New Roles for Marriage in Urban Africa *

Nancy Luke† Kaivan Munshi‡

October 2003

Abstract

This paper explores two new and important roles that the marriage institution could potentially play in urban Africa: facilitating kin and affine networks that find jobs for their members when information problems are present in the labor market and regulating non-marital sexual activity in a high HIV/AIDS environment. Using new data from Kisumu, Kenya, and controlling for selection into marriage, we find that marriage significantly increases employment levels and incomes in our sample of migrants, while at the same time increasing the remittances that they send to their rural homes, consistent with the view that the benefits of the network come with additional social obligations. In contrast, entry into the marriage institution appears to have no effect on the number or the type of non-marital sexual partners, once we control for selection into marriage. This finding has implications for the spread of HIV/AIDS in a high prevalence setting like Kisumu, which reported 26 percent HIV prevalence in 1997.

*We thank Esther Duflo, Jan Eeckhout, Mark Gersowitz, Yuichi Kitamura, Gregory Kordas, Michael Kremer, Costas Meghir, Sendhil Mullainathan, Mark Rosenzweig, seminar participants at Michigan, Michigan State, UCSD, and the Population Studies Centers at Brown and Harvard for helpful comments. Aldo Colussi provided superb research assistance. Research support from the University Research Foundation, the Mellon Foundation, the Center for AIDS Research, and the Population Aging Research Center, all at the University of Pennsylvania, is gratefully acknowledged. The African Census Analysis Project at the University of Pennsylvania graciously provided us with the Kenyan census data. Survey Research Team, lead by Francis Ayuka, supervised and conducted the survey of urban males in Kisumu. We are responsible for any errors that may remain.

†Population Studies and Training Center, Brown University

‡Department of Economics, Brown University and NBER

1 Introduction

Urbanization in sub-Saharan Africa is a relatively recent phenomenon, which has resulted in large migrant populations in cities throughout that region of the world. Information problems are likely to be substantial in labor markets characterized by a steady stream of new entrants, and it is not surprising that community-based networks play an important role in matching workers to firms in African cities, as elsewhere in the Third World. This paper is part of a larger research program that seeks to understand how the marriage institution facilitates such networks in urban labor markets in developing countries. A companion paper (Munshi and Rosenzweig 2003) studies how marriage ties have kept caste-based networks in place in Bombay's labor market over the past one hundred years. Here we will explore how marriage ties expand the individual migrant's kinship network in one sub-Saharan setting - the city of Kisumu in Western Kenya - improving his labor market outcomes, while at the same time increasing his social obligations.

Rapid urbanization in sub-Saharan Africa has improved the employment opportunities and increased the incomes of millions of migrants. But this economic growth has recently been accompanied by a public health disaster of unprecedented magnitude: the HIV/AIDS epidemic. While this epidemic is a global disaster, its focus has been and continues to be sub-Saharan Africa; almost 70% of persons with HIV/AIDS, and over 80% of those who have died of the disease, belong to a region that accounts for less than 10% of the world's population (UNAIDS and WHO 1999).

Why has sub-Saharan Africa borne the brunt of this deadly epidemic? 80% of HIV transmission in sub-Saharan Africa is attributable to heterosexual activity (Ulin 1992), and while regional migration patterns and certain health conditions, such as the high prevalence of STDs, may have increased the transmission rate, it is evident that social organization has played an important role in the rapid spread of the disease. Caldwell et al. (1994), for instance, cite numerous studies from countries throughout the region that describe a traditional sexual culture characterized by substantial permissiveness and the absence of punishment for extra-marital relations. Given the high HIV/AIDS prevalence in sub-Saharan Africa, and in Kisumu, a second objective of this paper will be to assess whether a new role for the marriage institution, regulating non-marital sexual activity, has emerged.

At first glance, the two research objectives listed above appear easy to implement. We are interested in studying the effect of marriage on the individual migrant's labor market outcomes, social obligations, and non-marital sexual activity. But estimation of the marriage effect must account for selection into

the marriage institution. For example, it has been observed that married men fare better in the labor market than single men, in the U.S. and other economies. One explanation for this result is that entry into marriage increases the individual's productivity or provides access to new employment opportunities. But this result could be entirely spurious if more able individuals, who would fare better in the market in any case, have a greater propensity to marry. Similarly, there have been a number of previous studies from various parts of Africa that compare the sexual behavior of married and unmarried men (Caraël 1994 provides a summary of these studies). The general pattern seems to be that married men are less likely to engage in risky behavior and tend to display lower rates of HIV infection. These studies do not account for selection into marriage, however.

In this paper we think of ability as a collection of unobserved traits that determines a wide range of individual outcomes. In general we would expect high ability individuals to have superior labor market outcomes, as well as greater success in attracting (non-marital) sexual partners. Marriage is associated with entry into the wife's family network in Africa, which brings both benefits and new obligations. High ability individuals will tend to subsidize the other members of their network, providing accommodation, credit, and job assistance in the city, and sending large remittances to their rural homes, while requiring and receiving little support in return. Such individuals might then have a lower propensity to marry, in which case a spurious negative correlation between marriage and sexual activity could now be obtained.

Our strategy to avoid the selection biases described above exploits traditional marriage rules among the Luo, the ethnic group local to Kisumu, to construct an instrument for marital status. The fundamental marriage rule followed throughout much of sub-Saharan Africa, and among the Luo, is very simple; no individual is allowed to marry someone from a related clan (including his own). While this rule of exogamous marriage is the same for the entire ethnic group, its effect on the actual matching process varies widely across traditional Luoland (modern Nyanza Province). The Luo are a tribe of Nilotic origin who migrated south into Kenya between three and five hundred years ago. Areas lying directly in the path of the incoming migrants were settled by large numbers of unrelated clans, while more remote areas were settled later, often by related clans in a single wave. These patterns of historical migration generated wide variation in the local level of relatedness across Nyanza Province, which we will exploit in the empirical analysis. Relatedness determines the efficiency of the matching process, and we would expect areas with a lower proportion of related clans to be characterized by higher marriage prevalence (at each age) since there are more eligible partners to choose from.

Most of the data used in the empirical analysis described in this paper comes from a survey of male Luo migrants aged 21-45 that the authors conducted in Kisumu, the capital of Nyanza Province, in 2001. One advantage of this urban setting is that migrants in Kisumu are drawn from all over the Province, leaving sufficient variation in the relatedness variable to test the relationship between marriage and individual outcomes. But the relatedness instrument will only have bite if migrants continue to follow traditional marriage rules and continue to find their brides in their rural homes. Reassuringly, we find a strong negative correlation between relatedness in the origin location and marital status among the migrants in the sample. Our instrumental variable strategy also takes advantage of the fact that social rules often persist long after they have ceased to serve their originally intended purpose. Thus relatedness continues to determine marital status among the migrants, even though any economic motivation for the local relatedness pattern when it was first put into place is unlikely to be relevant in the city today. Consistent with this view, we will later provide evidence supporting the identifying assumption that the relatedness instrument only affects outcomes in the city through its effect on marital status.

The main results in this paper can be summarized as follows: First, marriage significantly increases employment, income, and remittances among the migrants in our sample (after controlling for the individual's age). This result is consistent with the presence of an underlying network organized around the marriage institution, and is obtained with and without instrumenting for marital status; the instrumental variable (IV) marriage effect is actually larger than the corresponding OLS effect with all three outcomes. Second, while a preliminary OLS regression provides the usual result that being married significantly lowers the number of sexual partners (after controlling for the individual's age), this marriage effect disappears when we instrument for marital status with relatedness. Additional analysis suggests that marriage does not affect the type of partner that the individual matches with as well. Both results described above are indicative of negative selection into marriage, with able individuals deferring entry into the marriage institution and biasing the OLS estimates downward. We will later provide theoretical justification for such selection when networks are active as well as empirical support for its presence using observed characteristics that are plausibly correlated with individual ability.

The HIV/AIDS epidemic has threatened the Luo for two decades now, and HIV prevalence in Kisumu was estimated to be as high as 26 percent by 1997 (Glynn et al. 2001). There is very possibly a negative externality associated with sexual activity involving multiple partners in this

environment, and marriage appears as a natural institution around which to situate social norms restricting such activity. But the HIV/AIDS crisis, despite its severity, appears to have done little to change the historical role of the marriage institution with respect to the regulation of male sexual activity. The marriage institution has adapted its traditional role of smoothing economic risk and facilitating exchange to the new urban environment, where it now organizes job recruitment networks and provides other forms of social support. Jacoby's (1995) careful analysis of marriage in Côte d'Ivoire also shows that the social structure of the marriage institution, as measured by the prevalence of polygyny, can be remarkably responsive to changes in the economic environment. Understanding why an institution as important as marriage can show flexibility along some dimensions but not others, despite the serious social costs that could potentially be associated with such rigidity, would thus seem to be a useful area for future research.

The paper is organized in five sections. Section 2 describes the institutional setting. Section 3 presents a simple model of marriage in a network-based economy, which describes the selection into marriage and discusses solutions to this identification problem. Section 4 reports the empirical results, and Section 5 concludes.

2 The Institutional Setting

We begin this section by describing the historical structure of the marriage institution in sub-Saharan Africa, with a special emphasis on extra-marital sexual activity. Subsequently we discuss the contemporary role for marriage, with its associated kinship networks, in the city. And, finally, we describe the marriage institution among the Luo in some detail, with a particular emphasis on the relationship between marriage and relatedness. The discussion in this section prepares us for the theoretical model of marriage that follows in Section 3 and the empirical analysis in Section 4.

2.1 Marriage and Sexual Activity

Explanations for differences between African and Eurasian marriage systems typically begin with the historically low population density in sub-Saharan Africa (Goody 1971, Goode 1970). Since labor was in short supply, women were valued for both their productive and reproductive capabilities. This demand for female labor gave rise to the institution of bridewealth (payment from the husband's family to the bride's family) in the marriage market. Jacoby (1995) argues along the same lines that the absence of an active female labor market gave rise to the institution of polygyny, as one of the few

ways in which a man could augment his (female) labor supply.

Polygyny instilled a sexual culture in which men were permitted to acquire numerous sexual partners (Adepoju and Mbugua 1997, Obbo 1995, Goode 1970).[1] As Orubuloye et al. (1997:1196) explain, "[Polygyny's] central position in society inevitably means widespread social and ethical acceptance of the proposition that many men need more than one wife, and by implication and extension, that many - perhaps most - men need sexual relations with more than one woman. This remains the situation [even] in Southern Africa, for, although a century or more has been long enough to dismantle polygyny, its behavioral implications, built up over aeons, will not disappear so easily." Such values were internalized by women in polygynous societies as well. For example, a Luo elder with whom we spoke in Kisumu noted that a Luo wife may actually encourage her husband to see other women and ask him to bring his girlfriends home so that she could assess their potential as future co-wives. The threat of divorce by a wife, to deter extra-marital sexual activity, is in any case unlikely to be credible in many African societies. Divorce or separation typically leaves the woman with nothing in a polygynous society - the husband keeps the children and the bridewealth must be returned by the wife's family (Goody 1969, Potash 1978, Caldwell et al. 1994).

In contrast, the monogamous Eurasian system prescribes extremely severe social, religious, and legal sanctions against extra-marital relationships. The traditional system of African sexuality thus clashed with Western values when the European colonialists arrived in the nineteenth century. European schools set forth models of family conduct to their students, while missionaries attempted to inculcate Christian values both in their schools as well as among those they converted. The colonial administrations in turn legislated or adjudicated against many traditional African practices, among them bridewealth and polygyny, and these laws have been maintained and strengthened by many modern African governments (Goode 1970).

African social institutions proved to be remarkably resilient to these attacks. Social rules surrounding marriage, sexual behavior, and fertility tend to change very slowly, and we will later see that while almost all the migrants in our sample are Christian, most still follow the traditional rules and rites of Luo marriage. Despite legal restrictions, the traditional patterns of polygyny and divorce settlement have also been retained in many modern African societies; marriage to more than one wife remains common, the husband's family retains custody of the children, and alimony is rarely paid

[1] In contrast, female sexuality was more tightly controlled than male sexuality in many parts of Africa (Caldwell et al. 1992, Oppong 1992). A detailed exploration of these gender differences is beyond the scope of this paper.

(Parkin 1978, FIDA Kenya n.d.). Thus, while the advent of Christianity and the exposure to Western values might have nominally changed the structure of the marriage institution, there is no research that we are aware of that rigorously tests whether marriage has a causal effect on male non-marital sexual activity in sub-Saharan Africa.

2.2 Marriage and Labor Market Outcomes

To understand the role that marriage plays in determining labor market outcomes in urban Africa, it is useful to begin with the classical anthropological theory of "marriage as exchange" (Lévi-Strauss 1969). In this theory, exchange in traditional societies is dominated by reciprocal gifts between individuals, families, and clans (or castes). The marriage institution serves to facilitate this exchange.

Marriage throughout most of sub-Saharan Africa is exogamous, which implies that a man is not allowed to marry anyone from his own clan or from any clan that has been designated as being related to his clan. The individual is born into a network, drawn from his father's family, and then subsequently enters into a new network, drawn from his affines (wife's family), when he marries. Exogamous marriage is thus associated with the acquisition of a new network.

Marriage is also associated with the strengthening of the individual's birth-network. Marriages were, and continue to be, arranged by relatives and friends in many African settings (Goode 1970). For example, arranged marriages have traditionally been organized among the Luo by an intermediary, or *jagam*, who is usually one of the man's sisters, sisters-in-law, or another extended relative (Ocholla-Ayayo 1976, Ndisi 1974). Marriage arrangement continues to be prevalent in Nyanza Province today. Potash (1978) found that 79 percent of the marriages in a rural Luo community that she studied were arranged. High levels of marriage arrangement are observed in urban areas as well, with 44 percent of the married migrants in our Kisumu survey reporting that their marriages involved a *jagam*.

Marriage arrangement is a service that is usually matched by a reciprocal transfer from the individual to the *jagam*, who is typically connected to the individual's birth network. For instance, more than 72 percent of our married respondents who utilized the services of a *jagam* reported that they had subsequently provided some sort of service to them. Marriage in a traditional economy, such as among the Luo, thus should not be seen as simply a match between two individuals. Marriage also strengthens and widens existing network ties, expanding the services and support that the individual receives from the community, while at the same time increasing his social obligations (Shipton 1989).

Although urbanization in Africa is a relatively recent phenomenon, the marriage institution has

evolved so that it improves the individual's opportunities, while at the same time increasing his obligations, even in the city. Parkin's (1978:88) ethnography of Luo migrants in Nairobi describes how "A household head is subject to a barrage of requests for accommodation, many of them by job-seekers. All Luo who have 'spare' room in their houses are under some obligations to provide accommodation to a wide range of kin and affines." Consistent with this observation, over 60 percent of urban migrants interviewed in Kenya in the early 1990s reported that they were staying with kin or affines (Ocholla-Ayayo 2000). Over 90 percent of the respondents had also assisted kin or affines at some point, where assistance was specified to include paying school fees, providing housing, and job placement. While we do not attempt to formally identify network effects in this paper, one simple explanation for the positive effect of marriage on labor market outcomes and remittances that we will later observe is therefore that marriage strengthens existing network ties and builds new ones.[2] The link between marriage and the network will also later allow us to explain the negative selection into marriage that we observe with our sample of urban migrants.

2.3 Marriage Among the Luo

The Luo are one of the largest ethnic groups in Kenya today, numbering approximately three million, and they reside primarily in Luoland (modern Nyanza Province) in Western Kenya. Turning to Figure 1, we see that Luoland can be divided into two broad regions; Central Nyanza which consists of districts to the north and the east of Lake Victoria, and South Nyanza, composed of several districts south of the Lake.

Insert Figure 1 here.

The Luo migrated southward into Kenya from Egypt and Sudan, via Uganda, in three waves between 1490 and 1790 (much of the history that follows is based on Ogot 1996, 1967 and Ocholla-Ayayo 1976). Luo settlement in Kenya was initially restricted to Central Nyanza, which lies directly in the path of the incoming migrants. Pressure on the land, after the third wave of migrants, generated a movement to South Nyanza around 1790-1820. South Nyanza was settled for the most part by splinter groups from the clans in Central Nyanza, and the relatively uninhabited territory allowed groups to

[2]Other explanations for the marriage effect are also available. For example, entry into marriage, and the arrival of children, could increase the individual's responsibilities and so induce additional effort in the labor market. Alternatively, increased specialization in household tasks after marriage could make the man more productive in the workplace. But only the network story can easily explain why we also see increased remittances, as a fraction of total income, by married men.

settle in a more dispersed manner, with individual families frequently claiming large territories for themselves and their descendants. This internal migration into South Nyanza was completed around 1850.

Over time, Luo clans were united politically under tribal chiefs into areas known as *pinje* (territories). Property rights were historically poorly defined, and over the course of the centuries there came to be substantial spatial overlap among the clans within a *pinje*. With the arrival of the British in the late 1800s, these settlement patterns were frozen and the *pinje* were formally incorporated into administrative areas known as *locations*. The 30 "traditional locations" depicted in Figure 1 roughly correspond to administrative divisions today and the shading in the Figure represents the degree of clan relatedness in each location, which varies quite substantially across Nyanza Province.[3]

The idea for the relatedness instrument emerged during the course of a conversation with a Luo professor at the University of Nairobi in 1999. He noted that the Luo continue to adhere to many traditional practices and, as an example, he described the rule of marital exogamy, in which members of clans that were designated as "related" had not married each other for up to 15 generations.[4] He also mentioned that individuals in certain locations with many related clans apparently found eligible partners with great difficulty. Following up on this conversation, we collected relatedness patterns for Central and South Nyanza from Luo elders in Kibera, a predominantly Luo neighborhood in Nairobi. Subsequently a Luo historian independently verified the relatedness patterns in 2001. Although many Luo have a general knowledge of relatedness patterns throughout Nyanza Province, and more specific knowledge about their own traditional location, we believe that this is the first time that these patterns have been systematically documented for the entire area, facilitating the construction of a region-wide instrument for marital status.

We will effectively treat each traditional location as a distinct marriage market in this paper. Typically, each location consists of a number of *clusters* of related clans, as well as a number of independent clans. Clans within a cluster cannot marry each other, but can marry with any other

[3]Figure 1 shows the 30 traditional locations in present-day Nyanza Province. Districts in Nyanza Province comprised of other ethnic groups, primarily Kisii, are not included on the map. The two areas with no shading are recent settlement schemes that have an artificial arrangement of clans. These areas are excluded from the empirical analysis.

[4]Breaking rules of relatedness has connotations of incest, which would explain why the Luo continue to follow these rules today. But why would a society choose to put rules that are so rigid in place? The rules governing marriage presumably emerged to take account of the externality associated with the individual marriage decision, which affects the performance of multiple networks in a traditional economy. These rules had to be strong, and hence rigid, to ensure that individuals did not deviate from the social norm. Such rigidity may not have been particularly costly in an economic environment that was essentially static for centuries.

clan in the location. Independent clans, in contrast, can marry with any other clan in the location.

Let the number of *groups* in a location be the number of clusters plus the number of independent clans. Our measure of relatedness will be the ratio of the number of clans to the number of groups. Intuitively, the level of relatedness must grow as a greater fraction of clans are incorporated into clusters, which reduces the number of groups. When all the clans are independent, the number of clans is equal to the number of groups, and the degree of relatedness is one. When all the clans are incorporated in clusters, the degree of relatedness is the average number of clans per cluster. The relatedness statistic takes on values from one to 10 in our data, as described in Figure 1.

One of the important steps in the empirical analysis will be to show that high relatedness reduces the probability of finding an eligible partner, which leads to lower marriage prevalence. We now proceed to verify that the relatedness statistic that we have chosen does indeed satisfy the first part of this condition. Assume that the individual matches randomly with a partner from his own location in each period. Let N be the number of clans in the location, K the number of clusters, and n the number of clans within a cluster. For simplicity, we take it that all clans are of equal size.[5] Under these conditions, the probability of matching with an eligible (unrelated) partner in any period, for an individual drawn at random from the location, is obtained as

$$\frac{K}{N} \cdot n(1 - \frac{n}{N}) + \left(1 - \frac{K}{N} \cdot n\right)(1 - \frac{1}{N}).$$

The relatedness statistic can in turn be expressed as

$$\frac{1}{1 - \frac{K}{N}(n-1)}.$$

Relatedness is evidently increasing in K/N, while the probability of an eligible match is decreasing in K/N, since $n > 1$. Thus relatedness and this probability must be *negatively* correlated. Later in Section 3 we will complete the connection between relatedness and marital status by showing that an increase in the probability of an eligible match does in fact lead to a rise in marriage prevalence in equilibrium under reasonable conditions.

3 A Simple Model of Marriage

We now proceed to lay out a simple model of marriage in a network-based economy. In the previous section we saw that our measure of relatedness was negatively correlated with the probability of

[5] In practice, the relatedness statistic is computed using large clans only.

matching with an eligible partner. But individuals in high relatedness locations could compensate for this disadvantage by adopting less stringent matching strategies or by entering the marriage market earlier. In the discussion that follows we will derive conditions under which relatedness is negatively correlated with marital status in equilibrium.

The model is solved (backwards) in two stages. In the first stage each individual (man) decides whether or not to enter the marriage market. In the second stage, individuals in the marriage market match with (wives') family networks. The individual's ability is shown to determine both the age at which he enters the marriage market as well as the (expected) delay before he matches with a suitable partner once he is on the market. Marital status is thus a function of the individual's ability, at each age, and we will show that negative selection into marriage is obtained under reasonable conditions when networks are active. Since unobserved ability determines both marital status and individual outcomes such as employment, income, and sexual activity, a spurious correlation between marriage and these outcomes could be obtained. This section concludes with a discussion of this identification problem, in which relatedness is proposed as an instrument for marital status.

3.1 Population, Preferences, and the Matching Technology

Consider a marriage market in which individuals (men) match with (wives') family networks. Men are characterized by an index of ability u, while family networks are characterized by an index w, which denotes the average ability of their members. Both u and w are continuous distributions on $[0, 1]$.

The value of a network is measured by the assistance that it can provide in the labor market, which depends on the ability of its members. Thus the payoff to an individual from matching with a network whose average member has ability w will be $g(w)$, where g is an increasing function of w. Networks similarly benefit more from individuals with high ability.[6]

All participants in the marriage market are risk neutral utility maximizers who discount the future with a common and constant discount factor $\delta \in [0, 1)$. Individuals and networks meet randomly, once every period, and a successful match only occurs if both parties find their partners acceptable. Participants in the marriage market adopt a reservation strategy in which partners above a threshold ability (or average ability) level are accepted. Search costs are represented in this model by the discount factor δ, which reflects the participants' impatience to be matched. Lowering the ability

[6]The discussion in this section focusses on the individual's behavior since the empirical work that follows is restricted to male activity. Thus, it is not necessary to explicitly specify the benefit that the network derives from the individual, here or in the Appendix. All that we require is that this benefit be increasing in the individual's ability u.

threshold reduces these search costs by increasing the probability of a successful match, which avoids having to return to the market in the future. But the expected ability of the match will also be lower as the threshold declines. In practice, participants in the marriage market trade off these opposing pressures when choosing the ability threshold that is optimal for them. In general both individuals and networks will be more stringent in their choice (specify a higher ability threshold) as the discount factor grows.

Finally, individuals and networks who match in any period are replaced by individuals and networks of the same ability, preserving the stationarity of the ability and average ability distributions in the marriage market.

3.2 The Second Stage: Matching in the Marriage Market

Under the conditions described above, the properties of the two-sided search equilibrium are well known and have been described elsewhere (Burdett and Coles 1997, Eeckhout 1999, Bloch and Ryder 2000). In general, there will be (imperfect) positive assortative matching, with individuals of high ability matching with networks of high average ability. Moreover, the set of individuals and networks is partitioned into blocks, such that all individuals (networks) in the same block follow the same reservation strategy.

The intuition for this result is straightforward, and is well described in Bloch and Ryder. Begin with the highest ability individual ($u = 1$). Taking into account the fact that any network $w \in [0, 1]$ accepts him, and the cost (delay) associated with a mismatch, he will choose a threshold W_1 above which he accepts any network that he is paired with. All networks in the interval $[W_1, 1]$ are accepted by the highest quality individual, and therefore by all individuals. They make a similar calculation, to choose any individual in $[U_1, 1]$. This generates the first set of blocks.

Next, individuals in $[0, U_1]$ and networks in $[0, W_1]$ go through the same exercise to generate the second set of blocks; $[U_2, U_1]$, $[W_2, W_1]$. This process continues until the entire distribution has been partitioned into blocks, leaving us with the unique search equilibrium.

As noted, our first objective in this section is to derive conditions under which lower relatedness translates into an increased probability of being married. To derive these conditions, we will see that we must first study how block-size varies as we move down the ability distribution. Under the search equilibrium just described, an individual with ability u, belonging to the n^{th} block, takes the highest quality network that accepts him W_{n-1} as given when he chooses the threshold average ability $w(u)$

that maximizes his utility. All individuals in a block face the same choice problem and so $w(u)$ can be replaced by W_n. Denote the block size as $\Delta_n \equiv W_{n-1} - W_n$. Under the assumption that w is uniformly distributed, we show that

Proposition 1 *The block size Δ_n is increasing (decreasing) in n if $g'(w)/g(w)$ is increasing (decreasing) in w.*

While the proof is derived more formally in the Appendix, the intuition for this result is straightforward. If network quality $g(w)$ is a convex function of the average ability of its members w, then high ability individuals who end up matching with high quality networks in equilibrium tend to be more picky, since they have more to lose by choosing a slightly lower w network. While convexity is clearly necessary for the block size to be increasing as we move down the ability distribution, we require the stronger condition that the change in the slope of the g function should not only be positive, but also more rapid than the change in its level.[7]

This is not the sort of economy in which a Luo man with ability, or connections, can do exceptionally well. While Kisumu was a major port until the 1970s, moving goods between Kenya, Uganda, and Tanzania, the subsequent decline of East African trade and the poor performance of the Kenyan economy in recent years have seen a marked decline in its fortunes. The major trading houses and businesses are owned and controlled by the Indian community, and so opportunities are relatively limited for the Luo in any case. In these circumstances, we do not expect the strong convexity condition derived above to hold in practice. It seems reasonable to assume that $g'(w)/g(w)$ will be decreasing in w, which implies in turn that Δ_n will be decreasing in n. Under the assumption that w is uniform, block size maps directly into the probability of marriage. This tells us that the probability of being married will be declining in ability; high ability individuals find a match more quickly once they enter the marriage market.

Up to this point in the discussion we have assumed that all participants in the marriage market are free to match with each other. In practice, different marriage markets are characterized by different levels of relatedness. A mismatch occurs in any period if the individual's partner belongs to a related clan or lies outside the block that he belongs to. Earlier in Section 2.3 we derived the negative relationship between our measure of relatedness and the probability of meeting an unrelated partner,

[7]For example, it is easy to verify that $g(w) = e^w + a$, $w \in [0,1]$, satisfies this condition for $a > 0$. Intuitively, the increase in the intercept of the g function due to the $a > 0$ term slows down the rate of change in its level, leaving the change in the slope unchanged.

which we denote by λ. We now proceed to establish a positive relationship between λ and marriage prevalence in equilibrium, which completes the link between relatedness and marital status. Continuing with the assumption that $w \in [0,1]$ is uniformly distributed, the probability that an individual belonging to the n^{th} block will match successfully in any period is given by $\lambda(W_{n-1} - W_n)$, the joint probability that his randomly assigned partner is both unrelated, and is endowed with an ability level that results in both parties accepting the match. We can then show that

Proposition 2 *The probability of a successful match is unambiguously increasing in λ as long as Δ_n is decreasing in n.*

While the proof is relegated to the Appendix, the intuition for this result is once more straightforward. Differentiating the probability expression above with respect to λ we obtain

$$\Delta_n + \lambda\left(\frac{dW_{n-1}}{d\lambda} - \frac{dW_n}{d\lambda}\right).$$

Fixing the block size, an increase in λ reduces frictions in the marriage market and increases the probability of a successful match; this is the first term Δ_n in the expression above. But an increase in λ will also induce a change in participants' reservation strategies. Individuals become more picky when accepting a match since frictions in the marriage market have been reduced, which shifts W_n, W_{n-1} upward. How does this affect the equilibrium block size? Under our previous assumption that Δ_n is decreasing in n we show that $\frac{dW_{n-1}}{d\lambda} > \frac{dW_n}{d\lambda}$, and so the block size must expand when λ increases, reinforcing the direct (positive) effect of an increase in λ on marital status. Ultimately we will verify empirically that relatedness is negatively correlated with marital status. What the preceding discussion tells us is that we should expect to obtain this relationship under reasonable conditions on the network technology; as long as $g(w)$ is not too convex.

3.3 The First Stage: Entry into the Marriage Market

The downward bias in the marriage effect in the OLS labor market and sexual activity regressions that we later report will be interpreted as indicative of negative selection into marriage. Our next objective is thus to derive conditions under which negative selection into marriage is obtained. We noted above that high ability individuals match faster, under reasonable conditions, once they are on the marriage market. But this pattern of selection into marriage might be reversed when we allow individuals to choose when to enter the market.

For negative selection to be obtained, the payoff from marriage versus remaining single must be declining with ability. It is easy to see why this could be the case in a network-based economy. While high ability individuals do match with networks of higher average ability, we noted earlier that superior network quality might not translate into substantially higher benefits in this economy. A high ability migrant has easier access to jobs in the city, and so has less demand for the services that the network provides in any case. But perhaps more importantly, the high ability migrant might end up contributing disproportionately to the network, by way of housing assistance, job support, and transfers to kin and affines.[8] We will later provide evidence that individuals with observed characteristics associated with high ability do indeed transfer a greater fraction of their income to their rural homes.

Another feature of the marriage institution that we must account for when characterizing the entry decision is that ultimately all Luo men marry. We will see later that marriage prevalence starts to increase in the early twenties and continues to grow until age 40 in the city. A simple explanation for this observation is that marriage provides two benefits in an economy in which markets function imperfectly; access to a network, which has been the focus of the discussion this far, and access to children as old age support. While Luo men are permitted to have multiple extra-marital partners, the culture generally discourages children born out of wedlock. The demand for old age support will increase as the individual grows older, and so the incentive to marry and produce the children that will provide this support in the future must also be increasing with age.

The payoff from marriage that we specified earlier can now be augmented, based on the preceding discussion, as $h(\tau, u) \equiv g(w(u)) \cdot G(\tau, u)$. The first term, $g(w(u))$, is the payoff from the network, with network average ability w mapping into individual ability u through the second-stage matching process described above. The second term, $G(\tau, u)$, represents both the ability tax, represented by the transfers and services that flow from the high ability individual to the network, as well as the benefit from children, which grows with his age τ; $G_u(\tau, u) < 0$, $G_\tau(\tau, u) > 0$. It is easy to verify that the new multiplicative term has no effect on the second-stage matching that we describe in the Appendix. Matching in the marriage market continues to be based on ability alone. And the delay in matching continues to be declining with ability and increasing with relatedness, under the conditions derived earlier. But we will see that this new term now opens up the possibility for negative selection into

[8] It is unlikely that networks will be perfectly partitioned by ability in this economy, as these are family networks. Thus, even though there is positive assortative matching in the marriage market, high ability individuals will still end up having higher than average ability *within* their networks.

marriage.

Let the payoff from remaining single be $f(u)$.[9] Then negative selection into marriage requires that the returns to ability must increase more steeply for single than for married men:

Condition 1: $h_u(\tau, u) < f_u(u) \ \forall \ \tau, u$.

We account for the observation that marriage prevalence increases steadily over a wide age range, and that all Luo men ultimately marry, by imposing the additional restriction:

Condition 2: $h(\tau, 0) > f(0), \ \ h(\tau, 1) \leq f(1) \ \forall \ \tau$.

This condition specifies that the lowest ability individuals, who are subsidized most by the network, always prefer marriage to remaining single. In contrast, the highest ability individuals are at least as well off being single, at any age. Condition 1 and Condition 2 together ensure that a unique cutoff ability $u^*(\tau) \in (0, 1]$ is obtained, for any age τ, such that all individuals with $u < u^*(\tau)$ choose to marry. We thus have some amount of marriage at every age, with negative selection into marriage.

Whether we observe positive or negative selection into marriage is ultimately an empirical question. For instance, a large literature on the marriage premium in the U.S. labor market uses longitudinal data to control for selection into marriage (see Korenman and Neumark 1991, Gray 1997, Ginther and Zavodny 2001, and the references cited in those papers). While the uncorrected and corrected marriage effects are similar in a few studies, the usual result is that the marriage premium declines, but continues to be significant even after controlling for unobserved differences between married and single men. These results are indicative of positive selection into marriage, which requires from Condition 1 that the returns to ability should be steeper for married than for single men. This could well be the case in the U.S. economy if specialization in the household through marriage disproportionately increases productivity in the workplace among high ability men, particularly since there is no marriage-network tax in the modern economy. Among the Luo, men traditionally entered the marriage market as soon as they could pay the bridewealth, which also leads to positive selection into marriage. But high ability Luo men could well have a lower propensity to marry in the city today, to avoid being inundated by requests for assistance from kin and affines with little benefit in return. If the tax on ability in the network-based economy is sufficiently severe - $G_u(\tau, u)$ is sufficiently negative - then the returns to ability in marriage could well be shallower than the returns to ability from being single. Consistent

[9] We normalize so that $f(u)$ equals zero while the individual is searching in the marriage market.

with such negative selection, we will later see that individuals with observed characteristics that are associated with higher ability are significantly less likely to be married after controlling for age.

Why does the network not make a lump sum transfer to the high ability individual, which is effectively a dowry, to compensate him for the services and transfers that he will subsequently provide? Although the bride price appears to have declined in recent years, particularly in the city, there is no indication that bridewealth has been replaced by the dowry system anywhere in sub-Saharan Africa. This rigidity contrasts with Botticini and Siow's (2003) description of the emergence of the dowry system among South Indian Brahmins in the 1930s, as the demand for educated men holding salaried jobs under the British suddenly grew. One explanation for the inability of the African marriage market to shift to the dowry system is that the decentralized affine network may be unable to coordinate the requisite transfers to the high ability husband. Another explanation is that liquidity constraints could prevent the network from making a lump sum payment, which is necessary if subsequent commitment problems are to be avoided.

While high ability men might have a lower propensity to marry at any age, we noted above that almost all Luo men ultimately marry. We account for this stylized fact by allowing the (future) value from children, and hence the value from marriage, to grow as the individual ages; $G_\tau(\tau, u) > 0$, which implies $h_\tau(u, \tau) > 0$. As the h function shifts upward, it is easy to verify that the cutoff ability u^* shifts up with it; $u^*_\tau(\tau) > 0$. Thus the proportion of married men in the population is increasing with age. But at any age, married men continue to have lower ability than single men.[10]

The discussion on entry up to this point has ignored the delay associated with subsequent matching in the marriage market. Individuals will anticipate this delay, which we suggested will be greater for low types, in order to marry at the age that is appropriate for them. If there are no costs to entering the marriage market, then the differential delays (by ability) that we derived in the second stage equilibrium will wash out completely. The same argument also tells us that relatedness could have no effect on marriage prevalence if individuals adjust their time of entry into the marriage market.

There are two reasons why we expect relatedness to continue to determine marital status, even

[10]Note that the stationarity of the ability distribution in the marriage market, which we require to derive the second-stage matching equilibrium, is unaffected by the fact that individuals of different ability enter the marriage market at different ages. All that we need is that the underlying population distribution be stationary. By the same reasoning, differential exit by ability will not affect the stationarity of the distribution at the beginning of each period. For example, suppose that we begin with 100 low types and 100 high types in the marriage market in period 0, and suppose that 25 low types and 50 high types enter the market at the beginning of each subsequent period. If the probability of exit is 0.25 for the low types and 0.50 for the high types, then it is easy to verify that a stationary (and uniform) ability distribution in this marriage market is obtained.

when entry into the marriage market is endogenous. First, many of the responsibilities and the obligations associated with marriage appear as soon as the individual enters the marriage market. As van de Walle and Meekers (1994:57) explain, "African marriage is often a process with a varying number of potential stages, from the beginning of courting to the customarily sanctioned wedding, through periods where sexual relations are allowed, and periods of full cohabitation. Payment of bridewealth may also occur in stages. If all goes right, subsequent stages confirm the legitimacy of earlier ones ... but if the couple drift apart, the marriage, retrospectively, has never taken place." This tells us that the prospective groom must often incur substantial costs long before he marries. The prospective groom must also start to behave more responsibly, and his obligations to the community, particularly to the *jagams* that are searching on his behalf, will increase immediately. While there might be some adjustment in the time of entry, these costs to entering the marriage market imply that this adjustment will not be complete. Second, the very lowest ability individuals who enter the marriage market as soon as they reach adulthood cannot move their entry time forward. Later we will empirically verify that relatedness does indeed affect marriage prevalence, both in rural Nyanza Province as well as in urban Kisumu.

To summarize the discussion this far, the selection bias that creates problems for consistent estimation of the marriage effect is seen to arise at the first (entry) stage of the model. While higher ability individuals might match faster in the marriage market, we would expect the delayed entry by such individuals to dominate, particularly since the frictions in the marriage market can be adjusted for to some extent. In contrast, the instrument that we propose to correct the selection bias affects marital status in the second (matching) stage of the model.

3.4 Identifying Marriage Effects

Our basic objective in the empirical analysis will be to estimate the effect of marriage on labor market outcomes, remittances, and sexual activity. These outcomes will in general depend on observed and unobserved individual characteristics, as well as marital status, so we will estimate regressions of the form

$$y_i = \alpha M_i + X_i\beta + s(u_i) \tag{1}$$

where y_i is the outcome of interest for individual i, $M_i = 1$ if the individual is married and $M_i = 0$ otherwise. X_i is a vector of observed individual characteristics, such as the individual's age. $s(u_i)$

measures the effect of the individual's unobserved ability u_i; $s'(u_i) > 0$. The idea here is that high ability individuals are endowed with traits that lead to superior labor market outcomes as well as to greater success in the market for non-marital sexual partners. We take it that individuals participate in both these markets at every age.

Following the discussion in the previous section, the probability of being married is decreasing in relatedness and unobserved ability u_i (if the ability-tax is sufficiently severe), at any given age. In general, individual characteristics that determine the outcomes of interest in equation (1), such as age and ability, will also determine marital status. In the discussion that follows, we collect the observed determinants of marriage in a vector Z_i, that includes all the variables in X_i as well as relatedness. It will also be convenient to ignore the uncertainty associated with the search process and treat marriage as a deterministic outcome

$$M_i = 1 \qquad if \;\; Z_i\gamma - u_i \geq 0 \tag{2}$$

$$M_i = 0 \qquad if \;\; Z_i\gamma - u_i < 0.$$

Notice that M_i is negatively correlated with u_i, which would be the case with negative selection into marriage. This tells us immediately that the OLS estimate of the marriage effect in equation (1) will be biased downward, since M_i and the unobserved $s(u_i)$ are negatively correlated.

A standard solution to such selectivity bias applies a two-step estimator, which essentially introduces a consistent estimate of $s(u_i)$ in equation (1). When u_i is normally distributed, it is well known that a consistent estimate of the marriage effect is obtained by replacing M_i with $\Phi(Z_i\gamma)$, the predicted value from a first stage probit regression (Madalla 1983). But this estimator only provides consistent estimates if the distributional assumptions on u_i are correct. A more robust estimation strategy utilizes the predicted value from a first stage marriage regression, using either the probit or the linear probability model, as an *instrument* for M_i in equation (1) (Angrist 2001).[11] We prefer

[11] Under the set up of the model, a fixed effects estimator using repeated observations on y_i, M_i over time would allow us to difference out the $s(u_i)$ term in equation (1). But the assumption that individual ability u_i is constant over time is made for convenience, and is unlikely to be satisfied in practice. What we would expect to see instead is that shocks to ability (unexpected exogenous opportunities) determine entry into marriage as well as changes in the outcomes of interest. The fixed effects estimator no longer provides consistent estimates of the marriage effect once we allow for such shocks. For example, a negative shock could induce entry into marriage, to take advantage of the support that the network provides, while simultaneously reducing non-marital sexual activity. A spurious marriage effect would continue to be obtained in that case. Note that our instrumental variable estimates are unaffected by such shocks since patterns of relatedness were determined long ago and are in any case uncorrelated with idiosyncratic changes in ability.

to use the probit model to construct this instrument for the regressions that we report in this paper because it provides more precise estimates, consistent with Heckman's (1978) conjecture.

Equation (1) specifies that the marriage effect α does not vary by ability. But we could easily imagine, following the earlier discussion, that low ability individuals benefit more from the network, and hence from marriage. We see in the Appendix that the response in the probability of marriage $\lambda(W_{n-1} - W_n)$ to relatedness λ is a function of block size Δ_n. Since Δ_n varies with n, and hence with ability, the response to the relatedness instrument will vary by ability. If the marriage effect varies by ability as well, then the OLS and the instrumental variable estimates are not directly comparable. We will take note of this possibility later when interpreting the regression results.

The identifying assumption in our estimation strategy is that the relatedness instrument determines marital status, but does not directly determine the outcomes of interest in the city. In other words, relatedness is presumed to be uncorrelated with u_i, and hence with the ability distribution (broadly defined) in the traditional location. Ogot (1967:153) tells us that the initial settlement of Nyanza by the Luo "was not a united invasion, planned and executed deliberately. The whole operation was diversified, irregular and unorganized." From our knowledge of the cultural history of individual locations, it also appears that patterns of relatedness were put in place very soon after the Luo arrived in an area, often within a couple generations (Ayot 1979). Once these patterns had crystallized, there appears to have been little subsequent alteration over the centuries. This suggests that historical accident when the Luo first arrived might have played a role in determining the relatedness patterns that we see today. But at the same time it is possible that local conditions, such as climate and soil type, as well as the ability distribution among the arriving settlers, gave rise to particular local relatedness patterns.

In Lévi-Strauss'(1969) view, endogamy restricts marriage within a narrow social group, such as the caste, to strengthen network ties. In contrast, exogamy exploits gains from trade and risk diversification by spreading network ties. Extending this view, the different levels of relatedness that we observe across Nyanza Province could in principle have arisen as an optimal local response to the trade-off between network enforcement and the gains from trade, at the time when the Luo first settled in the area. Individuals in areas with high relatedness will in general travel further to find a wife, which effectively increases the level of exogamy. Thus we would expect to see high relatedness in areas where the gains from exchange dominated the gains from network enforcement.

The trade-off that we just described would in general depend on local economic conditions, such

as the level of risk, as well as the distribution of ability in the population.[12] While initial differences in ability may have disappeared over multiple generations, some local conditions such as climate and soil type are relatively permanent. These local conditions could have determined the patterns of relatedness as noted above, and they could also determine the nature of local economic institutions, the individual's incentive to invest in education, the distribution of wealth, and other characteristics of the local economy today. Relatedness would be correlated with unobserved determinants of the migrant's outcomes in the city in that case, to the extent that conditions at the origin determine the migrant's outcomes at the destination, and would no longer be a valid instrument.[13] Later in Section 4 we will verify that relatedness is uncorrelated with observed individual characteristics that are associated with ability or conditions at the origin, such as the migrant's inherited wealth and education, as well as demographic structure, measured by the size of the family.

As an additional validation test, we will show that the relatedness instrument only affects outcomes for individuals whose marital status is affected by that variable. We noted earlier that the relatedness instrument only has bite in Kisumu if the migrants continue to find their brides at home. Partitioning the sample into "early" migrants, who arrived in Kisumu as children or adolescents (before age 21) and "late" migrants who arrived as adults, we will later see that relatedness only affects marital status among the late migrants. Reassuringly, it turns out that relatedness only affects labor market outcomes and remittances among the late migrants as well. This useful result rules out the possibility that relatedness simply proxies for unobserved economic conditions at the origin, or individual characteristics that are common to migrants from the same origin location, that independently determine outcomes in the city. Relatedness appears to affect those outcomes exclusively through its effect on marital status, satisfying the conditions for a valid instrument.

4 The Empirical Analysis

Most of the analysis in this paper uses data from a survey of Luo migrants that we conducted in Kisumu. This section begins with a brief discussion on the survey design and the collection of the data. Subsequently we present descriptive statistics and the regression results.

[12] For example, we could imagine that the potential for default would matter more in a location that was settled by less reliable individuals, leading in turn to higher levels of endogamy (lower relatedness).

[13] With regard to sexual activity, no stigma is attached to non-marital relationships between members of related clans, or even the same clan, as long as these relationships do not lead to marriage. Thus, there is no reason for relatedness to have a direct effect on the sexual culture in the traditional location and, by extension, on non-marital sexual activity in the city.

4.1 The Data

We use two sources of data in this paper: a 5% random sample of Luo men residing in rural Nyanza Province from the 1989 Kenyan census, and information on 2,300 male Luo migrants collected from a survey of Luo men that we conducted in Kisumu in July-August 2001. The survey data will be used for most of the analysis, so we begin with a description of the data collection.

Kisumu is divided into 13 sub-locations, which are further divided into Enumeration Areas (EAs) by the Central Statistics Bureau of Kenya. Leaving aside EAs with almost no Luos and a few rural areas at the outskirts of the town which attract almost no migrants, we were left with 442 EAs to serve as the sampling frame. Eligible respondents were identified as Luo men, 21-45 years old, who had migrated to Kisumu after birth. All eligible respondents in a selected EA were interviewed. EAs were drawn randomly (without replacement) from the list of 442 EAs, until we had reached the targeted sample size.[14]

Before turning to a detailed description of the data, we verify that marriage patterns among the sampled migrants in Kisumu are roughly comparable to those in rural Nyanza Province, obtained from the 1989 Kenyan census. Figure 2 plots marriage prevalence and divorce rates, over the 21 to 45 age range, using both the urban Kisumu data and the rural census data. The 1989 census provides information on current marital status for each individual. With the Kisumu data we can determine in addition whether a migrant had ever been divorced or separated, since a complete marital history was collected from each respondent.

Insert Figure 2 here.

Marriage rates in urban Kisumu and rural Nyanza Province track fairly closely, particularly from age 21 to 30. We also see in Figure 2 that very few individuals (less than 2 percent of the population) are currently divorced/separated in both the urban and the rural samples, and there does not appear to be an appreciable age trend in this statistic. Notice, however, that the proportion of ever divorced/separated migrants is significantly higher than the corresponding statistic for currently divorced/separated migrants, which tells us that remarriage must be common, at least in the city. We use the full sample of migrants, aged 21-45, in the descriptive statistics and the regressions that

[14]Both migrants and locals (those born in Kisumu) were interviewed in the first 25 EAs that we selected. Subsequently, only migrants were interviewed in the remaining 76 EAs that were covered. Information on the local men will be used for analysis that is unrelated to the topic of this paper, so the discussion that follows will focus exclusively on the 2,300 migrants.

follow because the relatedness instrument determines marital status in our sample among the young men (aged 21-30) and among the older men (aged 31-45). One explanation for the persistence of the relatedness effect is that the high rates of divorce/separation and remarriage noted above make men return to the marriage market at older ages in the city.[15]

4.2 Descriptive Statistics

The discussion in this section begins by describing the individual characteristics of the migrants in the Kisumu sample. Next we study their ties to the community, which must be strong if relatedness is to affect marriage in the city. Subsequently we study the organization of the marriage institution, before concluding with a comparison of sexual activity between currently married and single men.

Once the individual is in the city, most women that he meets will belong to a different origin location and will therefore be unrelated to him. Relatedness and marriage will only be linked to the extent that migrants find their wives at home. And, not surprisingly, the relationship between relatedness and marital status is indeed stronger for individuals who move to Kisumu at a later age. As noted, the empirical analysis distinguishes between early migrants, who arrived in Kisumu as children and adolescents, and late migrants who arrived as adults. The descriptive statistics that follow will consequently compare currently married and single men, separately for early and late migrants.

The descriptive statistics that we report do not control for the individual's age (married men tend to be older) or for selection into the marriage institution. We will see later that not all the differences between married and single men reported below hold up to more careful scrutiny.

4.2.1 Individual Characteristics

We begin in Table 1, Panel A, with the individual's background. Married men are obviously older than single men, and late migrants tend to be older than early migrants, which is not surprising since they have older arrival ages by construction. Education levels are fairly high, around 10 years of schooling, for all groups of migrants.[16]

Insert Table 1 here.

[15] In contrast, the relatedness effect estimated with the 1989 Kenyan census drops substantially among the older men in rural Nyanza Province.

[16] Education levels tend to be generally high in Kenya, and among the Luo, as compared with other developing countries. For example, average schooling attainment for Luo men aged 21-45 residing in rural Nyanza Province was seven years in the 1989 census.

Notice, however, that married men have significantly lower schooling than single men. If low ability individuals marry early, then married men will tend to have lower ability on average than single men over the 21-45 year age range that we consider here. Ability in this paper covers a wide range of unobserved traits, one of which is intelligence, which in turn will determine the individual's education level. The difference in education levels that we observe could in that case be generated by selection into the marriage institution. Alternatively, if education levels have been increasing over time, then differences in education between married and single men could arise simply because married men are older. Regressions that we discuss later that control for the individual's age continue to provide a negative and significant correlation between marital status and education.

Turning to labor market outcomes in Panel B, married men work for roughly three more months in the year, and are 20 percent more likely to be employed. Consistent with this observation, married men earn on average 25-35 thousand Kenyan shillings more than unmarried men (all these differences are statistically significant).[17] The instrumental variable estimates of the marriage effect that we report later control for the individual's age as well as for selection into marriage, providing us with results that are qualitatively similar to what we see in Table 1.

We complete the description of the individual's characteristics by studying migration patterns among the survey respondents in Table 1, Panel C. Late migrants have obviously spent less time in Kisumu than early migrants. And among both groups of migrants, married men have spent significantly more time in the city than single men, presumably because they are older. While it has been quite a few years on average since all these groups arrived in Kisumu, a very large proportion of the migrants (over 80 percent) report that they have resided continuously in Kisumu since the time they arrived. Migration does not appear to be seasonal, or recurrent, as it is in many parts of the developing world. Consistent with this observation, the wives of migrants (even the late migrants) are likely to cohabitate with their husbands, and spend nine months of the year on average with them in Kisumu.

4.2.2 Connection to the Community

We noted earlier that the link between relatedness and marriage among the migrants could only be sustained if they continued to maintain close ties to their origin community and found their wives at

[17] At the time of the survey, one US dollar was approximately equal to 70 Kenyan shillings. We divided the 63 occupations reported by the migrants in our sample into four job categories: unskilled manual, skilled manual, business, and professional/white collar. The mean income in thousands of Kenyan shillings, with standard deviations in parentheses, for these four categories, is 43.68 (38.70), 71.32 (55.11), 74.02 (89.11), 116.20 (106.25).

home. Our characterization of the marriage institution in sub-Saharan Africa also ties married men more closely to family networks.

To verify that such community connections are indeed maintained, particularly among the married men, we begin by studying the pattern of home visits among the migrants in our sample in Table 2, Panel A. A sizeable proportion of the migrants visited their rural homes at least a few times per month over the last year, with the frequency of these visits being substantially higher among the married men. Late migrants also appear to visit more often than early migrants, suggesting that they have stronger ties to their home communities and will follow traditional marriage patterns more closely.

Insert Table 2 here.

Turning to remittances in Table 2, Panel B, most of the transfers seem to flow to the migrant's family, which includes affines, rather than to the home community more generally. This might be because patterns of exogamous marriage build ties between families belonging to different clans, rather than within the same clan. Remittances to the family are as high as 28 percent of the migrant's income, which is in line with other studies from sub-Saharan Africa that report remittances varying from 15 percent (Findley 1997) to 60 percent (Adepoju and Mbugua 1997) of the total income. Once more, married men remit significantly more than single men. Later we will see that this difference between married and single men holds up even when we control for the individual's age and selection into marriage with the relatedness instrument.

Apart from the family networks that each individual has access to, Luo men may also join community associations in most Kenyan cities. These associations were established as early as the 1920's, and are typically organized at the level of the clan, or the traditional location (Parkin 1978). They serve many of the same functions as the family networks: finding marriage partners, jobs, and housing for their members, and in general providing economic and social support. Turning to Table 2, Panel C, we see that both early and late migrants contribute to these associations and attend their meetings. The individuals in our sample, particularly the married men, are evidently closely tied to the origin community.

4.2.3 The Marriage Institution

Another important requirement for successful implementation of the instrumental variable procedure is that migrants in the city should continue to follow the traditional rules of marriage. In the discussion

that follows we will provide evidence that migrants continue to marry at home, and that the traditional organization of the marriage institution continues to be maintained.

We begin by studying the basic marriage patterns among the migrants in Table 3, Panel A. 72 percent of the late migrants and 54 percent of the early migrants are currently married. 15 percent of the migrants have been married more than once, which suggests a high level of divorce, separation, and/or polygyny. Indeed we see that 9 percent of the migrants have been ever divorced/separated, although only 2 percent are currently divorced/separated. Further, 6 percent of currently married migrants are polygynous, which is however lower than the corresponding figure of 9 percent obtained from the 1989 census.[18] With the exception of current divorce/separation and polygyny, all the statistics in Panel A are significantly different for early and late migrants at the 5 percent level.

Insert Table 3 here.

Next, we study the role of the home community in organizing marriage among the migrants. We see in Table 3, Panel B, that a substantial proportion of migrants met their wives in their rural homes, particularly among the late migrants. Continuing further in Panel B we see that this is because late migrants are more likely to have married before they left, and to have used a *jagam* (matchmaker) to find a wife. Recall from Table 2, Panel A, that late migrants visit home more frequently, which also increases the probability that they will independently find an acceptable match in the rural area. All of these statistics, taken together, tell us that relatedness should affect marriage more strongly among the late migrants, and later we will verify that this is indeed the case.

Looking finally at the marriage rites, we see in Table 3, Panel C, that over 80 percent of the migrants married in traditional Luo fashion (in contrast to a religious or legal wedding), despite the fact that most are Christian. A large proportion also paid some bridewealth, which was traditionally supposed to legitimate the union. Given that the migrants follow the traditional marriage rituals so faithfully, we certainly expect that they would adhere to the rules of relatedness as well. Over 90 percent of the married respondents reported that they actively checked that their wives were from unrelated clans.[19]

[18] Polygyny has declined very sharply over time among the Luo. While only 6 percent of the migrants currently have more than one wife, 48 percent of their fathers were polygynous.

[19] This statistic most likely understates the extent to which the Luo follow the rule of marital exogamy. Many of the 10 percent who reported that they did not check their wive's relatedness status met their wife in town, and probably knew from her traditional location that they could not possibly be related.

4.2.4 Sexual Activity

We conclude our description of the data by looking at non-marital sexual activity among the migrants, paying particular attention to activity that could potentially contribute to the spread of HIV/AIDS. Female partners among the Luo can be broadly divided into four categories: wives, *jodiya* (long-term girlfriends), casual partners, and commercial sex workers. The individual will often have multiple *jodiya*, from among whom he may ultimately choose a wife. Our first measure of non-marital sexual activity is simply the total number of non-marital sexual partners in the past year, without regard to their type. Our second measure acknowledges that certain types of partners can be more risky than others. It thus computes the number of *risky* non-marital partners in the past year, measured as the difference between the total number of non-marital partners and the number of *jodiya*.[20] Although the number of partners does provide information on the risk that the individual faces, an alternative, perhaps more precise, measure looks at his frequency of contact with risky partners. The survey instrument included a question on the frequency with which the individual matched with casual partners and commercial sex workers over the past year: the available categories in that question were a few times a week, a few times a month, a few times a year, once only, and never. For our third measure of non-marital sexual activity, the individual was coded as being $risky = 1$ if he matched a few times a week or a few times a month, $risky = 0$ otherwise.

Under-reporting of non-marital sexual activity is clearly a cause for concern, and the questionnaire was designed to avoid this problem by leading gradually to the section on sexual activity. The interviewers were also specially trained to ask these sensitive questions in a manner that would elicit accurate responses, and to probe (without leading) when the respondent appeared to be reluctant to reveal the full extent of his sexual activity.

As a check on the quality of the data, we computed the probability of having a risky partner in two different ways. First, the number of risky partners in the past year, as described above, allows us to construct a dummy variable for the presence of any risky partners at all. Second, the question on the frequency of encounters with risky partners allows us to once more construct a dummy variable for the presence of risky partners in the past year. The correlation between the two binary variables

[20] One way to assess the riskiness of a partner would be to compute the probability of using a condom with that partner, since we would expect condom use to be positively correlated with riskiness in equilibrium. The survey elicited condom-use information on the respondent's last five non-marital partnerships. Among all the partnerships reported in our sample, condoms were used 61 percent of the time with casual partners, 80 percent of the time with commercial sex workers, and 42 percent of the time with *jodiya*.

is 0.96, which at least verifies the internal consistency of the reported sexual activity, and in addition provides support for the accuracy of the responses.[21]

Starting with the first measure of sexual activity in Table 4, we see that single men had approximately two non-marital partners on average in the past year, while married men had one. These differences are statistically significant and are roughly the same for early and late migrants.

Insert Table 4 here.

There are a number of reasons why married men might have fewer non-marital partners than single men. First, married men already have at least one permanent partner, and so may have a physiological need for fewer additional partners. However, we take it that individuals have a strong preference for variety. The individual will always add another partner at the margin, if one is available, adjusting his sexual activity with other partners.

A second explanation for the lower number of non-marital partners among the married men is based on social norms restricting male extra-marital relationships. We have already argued in Section 2 that traditional African culture was not associated with such norms for historical reasons, but the advent of Christianity, colonialism, and exposure to Western values in the twentieth century could have altered the structure of the marriage institution.

A third explanation is simply an age effect; older men have fewer partners and are more likely to be married. Finally, a fourth explanation is based on selection into the marriage institution. Low ability individuals may be more likely to marry when networks are active, and these individuals could attract fewer partners in any case, generating a spurious negative correlation between marriage and sexual activity. The instrumental variable regression results that we present later will provide strong support for this last explanation of the patterns in Table 4.

Continuing with the first measure of sexual activity, we next report the proportion of men who had a non-marital partner in the past year in Table 4. This is followed by the proportion who had four or more partners in the past year. Sexual activity among single migrants is very similar to what we would expect to see in the United States. 84 percent of (currently) single Luo men had at least one partner in the past year, and 15 percent had four or more partners. The corresponding statistics for a representative sample of single (never married) U.S. men between 20-39, drawn from the 1991 National

[21] As an independent check, the survey supervisors re-visited four percent of the respondents and collected information on the total number of sexual partners in the past year. 96 percent of these selected respondents reported exactly the same number as they had earlier in the survey.

Survey of Men (Bill et al. 1993), are 90 percent and 18 percent. The major difference between the two groups of men is in the prevalence of extra-marital activity among married men; 44 percent of married migrants in Kisumu report such activity in the past year, whereas the corresponding U.S. statistic is less than 5 percent.[22] It is entirely possible that underreporting of sexual activity is more severe in the U.S., given the negative connotations associated with extra-marital relationships in this country. These comparative statistics also do not take account of selective entry into marriage, which may vary across countries. Nevertheless, the differences in extra-marital activity are striking, providing some *prima facie* evidence that the marriage institution regulates sexual activity quite differently in the two countries.

Comparing the number of *risky* partners among single and married men in Table 4 we see, once more, that married men appear to be less likely to engage in risky activity. And the same pattern continues to be obtained when individuals are classified as risky, based on the frequency of their interactions with casual partners and commercial sex workers, and then compared by marital status. But all the patterns in Table 4 could be spuriously generated if less able individuals, who are more likely to be married, lack the pecuniary resources or the skills to attract (risky) non-marital partners. Consistent with this alternative interpretation of the data, we will see that the negative relationship between marital status and risky sexual activity disappears when we control for selection into marriage.

4.3 Regression Results

We saw in the previous section that marriage has a strong effect on labor market outcomes and sexual activity. We now proceed to subject these relationships to more careful scrutiny. The regression analysis controls for the individual's age, which is correlated with marriage and could independently determine the outcomes of interest. The age effect is captured by a young-age dummy, with the cut off at 30 years; recall from Figure 2 that marriage prevalence increases rapidly prior to age 30, both in urban Kisumu and in rural Nyanza Province, before it starts to flatten out. While not reported here, the migrant's age in Kisumu is uncorrelated with relatedness in his origin location, and so the instrumental variable results that we report below are unchanged when a more flexible age specification that partitions the sample into three categories is introduced. However, the age effects are no longer

[22] Our estimates of non-marital sexual activity are also similar to those reported elsewhere in Africa. For example, a 1989-90 study of sexual activity conducted in the Ekiti district of Southwest Nigeria found that 75 percent of single males and 50 percent of married males had participated in non-marital sexual relations during the previous year (Orubuloye et al. 1991).

significant in the income, remittances, and sexual activity regressions. These additional tests, together with a number of robustness tests that we conducted, will be discussed in greater detail at the end of this section.

The regressions also control for the individual's experience in Kisumu, which could independently determine his performance on the labor market and the market for non-marital partners, by including a late-migrant dummy. Conditional on age, late migrants will evidently have less experience in Kisumu. Looking back at Figure 1 notice also that relatedness tends to be higher in South Nyanza. This region is situated far from the urban center, Kisumu, and has historically been relatively isolated. A South Nyanza dummy is thus included in all the regressions in this paper, to ensure that the relatedness instrument does not proxy for unobserved regional effects. Finally, we account for selection into the marriage institution by instrumenting for marital status. We will see that not all the differences between married and single men reported earlier hold up to more stringent testing.

4.3.1 Validating the Relatedness Instrument

Our instrumental variable strategy is valid if relatedness is uncorrelated with individual ability, broadly defined, which independently determines marital status, labor market outcomes, and non-marital sexual activity. For this we must rely on the idea that relatedness patterns were determined partly by historical accident, and that any economic motivation for the relatedness pattern in any given location when it was first put in place is no longer relevant, particularly for men in the city. Relatedness continues to determine marriage patterns, due to the inflexibility of social rules, but does not directly determine the other individual outcomes listed above.

To provide some support for this view, we verify that relatedness is uncorrelated with observed characteristics that are associated with the individual's ability, such as education, family size, and inherited wealth among the migrants in Kisumu, in Table 5. The regressions in Table 5, and most of the regressions that follow, include the young-age dummy, the late-migrant dummy, and a South Nyanza region dummy as controls.

Including all the migrants in the sample in Table 5, Columns 1-3, we see that relatedness has no effect on the migrant's years of schooling, the number of siblings, or the amount of land that he inherits from his father.[23] Note that the relatedness effect in these regressions is both statistically and

[23]Inherited land effectively measures average landholdings in the location, which is determined by the quality of the soil, the production technology, and the nature of institutions in the local economy. The absence of any relationship between inherited land and relatedness thus allows us to indirectly verify that relatedness is uncorrelated with a wide

economically insignificant. The relatedness variable takes values from one to ten in our data. The point estimate in Column 1 tells us that variation in relatedness, over its entire range, would translate into a corresponding variation in schooling of less than half a year. Relatedness can similarly explain little of the variation in family size or inherited land. In contrast, age, which can be interpreted as a cohort effect in these regressions, is negatively correlated with education, while the sign of the relationship is reversed for family size and inherited land.[24]

Inert Figure 5 here.

4.3.2 Relatedness and Migration

In general the migration decision depends on the individual's ability, conditions at the origin, and opportunities at the destination. Thus the individual's expected ability, conditional on migration, will be a function of conditions at the origin and opportunities at the destination. If relatedness is correlated with conditions at the origin that determine the migration decision, then ability in the city will be correlated with relatedness even if it is orthogonal to the ability distribution in the rural location.

The preceding results, demonstrating that relatedness is uncorrelated with schooling, family size, and inherited wealth, are particularly useful since they are obtained with a sample of migrants in the city. But these results are based on a limited number of observed characteristics. More generally we would like to establish that the level of migration is uncorrelated with relatedness, as a necessary condition for relatedness to be orthogonal to the ability distribution in the city.

We verify that migration does not depend on relatedness by merging the 1989 Kenyan census data with the 2001 Kisumu survey data. As noted, the young-age dummy takes on a value of one if the migrant is 30 years or younger, zero otherwise. The number of migrants in Kisumu from a given location, for each of the two age categories, divided by the total population of that location and age category in rural Nyanza Province provides us with a measure of the level of migration for that age category-location. South Nyanza lies further away from Kisumu than Central Nyanza, and while

range of economic conditions at the origin.

[24] As a robustness check we verified that relatedness is uncorrelated with education and wealth in rural Nyanza Province as well, using the 1989 census data. We use roof material as the measure of wealth in these regressions. This measure takes a value of one if the roof is constructed with iron, tile, concrete, or asbestos, zero if it is thatched. Controlling for regional differences with a South Nyanza dummy, we also verified that relatedness is uncorrelated with access to electricity and sanitation, road surface area and population density in the traditional location, as well as the Euclidean distance to Kisumu.

South Nyanza accounts for 40 percent of the rural Luo population of Nyanza Province, migrants from that region comprise only 25 percent of the urban sample. Controlling for this regional effect with the South Nyanza dummy (which is negative and significant), we see that relatedness has no effect on migration in Table 5, Column 4. The coefficient on the young-age dummy is positive and significant, which tells us that younger men dominate in the population of migrants.

Some of the regressions that we report in this paper will restrict attention to late migrants, who arrived in Kisumu as adults. We consequently go through the same exercise that we described above to verify that levels of late migration are uncorrelated with relatedness. Late migration is measured by the number of migrants in the sample who arrived after age 20 divided by the rural population for the corresponding age category-location. We see in Table 5, Column 5, that the same patterns that we obtained with all the migrants in Column 4 continue to be obtained with the late migrants; in particular, the relatedness effect is completely absent. In a related exercise, using individual level data, we regress the late-migrant dummy on relatedness, the young-age dummy, and the South Nyanza dummy in Table 5, Column 6. The age at arrival and the migrant's current age are mechanically positively correlated, which explains the negative coefficient on the young-age dummy. But the relatedness effect is completely absent, justifying our use of the arrival age (with the cut off at 21 years) to truncate the sample in some of the regressions that follow.

Finally, when studying the effect of marriage on sexual activity in a migrant population, it is important to consider whether wives live in the city with their husbands. For example, suppose that locations with low relatedness are characterized by lower levels of cohabitation. Now the higher marriage prevalence in these locations could be counter-balanced by the lower cohabitation, erroneously suggesting that marriage effects are absent in the instrumental variable regressions. To rule out this possibility, we proceed to verify that relatedness is uncorrelated with cohabitation. Cohabitation, measured by the number of months in the year that the migrant's wife spends in the city, is regressed on relatedness, the young-age dummy, the late-migrant dummy, and the South Nyanza region dummy in Table 5, Column 7. Relatedness has no effect on cohabitation, and the wives of early migrants spend *more* time with their husbands in Kisumu, which turns out to be helpful in interpreting some of the regressions that we later report.

4.3.3 Relatedness and Marriage

Once we have verified that relatedness has no effect on observed individual characteristics, migration, and cohabitation, the next step in the analysis is to establish the link between relatedness and marriage, which is the first stage of the instrumental variable regressions. Individuals in locations with mostly unrelated clans are more likely to meet an eligible partner (in terms of our model, λ is high), and we would therefore expect to see a negative correlation between relatedness and marriage prevalence under reasonable conditions, from Proposition 2 in Section 3.2. We begin by verifying this relationship in rural Nyanza Province, using data from the 1989 Kenyan census, in Table 6, Column 1.

Restricting attention to Luo men between 21 and 45 residing in rural Nyanza Province, we see using the 1989 census data that the probability of being currently married is negatively correlated with relatedness as expected. Younger men are significantly less likely to be married, which is consistent with the marriage model that we laid out in Section 3. Notice also the higher marriage prevalence in South Nyanza, which might arise because this region is more isolated and rural than the rest of Nyanza Province.

Insert Table 6 here.

While the results just reported from rural Nyanza Province are reassuring, we must still verify that relatedness affects marital status in the city. The marriage regression in Table 6, Column 2, uses late migrants who arrived as adults (after age 20) only; truncating the sample in this manner is not a problem since we previously established that the relatedness instrument is uncorrelated with the late-migrant dummy. We saw earlier in Table 3 that these migrants are more likely to marry women from the rural area, and we know that the relatedness instrument will only have bite in the city among migrants who find their wives at home. The regression estimates with the late migrants are qualitatively similar to what we obtained with the rural sample earlier. The relatedness coefficient, in particular, continues to be negative and is very precisely estimated.

Next, we report linear probability estimates of the marriage regression in Table 6, Column 3. Marriage and relatedness are negatively correlated with the linear probability model as well. While the estimated relatedness effect in Column 2 is statistically significant, the linearity in Column 3 allows us to easily verify that this effect is economically significant as well. Relatedness varies from one to ten in the data, so variation along the entire range of values for this variable would translate into a corresponding variation in marriage prevalence of as much as 17 percentage points.

A strong relatedness effect in Columns 2-3 would continue to be obtained if we lowered the cut off for the arrival age from 21 to 18 years, but the relatedness effect is much weaker and statistically insignificant for younger arrival ages. This is apparent in Column 4, which includes the full sample of migrants, with a late-migrant dummy and relatedness interacted with the late-migrant dummy as additional regressors. Note that inclusion or omission of the late-migrant dummy in the full sample regression has no bearing on the estimated relatedness effect since these variables are uncorrelated. The relatedness coefficient, which now measures the effect for early migrants, is very small and statistically insignificant. The relatedness-late migrant coefficient, which in that case effectively measures the relatedness effect for late migrants, is not surprisingly similar in magnitude to what we obtained in Column 2. Relatedness affects marital status for late migrants only, and we will take advantage of this observation later to provide additional validation for the relatedness instrument.

Finally, Column 5 reports Linear Probability estimates of the marriage regression with the full sample of migrants. The basic pattern reported in Column 4 continues to be obtained, and the magnitude of the relatedness effect, measured by the coefficient on the relatedness-late migrant interaction is once more similar in magnitude to the effect reported for the late migrants in Column 3.

4.3.4 Marriage, Labor Market Outcomes and Remittances

The preceding discussion linked relatedness in the origin location to marital status among the migrants in Kisumu. We now proceed to study the effect of marriage on the migrant's employment status, income, and the remittances that he sends home to his kin and affines. The presumption here is that marriage will improve the individual's labor market outcomes, while at the same time increasing his social obligations, if marriage-based networks are active. Some of these obligations will be fulfilled in the city, but the married migrant will also increase his transfers to the rural home when the network is organized around kin and affines. We begin with an OLS regression, with each outcome, before turning to instrumental variable estimates that control for selection into marriage.

We saw in Table 1 that married men are employed on average for about three months longer in the year than single men. This difference holds up in the OLS regression in Table 7, Column 1, which controls for the individual's age, the age at arrival in Kisumu, and the South Nyanza region effect. The marriage effect grows larger in Table 7, Column 2, and continues to be very precisely estimated, when we instrument for marital status with the predicted value from the probit marriage regression (reported earlier in Table 6, Column 4). While the coefficients on the young-age dummy and the late-

migrant dummy are also very precisely estimated, notice that the young-age coefficient reverses sign when we instrument for marital status, emphasizing the apparent bias in the OLS estimates reported in Column 1.

We saw in Table 6, Column 4 that relatedness affects marital status among the late migrants only, and so the instrumental variable estimates of the marriage effect in Table 7, Column 2 are identified off the late migrants. Restricting the sample to late migrants in Table 7, Column 3, the coefficient on marital status is not surprisingly very similar to what we obtained with the full sample.

Insert Table 7 here.

Similar results are obtained when we replace employment with annual income (in logs) in Table 7, Columns 4-6 and subsequently by remittances (measured as the percent of annual income) in Columns 7-9. We saw in Table 1 that married men earn on average 24 thousand Kenyan shillings more than single men, and that marriage increases remittances (as a fraction of annual income) by 7 percentage points in Table 2. These differences hold up in the corresponding OLS regressions in Column 4 and Column 7. And as before, the instrumental variable estimates of the marriage effect with relatedness as the instrument are substantially larger than the OLS estimates, with the full sample and with a truncated sample of late migrants. This increase in the marriage effect and the reversal in the sign of the young-age coefficient when we instrument for marital status indicate once again that the OLS estimates are biased.

The pattern of OLS and instrumental variable estimates in Table 7 is easily explained by our model of selective marriage. Individuals with lower ability, who perform poorly on the labor market and contribute less to the community, also have a greater propensity to marry. Such selection into marriage biases the marriage effect downward across all the OLS regressions in Table 7. However, the model laid out in Section 3 also tells us that individuals who respond to the relatedness instrument might be endowed with ability that differs from the average ability in the population; the OLS and IV estimates are not directly comparable in that case.

Here we can exploit the pattern of coefficients on the late-migrant variable in Table 6 and Table 7 to provide direct support for the interpretation of the results based on negative selection. Early and late migrants were seen to be generally comparable in terms of education levels, inherited wealth, and family size in Table 5. But conditional on age, late migrants have less exposure to the labor market in Kisumu. When the worker's ability is not observed *ex ante*, this lack of exposure can adversely affect

labor market outcomes. Consistent with this view, late migrants have significantly lower employment levels, and lower income as well (although the late-migrant coefficient in the income regression is not significant at the 5 percent level). Looking back at Table 6, Column 4 we see in contrast that late migrants are significantly *more* likely to be married.[25] Thus, we obtain what appears to be negative selection into marriage, with regard to one observed characteristic that is plausibly correlated with individual ability.

As a further check (not reported here), we included years of schooling as an additional regressor. We have already verified in Table 5 that schooling is uncorrelated with relatedness, and so not surprisingly the regression estimates were hardly affected by the inclusion of this variable. Schooling is plausibly positively correlated with the individual's ability. And consistent with negative selection into marriage, once more, we find that schooling is *negatively* correlated with marital status and *positively* correlated with employment and income (although the schooling coefficient is imprecisely estimated in the employment regression).[26]

Negative selection in the network-based model of marriage is generated by an ability-tax. Direct evidence of such a tax can once more be obtained from the estimated late-migrant dummy in the remittances regression. Late migrants, who are less likely to be employed and earn less when they are employed, remit a *lower* fraction of their annual income to their rural homes than the early migrants (Columns 7-8).[27] The apparent downward bias on the marital status coefficient in the OLS remittances regression (Column 7 versus Column 8-9) is also consistent with the view that lower ability individuals, who show a greater propensity to marry, remit a smaller fraction of their incomes. While such subsidies to low ability members of the network may arise out of fairness concerns - these are family networks after all - we will argue below that the negative selection into marriage that they generate could threaten the integrity of the urban kinship networks in the future.

[25] Our interpretation of the positive late-migrant coefficient in Table 6, Column 4 is that late migrants enter the marriage institution to take advantage of the networks that are organized around it. But late migrants could also have higher marriage rates if marriage prevalence is higher in the rural areas (a significant number of these migrants are married before they arrive in Kisumu) or if marriage induces subsequent migration (at an older age). Marriage rates in rural Nyanza Province and Kisumu track together over the entire 21-45 age range in Figure 2, and we established in Table 5 that relatedness and hence marriage prevalence are not correlated with migration, ruling out these alternative explanations.

[26] An alternative interpretation of the negative schooling coefficient in the first stage marriage regression is that individuals delay entry into the marriage market to complete their schooling. But we verify that the same result is obtained when the age range is restricted to 25-45 years; very few individuals study beyond the age of 25 in this economy.

[27] This result does not arise because the early migrants are more closely tied to their home communities or because their wives and children are more likely to remain at their rural homes. Later in Table 9 we will see that early migrants visit their rural homes less frequently than the late migrants. And the cohabitation regression in Table 5, Column 7 tells us that the wives of early migrants spend more time on average with their husbands in Kisumu.

We conclude the description of the relationship between marriage, labor market outcomes, and remittances by discussing a number of robustness tests not reported in the paper. We experimented with a more flexible specification of the age variable, with cut offs at 27 years and 35 years, dividing the sample into three age categories. The marriage effect and the coefficient on the late-migrant dummy are very similar to what we obtain in Table 7. But the age effects are now imprecisely estimated, except with employment as the dependent variable. The results are also generally unchanged when we use alternative age cut offs to separate early and late migrants. The point estimates are very stable when the cut off is lowered sequentially from 21 years to 19 years, but the marriage effect does decline slightly when it is increased (sequentially once again) to 23 years. Finally, the results are robust to the omission of older men, aged greater than 40, from the sample.

4.3.5 Validating the Relatedness Instrument: Early versus Late Migrants

One concern with the relatedness instrument is that it might be correlated with conditions at the origin that determine individual characteristics and access to resources, and which in turn determine sexual activity and labor market outcomes in the city. We showed earlier that relatedness is uncorrelated with observed individual characteristics such as education, inherited land, and family size. An alternative test to validate the relatedness instrument takes advantage of the observation in Table 6, Column 4, that relatedness affects marital status for late migrants, but not for early migrants. The reduced form regression corresponding to the IV regressions in Table 7 should provide us with the result that relatedness determines employment, income, and remittances among the late migrants. But we expect to see no relatedness effect for the early migrants if relatedness only affects the outcomes of interest through marital status.

We report reduced form employment, income, and remittance regressions, separately for early and late migrants, in Table 8. Relatedness has a strong effect on each of the dependent variables among the late migrants, consistent with the instrumental variable results in Table 7, whereas the relatedness effect is absent for early migrants.[28] The difference between early and late migrants in Table 8 would seem to rule out the possibility that relatedness in the instrumental variable regressions simply proxies for unobserved conditions at the origin that directly determine the outcomes of interest for all migrants from the traditional location. But could it be that the differences in the relatedness effect for early

[28]The coefficient on the relatedness variable for the late migrants is roughly 10 times larger than the corresponding coefficient for the early migrants. It is also fairly precisely estimated for the late migrants, except with income as the dependent variable.

and late migrants in Table 8 arise because those two groups of individuals are inherently different?

Insert Table 8 here.

Table 9 explores this possibility by comparing education, family size, and inherited land for early and late migrants in Columns 1-3. Origin location dummies are included in all the regressions in Table 9 to compare early and late migrants within each location. A young-age dummy is included as usual, to control for cohort effects. The coefficient on the late migration dummy is small in magnitude and statistically insignificant in Columns 1-3. This tells us that early and late migrants do not differ, at least with regard to permanent characteristics that would seem to be correlated with individual ability. The lower employment and income for the late migrants that we noted earlier in Table 7 arises presumably because they have less experience in Kisumu, and hence lower Kisumu-specific ability.

Insert Table 9 here.

While early and late migrants have similar characteristics, the fact that late migrants arrived in Kisumu at an older age implies that they might be more tied to their communities, even in the city. One way to assess the migrant's ties to his community in the city would be to study his participation in the community associations discussed earlier in Section 4.2.2. We include attendance in association meetings, and contribution to the association, as the dependent variables in Table 9, Columns 4-5. Reassuringly, we see no difference in the level of participation between early and late migrants in these associations.

Although early and late migrants might participate equally in the formal associations, they might still be tied differently to informal community or kinship networks. The pattern of home visits in Table 2, Panel A suggests that late migrants visit their rural homes more often than early migrants, and this difference holds up within the location, after controlling for age, in Table 9, Column 6. This result opens up the possibility that relatedness in Table 8 is associated with the quality of urban networks, formed among individuals from the same origin location, rather than around marriage. If late migrants, but not early migrants, participate in these networks on account of their closer rural ties, then that would explain why relatedness affects labor market outcomes and remittances exclusively among the late migrants.

The problem with this alternative interpretation of the results in Table 8 is that it requires that *only* the late migrants participate in the network. While late migrants might visit their rural homes

more frequently, the frequency of home visits for the early migrants in Table 2, Panel A is nevertheless quite high. Moreover, we noted in Table 7 that the early migrants remit a *greater* fraction of their income to their kin and affines at home. Thus, we are left with different marriage market behavior as the only available explanation for the difference between early and late migrants in Table 8. We see in Table 9, Columns 7-8, that late migrants are much more likely to be married before they arrive in the city, and in general are much more likely to have met their wives at home. This explains the strong link between relatedness and marital status for the late migrants, and the *absence* of this link for the early migrants, noted earlier in Table 6. And it also explains why relatedness affects labor market outcomes and remittances for the late migrants only; because it only affects those outcomes through marital status.

4.3.6 Marriage and Sexual Activity

We finally proceed to estimate the effect of marriage on each of the three measures of sexual activity. As before, we begin with an OLS regression that includes the young-age dummy, the late-migrant dummy, and the South Nyanza region dummy, before turning to instrumental variable estimates that control for selection into marriage.

Our first measure of sexual activity is the total number of non-marital partners in the past year, and we see that married men have significantly fewer partners in Table 10, Column 1, matching the descriptive statistics that we presented earlier in Table 4. But this OLS regression does not account for selection into marriage, and the marriage effect disappears when we instrument for marital status with relatedness in Column 2.[29] The same result is obtained with a restricted sample of late migrants in Column 3.

Our second measure of sexual activity is the number of risky partners, computed as the difference between the number of non-marital partners and the number of *jodiya* (long-term girlfriends). Once again marriage reduces the number of risky partners in the OLS regression in Table 10, Column 4, consistent with the patterns reported earlier in Table 4. But the marriage effect disappears when we instrument for marital status, both with the full sample and restricting the sample to late migrants. If anything, marriage appears to (weakly) *increase* the number of risky partners.[30]

[29] A few individuals at the very top of the non-marital partner distribution are dropped from these regressions. The instrumental variable estimates of the marriage effect are somewhat unstable (actually positive) when these outlying individuals are included in the sample.

[30] One explanation for this positive marriage effect is that *jodiya* are less interested in matching with married men, since the probability of ultimately marrying them is relatively low. Married men are thus channelled to a pool of riskier

Finally, our third measure of sexual activity defines the individual as being *risky* if he matched with casual partners or commercial sex workers a few times a week or a few times a month in the past year. As with the previous measures of sexual activity, we start with a decline in the probability of being risky with marriage in the OLS regression, consistent with the results in Table 4, but the sign of the marriage effect reverses when we instrument for marital status in Columns 8-9. The marriage effect with the full sample IV regression, in particular, is large, positive, and fairly precisely estimated.

The pattern of young-age coefficients and late-migrant coefficients is quite consistent across the regressions in Table 10. Young men have more partners, more risky partners, and are more likely to be classified as risky. The young-age coefficient is precisely estimated (with the full sample) and actually grows larger when we instrument for marital status. The late-migrant coefficient has the opposite sign of the young-age coefficient in all the regressions. It is also precisely estimated with the full sample, and grows larger (in absolute magnitude) when we instrument for marital status.

The sexual activity measures in Table 10 clearly have some information content, judging from the sensible and precise young-age and late-migrant effects. And the relatedness instrument provided us with fairly precise estimates of the effect of marriage on labor market outcomes and remittances in Table 7. Thus, although the IV marriage coefficients are not very precisely estimated, we can nevertheless conclude that entry into marriage does not appear to restrict unsafe sexual activity in this environment.[31] The marriage effect continues to be absent when we allow for a more flexible age specification, with three categories, when we use alternative age cut offs to separate early and late migrants, and when we drop older men (above age 40) from the sample.

The pattern of OLS and IV estimates in Table 10 can once more be explained as a consequence of negative selection into marriage. Low ability men fare worse on the market for non-marital partners because they have lower income (pecuniary resources) and perhaps because they are less skilled at matching with partners. This biases the OLS estimates downward when low ability men have a greater propensity to marry. Direct evidence of such negative selection can once more be obtained by inspecting the coefficients on the late-migrant dummy in Table 10. Recall that late-migrants were more likely to be married in Table 6, but had lower employment and income in Table 7. We explained that pattern by arguing that late-migrants have lower ability in the Kisumu labor market due to their

potential partners.

[31] Marriage was previously seen to substantially increase the individual's income, which will translate into a change in his personal expenditure once expenses on the family and transfers to the network are accounted for. What we are thus picking up in Table 10 is the net effect of changes in income and any social restrictions through marriage that may exist.

limited local exposure. By the same sort of argument we could imagine that late migrants have less skill on the market for non-marital partners in Kisumu which, together with their lower pecuniary resources, results in a lower level of non-marital sexual activity.[32]

5 Conclusion

This paper explores two new and important roles that the marriage institution could potentially play in urban Africa: facilitating kin and affine networks that find jobs for their members when information problems are present in the labor market and regulating non-marital sexual activity in a high HIV/AIDS environment. We find that marriage significantly increases employment levels and income in our sample of migrants, but appears to have no effect on the number or the type of non-marital sexual partners.

What role do we expect this institution to play in the future? The results in this paper suggest that high ability individuals are deferring entry into marriage in the city, to avoid the progressive ability-tax that is imposed on them by the network. The returns to individual ability will inevitably grow in this economy in the future, and at some point we expect that these high ability individuals will choose to bear the cost of defection from the network. Such defections will trigger further defections, worsening the quality of the network and threatening its integrity.[33] If the marriage-based network does not display some flexibility, perhaps by offering a dowry to high ability men or by lowering the current subsidy to low ability men, then we expect that it will not be viable in the long run.

There is also no evidence that the marriage institution will begin to regulate non-marital sexual activity in the future. If anything, it appears that the church has stepped in to assume this important role in many areas of East and Southern Africa (Watkins 2003). What is puzzling is that the marriage institution has shown flexibility along other dimensions in the past; the urban networks organized around marriage are a relatively new phenomenon, and the prevalence of polygyny has declined dramatically in Kenya and other parts of sub-Saharan Africa in response to economic change. Understanding why institutions can be flexible along some dimensions but not others, despite the consequent social cost, would thus seem to be a useful area for future research.

[32] Education has a weak effect on sexual activity; perhaps the positive effect of having higher ability is balanced by a greater concern for safety among educated men. Thus, we cannot use education to provide additional support for negative selection into marriage as we did with the labor market outcomes.

[33] Munshi and Rosenzweig (2003) uncover similar dynamics in Bombay's caste-based job networks, with high ability individuals defecting from the network as the returns to non-traditional occupations grow.

6 Appendix: Matching in the Marriage Market

Taking W_{n-1} as given, and assuming an equal measure of men and networks in the marriage market, an individual with ability u belonging to the n^{th} block chooses $w(u)$ to maximize

$$V(u) = \lambda \int_{w(u)}^{W_{n-1}} g(w)f(w)dw + \delta \left[1 - \lambda \int_{w(u)}^{W_{n-1}} f(w)dw\right] V(u), \tag{3}$$

where λ represents the probability that the network that the individual pairs with in any period is eligible (unrelated). $f(w)$ is the density of the average ability distribution at w, and δ is the discount factor. After some manipulation, the first-order condition for this choice problem can be expressed as

$$g[w(u)](1-\delta) = \delta\lambda \int_{w(u)}^{W_{n-1}} \left(g(w) - g[w(u)]\right) f(w)dw. \tag{4}$$

At this point we make the assumption that w is uniformally distributed, $f(w) = 1$, which will turn out to be very convenient below when studying selection into marriage. $g(w)$ is independent of u, and so all individuals in the n^{th} block face the same choice problem. It will consequently be convenient to replace $w(u)$ with W_n. The first-order condition can then be written as

$$\frac{1-\delta}{\lambda\delta} = \frac{\int_{W_n}^{W_{n-1}} g(w)dw}{g(W_n)} - (W_{n-1} - W_n). \tag{5}$$

Taking a linear approximation to the g function at each of the cut off points W_1, W_2, ..., W_N, the first order condition can be rewritten as

$$\frac{2(1-\delta)}{\lambda\delta} = \frac{g'(W_n)}{g(W_n)} \cdot \Delta_n^2, \tag{6}$$

where $\Delta_n \equiv (W_{n-1} - W_n)$ is the size of the n^{th} block, and $g'(W_n)$ is the slope of the g function evaluated at W_n.

Proposition 1: *The block size Δ_n is increasing (decreasing) in n if $g'(w)/g(w)$ is increasing (decreasing) in w.*

Proof. To prove this Proposition we make use of the following expression:

$$\frac{d}{dn}\left[\frac{g'(W_n)}{g(W_n)}\right] = \frac{d}{dW_n}\left[\frac{g'(W_n)}{g(W_n)}\right] \cdot \frac{dW_n}{dn}.$$

If $g'(w)/g(w)$ is increasing in w, then $g'(W_n)/g(W_n)$ must be increasing in W_n. Moreover, $dW_n/dn < 0$ by construction.

This tells us that $\frac{d}{dn}\left[\frac{g'(W_n)}{g(W_n)}\right] < 0$, which implies in turn that Δ_n must be increasing in n to satisfy the equality condition in equation (6). The analogous argument holds when $g'(w)/g(w)$ is decreasing in w.

Proposition 2: *Marriage prevalence is increasing in the probability of an eligible match λ when Δ_n is decreasing in n.*

Proof. Under the assumption that w is uniformally distributed, the probability that an individual belonging to the n^{th} block will match successfully in any period is given by $\lambda(W_{n-1} - W_n)$. This probability of being married is increasing in λ if

$$\Delta_n + \lambda\left(\frac{dW_{n-1}}{d\lambda} - \frac{dW_n}{d\lambda}\right) > 0. \tag{7}$$

$\Delta_n > 0$. To sign the second term on the left side of equation (7), we write

$$\frac{dW_n}{d\lambda} = \frac{\partial W_n}{\partial \lambda} + \frac{\partial W_n}{\partial W_{n-1}} \cdot \frac{\partial W_{n-1}}{\partial \lambda}. \tag{8}$$

Implicitly differentiating the first order condition, equation (5), and then taking a piece-wise linear approximation to the g function, we obtain

$$\frac{\partial W_n}{\partial \lambda} = \frac{\delta \Delta_n^2}{2[(1-\delta) + \lambda\delta\Delta_n]} > 0.$$

It is easy to verify from the expression above that $\partial W_n/\partial\lambda$ is increasing in Δ_n. A similar exercise derives

$$\frac{\partial W_n}{\partial W_{n-1}} = \frac{\lambda\delta\Delta_n}{(1-\delta) + \lambda\delta\Delta_n} > 0.$$

Once more it is easy to verify that $\partial W_n/\partial W_{n-1}$ is increasing in Δ_n.

The comparative statics that we have just derived tell us that all three terms on the right side of equation (8) are positive, which tells us in turn that all the cut offs W_1, W_2, ..., W_N shift upward when λ increases. Since all the terms on the right side of equation (8) were shown to be increasing in Δ_n above, they must be decreasing in n as well if Δ_n is decreasing in n. This tells us in turn that $dW_{n-1}/d\lambda - dW_n/d\lambda > 0$, completing the proof.

References

[1] Adepoju, Aderanti and Wariara Mbugua. 1997. The African Family: An Overview of Changing Forms. In Aderanti Adepoju, ed. *Family, Population and Development in Africa.* London: Zed Book Ltd.

[2] Angrist, Joshua D. 2001. Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice. *Journal of Business and Economic Statistics* 19(1):2-16.

[3] Ayot, Henry Okello. 1979. *A History of the Luo-Abasuba of Western Kenya.* Kenya Literature Bureau: Nairobi.

[4] Billy, John O.G., Koray Tanfer, William R. Grady and Daniel H. Klepinger. 1993. The Sexual Behavior of Men in the United States. *Family Planning Perspectives* 25(2):52-60.

[5] Bloch, Francis and Harl Ryder. 2000. Two-Sided Search, Marriages, and Matchmakers. *International Economic Review* 41(1):93-115.

[6] Botticini, Maristella and Aloysius Siow. 2003. Why Dowries? *American Economic Review* 93(4):1385-1398.

[7] Burdett, K. and M. Coles. 1997. Marriage and Class. *Quarterly Journal of Economics* 112: 141-168.

[8] Caldwell, John C. 2000. Rethinking the African AIDS Epidemic. *Population and Development Review* 26(1):117-135.

[9] Caldwell, John C., Pat Caldwell, and I.O. Orubuloye. 1992. The Family and Sexual Networking in Sub-Saharan Africa: Historical Regional Differences and Present-Day Implications. *Population Studies* 46:385-410.

[10] Caldwell, John C., Pat Caldwell, and Pat Quiggin. 1994. The Social Context of AIDS in Sub-Saharan Africa. In I.O. Orubuloye, John C. Caldwell, Pat Caldwell, and Gigi Santow, eds. *Sexual Networking and AIDS in Sub-Saharan Africa: Behavioural Research and the Social Context.* Canberra, Australia: Health Transition Centre, The Australian National University.

[11] Caraël, Michel. 1994. The Impact of Marriage Change on the Risks of Exposure to Sexually Transmitted Diseases in Africa. In Caroline Bledsoe and Gilles Pison, eds. *Nuptiality in Sub-Saharan Africa.* Oxford: Clarendon Press.

[12] Eeckhout, Jan. 1999. Bilateral Search and Vertical Heterogeneity. *International Economic Review* 40(4):869-887.

[13] Federation of Women Lawyers (FIDA) Kenya. no date. *The Women's Guide to Law.* Nairobi: FIDA Kenya.

[14] Findley, Sally E. 1997. Migration and Family Interactions in Africa. In Aderanti Adepoju, ed. *Family, Population and Development in Africa.* London: Zed Book Ltd.

[15] Ginther, Donna K. and Madeline Zavodny. 2001. Is the Male Marriage Premium due to Selection? The Effect of Shotgun Weddings on the Return to Marriage. *Journal of Population Economics* 14:313-328.

[16] Glynn, J.R., M. Caraël, B. Auvert, M. Kahindo, J. Chege, R. Musonda, F. Kaona, A. Buve and the Study Group on the Heterogeneity of HIV Epidemics in African Cities. 2001. Why Do Young Women Have a Much Higher Prevalence of HIV Than Young Men? A Study in Kisumu, Kenya and Ndola, Zambia. *AIDS* 15(suppl):S51-S60.

[17] Goode, William J. 1970. *World Revolution and Family Patterns.* New York: The Free Press.

[18] Goody, Jack. 1969. Inheritance, Property, and Marriage in Africa and Eurasia. *Sociology* 3:55-76.

[19] Goody, Jack. 1971. Class and Marriage in Africa and Eurasia. *American Journal of Sociology.* 76: 585-603.

[20] Gray, Jeffrey S. 1997. The Fall in Men's Return to Marriage: Declining Productivity Effects or Changing Selection? *Journal of Human Resources.* 32(3):481-504.

[21] Heckman, James J. 1978. Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica.* 46(6):931-959.

[22] Jacoby, Hanan. 1995. The Economics of Polygyny in Sub-Saharan Africa: Female Productivity and the Demand for Wives in Côte d'Ivoire. *Journal of Political Economy.* 103(5):938-971.

[23] Korenman, Sanders and David Neumark. 1991. Does Marriage Really Make Men More Productive? *Journal of Human Resources.* 26(2):282-307.

[24] Lévi-Strauss, Claude. 1969. *The Elementary Structures of Kinship.* Boston: Beacon Press.

[25] Madalla, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Econometric Society Monographs, No. 3. Cambridge: Cambridge University Press.

[26] Munshi, Kaivan and Mark Rosenzweig. 2003. Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy. M.I.T. Department of Economics Working Paper No. 03-23.

[27] National AIDS and STD Control Programme (NASCOP). 1999a. Edited by G.M. Baltazar, J. Stover, T.M. Okeyo, B.O.N. Hagembe, R. Mutemi, C.H.O. Olola. *AIDS in Kenya: Background, Projections, Impact and Interventions.* Nairobi: NASCOP. 5th Edition.

[28] National AIDS and STD Control Programme (NASCOP). 1999b (draft). Report on the Workshop, Priorities for HIV/AIDS Research on Areas of Socio-Economic, Cultural, Demographic, and Behavioural Issues. Nairobi, 29 April.

[29] National Research Council (NRC). 1996. *Preventing and Mitigating AIDS in Sub-Saharan Africa.* Washington, DC: National Academy Press.

[30] Ndisi, John W. 1974. *A Study in the Economic and Social Life of the Luo of Kenya.* Lund: Berlingska Boktryckeriet.

[31] Obbo, Christine. 1995. Gender, Age and Class: Discourses on HIV Transmission and Control in Uganda. In Han ten Brummelhuis and Gilbert Herdt, eds. *Culture and Sexual Risk.* New York: Gordon and Breach Publishers.

[32] Ocholla-Ayayo, A.B.C. 1976. *Traditional Ideology and Ethics among the Southern Luo.* Uppsala: Scandinavian Institute of African Studies.

[33] Ocholla-Ayayo, A.B.C. 2000. The African Family in Development Crisis in the Second Millennium. *The African Anthropologist* 7(1):84-113.

[34] Ogot, Bethwell A. 1996. *The Jii-Speaking Peoples of Eastern Africa.* Kisumu: Anyange Press Ltd.

[35] Ogot, Bethwell A. 1967. *History of the Southern Luo.* Vol. 1. Nairobi: East African Publishing House.

[36] Oppong, Christine. 1992. Traditional Family Systems in Rural Settings in Africa. In Elza Berquo and Peter Xenos, eds. *Family Systems and Cultural Change.* Oxford: Clarendon Press.

[37] Orubuloye, I.O., John C. Caldwell and Pat Caldwell. 1997. Perceived Male Sexual Needs and Male Sexual Behavior in Southwest Nigeria. *Social Science and Medicine.* 44(8):1195-1207.

[38] Parkin, David. 1978. *The Cultural Definition of Political Response: Lineal Destiny Among the Luo.* London: Academic Press.

[39] Potash, Betty. 1978. Some Aspects of Marital Stability in a Rural Luo Community. *Africa* 48(4):380-397.

[40] Shipton, Parker. 1989. *Bitter Money: Cultural Economy and Some African Meanings of Forbidden Commodities.* American Ethnological Society Monograph Series, No. 1.

[41] Ulin, Priscilla R. 1992. African Women and AIDS: Negotiating Behavioral Change. *Social Science and Medicine* 34(1):63-73.

[42] UNAIDS and WHO. 1999. *AIDS Epidemic Update: December 1999.* Geneva: Joint United Nations Program on HIV/AIDS, and World Health Organization.

[43] van de Walle, Etienne and Dominique Meekers. 1994. Marriage Drinks and Kola Nuts. In Caroline Bledsoe and Gilles Pison, eds. *Nuptiality in Sub-Saharan Africa.* Oxford: Clarendon Press.

[44] Watkins, Susan Cotts. 2003. Navigating AIDS in Rural Malawi. Paper presented at the annual meeting of the Population Association of America, Minneapolis, May 1-3, 2003.

Table 1: Individual Characteristics

Migrant type:	Early Migrants		Late Migrants	
Current marital status:	Single	Married	Single	Married
	(1)	(2)	(3)	(4)
Panel A: Individual Background				
Age	23.47	28.76	25.83	33.95
	(0.13)	(0.23)	(0.27)	(0.24)
Education	10.57	9.37	10.42	9.71
	(0.11)	(0.09)	(0.15)	(0.10)
Panel B: Labor Market Outcomes				
Months worked	7.73	10.93	7.34	10.69
	(0.20)	(0.09)	(0.27)	(0.10)
Employment dummy	0.78	0.99	0.82	0.98
	(0.02)	(0.004)	(0.02)	(0.004)
Income	36.72	60.48	37.79	71.13
	(1.63)	(1.64)	(2.51)	(2.02)
Panel C: Migration Patterns				
Years since migration	9.75	14.03	2.23	6.73
	(0.28)	(0.30)	(0.16)	(0.20)
Continuous residence dummy	0.86	0.82	0.92	0.91
	(0.01)	(0.02)	(0.02)	(0.01)
Months cohabited	--	9.79	--	8.94
		(0.13)		(0.15)
No. of observations	575	673	292	781

Note: Means reported for all variables, with standard errors in parentheses.

Early migrants arrived in Kisumu before age 21, late migrants arrived after age 20.

Income is measured per year, in thousands of Kenyan shillings. $1 was approximately 70 Kshillings in 2001.

Top 1% of incomes among early and late migrants are dropped from the analysis.

Employment dummy=1 if worked for more than one month in the year, 0 otherwise.

Continuous residence dummy=1 if resided in Kisumu continuously after migrating, 0 otherwise.

Cohabitation measured as the number of months of the year during which the migrant's wife lived in Kisumu.

Cohabitation statistics are computed for currently married men only. For men with multiple wives this applies to the wife who spent the most time in the city.

All means for single and married men are statistically different at the 5% significance level except Continuous residence dummy.

Table 2: Connection to the Community

Migrant type:	Early Migrants		Late Migrants	
Current marital status:	Single	Married	Single	Married
	(1)	(2)	(3)	(4)
Panel A: Pattern of Home Visits				
Frequency of home visits in the last year (%)				
Few times in a week	4.55	6.10	3.81	9.62
Few times in a month	27.97	41.37	33.56	48.08
Few times a year	47.90	40.77	42.56	35.51
Once only	11.89	6.85	9.34	5.13
Never	7.69	4.91	10.73	1.67
Total	100.00	100.00	100.00	100.00
No. of observations	572	672	289	780
Panel B: Remittances				
Remittances to home community	3.34	4.58	4.09	4.87
	(0.27)	(0.26)	(0.50)	(0.28)
Remittances to family	16.35	28.18	17.58	25.32
	(0.84)	(1.02)	(1.15)	(0.84)
No. of observations	489	651	249	757
Panel C: Participation in Clan Associations				
Attendance dummy	0.14	0.32	0.10	0.34
	(0.01)	(0.02)	(0.02)	(0.02)
Contribution dummy	0.26	0.51	0.24	0.55
	(0.02)	(0.02)	(0.03)	(0.02)
No. of observations	573	668	289	781

Note: Means reported for all variables, with standard errors in parentheses.
Early migrants arrived in Kisumu before age 21, late migrants arrived after age 20.
Remittances to community/family measured as percent of total income.
Top 1% of remittances among early and late migrants are dropped from the analysis.
Attendance dummy=1 if attends association meetings regularly, 0 otherwise.
Contribution dummy=1 if contributes to association, 0 otherwise.
All means for single and married men in Panel B and Panel C are statistically different at the 5% level except for Remittances to community - Late migrants.

Table 3: The Marriage Institution

Migrant type:	Early migrants	Late migrants
	(1)	(2)
Panel A: Marriage Structure		
Currently married	0.54	0.72
	(0.01)	(0.01)
Multiple marriages	0.11	0.17
	(0.01)	(0.01)
Ever divorced/separated	0.07	0.10
	(0.007)	(0.01)
Currently divorced/separated	0.02	0.02
	(0.003)	(0.004)
Currently polygynous	0.05	0.07
	(0.01)	(0.01)
Panel B: Marriage and the Community		
Met wife at home	0.39	0.55
	(0.02)	(0.02)
Married before migrating	0.03	0.55
	(0.01)	(0.02)
Used matchmaker	0.40	0.48
	(0.02)	(0.02)
No. of observations	672	781
Panel C: Marriage Rites		
Traditional marriage	0.87	0.82
	(0.01)	(0.01)
Paid bridewealth	0.48	0.69
	(0.02)	(0.02)
Checked relatedness	0.91	0.91
	(0.01)	(0.01)
No. of observations	673	780

Note: Means reported for all variables, with standard errors in parentheses.
Early migrants arrived in Kisumu before age 21, late migrants arrived after age 20.
Currently married=1 if married at the time of the survey, 0 otherwise.
Multiple marriages=1 if the respondent had more than one wife over his lifetime, 0 otherwise.
Polygynous=1 if respondent currently has more than one wife, 0 otherwise.
Traditional marriage=1 if respondent had traditional marriage, 0 if religious or legal marriage.
Currently married, multiple marriages, and divorce/separation are computed using the full sample of migrants (1248 and 1081 obs, respectively), while Polygynous is measured using currently married men (673 and 781 obs, respectively).
All statistics in Panel B and Panel C apply to currently married men. For men with more than one wife, these statistics are computed for the first wife.
All variables in Panel B and Panel C are binary variables.
All means for early and late migrants are statistically different at the 5% significance level except:
Currently divorced/separated, Polygynous, and Checked Relatedness.

Table 4 : Sexual Activity

Migrant type:	Early Migrants		Late Migrants	
Current marital status:	Single	Married	Single	Married
	(1)	(2)	(3)	(4)
Number of non-marital partners	2.21	0.97	1.85	0.74
	(0.10)	(0.08)	(0.09)	(0.08)
One or more non-marital partners	0.86	0.46	0.81	0.42
	(0.01)	(0.02)	(0.02)	(0.02)
Four or more non-marital partners	0.16	0.05	0.13	0.02
	(0.02)	(0.01)	(0.02)	(0.005)
Number of risky partners	1.15	0.58	0.84	0.43
	(0.10)	(0.06)	(0.08)	(0.08)
Risky status	0.12	0.07	0.07	0.03
	(0.01)	(0.01)	(0.01)	(0.01)
No. of observations	576	672	300	781

Note: Means reported for all variables, with standard errors in parentheses.
Early migrants arrived in Kisumu before age 21, late migrants arrived after age 20.
Number of non-marital partners is measured over the past year.
One/Four or more non-marital partners are binary variables constructed from the number of partners.
Number of risky partners is the difference between the number of non-marital partners and jodiya.
Risky status=1 if the individual matched with risky partners a few times a week or month, 0 otherwise.
Risky partners are defined as casual partners and commercial sex workers.
All means for single and married men are statistically different at the 5% significance level.

Table 5: Validating the Relatedness Instrument

Sample:	all migrants			rural+migrants		all migrants	
Dependent variable:	education	siblings	inherited land	migration	late migration	late migrant	cohabitation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relatedness	0.050	-0.043	-0.073	-0.003	-0.001	-0.0001	0.082
	(0.062)	(0.108)	(0.081)	(0.009)	(0.004)	(0.007)	(0.090)
Young age	0.642	-1.243	-0.559	0.139	-0.001	-0.338	0.014
	(0.137)	(0.260)	(0.257)	(0.036)	(0.016)	(0.027)	(0.227)
Late migrant	0.188	0.030	-0.254	--	--	--	-0.808
	(0.108)	(0.284)	(0.258)				(0.214)
South Nyanza	0.156	0.743	0.854	-0.160	-0.065	0.030	-0.512
	(0.308)	(0.509)	(0.454)	(0.042)	(0.019)	(0.025)	(0.379)
R^2	0.057	0.017	0.035	0.392	0.227	0.104	0.034
Number of observations	2,295	2,298	2,275	60	60	2,307	1,443

Note: Standard errors in parentheses.

Standard errors are robust to heteroscedasticity and clustered residuals within each traditional location.

Kisumu is divided into 13 administrative sublocations. All regressions include a full set of sub-location dummies.

Rural sample refers to a 5 percent random sample of Luo men 21-45 residing in rural Nyanza Province obtained from the 1989 census.

Early and late migrants are drawn from the 2001 Kisumu survey. Early migrants arrived in Kisumu before age 21, late migrants arrived after age 20.

Column 1 uses the migrant's years of schooling as the dependent variable.

Column 2 uses the migrant's number of siblings to measure the migrant's family size.

Column 3 uses inherited land as a measure of the migrant's wealth at home. This variable is the ratio of father's land (in acres) to the migrant's number of brothers.

Columns 4-5 study the level of migration, by traditional location, among late migrants and all migrants.

Migration is measured as the number of migrants from a given location in the Kisumu sample divided by the population of that location from the census.

This statistic is computed separately for Young and Old individuals, with the cut off at 30 years.

Late migration computes the corresponding statistic for late migrants, who arrived in Kisumu after age 20.

Column 6 uses the late migrant dummy as the dependent variable.

Column 7 treats cohabitation as the dependent variable, where cohabitation measures the number of months in the year that the migrant's wife spends in Kisumu.

This statistic is computed for currently married men. For men with multiple wives, this statistic applies to the wife who spent the most time in the city.

Young-age dummy=1 if the individual is less than or equal to 30 years, 0 otherwise.

Late-migrant dummy=1 if the migrant arrived in Kisumu after the age of 20, 0 otherwise.

South Nyanza dummy=1 if the location is in South Nyanza, 0 otherwise.

Table 6: Relatedness and Marriage

Dependent variable:	current marital status				
Sample:	rural	late migrants		all migrants	
Model:	Probit		LP	Probit	LP
	(1)	(2)	(3)	(4)	(5)
Relatedness	-0.025	-0.064	-0.017	0.004	0.003
	(0.009)	(0.019)	(0.005)	(0.023)	(0.008)
Relatedness-late migrant	--	--	--	-0.042	-0.013
				(0.024)	(0.007)
Young age	-1.095	-1.465	-0.397	-1.499	-0.427
	(0.027)	(0.131)	(0.026)	(0.091)	(0.019)
Late migrant	--	--	--	0.294	0.092
				(0.075)	(0.024)
South Nyanza	0.329	0.309	0.078	0.144	0.041
	(0.050)	(0.108)	(0.028)	(0.096)	(0.030)
R^2	0.130	0.211	0.220	0.187	0.209
Number of observations	8,535	1,061	1,063	2,307	2,307

Note: Standard errors in parentheses.

Standard errors are robust to heteroscedasticity and clustered residuals within each traditional location. All regressions include Kisumu sub-location dummies.

Rural sample refers to a 5 percent random sample of Luo men 21-45 residing in rural Nyanza Province obtained from the 1989 Census.

Early and late migrants are drawn from the 2001 Kisumu survey. Early migrants arrived in Kisumu before age 21, late migrants arrived after age 20.

Young-age dummy=1 if the individual is less than or equal to 30 years, 0 otherwise.

Late-migrant dummy=1 if the migrant arrived in Kisumu after age 20, 0 otherwise.

South Nyanza dummy=1 if the location is in South Nyanza, 0 otherwise.

Table 7: Marriage, Labor Market Outcomes and Remittances

Dependent variable:	employment			ln (income)			remittances		
Sample:	all migrants		late migrants	all migrants		late migrants	all migrants		late migrants
Model:	OLS	IV		OLS	IV		OLS	IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Current marital status	2.875	10.261	11.114	0.496	1.669	1.469	10.213	30.202	39.487
	(0.181)	(2.096)	(3.309)	(0.045)	(0.451)	(0.794)	(0.828)	(10.489)	(22.245)
Young age	-0.737	2.406	2.507	-0.303	0.156	0.028	-0.631	7.253	10.356
	(0.136)	(0.929)	(1.183)	(0.045)	(0.169)	(0.287)	(0.986)	(4.153)	(7.964)
Late migrant	-0.450	-0.908	--	-0.004	-0.080	--	-1.675	-2.962	--
	(0.209)	(0.259)		(0.057)	(0.059)		(1.026)	(1.107)	
South Nyanza	-0.156	-0.405	-0.365	0.034	-0.013	-0.008	0.494	-0.292	-1.847
	(0.176)	(0.228)	(0.275)	(0.043)	(0.042)	(0.066)	(1.289)	(1.588)	(2.042)
R^2	0.185	--	--	0.116	--	--	0.065	--	--
Number of obs	2,305	2,305	1,059	2,165	2,165	1,007	2,132	2,132	994

Note: Standard errors in parentheses.

Standard errors are robust to heteroscedasticity and clustered residuals within each traditional location. All regressions include Kisumu sub-location dummies.

Current marital status=1 if married, 0 if single.

Young-age dummy=1 if the individual is less than or equal to 30 years, 0 otherwise.

Late-migrant dummy=1 if the migrant arrived in Kisumu after age 20, 0 otherwise.

South Nyanza dummy=1 if the location is in South Nyanza, 0 otherwise.

Employment is measured as the number of months worked in the last year.

Income in last year is measured in thousands of Kenyan shillings.

Remittances measured as the percent of total annual income sent home to family. Top 1 percent of the remittances distribution is dropped from the sample.

Instrumental variable (IV) regressions use predicted value from probit marriage regression as instrument for marital status.

Table 8: Validating the Relatedness Instrument: Early versus Late Migrants

Dependent variable:	employment		ln (income)		remittances	
Sample:	late migrants	early migrants	late migrants	early migrants	late migrants	early migrants
	(1)	(2)	(3)	(4)	(5)	(6)
Relatedness	-0.216	-0.023	-0.031	-0.003	-0.802	-0.035
	(0.079)	(0.127)	(0.029)	(0.042)	(0.463)	(0.408)
Young age	-1.895	-2.028	-0.499	-0.498	-3.847	-5.478
	(0.209)	(0.194)	(0.054)	(0.050)	(1.218)	(1.169)
South Nyanza	0.577	-0.042	0.121	0.079	1.476	2.592
	(0.316)	(0.480)	(0.124)	(0.120)	(2.305)	(2.215)
R^2	0.094	0.089	0.076	0.058	0.034	0.041
Number of obs.	1,061	1,246	1,009	1,158	996	1,138

Note: Standard errors in parentheses.

Standard errors are robust to heteroscedasticity and clustered residuals within each traditional location. All regressions include sub-location dummies.

Young-age dummy=1 if the individual is less than or equal to 30 years, 0 otherwise.

Late-migrant dummy=1 if migrant arrived in Kisumu after the age of 20, 0 otherwise.

South Nyanza dummy=1 if the location is in South Nyanza, 0 otherwise.

Columns 1-2: number of months worked in the last year.

Column 3-4: income in last year is measured in thousands of Kenyan shillings.

Columns 5-6: remittances to the family, measured as the percent of total annual income. Top 1 percent of the remittances distribution dropped from the sample.

Table 9: Comparing Early and Late Migrants

Sample:	all migrants							ever married migrants
Dependent variable:	education	siblings	inherited land	attendance	contribution	visits	married at arrival	met wife at home
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Late migrant	0.175	0.001	-0.303	-0.008	0.007	0.077	0.343	0.128
	(0.106)	(0.265)	(0.271)	(0.018)	(0.024)	(0.028)	(0.015)	(0.027)
Young	0.632	-1.303	-0.572	-0.136	-0.190	-0.168	-0.200	-0.092
	(0.136)	(0.265)	(0.266)	(0.024)	(0.021)	(0.018)	(0.018)	(0.033)
R^2	0.080	0.033	0.089	0.062	0.082	0.096	0.311	0.079
Number of obs.	2,306	2,310	2,287	2,309	2,309	2,311	2,319	1,505

Note: Standard errors in parentheses.

Standard errors are robust to heteroscedasticity and clustered residuals within each traditional location.

All regressions include Kisumu sub-location dummies and traditional location dummies.

Late-migrant dummy=1 if migrant arrived in Kisumu after the age of 20, 0 otherwise.

Young-age dummy=1 if the individual is less than or equal to 30 years, 0 otherwise.

Column 1 uses the migrant's years of schooling as the dependent variable.

Column 2 uses the migrant's number of siblings to measure the migrant's family size.

Column 3 uses inherited land as a measure of the migrant's wealth at home. This variable is the ratio of father's land (in acres) to the migrant's number of brothers.

Column 4: attendance=1 if the migrant attends clan association meetings regularly, 0 otherwise.

Column 5: contribution=1 if the migrant contributes to the clan association, 0 otherwise.

Column 6: visits=1 if the migrant visits his rural home a few times a week or a few times a month, 0 otherwise.

Column 7: married at arrival=1, single=0.

Column 8: met wife at home=1, met in city=0.

Table 10: Marriage and Sexual Activity

Dependent variable:	number of non-marital partners			number of risky partners			risky status		
Sample:	all migrants		late migrants	all migrants		late migrants	all migrants		late migrants
Model:	OLS	IV		OLS	IV		OLS	IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Current marital status	-1.064	0.015	0.044	-0.407	0.518	0.337	-0.051	0.178	0.011
	(0.076)	(0.602)	(1.105)	(0.056)	(0.504)	(1.033)	(0.012)	(0.101)	(0.099)
Young age	0.427	0.886	0.858	0.207	0.600	0.521	0.012	0.109	0.031
	(0.065)	(0.274)	(0.487)	(0.053)	(0.227)	(0.450)	(0.015)	(0.044)	(0.037)
Late migrant	-0.033	-0.101	--	-0.060	-0.118	--	-0.027	-0.041	--
	(0.061)	(0.066)		(0.046)	(0.055)		(0.010)	(0.011)	
South Nyanza	-0.018	-0.056	-0.122	-0.034	-0.066	-0.087	-0.006	-0.013	-0.022
	(0.050)	(0.048)	(0.080)	(0.041)	(0.046)	(0.066)	(0.011)	(0.011)	(0.011)
R^2	0.228	--	--	0.085	--	--	0.036	--	--
Number of observations	2,272	2,272	1,059	2,272	2,272	1,059	2,266	2,266	1,056

Note: Standard errors in parentheses.

Standard errors are robust to heteroscedasticity and clustered residuals within each traditional location. All regressions include sub-location dummies.

Current marital status=1 if married, 0 if single.

Young-age dummy=1 if the individual is less than or equal to 30 years, 0 otherwise.

Late-migrant dummy=1 if the migrant arrived in Kisumu after age 20, 0 otherwise.

South Nyanza dummy=1 if the location is in South Nyanza, 0 otherwise.

Number of non-marital partners is measured over the past year.

Number of risky partners over the past year is computed as the difference between the number of non-marital partners and the number of jodiya.

Risky status=1 if the individual matched with casual partners or commercial sex workers a few times a week or a few times a month over the past year, 0 otherwise.

Instrumental variable (IV) regressions use predicted value from probit marriage regression as instrument for current marital status.

Figure 1: Relatedness across Traditional Locations

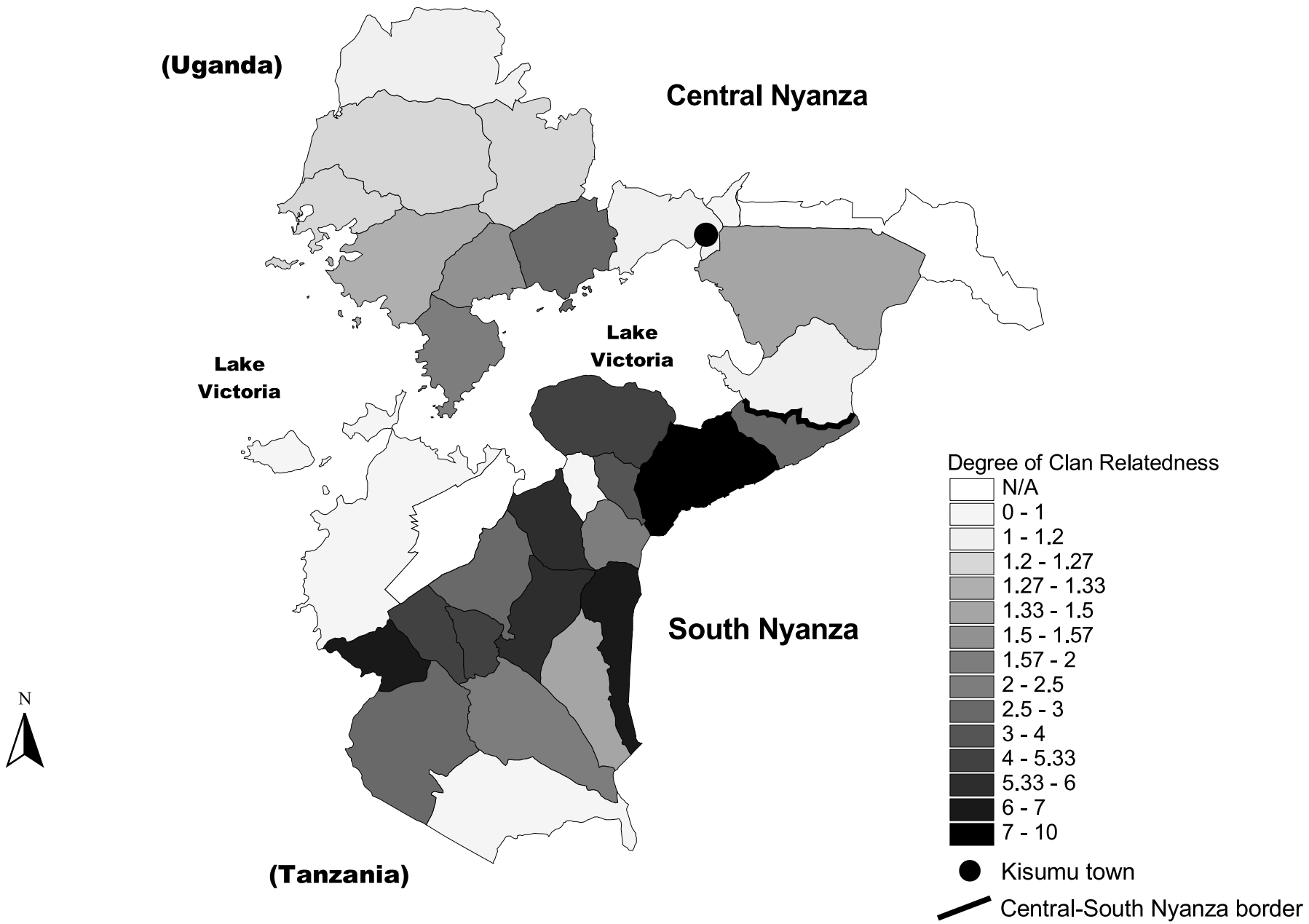

Figure 2: Marriage and Divorce