts of two months of protests and violent, primarily ethnic-based clashes that erupted across central and western Kenya in late 2007 following the controversial conclusion of a heavily-contested presidential election. Although not an epicenter of the conflict, Busia District experienced sporadic unrest, an influx of refugees from other parts of Kenya, inflation, supply shortages, and local market closures. Unique and timely survey data collected from young adults living primarily in this district of rural western Kenya in the months surrounding the election permits the use of both differences-in-differences and propensity score matching methodologies to estimate the short- to medium-run impacts of this conflict, and both approaches yield broadly similar findings. Despite little support for lasting effects on labor, migration and nutritional outcomes within weeks of cessation of the violence, there do appear to be persistent consequences for social cohesion and informal financial activities. While there is little indication of change in survey respondents' self-reported attitudes regarding trust of others, analysis confirms large declines in attendance at religious services, participation in community and bible groups, and utilization of non-family members as points of contact for future survey enumeration efforts. These findings highlight a disconnect between *reported* attitudes and *observable* behavior. Furthermore, respondents are between 29 and 53 percent less likely to engage in informal lending and transfers post-conflict. Given the key role played by social networks in informal financial markets in less-developed countries, these results indicate that even brief civil unrest may have lasting negative consequences.

1. Introduction

In recent years there has been a proliferation of research using micro-level data to examine the effects of civil war on economic development in Sub-Saharan Africa. The resulting evidence is mixed, with some studies finding reductions in nutrition, human capital accumulation, and labor market outcomes, and other studies finding no such impacts.[51] Results related to the effects of civil war on social behaviors such as political and community group participation are similarly mixed.[52]

This paper explores the socioeconomic impacts of civil unrest, which is generally much briefer and milder than civil war.[53] There are many reasons to think that civil unrest would have a different legacy than civil war, given its shorter duration, more limited area of coverage and lower overall level of destruction. However, due to a paucity of data, opportunities to study the consequences of this type of conflict have been extremely limited. This study aims to help fill this void using individual-level survey data spanning a brief period of civil strife in Kenya.

This paper makes three main contributions. First, employing data from a unique and timely survey permits the use of both differences-in-differences and propensity score matching methodologies to estimate the short- to medium-run impacts of the conflict, where the "treatment" is the timing of survey enumeration. This presents a natural opportunity to compare both approaches in a real life setting, and to leverage their respective benefits and pitfalls in order to obtain useful insights. Both methods yield broadly similar findings, lending credibility to the results. Second, while regression analysis indicates little change in respondents' self-reported attitudes regarding trust of others, results confirm large declines in attendance at religious services and bible groups, participation in community groups, and utilization of non-family members as points of contact for future survey enumeration efforts. These findings indicate a disconnect between *reported* beliefs and *observable* outcomes, suggesting important limitations for direct inquiry of individuals' opinions. Finally, regression analysis reveals that respondents are between 29 and 53 percent less likely to engage in informal lending and transfers post-conflict. Given the key role played by social networks in informal financial markets in less-developed countries, these changes indicate that even brief civil unrest may have persistent negative consequences.

In particular, this study examines the impacts of violent conflict that erupted across Kenya in late December 2007 following the conclusion of a heavily-contested presidential election. The proximate causes of the unrest can be tied to the election itself as feelings of mistrust swelled over reports of vote tampering, delays in poll announcements, and discrepancies between results broadcast first locally and then nationally. The incumbent candidate Mwai

---

[51] For example, Blattman and Annan (forthcoming) find large negative effects of civil war in Uganda on education, employment and earnings of youth up to ten years after initial exposure to violence. Bundervoet *et al.* (2009) and Akresh *et al.* (forthcoming) find substantial negative consequences of civil war on child nutrition in Burundi up to four years after first exposure to war and Rwanda up to two years after first exposure, respectively. In contrast, Bellows and Miguel (2009) find no impact on household asset ownership, consumption or child nutrition in post-conflict Sierra Leone two to three years after the end of the conflict.

[52] For instance, Blattman (2009) reports that exposure to civil war in Uganda increases individual political participation, but has no effect on participation in other types of community groups. Meanwhile, Bellows and Miguel (2009) conclude that exposure to civil war in Sierra Leone is associated with a significant increase in community and political participation three to five years after the end of the conflict.

[53] According to Blattman and Miguel (forthcoming), classification of an event as a "civil war" requires at least 1,000 battle deaths within a single year, and classification as "civil conflict" requires 25 battle deaths. The term "civil unrest" is used in this paper to represent tension within a country that does not manifest itself in outright battle, but nevertheless can be quite violent and destructive.

Kibaki's victory over Raila Odinga, who had been ahead in the polls leading up to the election, was a surprise to many. Furthermore, it appears that deep-seated ethnic frictions further fueled the unrest in Kenya, consistent with findings from related literature examining elections in Africa which has found that ethnic tensions tend to run high close to elections.[54] Conflict broke out immediately upon announcement of the election results, and continued until a government power-sharing agreement was brokered in late February 2008. By this time, estimates put the death toll well above 1,000 and the number of internally displaced persons as high as 500,000 (Bayne, 2008).

Anecdotal accounts in the popular media suggest that the election crisis reduced social cohesion within and across communities. For instance, reports describe neighbors belonging to different ethnic groups turning on one another in the days and weeks following the election (Elkington, 2009). In addition, there is evidence that this may persist long after the measurable end of the conflict with displaced persons refusing to return home more than eighteen months after the signing of the power-sharing agreement, for fear that members of their community will harass them once again (*BBC News,* 2009).

If this brief episode of civil unrest did damage social cohesion in Kenya, there may be significant and lasting implications. Miguel and Gugerty (2005) find that regions of rural western Kenya with less social capital (defined by the level of local ethnic diversity) have lower public goods provision. Narayan and Pritchett (1999) explore social capital (defined as attitudes toward others and participation in local groups) in rural Tanzania, and results indicate a positive effect of village-level social capital on the income of households within that village. The authors suggest that one possible mechanism through which social capital may affect incomes is by augmenting informal credit arrangements.[55]

In regions where formal credit and insurance markets are thin, relatives, friends, neighbors and other acquaintances frequently provide assistance to one another in the form of transfers and informal lending. Because of the value of social connections in this context, conflict affecting such relationships can have potentially large negative socioeconomic impacts. Where households are unable to obtain outside assistance to smooth their consumption, they may be forced to pull children out of school (potentially reducing human capital formation) or to engage in risky commercial sexual behaviors (which may increase rates of HIV transmission, and carry other physical and psychological risks).[56]

This study exploits unique individual-level survey data collected between August 2007 and September 2008 to study the short- to medium-run social and economic impacts of the Kenyan election crisis. Nearly 80 percent of individuals included in this analysis live in Busia, a rural district of western Kenya. The market center of this district is Busia Town, a populous municipality located on the border with Uganda which serves as a stop on the major trucking route running between Nairobi and Kampala. Although just 15 percent of respondents interviewed after the election report being directly affected by the violence through personal assault, theft or arson, all were indirectly influenced by the conflict in their area and beyond.

[54] See Eifert *et al.* (forthcoming) for a cross-country analysis, and Lewis (2007) on Nigeria.

[55] However, it should be noted that there is a debate over which aspects of social capital are important for development. Employing a cross-country framework, Knack and Keefer (1997) find that trust and "civic norms" (an index from various questions related to social honesty, such as how likely one would be to cheat on his taxes or avoid paying fare on public transport) are significantly associated with positive per capita income growth, while community group membership is not associated with either trust or income growth gains. (Note that the sample of countries utilized in this analysis only contains two from Sub-Saharan Africa.)

[56] See, for example, Jacoby and Skoufas (1997) and Dupas and Robinson (2009).

Busia District witnessed unrest (anecdotal accounts cite piles of burning tires blocking the main road and businesses set aflame), refugees arriving from other parts of Kenya, inflation, supply shortages, and local market closures. A number of the survey respondents who had previously migrated outside Busia primarily inhabit urban areas elsewhere in Kenya, many of which were hotspots of protest and violence. Seventy-one percent of all individuals interviewed post-election describe feeling "very" concerned for their safety, and 95 percent report being at least "somewhat" concerned.

Combining this survey data with information collected from these same individuals in 2003-2004 permits the use of both differences-in-differences and propensity score matching methodologies to estimate effects of the crisis. Results obtained from both approaches are broadly similar. Surprisingly, there is no discernable evidence of a decline in individual employment, migration and nutrition up to seven months after the conclusion of the violence. Among the subset of individuals interviewed starting just one month after the conclusion of the crisis, 87 percent report that "things have returned to normal." Related work in the same district found large decreases in income, expenditures and consumption *during* the conflict, particularly in January, but with these declines tapering off and even reversing in February and March (Dupas and Robinson, 2009). Together, these findings suggest that there may have been a quick recovery across these particular outcomes.[57]

Despite a dearth of evidence of persistent impacts of the election on these outcomes, there is substantial support for enduring changes in social behaviors. Nearly 96 percent of the sample belongs to the same ethnic group (tribe) – the largest in this region of Kenya – so this analysis examines effects of the crisis within a single ethnic group, rather than across groups.[58] In spite of this, results still indicate sizeable decreases in participation in community and religious groups, and in the probability of listing non-family members as points of contact for future data collection efforts. Moreover, results suggest considerable reductions in the level of informal borrowing, lending and transfers between households. In a region with extremely limited access to formal credit institutions, and where informal credit plays an important role in household investment and consumption smoothing activities, such findings could suggest important negative long-term effects of political violence.

This paper proceeds as follows. Section 2 describes the data employed in the analysis. Section 3 outlines the 2007 Kenyan presidential election crisis, with specific emphasis on the primary study area in rural western Kenya. Section 4 presents the empirical analysis and section 5 provides conclusions and recommendations for future work.

2. The Kenyan Life Panel Survey

This analysis utilizes data from the Kenyan Life Panel Survey (KLPS), a longitudinal survey following 7,500 individuals in three rounds of data collection over the period 1998-2009. The KLPS sample was drawn from all youth confirmed enrolled in 75 primary schools, grades 2 through 7 in two divisions of Busia District, a rural district of western Kenya, in 1998. This sample is fairly representative of the adolescent population in western Kenya: according to a Kenya Demographic and Health Survey, 85 percent of children in Western Province aged 6-15

[57] It should also be pointed out, however, that Dupas and Robinson (2009) focus their analysis on a sample of wage-earners living in Busia town, while the present analysis includes a representative subset of young adults living in the district as a whole (many of whom are still in school, or engage primarily in subsistence farming for their household).

[58] There is more variation in ethnic grouping at the sub-tribe and clan level.

were enrolled in school in 1998. However it should be noted that eighteen schools in Busia District were excluded from the sample, mainly because they were either economically and geographically unrepresentative of schools in the district or they had already received health and worm treatments under prior programs prior to the PSDP (see below).

The first round of data collection occurred in 1998, and was undertaken in order to evaluate the introduction of an intestinal helminth treatment program known as the Primary School Deworming Project (PSDP) that was introduced in Busia District.[59] Detailed information was collected on individual and household characteristics for nearly 30,000 primary school students. Five years later a follow-up survey entitled the Kenyan Life Panel Survey Round 1 (KLPS-1) was launched, in an effort to track a representative subset of 7,500 individuals involved in the PSDP. Enumerators traveled throughout Kenya and into neighboring Uganda to locate individuals who had migrated from their 1998 location. Between 2003 and 2005, data on a wide range of outcomes was successfully collected for over 5,200 of these young adults, including information on education, migration and labor market participation, as well as ethnic and religious identification, community group participation and trust of others.

In mid-2007, a second round of the Kenyan Life Panel Survey (KLPS-2) went to the field, and the entire sample of 7,500 adolescents was sought once again. Individuals were randomly divided into two groups, to be tracked in two separate waves of data collection, both of which are individually fully representative of the main sample. The timing of the first wave of KLPS-2 enumeration is particularly fortuitous with respect to our ability to analyze the election violence. Specifically, Wave 1 of the KLPS-2 continued from August 2007 through September 2008, encompassing the months plagued by violence related to the December 2007 national election. The final KLPS-2 Wave 1 dataset contains survey information for nearly 2,500 individuals.[60] In addition to collecting a second round of information on education, migration, labor market participation, ethnic and religious identification, community group participation, and trust of others, the KLPS-2 survey contains a rich set of questions on informal credit and transfers.

The 2007 Kenyan presidential election occurred during the middle of KLPS-2 Wave 1 enumeration, and approximately half of respondents were surveyed on either side of the election. This provides a rare opportunity to explore the effect of election-related violence on the short- to medium-term outcomes of individuals. To this end, the present analysis utilizes data from the original PSDP, the KLPS-1 and the first wave of the KLPS-2, with particular focus on two groups: all 2,500 respondents of the KLPS-2 Wave 1 survey, and the sub-set of over 1,900 respondents interviewed in both KLPS-1 and KLPS-2. The former is of interest because it allows for matched propensity score analysis of survey responses just before and after the crisis. The latter is of particular value because it provides a pre- and post-crisis panel for some PSDP individuals, and permits differences-in-differences estimation methods. Consistency of findings across both econometric methodologies lends further credibility to the results.

3. 2007-2008 Election-Related Violence in Kenya

Just over four months into the first wave of KLPS-2 survey enumeration, a presidential election took place in Kenya. The incumbent candidate Mwai Kibaki's heavily-contested victory over Raila Odinga in the election of December 27, 2007 was followed by widespread protest, violence

[59] Miguel and Kremer (2004) provide more background information on the PSDP.

[60] KLPS-2 Wave 2 data collection is currently underway, and will be completed in late 2009 and included in future analysis.

and displacement. While previous national elections in the 1990s and in 2002 were catalysts for smaller outbreaks of localized unrest, the scale of fallout from the 2008 election was broadly unanticipated. Reputable estimates put the death toll well above 1,000 and while estimates vary, the number of internally displaced persons has been cited as high as 500,000 (Bayne, 2008, Gibson and Long, 2009).

Kenya encompasses 42 distinct ethnic groups, and although none have majority power in national elections, five groups account for approximately 70 percent of the population (Kanyinga and Nyamweya, 2008). In the 2007 election, the presidential race was dominated by two political parties, one commonly associated with the Kikuyu tribe, and the other with the Luo tribe. President and Kikuyu leader of the Party for National Unity (PNU) Mwai Kibaki was pitted against Raila Odinga, a Luo and candidate for the Orange Democratic Movement (ODM). Ethnic differences between the candidates and in their support bases helped spark the fervent response to the contested election.

In the hours and days immediately following the election, suspicions regarding the legitimacy of the process grew. Media coverage detailed delays in reporting of results, discrepancies between local and national vote counts, the high numbers of voters participating in the presidential race compared to local races, and a higher than 100 percent turnout in one constituency. Skepticism was reinforced when a majority of seats in Parliament went to the ODM (nearly two-thirds for the party and its supporters), and Kibaki was announced winner by only a small margin of victory (230,000 votes out of nearly 10 million cast) even though Odinga had led in opinion polls prior to the election (Thompkins, 2007, *BBC News,* 2008). Indeed, recently published results utilizing exit poll data conclude that Odinga won the election with a vote share of 46.1 to 40.2 percent (official results showed a loss for Odinga with a vote share of 44.1 to 46.4 percent) and conclude "…the [Electoral Commission of Kenya] results are off by enough of a margin to have declared the wrong winner" (Gibson and Long, 2009, p. 499).

Meanwhile, underlying the contested ballot box exercise was a deeper context of tension, springing from historical land disputes and other grievances as well as from perceived income and political inequality across ethnic groups (Doyle, 2008, Kanyinga and Nyamweya, 2008). Reports have surfaced of radio station broadcasts and text messages exploiting these feelings (among several ethnic groups) in the wake of the election (Bayne, 2008). Nonetheless, the timing of the violence suggests that the primary catalyst for the outbreaks were election-related. Although some violence preceded the election – according to Thompkins (2007), there were approximately 20 deaths and a few hundred individuals displaced in the weeks leading up to the contest – the most extreme unrest occurred after the announcement of outcomes.

Civil unrest subsequent to the election was widespread. Immediately following announcement of the results and continuing in waves throughout January 2008, violence broke out in ODM strongholds, including the slums of Nairobi as well as towns across Nyanza, Western and Rift Valley Provinces. This hostility, targeted against Kikuyus and other PNU supporters, largely took the form of burning, looting, and placing road blocks. These activities resulted in a number of fatalities. At the same time, more systematic violence erupted in the Rift Valley Province, with organized groups of Kalenjin youth, possibly coordinated and financially supported by local leaders, deliberately driving Kikuyu and other PNU supporters out of their homes.[61] By late January, reports of retaliatory attacks led by groups of Kikuyu youth against ODM supporters surfaced as well. Attempted police suppression of protests also resulted in

[61] A nationwide pre-election survey indicated that Kalenjins largely supported Odinga (Dercon, 2008).

several fatalities (Bayne, 2008). Violence largely came to a halt after a government power-sharing agreement was reached at the end of February 2008.

As a result of the conflict, the total number of internally displaced persons (IDPs) reached nearly half a million. Some of these individuals were migrant workers, who simply returned home to weather the turbulent months of January through March 2008. Others were persons forced out of their permanent homes and moved to sponsored refugee camps or other, more sympathetic communities. Caution over resettlement persisted for several months after the end of the conflict. In April 2008, over 150,000 individuals were still in refugee camps across the country (Kenya Red Cross, 2008a). Although reports are mixed, sources indicate that resettlement was more rapid in ensuing months – the number of IDPs had reduced by one-half as of late June 2008 (Kenya Red Cross, 2008b), and by an additional two-thirds by the end of July 2008 (OCHA-Kenya, 2008a). However, a recent news article suggests that many displaced persons, fearful of their personal security, are still refusing to return to their pre-election residence more than a year and half after the signing of the government power-sharing agreement (*BBC News*, 2009).

Western Kenya, and Busia District in particular, was not spared the crisis. Respondents of the KLPS-2, residing predominantly within the Busia area, were affected in a number of ways. The first was violence. Anecdotal reports of hostility, burning of businesses, and blockages along the road leading to Uganda abound. DELPHA (2008b) classifies Busia as a "hotspot" for violence between late December 2007 and late January 2008. Although just 15% of KLPS respondents interviewed after the election report being directly affected by the violence through personal assault, theft or arson, individuals were clearly influenced by the unrest in their area and beyond – 71 percent were "very" concerned for their safety, and an additional 24 percent "somewhat" concerned. Two IDP camps were established within Busia District, one at the police station in Busia Town, housing over two hundred IDPs by late January 2008, and another in the central part of the district (UNHCR, 2008, DEPHA, 2008a). Busia was also the departure site for many refugees crossing the border into Uganda (ReliefWeb, 2008). Because of this, Busia residents interacted daily with refugees from more distant regions of Kenya. Furthermore, survey evidence collected from residents of Busia in 2008 suggest that civil unrest led to local market closures, reduced the supply of goods, sparked inflation, and lowered incomes (Dupas and Robinson, 2009, and author's own findings using KLPS-2 survey data).

Due to concerns for the safety of project staff, KLPS-2 enumeration activities stopped during the several weeks immediately following the election. Because the bulk of violence ended on February 28, 2008, when a power-sharing deal was brokered between Kibaki and Odinga, and by the time KLPS-2 surveying recommenced in late March 2008, local markets were already reopened and prices were declining back toward pre-election levels. In fact, 87 percent of respondents interviewed following the post-election crisis between May and September of 2008 report that things had returned to normal by the time of their survey (even among those surveyed in earlier months).[62]

4. Empirical Analysis

4.1 Attrition in the KLPS-2 and Panel Samples

Before moving to the main empirical analysis, we must first address the issue of sample attrition. Searching for individuals in rural Kenya is an onerous task, and migration of target respondents

[62] It should be noted that the IDP camp in Busia Town was still open as of mid-June 2008 (OCHA-Kenya, 2008b).

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 late teens and early twenties – an age group likely to be extremely mobile due to marriage, schooling, and labor market opportunities. Nearly 81 percent of individuals sought for the first wave of KLPS-2 enumeration were surveyed, and an additional 4 percent were found but unable to be surveyed.[63,64] To the extent that attrition is non-random (in other words, the 19 percent of individuals not surveyed differ from those surveyed), analysis using the KLPS-2 data could be biased.[65]

Column (1) of Table 1 provides results from a regression of an indicator for being surveyed during the first wave of the KLPS-2 on a set of individual characteristics collected as part of the 1998 PSDP evaluation. There are no measurable differences across survey status along dimensions of gender, age, 1998 grade level, 1996 school average test score (a proxy for school quality), or length of assignment to deworming treatment. Individuals who attended school in Budalang'i Division, the southernmost portion of Busia District abutting Lake Victoria, are nearly 8 percent less likely to be surveyed. This could be because of the difficulty in locating respondents still living in this region (which includes remote islands in the lake as well as areas on the mainland that become swampy and difficult to travel during the rainy reasons), or because more of these respondents have migrated away from their 1998 home. Among surveyed individuals, migration rates are slightly lower for those initially attending school in Budalang'i than those attending school elsewhere – with a mean of 0.567 (standard deviation 0.496) versus 0.596 (0.491) – lending more credence to the former explanation.

Together these results fail to provide strong evidence of differential attrition in the KLPS-2 Wave 1 survey along observable characteristics, other than 1998 schooling location within the district. It is unclear what effect such differential attrition could have on the results of the analysis utilizing this sample. To the extent that individuals attending school in Budalang'i Division in 1998 still live there, and were more sheltered from the election crisis due to their remote location, estimates obtained in the main analysis may be biased slightly upward.

Column (2) of Table 1 performs this same regression, this time with the dependent variable defined as an indicator for whether the individual was surveyed in the first wave of KLPS-1 data collection. Again, the results suggest little evidence of differential attrition. Only age of individual appears to be significantly related to likelihood of survey, such that one additional year of age reduces the probability of being surveyed by 3 percent.

Column (3) provides analogous regression results for the sample of individuals sought in both KLPS survey rounds. These results suggest a greater degree of differential attrition in the panel – females are nearly 7 percentage points less likely to have been surveyed in both rounds,

[63] KLPS-2 Wave 1 tracking launched in Fall 2007 and ended in November 2008. During the first several months of data collection, 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 for the KLPS-2 sample 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).

[64] Typical reasons for not surveying an individual who had been located were: refusal of individual, refusal of parent, individual mentally incapable of being interviewed, individual physically incapable of being interviewed (e.g., imprisoned), or individual deceased.

[65] Similarly, 84 percent of individuals sought in the first wave of the KLPS-1 were located. For a more detailed analysis of survey data and attrition in the KLPS-1, see Baird *et al.* (2008).

and older children and those initially in higher grades are each 2 percentage points less likely. These findings are unsurprising, due to the higher rates of migration within these groups.[66] There is no differential attrition across initial deworming treatment assignment, but individuals attending better schools in 1998 are slightly less likely to be surveyed. Interestingly, adolescents attending school in Budalang'i Division are slightly *more* likely to be surveyed in the panel – perhaps because, as noted above, migration rates for this group are slightly lower, and hence these individuals are easier to find repeatedly.

Together the results presented in Table 1 provide limited evidence of differential survey attrition within each survey round, but suggest a significant amount of differential attrition in the longitudinal sample. Since information for individuals in the latter sample will be used in a differences-in-differences framework, a key concern is that results from that analysis will not be generalizable to the population of adolescents in rural western Kenya at large. Additional estimation results will be provided using the entire KLPS-2 sample (employing a propensity score estimation strategy) in order to assuage this fear.

4.2 Summary Statistics

Table 2 provides summary statistics for respondents in both samples of interest.[67] Column (1) summarizes characteristics for respondents of the KLPS-2 Wave 1. Just under half of the sample is female. Individuals are between the ages of 15 and 31, and are on average 22 years old, at the time of survey. Nearly all respondents are members of the Luhya tribe, the largest ethnic group in this part of Kenya. Only 3 percent are Luo, and none are Kikuyu. Therefore, this analysis will focus on the effects of the election crisis within a single ethnic group, rather than across groups. On average respondents were assigned to receive 3.5 years of deworming treatment through the PSDP. Just over one-third attended school in Budalang'i Division, the southernmost part of Busia District, in 1998.

The sample is almost entirely of Kenyan birth, though 10 percent report a natal district outside of Busia. Approximately 95 percent of respondents are living in Kenya at the time of interview, with the vast majority in Busia (results not shown). As described previously, this district experienced many aspects of the election crisis first-hand. Just 15 percent of respondents are living in an urban area at the time of survey, and most of these are in Nairobi – another hotspot for civil unrest. Thus, most individuals in the sample lived in close proximity to the conflict.

Nearly 42 percent of individuals were still in school in 2006. Average attainment at that time was 8.4 years (which just exceeds graduation from primary school). In fact, more than three-quarters of the sample had finished primary school, but less than half had either attended or completed secondary school. Average father's education is 8.7 years (nearly two years of secondary schooling under the previous educational system), and mother's is 5.6 (more than one year shy of completing primary school).

At the time of KLPS-2 survey, just over one-third of respondents are living in a house with a cement floor (the alternative typically being mud), and only 15 percent have access to electricity at their place of residence. Two-thirds of households have a telephone (generally of the mobile variety), and over 83 percent report ownership of a TV or radio (primarily the latter).

---

[66] For more on this, see Hamory and Miguel (2009).

[67] All information provided in this section characterizes individuals at the time of KLPS-2 enumeration, unless otherwise noted.

Column (5) of Table 2 provides summary statistics for the sample of individuals with panel information. Overall the characteristics of these individuals track those in the full KLPS-2 Wave 1 survey sample fairly well – the only departure of note appears to be that those with panel data are on average less likely to have migrated since 1998. This is not an unexpected finding, since those who have not moved are easier to find in each survey round.

4.3 Overview of Main Analysis

The goal of this analysis is to evaluate the short- to medium-term impacts of civil unrest on socioeconomic outcomes of KLPS-2 survey respondents. It is not possible to simply compare outcomes for individuals interviewed just prior to the outbreak of violence to those interviewed afterward, as such an approach assumes that survey respondents are randomly assigned with respect to the "treatment" (timing of enumeration). This comparison is not informative if individuals interviewed before the election have different characteristics from those interviewed afterward.

And indeed they do have different characteristics. During the first several weeks of KLPS-2 field work, survey enumerators traveled around Busia and nearby districts, locating respondents still living in the local area and gathering tracking information from family members and neighbors of individuals who had left. Once most respondents believed to be living in and around Busia had been sought, tracking was expanded into more distant parts of Kenya and Uganda. As a result of this tracking methodology, individuals found prior to the election are less likely to have migrated, and may be different across a range of other observable and unobservable characteristics as well.

Columns (2) and (3) of Table 2 provide selected summary statistics for individuals and their households according to the timing of survey enumeration, and the p-values from tests of equality of means listed in column (4) suggest that there are many statistically significant differences between these groups. These figures indicate that individuals interviewed after the election are on average more likely to be female, older, have fewer years of deworming treatment under the PSDP, and have attended school in Budalang'i Division in 1998. These respondents are much less likely to have been born in or around Busia District, and much more likely to have migrated during 1998-2006. They are substantially less likely to be living in Busia District at the time of interview and more likely to be found living in an urban area (results not shown). Finally, the homes of respondents located post-election are much more likely to have cement floors, electricity, and access to a telephone.

In order to properly identify the causal effect of civil unrest on individual behaviors in the presence of nonrandom treatment assignment, two different evaluation methods will be used – a differences-in-differences approach exploiting the panel data available for a subset of sample individuals, and a propensity score approach attempting to match individuals across treatment groups based on observable characteristics. Each methodology has its own benefits and drawbacks. The panel approach includes respondent fixed effects, and thus controls for any unobserved characteristics that may be jointly related to exposure to violence and the outcomes of interest. This is not possible in propensity score estimation, and in fact the latter approach is less informative in the presence of any such unobserved characteristics. Furthermore, panel estimation allows for the inclusion of month of survey fixed effects to account for any seasonal variation in outcomes – which is very useful if particular socioeconomic outcomes display a seasonal component. In contrast, propensity score analysis allows us to make use of the entire sample of KLPS-2 respondents, rather than forcing focus on a restricted sample of individuals

interviewed in both KLPS survey rounds. This is a plus because, as we have seen, this subset of individuals is not completely representative of young adults in rural western Kenya.

Given that neither econometric approach is entirely satisfactory on its own, both methods will be presented here in order to compare results. The following two sub-sections describe each methodology in turn.

4.4 Differences-in-Differences Approach

A subset of KLPS-2 respondents were also interviewed five years earlier in KLPS-1, and panel data exists on various social attitudes and behaviors for over 1,900 individuals (78% of KLPS-2 Wave 1 respondents). Among these surveys, nearly half of KLPS-2 interviews were performed prior to the December 2007 election, and the other half performed subsequent to the election. Such longitudinal information can be exploited to estimate the causal effect of election violence on individual behavior using a differences-in-differences estimation approach.

Consider the following specification:

$$y_{it} = \alpha + \beta_1 T_{it} + \beta_2 X_{it} + \delta_i + \gamma_t + \varphi_j + \varepsilon_{it} \quad (1)$$

where *i* indexes individuals and t indexes year of survey. y is an outcome of interest, such as an indicator for membership in a local community group or for the belief that individuals of other ethnic groups can be trusted. T is the treatment, defined as an indicator for a survey conducted after the December 2007 election. X is a vector of individual characteristics that vary over time, and includes age, school attainment, characteristics of the home (availability of electricity, latrine, telephone, and TV/radio), as well as indicators for residence in Kenya and/or in an urban area. δ_i is an individual fixed effect, γ_t is a survey round identifier (KLPS-1 or KLPS-2), and φ_j denotes month of survey (included to control for any seasonal differences in outcomes). Under the identifying assumption that, in the absence of the election, individual behaviors would not have evolved in a systematically different way, β_1 can be interpreted as the causal effect of election violence on an outcome y.

The left-hand side of Table 3 displays results using the specification defined in equation (1) for various measures of labor, migration and nutrition. The study area is largely rural, and small-holder subsistence farming is common. Ninety percent of individuals interviewed in KLPS-2 prior to the December 2007 election come from farming households, and only one-quarter of individuals work outside of the household farm, either for themselves or someone else. One-third of respondents have migrated at some point since initial observation in 1998, and 10% have migrated to an urban area. Finally, individuals consume on average just two meals per day, less than one of which contains any meat or fish.

Regression analysis indicates that there were no significant changes in any of these labor, migration, or nutritional outcomes associated with the election violence. Dupas and Robinson (2009) surveyed a sample of self-employed individuals in Busia District during the crisis, and found large reductions in income, expenditures and food consumption in January due to the unrest. However, by February and March, the authors show these effects had slowed and even began to improve, lending credence to the present findings.

The left-hand side of Table 4 provides more results using the panel estimation strategy, this time examining a range of social attitudes and behaviors for which longitudinal information is available. Fractions of the sample professing beliefs that religion and ethnic origin are important are extremely high, at 99 and 98 percent respectively, and corresponding behavior (church attendance) is also high. Further, the majority of individuals report trust of own ethnic

group and religion (over 80 percent) while trust of individuals in other ethnic groups and religions is low (28 and 44 percent, respectively).[68]

Results of regression analysis presented in panel A suggest that *stated attitudes* with regard to the importance of ethnicity and religion, and trust of others, show no statistically significant changes associated with the election crisis. However, evidence presented in panel B implies that while reported attitudes may not have changed, individual *behavior* does appear to have been altered as a result of the violence. Specifically, although respondents do not display any change in self-reported trust of others, there is a very large and statistically significant decrease in the frequency with which they list non-family members as a contact for future data collection efforts.[69] The coefficients resulting from regressions on church attendance, community and bible study group membership are not significant at traditional levels of confidence, but it should be noted that the signs of these coefficients are also negative.

Admittedly, there is a tension between the findings regarding stated social opinions and actual social behaviors of respondents. It may be that asking directly is not the best way to elicit respondents' true feelings on a matter. Dercon (2008) comes to a similar conclusion. In his own research, the author finds that prior to the election, Kenyans were very unlikely to associate themselves first and foremost with a particular ethnic group – 90 percent identified themselves as Kenyan – and more than three quarters of individuals claimed that ethnicity played no role in creation of personal friendships.[70] However, the author also asserts that "Voting intentions reflected a strong ethnic quality… It is clear that among the Kikuyu, support for Kibaki was overwhelming, while support for Odinga was similarly high amongst the Luo; each enjoyed more than 80 percent support from voters of their own ethnic groups." Together these findings lead him to the conclusion that "…while in practice ethnic considerations clearly shape the behavior of Kenyans in the voting booth, they seem to yearn for a society in which such considerations are less relevant" (p. 3).

The panel analysis results presented in Tables 3 and 4 suggest that although observable labor, migration and nutrition outcomes were not significantly affected by the heavily contested December 2007 national election, and individuals do not report any significant changes in social attitudes, there is some evidence of changing social behaviors. Because of the importance of social networks and informal credit in this rural region, these findings beg the question of how informal financial markets have been affected by the election violence. The KLPS-2 includes a rich set of questions on informal credit markets that were not asked in the KLPS-1. To utilize this cross-sectional data, the following section employs a propensity score methodology.

4.5 Propensity Score Matching Approach

The propensity score matching approach to estimation is extremely useful for observational studies, when estimating a treatment effect where treatment is randomly assigned conditional on a multitude of observables. This method entails classifying individuals according to their probability of treatment based on observable characteristics, and then comparing individuals across actual treatment groups within each subclass. The propensity score (probability of

[68] The exact question wording is: "Is your religion (ethnic/tribal origin) somewhat important, very important or not very important to your life?" and "In general, can you trust people of your tribe/religion (other tribes/religions)?"
[69] The exact question wording is: "In case you are not at your current address, can you tell us the names of two friends or family members who are sure to know where you are, and how to contact you?"
[70] Similar results were obtained from the 2003 round of the Afrobarometer survey, which was conducted just eight months after the 2002 national presidential election (Wolf *et al.,* 2004).

treatment) is chosen through an algorithm attempting to guarantee a parsimonious estimation scheme, and balance both among covariates in each grouping and across groupings. Such a score is considered a sufficient statistic for the relationship between the treatment and the controls, although the approach itself is valid only under the assumption that treatment assignment is indeed random conditional on observed characteristics (Rosenbaum and Rubin, 1984).

Given the tracking methodology utilized during KLPS-2 survey enumeration, this assumption is plausible – controlling for characteristics related to how difficult a person might be to locate, then the probability of being found before or after the election is arguably random. The propensity score measuring the likelihood of an individual being interviewed after the December 2007 Kenyan election is calculated here using a probit specification based on several different characteristics, including gender, age at survey, tribe, years assigned deworming treatment in 1998, indicators for birth in Kenya and in an urban area, an indicator for attending a school in Budalang'i Division in 1998, indicators for completing primary and secondary school, parent educational attainment, and indicators for household possession of a cement floor, telephone, and electricity.[71]

One way to judge whether the propensity score approach is reasonable in this context is to examine a box plot showing the distribution of scores between treatment and control groups (where "treatment" is defined as those surveyed post-election, and "control" includes those surveyed prior to the election). Such a plot is displayed in Figure 1. With random assignment of treatment groups, we would expect to see identical box plots. Although that is not the case here, there is good overlap in scores. For almost every individual interviewed prior to the election, there is a comparable (at least along observable covariates) individual interviewed after the election. Because of the rich set of controls available, the use of matching methodology in this case should reduce any bias in the estimation of treatment effects.

The figures presented in the right-hand columns of Tables 3 and 4 are estimates of the average treatment effect on the treated (ATT) using the "nearest neighbor" propensity score matching approach with bootstrapped standard errors.[72] These results do not exactly match those of the differences-in-differences estimation method in terms of magnitude, but do appear to tell a similar story.

Results in Table 3 indicate that there was a 12 percent decrease in participation in farming activities as a result of the election. The corresponding result using panel data is also negative, but not significant.[73] Matching estimation does not suggest any other statistically significant relationships between the election violence and labor market, migration or nutritional outcomes.

---

[71] To the extent that unobserved factors are correlated with both difficulty to locate an individual and their credit/transfer practices, results using propensity score matching methods will be biased. However, together the rich set of observable characteristics used to calculate propensity scores, and the comparability of results from the matching and differences-in-differences approaches (see below), serve to assuage concerns that this bias is large. Furthermore, to confirm that the approach outlined here is not merely comparing those who migrate to those who do not, the propensity score analysis was also run on the subset of individuals who have never migrated, and live in Busia or another very nearby rural area. Results are very similar (not shown).

[72] Becker and Ichino (2002) provide clear documentation and a STATA program to execute this approach.

[73] These two findings together may indicate some bias in the results obtained using matching methods, especially since individuals surveyed after the election crisis were more likely to have migrated (and more likely to be living in an urban area, where farming is rare). Footnote 21 supplies reasons why this bias is likely small.

The right-hand side of Table 4 displays results of the propensity score analysis for the set of variables exploring social attitudes and behaviors. With the exception of weak statistical significance for the coefficient on "members of own tribe can be trusted," none of the social attitudes explored in either the matching or panel analyses showed any significant relationship with exposure to the election crisis. However, in the top of panel B, both methodologies suggest fairly large decreases in attendance to religious services, participation in community and bible groups, and listing non-family members as contact points for future data collection efforts (and the results employing propensity score analysis are highly statistically significant).[74]

Overall, the results presented in panel A and the top of panel B indicate broadly similar patterns across both methods of analysis – although there is little evidence of changes in respondents' attitudes toward trust of others and identification with individuals like or unlike themselves, there are significant changes in their social behaviors. Knowing that informal lending and transfer mechanisms are supported by social relationships, these findings lead one to wonder about the effects of the December 2007 election violence on such mechanisms. Information on informal credit and transfer practices was collected in KLPS-2, and results using propensity score analysis to test this relationship are provided at the bottom of Table 4.[75]

These results indicate little effect of election violence on participation in informal savings and credit groups, such as Savings and Credit Co-operatives (SACCO), merry-go-rounds, or Rotating Savings and Credit Associations (ROSCA). Rates of participation in these groups are fairly low to start with, though, at 1 and 14 percent, respectively. In contrast, there do appear to be large decreases in the use of binary informal credit and transfer mechanisms between households, which prior to the violence were relatively common. Specifically, individuals interviewed after the election are between 41 and 52 percent less likely to lend to or borrow money from individuals outside of their household, and 53 percent less likely to receive transfers from others.[76] Combined with the results described previously on social behaviors, this suggests that election violence at least in the medium-run has led to substantial reductions in informal financial market participation.[77,78]

5. Conclusion

[74] Note that decreased participation in religious services and local groups is not likely due to lack of access. As previously described, what services may have been interrupted during the conflict were generally up and running pretty quickly after February 2008.

[75] Because participation in formal financial markets is extremely limited in the study area (only 3 percent of respondents have a savings account in a bank, and only 1% have taken a loan from a commercial bank or lender in the last 12 months), the analysis presented here focuses on informal financial activities only.

[76] The coefficient on "household has sent a transfer in the last twelve months" is not significant at traditional levels of confidence, but in magnitude displays nearly a 30 percent reduction in outgoing transfers.

[77] It could be that households switched from monetary credit and transfers to in kind assistance such as labor. This cannot be formally tested using the present dataset. However, given that Dupas and Robinson (2009) found *increased* rates of inter-household monetary transactions during the crisis, it is doubtful that households suddenly switched into nonmonetary transactions immediately after its end.

[78] Another possible explanation for the reduction in informal credit and transfers is that default rates increased during the election crisis, and individuals were less willing to lend to acquaintances afterward. Due to data limitations, this cannot be formally refuted in the present analysis. However, it does seem unlikely that this account can offer a complete explanation for the dramatic decrease in transactions immediately following cessation of the violence, as compared to pre-violence levels. Due to the informal nature of these transactions, repayment or obligated return transfers that are late by just a few weeks (the crisis lasted only two months) would likely not result in such large reductions witnessed immediately afterward.

Due to scarcity of data, very little research has been undertaken to date to analyze the socioeconomic impacts of civil unrest in Sub-Saharan Africa, a continent that has been plagued by such conflict in recent decades. This study utilizes a unique dataset collected in the months surrounding the 2007 Kenyan presidential election to explore the short- to medium-term consequences of civil unrest. Protest and violence following the election affected large areas of Kenya. Though the duration of conflict was brief – the bulk of clashes occurred between late December 2007 and late February 2008 – there were more than one thousand fatalities, and hundreds of thousands of individuals displaced from their homes. There is good reason to think that such events might have a damaging legacy for some time to come.

One exceptional feature of this analysis is its ability to utilize both differences-in-differences and propensity score matching methodologies, each approach with its own particular benefits and pitfalls, to reinforce findings. This is important for two reasons. First, this allows us to directly compare the two methodologies in a real life setting. Both approaches provide useful insights and are important in addressing unique empirical concerns. In addition, both are found to yield broadly similar findings suggesting concordance in the two methodologies and robustness in the findings.

Remarkably, both strategies provide little evidence of lasting declines in farming, employment, migration, and nutrition just one to seven months following the conflict. Furthermore, although there is no indication of changes in respondent attitudes toward members of own ethnic and religious groups or toward those of other groups, there is substantial evidence of changes in social behaviors associated with the violence. Individuals are less likely to participate in religious services, to be a member of a community or bible study group, and to list a non-family member as a contact for future tracking endeavors. These results are significant because they highlight a tension between self-reported beliefs and actual observable behaviors of respondents, suggesting that direct inquiry may not be best way to elicit respondents' true feelings on a matter.

There is a large literature examining the role of informal credit and transfers, particularly in rural areas of less-developed countries, in helping households smooth consumption and make capital-enhancing investments. Relationships with individuals outside of one's own household are a key element in this form of financing. To the extent that civil unrest disrupts informal credit market participation, the findings of this analysis suggest that there may be lasting consequences for Kenyan households' access to credit and investment, and thus to overall economic development in the regions affected. However, more research is needed to examine the longer-term effects that damage to social cohesion and informal credit markets have on the level of investment, and whether such findings generalize to other countries in Africa. If indeed there are such persistent impacts, then the rebuilding of community ties should become a priority in recovery efforts.

Future work will examine these issues utilizing the full KLPS-2 dataset (including Wave 2, which at the time of writing is ongoing in the field). This will allow for an examination of longer-run impacts of the election violence by extending available survey data through December 2009, an additional year of data collection and a doubling in sample size. Future analysis will also explore potential heterogeneity in impacts based on the gender and education level of respondents.[79] Finally, upcoming work will use a mean effects approach to study

[79] Note that it is not possible to explore differential impacts across ethnic groups, as the KLPS sample is primarily composed of a single ethnic group (tribe).

impacts of civil unrest on sets of outcomes, in order to incorporate broader definitions of social cohesion and informal financial market participation.

Chapter 3 References

Akresh, Richard, Philip Veriwimp and Tom Bundervoet (forthcoming). "Civil War, Crop Failure, and Child Stunting in Rwanda." *Economic Development and Cultural Change.*

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 Research Working Paper C08-151.

Bayne, Sarah (2008, April). "Post-election Violence in Kenya: An Assessment for the UK Government." Retrieved from www.hackenya.org.

BBC News (2008, January 8). "Kenya's Dubious Election." Retrieved from newsvote.bbc.co.uk.

BBC News (2009, September). "Kenya to shut down refugee camp." Retrieved from newsvote.bbc.co.uk.

Becker, Sascha O. and Andrea Ichino (2002). "Estimation of Average Treatment Effects Based on Propensity Scores." *The Stata Journal*, 2(4): 358-377.

Bellows, John and Edward Miguel (2009). "War and Collective Action in Sierra Leone." *Journal of Public Economics*, 93(11-12): 1144-1157.

Blattman, Christopher (2009). "From Violence to Voting: War and Political Participation in Uganda." *American Political Science Review*, 103(2): 231-247.

Blattman, Christopher and Edward Miguel (forthcoming). "Civil War." *The Journal of Economic Literature.*

Blattman, Christopher and Jeannie Annan (forthcoming). "The Consequences of Child Soldiering." *Review of Economics and Statistics.*

Bundervoet, Tom, Philip Verwimp and Richard Akresh (2009). "Health and Civil War in Rural Burundi." *Journal of Human Resources,* 44(2): 536-563.

DEPHA (2008a, April 14). "Displaced Person Sites." Retrieved from http://www.depha.org/Unhcr/Maps/IDP_locations.pdf.

DEPHA (2008b, January). "Post-Election Violence Hotpots – Kenya." Retrieved from http://www.depha.org/Post_election_Violence.asp.

Dercon, Stefan (2008). "Ethnicity and Violence in the 2007 Elections in Kenya." Afrobarometer Briefing Paper No. 48.

Doyle, Mark (2008, January 28). "Kenya's geographic and political rift" *BBC News.* Retrieved from http://news.bbc.co.uk/go/pr/fr/-/2/hi/africa/7213211.stm.

Dupas, Pascaline and Jonathan Robinson (2009). "Coping with Political Instability: Micro Evidence from Kenya's 2007 Election Crisis." Mimeo.

Eifert, Benn, Edward Miguel and Daniel N. Posner (forthcoming). *American Journal of Political Science.*

Elkington, Natasha (2009). "Wounds fester a year after Kenya election violence." Reuters AlertNet. Retrieved from http://www.alertnet.org/db/an_art/55868/2009/06/16-142727-1.htm.

Gibson, Clark C. and James D. Long (2009). "The Presidential and Parliamentary Elections in Kenya, December 2007." *Electoral Studies* 28(3): 497-502.

Hamory, Joan and Edward Miguel (2009). "Individual Ability and Selection into Migration in Kenya." Mimeo.

Jacoby, Hanan G. and Emmanuel Skoufias (1997). "Risk, Financial Markets, and Human Capital in a Developing Country." *The Review of Economic Studies,* 64(3): 311-335.

Kanyinga, Karuti and Pauline Nyamweya (2008). "Governance Assessments in Practices: Case Studies 4B Kenya." OECD DAC GOVENET Conference Governance Assessments and Aid Effectiveness, February 20-21.
Kenya Red Cross (2008a, April 25). "Kenya: Electoral Violence Operations Update No. 46/2008." Retrieved from http://wwww.reliefweb.int/rw/RWFiles2008.nsf/FilesByRWDocUnidFilename/SODA-7FR9AA-full_report.pdf/$File/full_report.pdf.
Kenya Red Cross (2008b, June 25). "Kenya: Electoral Violence Operations Update No. 63/2008." Retrieved from http://ocha-gwapps1.unog.ch/rw/RWFiles2008.nsf/FilesByRWDocUnidFilename/TUJA-7G49AH-full_report.pdf/$File/full_report.pdf.
Knack, Stephen and Philip Keefer (1997). "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics,* 112(4): 1251-1288.
Lewis, Peter (2007). "Identity, Institutions and Democracy in Nigeria." *Afrobarometer Working Paper No. 68.*
MEASURE DHS STATcompiler. (2007). Retrieved from http://www.measuredhs.com.
Miguel, Edward and Mary Kay Gugerty (2005). "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya." *Journal of Public Economics,* 89(11-12): 2325-2368.
Miguel, Edward and Michael Kremer (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, 72(1), 159-217.
Narayan, Deepa and Lant Pritchett (1999). "Cents and Sociability: Household Income and Social Capital in Rural Tanzania." *Economic Development and Cultural Change,* 47(4): 871-897.
OCHA-Kenya (2008a, July 30). "Humanitarian Update Vol 30." Retrieved from http://www.scribd.com/doc/4379385/Kenya-Humanitarian-Update-Vol-30.
OHCA-Kenya (2008b, June 10). "Kenya – IDP Transit Sites." Retrieved from http://ochaonline.un.org/OchaLinkClick.aspx?link=ocha&docId=1096749.
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.
ReliefWeb (2008, January 7). "Kenya: Presidential Election Violence (as of 6 Jan 2008)". Retrieved from www.reliefweb.net.
Rosenbaum, P., and Rubin, D. B. (1984). "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association*, 79, 516-524.
Thompkins, Gwen (2007). "Kenya's Upcoming Elections Bring Ethnic Tension" *NPR Morning Addition*. Published: 12/21/2007, Accessible: http://www.npr.org/templates/story/story.php?storyId=17495005. Accessed: 07/03/2009.
UNHCR (2008, February 4). "Kenya – IDP Situation Map." Retrieved from http://www.depha.org/Unhcr/Maps/Regional/KEN_IDP_SituationMap_A3PC_04Feb08.pdf.
Wolf, Thomas P., Carolyn Logan, Jaremiah Owiti and Paul Kiage (2004). "A New Dawn? Popular Optimism in Kenya After the Transition." *Afrobarometer Working Paper No. 33.*

Figure 1: Overlap of propensity score by treatment status

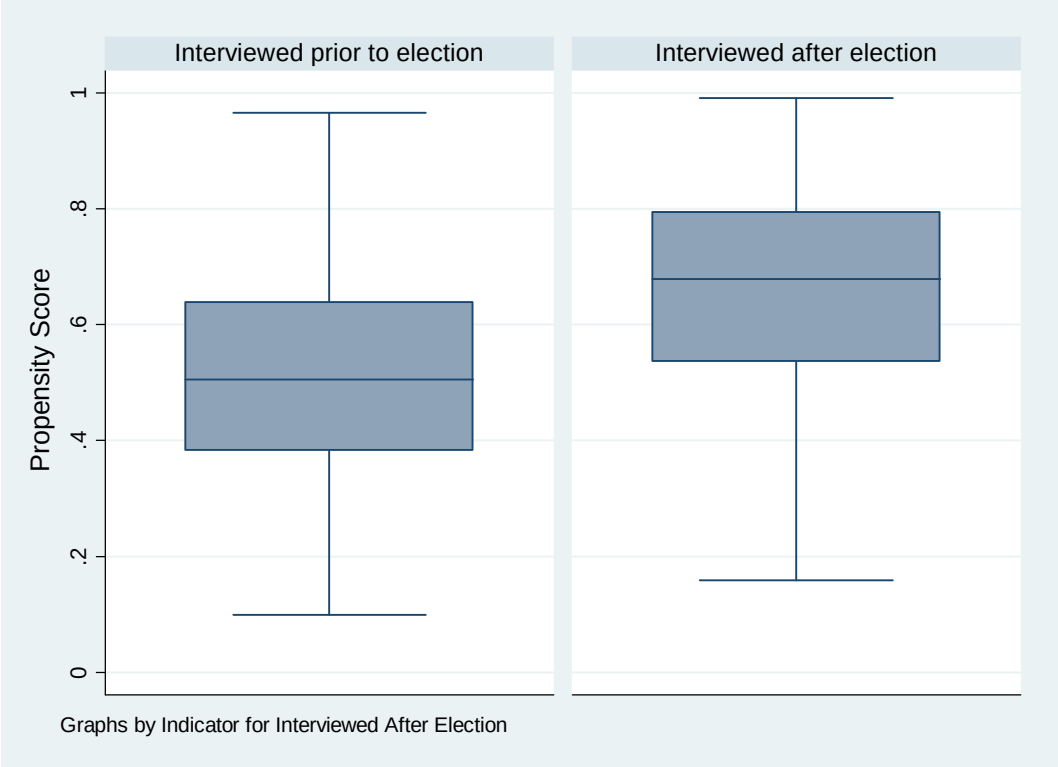


Table 1: Attrition analysis

	Dep Var: Indicator for Individual Surveyed in…		
	KLPS-2	**KLPS-1**	**KLPS-1&2**
	(1)	(2)	(3)
Female	-0.035	-0.033	-0.067***
	(0.024)	(0.024)	(0.019)
Age (1998)	-0.011	-0.025***	-0.024***
	(0.009)	(0.008)	(0.007)
Grade (1998)	-0.020	-0.013	-0.020*
	(0.013)	(0.012)	(0.012)
Years assigned deworming (1998)	-0.006	-0.002	-0.001
	(0.009)	(0.010)	(0.009)
Attended school in Budalang'i (1998)	-0.064***	0.004	0.036*
	(0.020)	(0.026)	(0.022)
School average of mock test score (1996)	-0.033	-0.017	-0.040*
	(0.033)	(0.023)	(0.024)
Observations	2749	2730	3745
Mean (std. dev.) of dependent variable	0.808	0.835	--
	(0.394)	(0.371)	--

Notes: Probit specification, with marginal effects evaluated at mean values. The sample used in column (1) contains all individuals who were surveyed, deceased, refused participation, found but unable to survey, or not found but searched for during intensive tracking for KLPS-2 Wave 1. Column (2) contains all individuals who were surveyed, deceased, refused participation, found but unable to survey, or not found but searched for during intensive tracking for KLPS-1 Wave 1. Column (3) contains all 3,745 individuals sought during the first wave of KLPS Rounds 1 and 2. Regressions are weighted in order to maintain initial population proportions (columns (1) and (2) employ the "effective" weighting technique described in the text, while column (3) employs standard population weighting techniques). Survey rates are not listed for column (3), as the traditional method of population weighting employed in this column does not account for tracking design in each round of the KLPS. Robust standard errors, corrected for clustering at the 1998 school level, are in parentheses. Age information is demeaned. Missing age data was replaced with mean values, and regressions include a control for missing data. 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. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 2: Summary statistics for various samples of interest

Means (except column 3)	KLPS-2 Sample				Panel Sample
	Total (1)	Before Dec '07 (2)	After Dec '07 (3)	P-value: (2) vs (3) (4)	Total (5)
Female	0.465	0.431	0.489	0.021	0.455
Age	21.6	20.8	22.1	0.000	21.3
Member of Luhya tribe	0.958	0.956	0.959	0.748	0.961
Years assigned deworming (1998)	3.54	3.68	3.44	0.020	3.62
Attended school in Budalang'i (1998)	0.360	0.328	0.382	0.033	0.403
Born in Kenya	0.993	0.994	0.993	0.910	0.994
Born in or near Busia District	0.897	0.910	0.889	0.013	0.896
Born in an urban area	0.082	0.073	0.087	0.286	0.085
Migrated during 1998-2006	0.530	0.411	0.611	0.000	0.486
Migrated to an urban area during 1998-2006	0.194	0.122	0.243	0.000	0.163
Attending school in 2006	0.418	0.475	0.379	0.000	0.457
Highest grade attained by 2006	8.42	8.34	8.48	0.197	8.51
Completed primary school by 2006	0.755	0.761	0.751	0.650	0.779
Attended secondary school by 2006	0.372	0.361	0.380	0.441	0.391
Completed secondary school by 2006	0.480	0.528	0.447	0.001	0.518
Attended post-secondary school by 2006	0.017	0.013	0.019	0.264	0.019
Father's educational attainment	8.66	8.59	8.72	0.596	8.57
Mother's educational attainment	5.64	5.59	5.67	0.745	5.62
Floor of home is cement	0.393	0.286	0.466	0.000	0.383
Home has electricity	0.151	0.071	0.206	0.000	0.142
Home has a latrine	0.888	0.912	0.872	0.021	0.882
Household has a phone	0.657	0.561	0.724	0.000	0.651
Household has a TV or radio	0.831	0.835	0.829	0.755	0.839
Number of observations	2492	1199	1293	2492	1939

Notes: The sample used in columns (1)-(3) includes all KLPS-2 Wave 1 respondents, where columns (2) and (3) are divided according to timing of interview. Column (4) is restricted to individuals with both KLPS-2 Wave 1 and KLPS-1 Wave 1 information. All figures are weighted in order to maintain initial population proportions (employing the "effective" weighting technique described in the text, based on KLPS-2 Wave 1 tracking).

[a] 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.

[b] Educational attainment is coded according to the post-1985 school system, which includes eight years of primary, four years of secondary, and four years of university. Attainment is censored at 13 years, such that 12 years of schooling encompasses those who attended through the last year of secondary school and any Ugandan A-level schooling, and 13 years encompasses those who attended any amount of post-secondary school (including college and university).

[c] Phone can be either mobile or a landline. In most cases it is a mobile phone.

Table 3: Results from regression analysis, part 1

	Panel Analysis				Propensity Analysis			
	Dependent Var		Post-Election Dummy		Dep Var		Post-Election Dummy	
Description of dependent variable	Mean	Std. Dev.	Coeff	Std. Error	Mean	Std. Dev.	Coeff	Std. Error
Household performs farming activities	0.90	0.31	-0.031	(0.049)	0.87	0.33	-0.108***	(0.030)
Self-employed or working for pay/family gain	0.28	0.45	0.039	(0.060)	0.19	0.39	0.055	(0.037)
Has ever migrated	0.33	0.47	-0.090	(0.054)	0.11	0.32	-0.014	(0.024)
Has ever migrated to an urban area	0.10	0.30	-0.001	(0.043)	0.02	0.14	0.010	(0.015)
Total number of meals eaten yesterday	2.24	0.63	-0.101	(0.103)	2.14	0.66	0.005	(0.054)
Total number of meals eaten with meat or fish yesterday	0.87	0.67	0.227	(0.139)	0.91	0.69	0.041	(0.053)

Notes: The sample used for the left side of the table includes individuals with both KLPS-2 Wave 1 and KLPS-1 Wave 1 survey information. Mean and standard deviation of dependent variable are calculated at time of Round 1 survey. Regressions are weighted in order to maintain initial population proportions (employing the "effective" weighting technique described in the text, based on KLPS-2 Wave 1 tracking). Additional controls include the following, measured at time of survey: age; school attainment; indicators for residence in Kenya and in an urban area; indicators for home has electricity, a latrine, a phone (mobile or landline), and a TV or radio. Number of observations ranges from 3,349-3,832, and R-squares range from 0.57-0.87. The sample used for the right-hand side of the table includes all KLPS-2 Wave 1 survey respondents. Summary statistics and regressions are unweighted. Mean and standard deviations of dependent variable are calculated for the sample of individuals surveyed prior to the election only. Regressions are performed using the ado file documented in Becker and Ichino (2002), with bootstrapped standard errors. "Has ever migrated" is defined in this sample to only include migrations occurring in the year of survey (2007 or 2008). Number of observations ranges from 1,306-1,307.

Table 4: Results from regression analysis, part 2

	Panel Analysis				Propensity Analysis			
	Dep Var		Post-Election Dummy		Dep Var		Post-Election Dummy	
Description of dependent variable	Mean	Std. Dev.	Coeff	Std. Error	Mean	Std. Dev.	Coeff	Std. Error
Panel A: Stated opinions								
Religion is very/somewhat important	0.99	0.10	0.006	(0.018)	1.00	0.03	0.000	(0.003)
Ethnic/tribal origin is very/somewhat important	0.98	0.15	0.003	(0.016)	1.00	0.07	0.004	(0.005)
Most people can be trusted	0.20	0.40	0.029	(0.058)	0.12	0.32	-0.035	(0.022)
Members of own tribe can be trusted	0.80	0.40	0.005	(0.077)	0.53	0.50	0.072*	(0.037)
Members of other tribes can be trusted	0.28	0.45	0.088	(0.080)	0.25	0.43	-0.001	(0.033)
Members of church/mosque can be trusted	0.83	0.38	0.096	(0.077)	0.69	0.46	0.001	(0.036)
Members of other churches/mosques can be trusted	0.44	0.50	0.031	(0.086)	0.37	0.48	0.000	(0.043)
Panel B: Actual Behaviors								
Social Behaviors								
Attends church/mosque regularly	0.76	0.43	-0.093	(0.066)	0.86	0.35	-0.063**	(0.026)
Member of a community group	0.45	0.50	-0.063	(0.086)	0.74	0.44	-0.110***	(0.037)
Member of a bible study group	0.16	0.37	-0.072	(0.067)	0.36	0.48	-0.076**	(0.038)
Lists a non-family member as a contact	0.11	0.32	-0.211***	(0.045)	0.07	0.26	-0.044**	(0.020)
Financial Behaviors								
Participates in a SACCO	--	--	--	--	0.01	0.09	0.001	(0.008)
Participates in a merry-go-round or ROSCA	--	--	--	--	0.14	0.34	-0.034	(0.031)
Taken loan from someone else outside hh in last 12 months	--	--	--	--	0.38	0.48	-0.198***	(0.039)
Lent money to someone outside household in last 12 months	--	--	--	--	0.33	0.47	-0.134***	(0.044)
Household has received a transfer in last 12 months	--	--	--	--	0.09	0.29	-0.048**	(0.020)
Household has sent a transfer in last 12 months	--	--	--	--	0.15	0.36	-0.043	(0.031)

Notes: The sample used for the left side of the table includes individuals with both KLPS-2 Wave 1 and KLPS-1 Wave 1 survey information. Mean and standard deviation of dependent variable are calculated at time of Round 1 survey. Regressions are weighted in order to maintain initial population proportions (employing the "effective" weighting technique described in the text, based on KLPS-2 Wave 1 tracking). Additional controls include the following, measured at time of survey: age; school attainment; indicators for residence in Kenya and in an urban area; indicators for home has electricity, a latrine, a phone (mobile or landline), and a TV or radio. Number of observations ranges from 3,758-3,832, and R-squares range from 0.58-0.68. The sample used for the right-hand side of the table includes all KLPS-2 Wave 1 survey respondents. Summary statistics and regressions are unweighted. Mean and standard deviations of dependent variable are calculated for the sample of individuals surveyed prior to the election only. Regressions are performed using the ado file documented in Becker and Ichino (2002), with bootstrapped standard errors. Number of observations ranges from 1,306-1,307.