Chatting at Church: Information Diffusion through Religious Networks*

David M. A. Murphy[1], David R. Lee[2], and Vesall Nourani[3]

[1]*Schroeder School of Business, University of Evansville.*

[2]*Charles H. Dyson School of Applied Economics and Management, Cornell University.*

[3]*Department of Economics, Massachusetts Institute of Technology.*

Version Date: September 18, 2018

JEL Codes: Z12, O33, D83

Abstract Religious institutions (RIs) can be an important platform for information diffusion through the creation of bridging social ties in communities. Using a data set from western Kenya, this study analyses the relationship between attending an RI with and receiving practical advice from a peer. To causally identify this relationship, we use a novel spatial instrumental variable strategy that incorporates the triangular distances between peers and RI locations within Kenyan villages. We find that shared attendance of two peers at an RI increases the likelihood of seeking out and receiving advice from their peer by 33 percentage points.

*Funding for this study provided by: The Atkinson Center for a Sustainable Future, Cornell University; a U.S. Borlaug Fellowship funded through USAID; and several grants and fellowships from Cornell University. Crucial support for this project also comes from the research project "Improving Bean Yields by Reversing Soil Degradation and Reducing Soil Borne Pathogens on Small-holder Farms in western Kenya," funded by a USDA National Institute of Food and Agriculture grant, and from the International Institute of Tropical Agriculture (IITA), Nairobi, Kenya.

Throughout much of history, religion has been the central platform for human interaction. In agrarian societies, such interactions become crucial avenues through which households engage in economic activity — either through informal gift-giving, risk-sharing, or exchange of information. It is perhaps no surprise, then, that participation in religious activity continues to play a large role in these contexts. Nevertheless, there is as yet little empirical evidence suggesting that participation in a religious community influences these types of interactions.

This paper investigates the causal linkages between affiliation in religious institutions (RIs) and information sharing, as well as trust and friendship. In many areas of the world, RIs are a focus of the community and often serve as centers for local government, conflict resolution, and agricultural and economic development (Jones, 2005; Bompani and Smith, 2013). Networks formed through RIs can therefore potentially be conduits for information sharing, especially due to the greater levels of trust and confidence that often exist between members of a shared religious community (Johansson-Stenman, Mahmud, and Martinsson, 2009; Asadullah, 2017).

Our analysis suggests that RIs may be central in the formation of "weak ties" within these communities. Unlike strong ties, which exist between family members and life-long friends, weak ties are necessary within social networks to serve as bridges among social cliques to diffuse information (Granovetter, 1973). Putnam (2000), for example, observed that the decline in membership in civic associations was correlated with the reduction of bridging ties in the United States and thus, he argues, a decay of social capital. Using a sample of individuals and RIs from western Kenya, our paper confirms the causal association of RIs in creating these bridging ties, causing these institutions to effectively become platforms for information diffusion. To show this, we use a novel dataset collected in 2016 that included questions on the religious affiliation of individuals and a detailed social network module. This network module asked randomly sampled individuals to indicate information exchange, trust, and friendship relationships with random subsets of other individuals within the same sample.

Additionally, we collected GPS coordinates of each household and RI in our study area. Using network dyads as our unit of observation (links between households i and j), we employ an instrumental variable strategy that regresses common RI affiliation between i and j against outcome variables such as information exchange. The core intuition of this strategy is the following: if household and RI locations within a village are exogenously determined (i.e., via inheritance of household plot), we can construct lines between the geographic locations of households i and j and link the endpoints of this line to any given RI in the village. The geometric properties of the ensuing triangles generate multiple distance-based instruments that predict whether the two households attend the same RI. We argue that, since such instruments only use information constructed using geographic distances, they are excludable from the second stage of the analysis as long as we control for distance between households i and j in addition to the unobservable characteristics of each village. In addition, we use a novel robustness check to show that the influence of nearby, within-neighborhood, social institutions are negligible. Indeed, the precise RI location is the key driver for the validity of our instruments. This instrumental variable strategy enables us to identify the impact of joint RI membership on information sharing: attendance at the same RI increases the probability of seeking out and receiving agricultural information from a particular peer by 33 percentage points — a significant effect.

However, it is not immediately clear what is driving information sharing from shared RI attendance. We thus explore several possible mechanisms, and our results indicate that social spaces formed through RIs increase trust and perceptions of information quality between individuals — although they do not necessarily lead to significantly stronger bonds of friendship. Thus, RIs create spaces that facilitate higher levels of information diffusion, perhaps due to their role in increasing opportunities for interpersonal connection. Finally, we show that RIs have the greatest effect on information sharing between individuals who, based on observable characteristics, would be *less* likely to be friends in the absence of the RI, demonstrating the importance of this social space in creating social links among disparate individuals.

This study makes three core contributions. The first is that it demonstrates the importance of social spaces formed through RIs for information diffusion across individuals. Within rural Sub-Saharan Africa (SSA), this result has important implications for the diffusion of information that can lead to, for example, agricultural productivity enhancements. Many studies have found peer networks to be instrumental in diffusing information (e.g., Munshi, 2004; Conley and Udry, 2010). Farmers not only learn about technologies and practices from their own experiences, but also rely on their peers as sources for information (e.g., Feder and Slade, 1984; Foster and Rosenzweig, 1995). This is at least partly due to a farmer's peer having "safety credibility," or trustworthiness, that can overcome bias towards the status quo in driving technology adoption (Rogers, 1995). The influence of a farmer's peers has been argued to affect adoption rates more strongly even than agricultural extension agents, who likely have greater technical expertise (Krishnan and Patnam, 2014). As a result, policymakers can attempt to use networks formed through institutional membership to aid in the diffusion of welfare-improving information. Several studies have analyzed methods to locate optimal injection points of information into peer networks for optimal rates of diffusion (Banerjee et al., 2013; Beaman et al., 2015). A major contribution of this research, then, is demonstrating the role of RIs as one institutional mechanism for introducing and disseminating productivity-enhancing information to rural smallholders through the construction of bridging ties.

Second, we introduce a novel instrumental variable (IV) strategy for dyadic data that can be used when the social proximity of the nodes in a dyad are determined, in part, by the dyad's geographic proximity to a nearby point of interest. Specifically, this strategy correlates measures of within-village variation in dyadic distance to RIs with the probability of mutual RI attendance. In many village settings around the world, residents of a village live in homes they have inherited, thereby reducing selection into homesteads within villages. By simultaneously controlling for the geographic proximity of households to each other, we isolate the causal channel of mutual RI attendance on other, social, dyadic outcomes. While in this study, we use this IV strategy to identify peer dyads and RIs, it can potentially be replicated in other contexts as well, such as for analysis of peer relationships influenced by mutual school, business, or civic center attendance.

Third, we provide evidence of a causal relationship between attendance at the same RI and greater levels of trust between peers within our rural Kenyan sample. While there has been a recognition for some time that joint membership in social institutions is related to stronger bonds between individuals and higher levels of social capital (Putnam, 2000), there have thus far been no studies that attempt to show a causal link between joint membership at religious institutions and increased trust and information sharing. The effect of shared religion (not necessarily RI) on interpersonal relationships has been explored by others in the field: both Johansson-Stenman, Mahmud, and Martinsson (2009) and Asadullah (2017) analyze data from Bangladesh, and find that co-religionists have greater levels of trust in each other compared to those from other religions. Using experimental data from Malaysia, Chuah et al. (2014) find greater levels of cooperation among subjects who know that their co-player is the same religion.

It should also be noted that there there are numerous studies that examine religion in an economic context (see Iannaccone (1998) and Iyer (2016) for reviews). Much of this literature asks whether the degree of religiosity of an individual or region leads to differential economic outcomes. Our study diverges from this focus by analyzing the extent to which religion, conceived as a social institution, influences interpersonal connections — which, in turn, have information-sharing and, ultimately, economic outcomes. In our conclusion, we discuss how future work might benefit from examining the extent to which economic outcomes that require coordination (social choice) are affected by religion. We also discuss potential negative effects on equitable economic growth as a result of RI membership. Religious groups can often be exclusive by definition, and we discuss whether social spaces formed by RIs are beneficial to villages overall if they lead to inequities between those inside the RI and those outside.

Religion in Kenya

Religion in Kenya historically was based on beliefs and customs that differed by tribe, although many religious affiliations shared common characteristics such as monotheism and ancestor veneration (Kenyatta, 1938). Islam became the majority religion in the coastal region by the 14th century, but over the centuries, Christianity has since grown to be the majority religion (83 percent of the population) (UN, 2009). Christianity in Kenya is not homogeneous, however; it represents many different traditions, often incorporating indigenous beliefs or practices (Meyer, 2004).

Religion in Kenya, however, varies significantly by region. Eastern Kenya, especially the northeast, is predominantly Muslim. Hinduism is common in the larger cities among the Indian population, while the majority of African traditionalists are among the Maasai in the south, the Turkana in the north, and among other pastoralists. Christianity is the predominant religion in the remainder of the country, especially in the more densely populated central and western regions of Kenya. Table 1 summarizes the most recent statistics for religious membership in Kenya.

The practice of religion is often a daily group activity in Kenya. Not only are there weddings and funerals that take place during the week, but many RIs have weekday prayer services, dance and singing practices, and clubs for particular groups (such as youths or mothers) within the RI. As Iannaccone (1992) shows, religions often impose mandatory activities or prohibitions that raise the cost of engaging in other activities and relationships, leading to a substitution towards RI activities and members' deriving increased reliance and support from one's RI. Based on anecdotal observations, this seems to be particularly evident in Kenya.

The recent rise of Charismatic, African indigenous, and "Neo-Pentecostal" (NP) churches in Kenya has had significant effects on the religious landscape and implications for peer associations. Such churches have experienced massive growth since the 1980s, usually at the expense of more established churches such as the Roman Catholic and Anglican churches. Anecdotal observations and survey data suggest that NPs likely constitute a plurality of the population in many areas of western Kenya. NP churches tend to be more numerous and smaller than established Christian churches. Moreover, membership or baptism into an NP church often means breaking with one's past and the realignment of one's previous peer network (Meyer, 2004; Jones, 2005). Members of NP churches do not typically withdraw from village-level groups or from interactions with other village members; rather, they refocus their peer group towards members of a new bounded religious community. The rise of NP and Charismatic churches has thus altered traditional information networks and decreased interactions between those attending different RIs, while deepening trust and reliance within a singular social space.

Religious Context within Sample Villages

We discuss the religious context in our study by first describing church characteristics by religious denomination. In Table 2, we include summary statistics regarding each religion represented in the sample. Each individual was asked the average number of people who attend the main weekly service (RI size) of their RI and how often s/he attends the RI (RI attendance). We find that Neo-Pentecostal and Charismatic churches are on average the smallest (aside from Other Christian and Other Religion RIs), with an average attendance of 117.2 people. These churches are also the most frequently attended, aside from those in the Other Religion category which constitute only a handful of observations (and include Muslim mosques and African traditionalists). Other Christians, among the Christian groups, on average have the smallest RI, while Catholics have the largest average RI size.

We next turn our focus to the religious attendance of individuals in our sample. Individuals often attend churches they live close to. The median distance to an individual's RI is

897 meters and individuals state that the median amount of time it takes to get to his/her RI is 20 minutes. While many attend church multiple times a week, at the median, individuals attend their RI on a weekly basis. Of the 884 individuals in the sample, only four reported having no religion. A plurality of individuals attend a Pentecostal or Charismatic church, while Anglicans are the second largest group, and Roman Catholics the third. The complete list of religions found in our sample is given in Table 3.

In this study, our focus is on within-village analysis of peers and RI attendance, thus we describe characteristics of RI distributions at the village-level. In our sample of 17 rural villages, each village contained on average 12 RIs (ranging from 5 to 17 RIs), with an average village containing 97 households.[1] The median year of RI establishment is 2000, with RIs in the sample ranging in the year they were established from 1921 to 2017. The median distance of each RI to within-village households in our sample is 1143 meters. Significant heterogeneity exists however among villages, and in Appendix Table A1 we present additional data on village level RI statistics.

To understand the religious context in greater detail, we interviewed a representative of each RI to gain qualitative insights into the history of each institution. Qualitative data collected from RIs in the sampled villages show that RI locations are not randomly determined. RIs often acquire land according to where founding members gifted land to the RI, while many others indicate that they chose their location to be convenient to members of the community or to be highly visible – for example, on permanent roads or in market centers. As one RI reported: the reason the RI located in this place was to "1) make converts, 2) be near to the road, and 3) to save souls." As detailed in following sections, we construct IVs based on the location of these RIs for the identification of our empirical model.

Model and Identification

In the survey underlying this study, we ask each individual i whether they have sought out and received agricultural information from individual j. The probability of individual i receiving agricultural information from individual j, denoted $\Pr(\theta_{ij} = 1)$, is represented by the following equation:

$$P(\theta_{ij} = 1) = f(Z_{ij}, X_{ij}, A_{ij}, D_{ij}, V_{ij}) \tag{1}$$

where Z_{ij} is a set of variables measuring the strength of the relationship between individuals i and j (such as family ties) and D_{ij} is the spatial distance between i and j's homesteads. X_{ij} are household and demographic characteristics of i and j; A_{ij} are characteristics of farmer i and j's farms; and V_{ij} represents other unobserved variables between i and j that can influence the probability of $\theta = 1$, such as reputation or outcomes of previous interactions. We further define the relationship strength, Z_{ij} as:

$$Z_{ij} = g(\zeta_{ij}, R_{ij}) \tag{2}$$

where ζ_{ij} is a binary variable that represents whether individual i states that j attends the same RI, and R_{ij} is an exogenous subset of other relationship variables contained in Table 5 in the following section.[2]

We hypothesize that mutual RI attendance increases the strength of a social tie, which in turn increases the probability of information exchange between i and j (e.g., the likelihood of seeking out and receiving advice θ_{ij}). Formally, this hypothesis is represented by the following comparative static:

$$\frac{\partial \theta_{ij}}{\partial \zeta_{ij}} = \frac{\partial f}{\partial g}\frac{\partial g}{\partial \zeta_{ij}} > 0 \tag{3}$$

For ease of interpretation, we estimate Equation 1 using a linear probability model due to our interest in applying two-staged-least squares (2SLS) estimation procedure in a context with binary outcome and explanatory variables, θ and ζ. This avoids inconsistent estimates that a probit or logit estimator would yield when using instrumental variables (Wooldridge, 2002). Moreover, estimates of the coefficients of binary variables in this situation are shown to be unbiased using a LPM and standard errors may be more accurate than with other methods (Deke, 2014).

Formally, then, we seek to estimate:

$$P(\theta_{ij} = 1) = \alpha + \beta_1\zeta_{ij} + \beta_2 R_{ij} + \beta_3 X_{ij} + \beta_4 A_{ij} + \beta_5 D_{ij} + \phi + \varepsilon_{ij} \tag{4}$$

where ϕ represents a vector of village, enumerator and survey-month fixed effects. Because we are using a dyadic regression, symmetry must be preserved as variables enter into the estimation (Fafchamps and Gubert, 2007). That is, as $P(\theta_{ij} = 1)$ examines the effect between i and j, so all the variables must also be in dyadic form and reflect a characteristic of the *dyad*, which is the unit of analysis. Specifically, we model this relationship as unidirectional, as we assume that $P(\theta_{ij} = 1)$ does not necessarily equal $P(\theta_{ji} = 1)$. Therefore, as Fafchamps and Gubert (2007) show, we can estimate the following equation to satisfy the symmetry requirement:

$$\begin{aligned} P(\theta_{ij} = 1) = \alpha &+ \beta_1\zeta_{ij} + \beta_2 R_{ij} + \beta_3(X_i + X_j) + \beta_4(X_i - X_j) \\ &+ \beta_5(A_i + A_j) + \beta_6(A_i - A_j) + \beta_7 D_{ij} + \phi + \varepsilon_{ij} \end{aligned} \tag{5}$$

where the sums and differences of variables X and A between individuals i and j are used as regressors.[3]

Spatial Distance of Dyad from RI as Instrument

As stated earlier, we are primarily concerned with establishing that an increase in mutual RI attendance (increase in ζ_{ij}) causes an increase in the strength of the social tie between i and j (increase in θ_{ij}) as measured through increased information sharing and perceptions of trust and information quality. There are many factors that simultaneously determine the strength of a connection between i and j and their mutual RI attendance, such as geographic proximity, age proximity, kinship relationship, and many other observable and unobservable characteristics embedded in X_{ij}, A_{ij}, D_{ij}, or V_{ij}. Therefore, we propose an instrumental variable identification strategy that takes advantage of nuanced features of the local context in our approach to identifying this causal relationship.

Formally, we must instrument for endogeneity in ζ_{ij} (knowledge of co-membership at an RI) to identify this causal relationship. To do this, we utilize two important features of the local context gleaned from interviews with RI leaders in our study area along with institutional knowledge of homestead tenure rights in Western Kenya to construct spatially induced variables that predict co-membership in an RI: 1) in a vast majority of cases, homestead locations are determined at birth (due to market frictions surrounding property rights for land) and 2) churches locate in areas that are locationally convenient to constituents. This knowledge suggests that a homestead's proximity to an RI is predictive of household i's attendance at a given RI. However, our unit of observation is the relationship between i and j, which implies that to predict mutual church attendance we need to construct a variable that takes the dyad ij's distance from RI_k, denoted k to reflect each church villagers attend, into account.

Denote this distance, for the time being, by $F_{ijk}(C_i, C_j, C_k)$ which is constructed using the GPS coordinates, C, of homesteads i, j, and RI k. Notice, first, that lines drawn to connect these three coordinates form a triangle — we exploit properties of the geometry of triangles in one of our instruments. Second, the spatial distance between households i and j, D_{ij} is a function of C_i and C_j. Third, the spatial distance between a homestead, i, and RI, k, can be denoted by D_{ik}.

With this notation, we describe three instruments used in our analysis: 1) the triangular median formed by connecting the mid-point of the line between C_i and C_j to the vertex C_k; 2) percentage of the nearest five RIs shared by both i and j, and 3) joint probability distribution of attendance at RIs based on the spatial distance between each i and RI_k, D_{ik}

and D_{jk}. The first instrument provides us with a measure of the dyad's distance to the nearest church (as opposed to each household's distance), the second, a measure of relative probability of church attendance, and the third provides a measure of the probability of church attendance based on absolute distance measures.

Our instruments must only influence outcomes through their ability to predict mutual RI attendance. In other words, our distance measures can not influence the probability that two households interact with each other apart from their influence on predicting mutual RI attendance. Thus, to identify the effect of interest we make the following identifying assumption:

Identifying Assumption. *Conditional on controlling for spatial distance between households i and j, measures of spatial distance between dyad ij and RI_k provide exogenous variation in the determination of mutual RI attendance within dyad ij.*

This follows from the exogenous determination of homestead location: 89% of households in our sample inherited their homestead (either from parents or spouse). Households do not select their location of residence endogenously and we are controlling for distances between households i and j. Thus, our instruments are identifying the effect of mutual RI attendance on social relationship outcomes off of the "residual RI distance," or the component of the RI's distance to the dyad that is uncorrelated with how proximate households i and j are to one another. In effect, the above assumption states that this residual RI distance component provides a source of exogenous variation in predicting mutual RI attendance.

Since we are also controlling for village fixed-effects, our analysis examines *within-village* variation in a dyad's distance from specific RIs in addition to controlling for the distance between the two homesteads in the dyads in the second stage (D_{ij}). While "neighborhood" effects are negligible in our context due to the small size of villages, we are nevertheless controlling for any effects stemming from churches being located in areas of the village where similar households are located (perhaps due to historical factors) through household distance between i and j.[4] Furthermore, it is conceivable that measures of RI attendance based on geographic distance are correlated with measures such as population density and proximity between households i and j, which also influence whether i and j form other types of social connections. We argue that this is not a concern in our study because we control for both covariates through X and these are the main channels through which our instruments potentially influence outcomes in the second stage – thus, our instruments are not correlated with the second stage error term. Below, we discuss how each instrument is calculated and present arguments that lead us to conclude that each instrument is excludable and valid and thus produces unbiased results from our estimation.

1. *Spatial distance triangular geometry.* For this instrument, we used the spatial distribution of minimum triangular medians between farmer homesteads and RIs within each village (using distance between homesteads as the base of the triangle). An example is shown in Figure 1. In the figure are the homesteads of two individuals, i and j. We calculated the physical distance between them, and determined the midpoint coordinate of this distance, $B_{ij}(C_i, C_j)$. We next calculated the distance between B_{ij} and every RI in the area — in other words, we calculate all distances between coordinates B_{ij} and C_k for all RI k in K_v where K_v denotes the set of RIs attended by households in village v. The set of the minimum of these distances for each individual is used as our instrumental variable. In the example presented in Figure 1, there are two RIs. Thus, there are two distances that we measure resulting in a single instrument constructed using $min_{ij}\{\overline{B_{ij}C}_1, \overline{B_{ij}C}_2\}$. In the case of Figure 1 we pick $\overline{B_{ij}C}_1$ as the instrument's value for the relationship between i and j.

The resulting instrument has a strong positive correlation with RI co-membership: the farther the nearest RI from two peers, the more likely they are to attend that nearest RI. RIs are sometimes clustered in more central locations. Thus, a short distance to the nearest RI generally means more RI choices near both homesteads, leading to a lower likelihood of mutual RI attendance.

2. *Percentage of nearest RIs shared by both i and j.* For this instrument, we rank the closest five RIs to each individual.[5] We then identify the total number of unique RIs shared in

each dyad (each i and j), dividing by the total number of ranked RIs. The resulting IV then ranges from 0 to 1, with 0 indicating that among each i and j's nearest five RIs, no RIs are shared, and 1 indicating that all five are shared between both i and j. The instrument should be positively correlated with attendance at the same RI, as the more proximate RIs that i and j have in common, the more likely i and j are to attend a particular RI together. As in the case of the prior IV, the validity of this instrument rests in the assumption that homestead locations are exogenous. Furthermore, since we are directly controlling for distance between i and j in the second stage, it is reasonable to assume that this instrument only influences our outcome variable of interest through the first-stage relationship.

3. *Joint probability distribution of attendance at RIs based on spatial distance between each i and RI.* Here, we assume that i's attendance at a particular RI is negatively correlated with i's distance from any given RI. Specifically, we denote the probability that i attends RI_k as $P(attend_i = RI_k) = \frac{1/D_{ik}}{\sum_{k=1}^{N} 1/D_{ik}}$. Assuming that i and j's probability distributions are independent (as is required for an excludable instrument), we can construct a probability measure for joint RI attendance for every RI in i and j's village. Specifically, our instrument is $P(attend_i = RI_k \cap attend_j = RI_k) = P(attend_i = RI_k)P(attend_j = RI_k) = \sum_{l=1}^{N} \frac{1/D_{il}}{\sum_{k=1}^{N} 1/D_{ik}} \frac{1/D_{jl}}{\sum_{k=1}^{N} 1/D_{jk}}$.

A final concern related to any analysis of the effect of RI attendance on outcomes of interest is the first-stage decision to belong to an RI in the first place. In rural Kenya attendance and membership at an RI is nearly universal. In our sample of 884 individuals, only four individuals reported that they had no RI membership, and another six had RI membership but never attended. This means that 99% of the sample attended RIs regularly, mitigating endogeneity concerns on this score. Of course, individuals may tell the enumerator they attend but in reality do not attend. This is certainly a possibility, but anecdotal evidence suggests that the vast majority of individuals do attend.

Data

Data for this study were collected in 2016 in partnership with the International Institute of Tropical Agriculture (IITA). Participants were randomly selected from official village rosters of household heads in 17 villages from three counties (Busia, Bungoma, and Kakamega) in western Kenya (we show these village locations in Appendix Figure A1). After the household heads were selected from each village roster, the fieldwork team visited the households to ensure that the household heads were available and willing to participate in the study. If so, a soil sample was taken from a plot of each sample household and analyzed by IITA.[6] Within a couple of months of the initial visit, the research team returned to the household and separately interviewed both the household head and spouse. In total, 884 individuals in 548 households were surveyed. Table 4 contains the summary statistics of these individuals, households, and villages.

As can be seen in Table 4, 58 percent of the sample were women. This is due to the fairly large number of widows in the sample (14 percent), and to the fact that many men have migrated to cities for work. The average age of the individuals in the sample (household heads and spouses) was 48 years old, with an average of nearly eight years of schooling. About 56 percent of the sample could do a simple multiplication problem given to them by the enumerator at the time of the survey. Farms are very small and generally poor in this sample: the average farm size was just over one acre, 72 percent of the households had farmland that was nitrogen deficient (<20mg NO_3 per kg soil), and another 84 percent had farmland that was phosphorus deficient (<0.5mg PO_4^{-3} per kg soil).[7] Of the plots tested in the sample, 33 percent also were deficient in organic matter (<350mg Active C per kg soil). This is despite 88 percent of households using inorganic fertilizer and 45 percent using organic inputs in the past year).

Included in the survey was also a detailed peer network module, which utilized a "random matching within sample" (RMWS) methodology to generate peer linkages (Conley and

Udry, 2010). Each respondent was asked network questions about one to three pre-identified progressive farmers from his/her village in addition to a random subset of ten other individuals from the village.[8] RMWS is convenient to use because it is generally not feasible to ask each respondent about every other individual in the village due to time constraints. Furthermore, this method can obtain dyadic results closer to that of the population than other methods (Santos and Barrett, 2008).

In the network module, the respondent was asked multiple questions about his/her peers in the random village sub-sample to determine if the peer linkage existed, and the strength of that link. Table 5 contains summary statistics of information obtained from questions asked to individual i regarding another individual j in their village. We find that, on average, about 69% of randomly selected individuals are known to i, and i has met 67% of the individuals in person. The survey also found that i states that j attends the same RI as him/herself in 34% of peer dyads (conditional on having met j), and that i has sought out and received some kind of agricultural advice from j in 23% of these linkages. However, despite most respondents believing individual j is a "good or close friend" (4 or 5 on a 5-point scale), in only 21% of the personal linkages would individual i trust j to watch a valuable item for one week. We therefore see that even among farmers who often have known each other for long periods of time and consider each other good friends, a trust deficit often exists that may impede information diffusion within the village network. What we show through this analysis is that it appears that individuals who attend the same social institution, e.g. church or mosque, are able to bridge this trust deficit and, as a result, are more likely to communicate important information among one another.

We add two additional columns to Table 5 that split the sample between those who stated that they have the same primary RI, and those who responded with different primary RIs. For every measure of relationship strength, those who were associated with the same primary RI had a stronger relationship than those who were not, with the differences significant (p-value of 0.00) in all but one case. This suggests a potential exclusionary effect from RI membership, as those attending the same RI have stronger linkages than those who do not attend the same RI, though these statistics alone are insufficient for any form of causal analysis.

Results

The estimation of Equation 5 uses 2SLS within a linear probability model to estimate the causal effect of i and j's attendance at the same RI on i receiving agricultural advice from j among agricultural households in western Kenya. The first stage results of the estimation are located in Appendix Table A2, and demonstrate the strength of the instruments: with full dyadic regressors, the first stage has an F-statistic of 46.33. Results from the OLS and 2SLS regressions are in Table 6. Dyadic standard errors in these data are not independent and are correlated across observations of the same individual. Without correction, this correlation can cause inconsistent estimates of the standard errors and possibly underestimates their magnitudes (Fafchamps and Gubert, 2007). To mitigate this concern, we estimate both OLS and 2SLS using two-way clustering of standard errors at the individual level (Cameron, Gelbach, and Miller, 2011). All regressions include fixed effects for village, enumerator, and survey month.

Based on the coefficient estimations from the 2SLS regression in Columns 2 and 4, it appears that the OLS coefficients are significantly downward biased (towards zero). This is likely due to unobserved factors positively correlated with both attending the same RI and sharing of agricultural information, such as a friendship formed outside of the RI. Looking at the 2SLS results in Column 4, we find that the belief that individual j attends the same RI as individual i leads to an increase in the probability of receiving agricultural advice from j by 33 percentage points and is statistically significant at the p=0.05 level. This suggests a significant causal impact on information transfer to i from attending the same RI as j.

In addition, in Columns 3 and 4 we find that the gender of i and j appear to not be related to sharing of advice. Estimations in Columns 3 and 4 also include the sums and differences of household and demographic characteristics (X_{ij}) and farm characteristics (A_{ij}) (results in Appendix Tables A3 and A4), which have little significant impact on i receiving advice from j in the 2SLS regressions and likewise little effect on the coefficient on ζ_{ij}. Only a few of these regressors are significant, including the sums of peers' ages, difference in peers' asset levels, and difference in (squared) farm size. The significance on age (and age squared) indicates that as the ages of the two peers increase, they are more likely to seek out and share information from one another (at a decreasing rate). On the other hand, as the difference in asset levels and farm size increase between peers, they are less likely to seek out and share information from one another - suggesting a social gulf between the relatively richer and poorer members of a village network.

In Table 6, the variable measuring the distance between the homesteads of i and j is exogenous, conditional on the "immediate family" variable in the estimation. The majority of households in the sample are on inherited land (78%), while a larger majority either live on inherited land or came into the land through marriage (89%). Since it is only a small minority of households that reside on purchased land (11%), it is likely that there are no confounding variables that affect the relationship between distance between i and j and receiving advice from j.[9] As we can see from Table 6, the effect of distance on receiving agricultural advice is negative, and decreases at a decreasing rate as the distance between the homesteads of i and j increases. We also see that i knowing j for at least ten years and whether j is an immediate family member have strong effects on the likelihood of i receiving agricultural information from j, increasing the likelihood of seeking out and receiving agricultural advice by eight and twenty percentage points respectively. Whether j is an extended family member has a statistically significant though smaller effect, increasing the probability of receiving this form of information from j by about six percentage points. These results point to the considerable importance of family ties in the sharing of agricultural information in western Kenya.

Robustness Checks

There are several challenges to the identification strategy driving our analysis. Recall, since we have included village fixed effects in the analysis, all challenges address issues that exist in *within-village* sources of variation.

Centrally located households versus distant households. First, centrally located households may be less likely to share RIs and are also less likely to mutually visit social other spaces where knowledge diffusion might occur. This is because centrally located households may have more potential social spaces to frequent and are thus less likely to frequent the same space at the same times.[10] To ensure this challenge is not significantly confounding our results, we calculate the average distance between each individual and his/her peers, and then create a dyadic variable τ_{ij} by summing this average distance with the average distance of his/her peer. Lower values of this measure indicate a relatively more centrally located dyad. We include these variables as regressors in our estimations in Table 7.[11] The results, shown in Columns 3 and 4 of Table 7, demonstrate that these regressors are not statistically significant, suggesting that more remote individuals do not have differential results compared to less remote individuals.

Church or other nearby social space. Second, the location of an RI can be correlated with the locations of other spaces where knowledge diffusion can occur. For example, if RIs are located in neighborhoods where individuals congregate for other reasons (e.g., near a shop, village common area, etc.), then our instruments will affect social interactions directly regardless of the RI's influence. To check for this possibility, we re-run our analysis with simulated locations in neighborhoods around actual RIs. Specifically, since neighborhoods are likely to be determined by the relative density of households within villages, we define a *neighborhood-radius* around each church equivalent to 1/9th and 1/6th of the standard deviation of within-village church distances. Appendix Figure A2 demonstrates how data are simulated in a representative village. Then, after calculating radii for each village, we generate 1,000 new datasets with simulated RI locations (for each radius measure). For each dataset, we construct all of our instruments and re-run our analysis.

If the precise location of the church matters more than its surrounding neighborhood, simulated RI locations should be less meaningful in both the first and second stages of our 2SLS estimation procedure. In the first stage, F statistics should weaken as the neighborhood-radius increases and, in the second stage, results should be less significant. Table 8 summarizes the results from this exercise. The top panel shows the mean p-value of the second stage coefficient under the 1/9th and 1/6th neighborhood-radius rules; the bottom panel summarizes the F-statistic. The mean p-value for our conservative neighborhood distance (1/9th) is 0.146, compared to the original p-value of 0.017 with the true RI location. Of the 1,000 simulations, only 10.7% have p-values less than 0.017, suggesting that these draws represent Type I errors and the null hypothesis that the coefficient is not statistically significantly different from zero should not be rejected. Simulated RIs are less statistically significant as the neighborhood-radius expands to 1/6th neighborhood-radius rule as expected. The F-statistic on the first-stage decreases to 5.45 (from 8.08) under the 1/9th neighborhood-radius rule and 4.40 under the 1/6th rule.[12] We conclude that the precise location of the church is a critical component of our analysis, thus the influence of neighborhood-related confounders is negligible.

Non-random distribution of households. Suppose that there are some unobservable variables that are uncorrelated with distance between households i and j, but correlated with our distance instruments and the structure of the social networks. If this is the case, a sufficient characteristic for identifying the causal relationship between church attendance and outcomes of interest is to have spatially randomly distributed households. From experience, we know that homesteads in our sample villages are inhabited by households who inherit the homestead and there is very little selection into a homestead. Furthermore, there is substantial mixing of households within villages. However, there may still be clusters of households that possess certain characteristics within villages that suggest households are not distributed in a spatially random manner. To determine whether spatial autocorrelation is potentially an issue, we construct measures of autocorrelation by village after Moran (1950) for variables including asset index, education, and religion (e.g. Anglican, Pentecostal, etc.). We show the results in Appendix Table A6, which demonstrate some potential concerns. Two villages appear to have strong spatial autocorrelation by asset index (Village 1 and Village 11), while several villages appear to have spatial autocorrelation by religion (notably, Villages 6, 9, and 18). For this reason, as an additional robustness check, we re-estimate our primary results (Table 6, Column 4) omitting these villages in two groups: first those that have potential spatial autocorrelation by assets, and then by those that have potential spatial autocorrelation by religious membership. These results are in Appendix Table A7, and demonstrate that the results change very little when the villages with potential spatial autocorrelation are removed from the analysis. The original results show that joint attendance at the same RI increases the probability of receiving agricultural advice from one's peer by 33 percentage points, significant at the p=0.05 level. In the first column of Table A7, with the two villages omitted that have potential spatial autocorrelation in assets, the results are essentially unchanged. In the second column, with three villages omitted that have potential spatial autocorrelation in religious membership, the magnitude of the effect decreases slightly, to 0.30, and the statistical significance also decreases slightly. This is most likely simply due to less power available as a result of the smaller sample size. Overall, the results suggest that spatial autocorrelation does not appear to be a factor that biases our results.

Correlation of family relationships with instruments. While unlikely, there is the possibility that our instruments are not correlated with RI attendance, but instead are correlated with another variable that is itself correlated with shared RI attendance. An obvious possibility is family connections: those in the same family are more likely to attend the same RI, and the instruments that we use could potentially be picking up this correlation. To explore this possibility, we set a measure of family relationship as the dependent variable and conduct 2SLS regressions with our standard set of instruments to determine whether we find any causal relationships of shared RI attendance on family connections. In Appendix Table A8, we show the results using both immediate family (Column 1) and extended family (Column 2) as dependent variables. As expected, we find no causal relationship, suggesting that our instruments are not picking up spurious correlations that are biasing our results.

Alternative Specifications and Potential Mechanisms

In our analysis, we were curious whether our results were heterogeneous in the type of RI. Because of data limitations, we are unable to test heterogeneity with respect to religion (e.g. Pentecostal, Anglican). However, we can test whether our results vary with size of RI, with the understanding that on average, certain types of RIs, such as Neo-Pentecostal and Charismatic churches are smaller than other RIs. Small social institutions may create more intimate social spaces and lead to more frequent interactions between same members. In Column 1 of Table 7, we explore this by including an indicator variable for whether the RI that i attends and the RI that j attends are both "small," defined as being less than the median RI size in the full sample. In Column 2, we interact this variable (RI_{sml}) with attendance of i and j at the same RI (ζ_{ij}) to inform us whether joint attendance at a small RI increases the likelihood of shared information.[13] In the first column, we find a small but statistically significant effect: if both i and j attend a small RI, this increases their likelihood of sharing information by about 3 percentage points. In Column 2 however, we find that the interaction term of small RI with attendance at the same RI is slightly negative and statistically insignificant, indicating that peers are not necessarily more likely to seek out and receive agricultural information if joint members of a small RI compared to a large RI.

Whether the particular type of RI has an additional impact on information sharing beyond size is an interesting question for future research.

To this point, we have demonstrated that those who attend the same religious institution are more likely to seek out and share information from one another, though this is not necessarily more prevalent in small RIs compared to larger RIs. However, we have yet to identify the primary mechanisms of this effect. Why does this take place in a social space formed by RIs in particular? In the remainder of this section, we show that attending the same RI increases both the trust between individuals and the perception of information quality from his/her peer, which in turn is likely the root cause behind the greater likelihood of information sharing. Joint attendance at RIs also increases the frequency of weekly interactions, compared to those who speak more or less frequently than weekly. In addition, when we predict the strength of friendship based on observable characteristics and then divide the sample between those more likely and less like to be friends with one another, the effect of joint RI attendance on information sharing is most strong among those who are *less* likely to be friends. This tells us that the social space formed by RIs is effective at diffusing information between those who would otherwise not necessarily interact with one another.

To explore these mechanisms, we first use the same estimation strategy as our primary estimations found in Table 6, but change the dependent variable. In our survey, we included several additional questions asked about the respondents' peers, including whether i trusts j to watch a valuable for him/her, the level of friendship with j, and the level of information quality of j. Statistics for these variables are included in Table 5. We first test whether attending the same RI increases the likelihood that i trusts j. Using the same identification strategy as before, Column 1 of Table 9 shows that joint attendance of i with j at an RI increases the likelihood of i trusting j by 24 percentage points, significant at the p=0.05 level. The estimation thus suggests that social spaces formed through RIs bridge trust deficits that are found in the broader village network and serve as a likely mechanism for our finding of greater information sharing among joint RI attendees.

We next test whether joint RI attendance of i with j increases i's perception of the quality of information coming from j, or whether it leads i to view j as a closer friend. These two variables are both categorical, ranging from 1 to 5: 1 is poor information quality or no friendship and 5 is great information or close friend for the respective variables. For this analysis, we transform the variables to binary given their skewed distributions, especially present in the friendship variable.[14] Again using a linear probability model with the same identification strategy, we find a statistically significant effect of joint attendance on the perception of information quality: joint attendance at an RI increases the perception of i that j's information is of good quality by 40 percentage points. However, we do not find an effect of joint attendance at an RI on the level of friendship of i with j (results for information quality and friendship shown in Columns 2 and 3 respectively of Table 9). This further supports our hypothesis that trust is a primary driver of information sharing: joint RI attendance does not necessarily lead to closer friendships, but does appear to increase levels of trust and perceptions of information quality between members, increasing information diffusion among RI network members.

It is also possible that joint religious membership has a non-linear relationship with frequency of interaction. Since most individuals attend their RI weekly, it is likely that attendance by i and j at the same RI is more closely correlated with weekly communication compared to less than or more than weekly interactions between i and j. To determine this statistical relationship, we use a multinomial logit estimation with frequency of interaction as the dependent variable divided into four categories: never, more than never but less than weekly, weekly, or more than weekly interactions. In these estimations, we are only interested in correlation and do not seek to prove causality, and therefore to not use instruments or fixed effects. The results are in Table 10, with more than weekly communication as the excluded category. We find results that differ significantly between the communication frequency categories: compared to those who communicate daily or semi-daily, joint attendance at an RI between i and j has a statistically significant correlation with weekly communication. On the other hand, joint attendance at an RI has a statistically significant negative correlation with communication less than weekly or never, compared to daily or semi-daily communication. Without implying causality, these results suggest that weekly communication and attendance at the same RI go hand-in-hand, demonstrating the impact of this social space for shared interactions between members.

A powerful aspect of social organizations in general and RIs in particular is that they bring together disparate individuals who may not communicate or have friendships with one another in their absence, creating weak (or bridging) ties between social cliques. We believe therefore that the impact of joint attendance at an RI on information sharing will be greater among those who are less likely to be close friends with one another than those more likely. To investigate this, we use both stated friendship ties between i and j and predicted levels of friendship based on observable dyadic variables.[15] We split the sample first by stated friendship levels (whether i states j is a close friend), and second, using the median value of the predicted friendship variable as a cut-off. Then, using the same identification and estimation strategy as our primary results, we estimate the effect of joint RI attendance on the likelihood of receiving advice from his/her peer for each segment of the sample. We present the results in Table 11. In the first two columns, we present results first for those who are not close friends and for those who are close friends. These results show that for those who state that j is not a close friend, attendance at the same RI increases their likelihood of receiving information from j by a statistically significant 28 percentage points. For those who state that j is a close friend, joint attendance increases likelihood of obtaining this information by 26 percentage points, though not statistically significant. However, given the differences in number of observations between these split samples (7,263 and 1,442), these results are not necessarily informative. More interesting though, among those predicted less likely to be friends (Column 3, Table 11), the impact of joint attendance at an RI on information sharing is 44 percentage points (significant at the p=0.1 level). Among those predicted more likely to be friends, the impact is a statistically insignificant

17 percentage points. These latter results present evidence that the social space created through RIs encourages information diffusion among those less likely to share friendships in the absence of joint attendance at an RI, indicating the importance of social spaces in these communities for the creation of bridging social ties that aid information diffusion within groups.

Conclusions and Implications

Because of the importance of social institutions in general and RIs in particular in Sub-Saharan Africa (SSA), RIs often serve as platforms for information exchange, creating bridging ties among peers that increase information diffusion. This study attempts to identify the causal relationship between shared attendance at a religious institution (RI) and seeking out and receiving practical advice among farmers in western Kenya. For our estimation, we use network data collected from individuals in 548 households (17 villages) in western Kenya in 2016. Using a novel spatial instrumental variable strategy to control for likely endogeneity, the results show that RI co-membership between individual i and j increases the probability of i seeking out and receiving agricultural advice from j by 33 percentage points – a substantial impact. This finding is also robust to two-way clustering of standard errors used to control for correlation among observations from the same individual.

Several studies have shown that members of the same faith have greater levels of trust and cooperation towards each other (Johansson-Stenman, Mahmud, and Martinsson, 2009; Chuah et al., 2014; Asadullah, 2017). Our findings suggest that co-membership in an RI inculcates trust among its members, strengthening ties between individuals. Attending the same RI leads to an increase in the likelihood that individual i trusts individual j to watch a valuable item for him/her by 24 percentage points, significant at the p=0.05 level after instrumentation. We believe that this increase in trust is likely the mechanism for the increased likelihood of sharing information when attending the same RI. In addition, we find that attendance at the same RI increases the perception that an individual's peer has high information quality by 40 percentage points, also statistically significant. On the other hand, we find no evidence that co-membership at the same RI increases reported levels of friendship between individuals or i's perception in the information quality of j.

These findings support the theory that social spaces formed through institutions significantly facilitate information diffusion in SSA. The policy conclusion is that social spaces formed through institutions can serve as platforms for information diffusion that contribute to agricultural and economic development. Bompani and Smith (2013), for example, argue that it is imperative to take social institutions such as RIs into account when seeking to diffuse technology in rural Kenya given the lack of local resources and the need for information to come from trusted institutions and individuals. As discussed earlier, an individual's peers have a safety credibility that encourages information sharing (Rogers, 1995), whereas information from an extension agent may be viewed with relative scepticism (Krishnan and Patnam, 2014). Moreover, peers within the social space of an RI possess links that are deepened through shared religious practices and activities – prayer services, dancing and singing, youth and women's groups – ties that often lead to a substitution towards church-related activities at the expense of other activities within the village (Iannaccone, 1992).

This study makes an additional contribution to the literature by developing a novel instrumental variable strategy that is effective in identifying causal relationships between peer dyads. As one of our instruments, we find the physical distance between peer homesteads, take the midpoint, then calculate the distance between this point and closest RI. Because this IV does not rely on the choice of RI for its relevance and the physical locations of the homesteads are plausibly exogenous, this instrument is valid. Because the validity of these instruments could be threatened by potential correlations between RI location and the location of other social or economic centers, we conduct robustness checks generating random RI locations at a particular radius from the true RI location. The results of this check illustrate that it is the precise location of the RI that is crucial for our identification strategy, allaying these possible concerns. Using these spatial instruments can potentially aid in identifying other dyadic relationships where GPS measurements are taken of peer dyads and other locations of interest.

From these results, it is apparent that technology diffusion in SSA occurs through the social spaces created by RIs. Oftentimes, significant amounts of time are spent by RI members on RI related activities, and as a result, RIs can potentially be effective tools for information diffusion through the intimate social spaces created through their activities. As we see in this context, innovative technologies or practices (in agriculture, for example) can potentially spread quickly through these networks given the strength of the connections that are created. Potentially utilizing these social spaces as injection points to diffuse information, policymakers can potentially spread productivity-enhancing information more effectively and aid in economic development.

On the other hand, this analysis begs the question as to whether RI-generated social spaces are beneficial to villages overall, or whether they aid information diffusion within the closed group at the expense of the wider community. In this analysis, due to data limitations we are unable to explore these potential equity issues, nor whether, as discussed by Fafchamps (2006) among others, increases in personalized trust between members of associations or clubs is positive if it comes at the expense of generalized community trust. Fafchamps (2006) is also sceptical of entrusting RIs with information or public goods provision, as they may have interests at odds with the majority of the local population. However, we see this study as important in demonstrating the existence of substantial information diffusion through individual RI networks with the expectation that future research will provide additional analysis on long-run effects of RI networks on information equity and village-level information diffusion.

Membership in social organizations, including religious institutions, continues to decline in the West. In SSA however, activities within social spaces created through these institutions have a central place in the everyday life of agrarian communities. This study demonstrates a strong link between the social space created through RIs and information sharing through peer networks. A large literature already exists analyzing optimal methods for technology diffusion in the developing world; further research on information diffusion within social spaces formed through RIs and other social institutions could add valuable insights to this research area.

Notes

[1]While we do not have data on the population of the villages, the average household size is 5.29 individuals, which would imply an average population of 513 per village.

[2]The variables included in the subset are exogenous, such as whether a family tie exists between i and j, while endogenously co-determined variables, such as trust and friendship levels between i and j, are excluded from the subset.

[3]For the remainder of the paper, we abbreviate $X_i + X_j$ and $X_i - X_j$ together as X_{ij}.

[4]Nevertheless, in our section examining the robustness of our approach, we examine the within-village spatial correlation of household characteristics along observable dimensions to see if residents are clustered in pockets within a village based on wealth, education, religious affiliation, and other characteristics. We show that such clustering is limited to a small subset of villages in our sample. Furthermore, when we exclude these villages and repeat our analysis, our results are consistent with those from the original estimation procedure.

[5]Five RIs were ranked as five is the maximum number of RIs in the village with the fewest total number of RIs in our sample.

[6]The soil sample was taken from the maize plot closest to the homestead. Soil samples were analyzed for nitrate, phosphate, potassium, sulphur, active carbon, soil acidity, and cation exchange capacity (CEC).

[7]Thresholds developed by IITA for this project.

[8]"Progressive" farmers were farmers that IITA had worked with in past projects. Most

villages had only one of these farmers.

[9]The exception is between married individuals or close family in the sample, as they live close together and are more likely to share information, but this is controlled for by the "immediate family member" variable.

[10]For example, consider two pairs of households (dyads). Each pair shares similar characteristics and live 100 meters apart; however, one pair is located in the village center and the other pair live in the outer edge of a village. Thus, the pair at the edge of the village is more likely to share a closest church even if RIs are uniformly randomly distributed throughout the village. This pair might also be more likely to interact with others given that the likelihood of frequenting the same social spaces is higher.

[11]We also include a specification in which we interact τ_{ij} with ζ_{ij}. This interaction term adds an endogenous variable to the model. Because of this, we also add the nearest RI distance interacted with τ_{ij} as an additional instrument in Column 4 of Table 7; however, this instrument is quite weak and results from such a specification should not be over-analyzed.

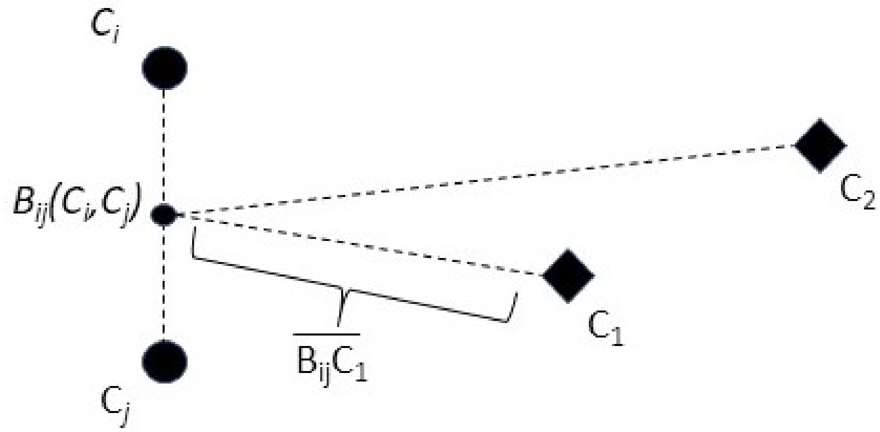
[12]We present the 95% confidence interval of the simulated F-statistic for full transparency.

[13]Because this interaction term adds an endogenous variable to the model, we also add IVs interacted with RI_{sml} to the first-stage estimation.

[14]The information quality variable ranged from 0 to 5. We coded 0-3 as 0 and 4 and 5 as 1. For the friendship variable, 5 (close friend) is coded as 1, while all other responses are 0.

[15]The predicted ordered logit results are in Appendix Table A5.

Figure 1: Illustration of Instrument 1: Triangular Geometry



Notes: C_i and C_j represent two, same-village farmers in the sample. Point $B_{ij}(C_iC_j)$ is the midpoint of the physical, straight-line distance from the homestead of i to the homestead of j. C_k denotes RIs, and C_1 and C_2 are the two nearest religious institutions (RIs). We calculated all distances from each B_{ij} to each C_k among all peer dyads. As an instrument, we use the distance of B_{ij} to the nearest RI, shown in Figure 1 as $\overline{B_{ij}C_1}$.

Table 1: Religion in Kenya

Religion	Population (thousands)	Population (percent)
Christian	**31,878**	**83**
Catholic	9,011	23
Anglican	5,000	13
Other Protestant (e.g., Pentecostal)†	13,307	35
Orthodox Christian	650	2
Other Christian	3,910	10
Muslim	**4,305**	**11**
Hindu	**53**	**<1**
Traditional	**635**	**2**
Other Religions	**557**	**1**
Baha'i	423	1
Other	134	<1
No Religion	**922**	**2**
Unknown	**61**	**<1**

† The majority of the "Other Protestant" group are Pentecostal and Charismatic Churches.
Percentages may not add up to 100 due to rounding. Population total: 38,412,088 (2009). Data from Kenya Population and Housing Census as reported by the UN (2009), from the Association of Religion Data Archives (2010), the Pew Research Center (2011), and from the World Council of Churches (2013).

Table 2: Religious Institution Summary Statistics from Sample

Neo-Pentecostal & Charismatic (n=356)	Mean	Std. Dev.	Minimum	Maximum
RI size	117.82	95.99	20	500
RI attendance (1=daily, 9=never)†	3.15	0.69	1	9
Anglican (n=247)				
RI size	162.56	112.62	20	500
RI attendance (1=daily, 9=never)†	3.28	0.79	1	8
Catholic (n=147)				
RI size	262.83	134.69	20	500
RI attendance (1=daily, 9=never)†	3.45	1.30	2	9
Other Christian (n=90)				
RI size	103.30	51.93	18	400
RI attendance (1=daily, 9=never)†	3.44	1.08	2	9
Other Religion (n=11)				
RI size	90.45	79.67	25	230
RI attendance (1=daily, 9=never)†	2.91	0.70	1	4

For this table, the sample was trimmed and observations with stated RI size in the top or bottom 5%, and those who did not state an answer for attendance frequency were dropped.
† 1=daily, 2=every few days, 3=weekly, 4=every few weeks, 5=monthly, 6=every few months, 7=twice a year, 8=yearly, 9=never.

Table 3: Religion in Sample

Religion	Individuals	Percent of Sample
Pentecostal/Charismatic	378	42.76
Anglican	255	28.85
Roman Catholic	146	16.52
Salvation Army	52	5.88
Quaker (Friends)	23	2.60
Seventh Day Adventist	14	1.58
Muslim	9	1.02
Baptist	2	0.23
Presbyterian	2	0.23
Traditional	2	0.23
Orthodox Christian	1	0.11
No Religion	4	0.45
	884	100

Table 4: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Individual (n=884)				
Age	48.29	16.09	19.00	109.00*
Years of education	7.95	3.80	0.00	26.00**
Yes=1:				
Basic math ability***	0.56	0.50	0.00	1.00
Female	0.58	0.49	0.00	1.00
Widow/er	0.14	0.35	0.00	1.00
Bukusu subtribe	0.37	0.48	0.00	1.00
Luhya tribe (except Bukusu)	0.31	0.46	0.00	1.00
Iteso tribe	0.29	0.45	0.00	1.00
Household (n=548)				
Household size§	5.29	3.27	0.00	40.00
Total land area (arable acres)	1.06	1.06	0.02	8.87
Yes=1:				
Soil nitrogen deficient	0.96	0.21	0.00	1.00
Soil phosphorus deficient	0.96	0.20	0.00	1.00
Soil carbon deficient	0.67	0.47	0.00	1.00
Organic inputs (within past two seasons)	0.45	0.50	0.00	1.00
Inorganic inputs (within past two seasons)	0.88	0.33	0.00	1.00
NGO contact	0.13	0.34	0.00	1.00
River as water source	0.43	0.50	0.00	1.00
Electricity (grid)	0.13	0.33	0.00	1.00
Solar panels	0.29	0.45	0.00	1.00
Metal roof	0.87	0.33	0.00	1.00
Mud walls	0.77	0.42	0.00	1.00
Village (n=17)				
Individuals (interviewed per village)	52.00	13.91	38.00	97.00
Households (sampled per village)	32.24	8.39	21.00	57.00

*There was one woman who claimed she was 109 years old. **In the sample there were three farmers who were also university professors with PhDs. ***Was able to do a basic multiplication problem. §Defined as the number of individuals who spent the night at that dwelling last night.

Table 5: Network Summary Statistics

Variable	Mean	Std. Dev.	Same Church	Different Church	P-value
Know individual j	0.69	0.46	0.78	0.66	0.00
Have met individual j in person	0.67	0.47	0.75	0.63	0.00
Conditional on having met j (n=5811 dyads)					
Friendship level (1=Not a friend, 5=Close friend)	4.02	0.81	4.15	3.96	0.00
Level of information quality from j (1=Poor, 5=Great)	3.61	1.00	3.71	3.57	0.00
Yes=1					
j is immediate family	0.05	0.23	0.11	0.03	0.00
j is extended family	0.37	0.48	0.41	0.34	0.00
Known j at least ten years	0.81	0.39	0.86	0.79	0.00
Speak at least weekly with j	0.81	0.40	0.85	0.79	0.00
Speak daily with j†	0.26	0.44	0.31	0.24	0.00
Attend same RI as j‡	0.34	0.47	0.76	0.16	0.00
Attend same non-religious organization as j	0.19	0.39	0.24	0.17	0.00
Fertilizer advice received from j	0.18	0.38	0.22	0.16	0.00
Planting advice received from j	0.16	0.37	0.21	0.14	0.00
Crop buyer advice received from j	0.09	0.29	0.14	0.07	0.00
At least one form of advice received from j	0.23	0.42	0.28	0.21	0.00
Worked together with j in past 12 months	0.33	0.47	0.37	0.32	0.02
Trust j to watch a valuable for you for one week	0.21	0.41	0.25	0.20	0.00

† "Speak daily with j" is a subset of "Speak at least weekly with j." ‡ The "Attend same RI as j" (the key independent variable in the analysis) asked the participant whether j attended the same RI as him/herself. The Same Church/Different Church column looks at the RI that i stated was his/her primary RI, and compares it with the RI that j stated was his/her primary RI. The table shows that in 76% of the links, i correctly identified j as attending his/her RI, but in 16% of the links, i stated that j attended the same RI, but the primary RIs of the two individuals did not match.

Table 6: Primary Results

Variable	Dep. Variable: Received Agricultural Advice from j (1)	(2)	(3)†	(4)†
	OLS	2SLS	OLS	2SLS
Religious institution co-membership (ζ_{ij})	0.108***	0.287**	0.107***	0.329**
	(0.017)	(0.135)	(0.016)	(0.138)
Known $j \geq 10$ years	0.129***	0.088***	0.125***	0.074**
	(0.014)	(0.033)	(0.013)	(0.034)
Immediate family of j	0.236***	0.202***	0.237***	0.194***
	(0.031)	(0.040)	(0.031)	(0.041)
Extended family of j	0.080***	0.065***	0.075***	0.058***
	(0.015)	(0.018)	(0.014)	(0.018)
Distance to j (kilometres)	-0.220***	-0.190***	-0.241***	-0.202***
	(0.037)	(0.044)	(0.037)	(0.045)
Squared distance to j (kilometres)	0.111***	0.096***	0.121***	0.102***
	(0.023)	(0.028)	(0.023)	(0.028)
Male_i-Male_j			-0.007	-0.000
			(0.015)	(0.016)
Male_i-Female_j			-0.006	-0.006
			(0.016)	(0.017)
Female_i-Male_j			0.002	0.002
			(0.011)	(0.012)
Fixed effects (enumerator, village, survey month)	✓	✓	✓	✓
X_{ij}, A_{ij}	-	-	✓	✓
Observations	8,704	8,704	8,704	8,704
Number of clusters	883/883	883/883	883/883	883/883
Kleibergen-Paap rk LM P-value	-	0.000	-	0.000
First-stage F-stat	-	46.33	-	19.03
Kleibergen-Paap rk Wald F-stat	-	10.33	-	8.04
Hansen J Statistic P-value (H_0: IVs are valid)	-	0.86	-	0.61

†Include sums and differences of household and soil characteristics. Coefficient estimates for X_{ij} and A_{ij} are shown in the Appendix Tables A3 and A4. 2SLS instruments are explained in text. Two-way clustered standard errors (at individual level) in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Table 7: Alternative Specifications

Variable	Dep. Variable: Rec. Advice from j (1)	(2)	(3)	(4)
	2SLS			
RI co-membership (ζ_{ij})	0.336**	0.312	0.296**	0.694
	(0.137)	(0.224)	(0.141)	(0.566)
RI co-membership × small RI ($\zeta_{ij} \times RI_{sml}$)		-0.126		
		(0.221)		
Small RI (RI_{sml})	0.028**	0.052		
	(0.013)	(0.036)		
RI co-membership × Avg dist. neighbours ($\zeta_{ij} \times \tau_{ij}$)				-0.303
				(0.319)
Avg. distance neighbours (τ_{ij})			0.040	0.135
			(0.036)	(0.102)
Known $j \geq 10$ years	0.072**	0.085*	0.082**	0.076
	(0.033)	(0.045)	(0.034)	(0.053)
Immediate family of j	0.190***	0.210***	0.198***	0.172**
	(0.040)	(0.042)	(0.041)	(0.067)
Extended family of j	0.057***	0.064***	0.060***	0.050*
	(0.018)	(0.019)	(0.018)	(0.029)
Distance to j (kilometres)	-0.204***	-0.219***	-0.210***	-0.195***
	(0.045)	(0.050)	(0.046)	(0.062)
Squared distance to j (kilometres)	0.104***	0.110***	0.099***	0.083**
	(0.028)	(0.030)	(0.027)	(0.039)
Male_i - Male_j	-0.001	-0.001	-0.001	0.001
	(0.016)	(0.017)	(0.016)	(0.018)
Male_i - Female_j	-0.007	-0.006	-0.006	-0.008
	(0.017)	(0.016)	(0.017)	(0.017)
Female_i - Male_j	0.002	0.002	0.002	0.001
	(0.012)	(0.012)	(0.012)	(0.013)
X_{ij}, A_{ij}	√	√	√	√
Observations	8,704	8,704	8,704	8,704
Number of clusters	883/883	883/883	883/883	883/883
Kleibergen-Paap rk LM P-value		0.000		0.012
Kleibergen-Paap rk Wald F-stat		8.19		2.71
Hansen J Statistic P-value (H_0: IVs are valid)		0.34		0.24

Two-way clustered standard errors (at individual level) in parenthesis. 2SLS instruments are explained in text. ***p<0.01, **p<0.05, *p<0.1.

Table 8: IV Robustness Check

	1/9 SD Village-level Random RI				
			Percent of iterations		
	Mean	StDev	>0.10	>0.05	>0.017
P-value of RI attendance (Original=0.017)	0.146	0.173	42.7	63.8	89.3
	1/6 SD Village-level Random RI				
			Percent of iterations		
	Mean	StDev	>0.10	>0.05	>0.017
	0.338	0.278	75.3	88.6	97.1
	1/9 SD Village-level Random RI				
	Mean	StDev	95% CI		
F-stat of IVs (Original=8.03)	5.45	2.02	5.32	5.57	
	1/6 SD Village-level Random RI				
	Mean	StDev	95% CI		
	4.40	2.00	4.27	4.52	

For each RI, we randomly generated 1000 RIs surrounding the original RI in a circle with a radius equal to a fraction of the standard deviation of the village-level RI distances to represent a reasonable "neighborhood" around the RI. With these randomly generated RIs, we ran 1000 estimations of both the first and second stages of the 2SLS regressions. This table presents the statistics for the p-value of the joint attendance variable and the f-statistic of the instruments in the first round. If the precise location of the RI matters more than its neighborhood, then the influence of simulated RIs should weaken. As expected, when the neighborhood is defined by 1/9 (1/6) Standard Deviation of the distance, the mean value of the simulated second stage P-value increases to 0.146 (0.338). The influence of the precise RI location has a P-value of 0.017. Similarly, the constructed instruments weakens, reflected in an F-statistic of 5.45 (4.40), as the neighborhood expands. Therefore, we can assert that the precise location of the RI is crucial in the construction of the instrument and that the influence of neighborhood-specific confounds are likely to be negligible.

Table 9: Potential Mechanisms

Variable	Trust (1)	Info Quality (2)	Friendship (3)
		2SLS	
RI co-membership (ζ_{ij})	0.241**	0.399**	0.120
	(0.122)	(0.156)	(0.146)
Known $j \geq 10$ years	0.090***	0.207***	0.106***
	(0.031)	(0.040)	(0.035)
Immediate family of j	0.259***	0.134***	0.242***
	(0.039)	(0.043)	(0.046)
Extended family of j	0.092***	0.140***	0.188***
	(0.017)	(0.021)	(0.019)
Distance to j (kilometres)	-0.170***	-0.131**	-0.217***
	(0.040)	(0.055)	(0.042)
Squared distance to j (kilometres)	0.098***	0.057*	0.098***
	(0.023)	(0.032)	(0.024)
Male_i - Male_j	0.030**	0.034*	0.034**
	(0.015)	(0.019)	(0.015)
Male_i - Female_j	0.009	-0.000	-0.006
	(0.015)	(0.021)	(0.015)
Female_i - Male_j	0.007	0.006	0.026**
	(0.012)	(0.016)	(0.012)
Fixed effects (enumerator, village, survey month)	√	√	√
X_{ij}, A_{ij}	√	√	√
Observations	8,704	8,704	8,704
Number of clusters	883/883	883/883	883/883
Kleibergen-Paap rk LM P-value	0.000	0.000	0.000
First-stage F-stat	19.03	19.03	19.03
Kleibergen-Paap rk Wald F-stat	8.04	8.04	8.04
Hansen J Statistic P-value (H_0: IVs are valid)	0.29	0.40	0.38

Two-way clustered standard errors (at individual level) in parenthesis. 2SLS instruments are explained in text. ***p<0.01, **p<0.05, *p<0.1.

Table 10: Communication Frequency

Variable	Never (1)	Less than Weekly (2)	Weekly (3)
		Multinomial Logit	
Religious institution co-membership (ζ_{ij})	-20.828***	-0.503***	0.267***
	(0.176)	(0.114)	(0.098)
Known $j \geq 10$ years	-22.784***	-0.700***	-0.143
	(0.176)	(0.120)	(0.119)
Immediate family of j	0.196	-0.679***	-0.893***
	(0.507)	(0.260)	(0.248)
Extended family of j	-21.040***	-0.392***	-0.638***
	(0.275)	(0.114)	(0.096)
Distance to j (kilometres)	4.222***	3.873***	2.855***
	(0.548)	(0.394)	(0.321)
Squared distance to j (kilometres)	-2.043***	-1.756***	-1.387***
	(0.338)	(0.256)	(0.209)
Male_i - Male_j	-0.126	-0.120	-0.247*
	(0.235)	(0.166)	(0.146)
Male_i - Female_j	0.037	0.083	-0.107
	(0.215)	(0.176)	(0.159)
Female_i - Male_j	-0.166	-0.278**	-0.235**
	(0.132)	(0.110)	(0.105)
Constant	4.177***	-0.502	-0.035
	(0.909)	(0.693)	(0.642)
X_{ij}, A_{ij}	✓	✓	✓
Observations	8,704	8,704	8,704
Number of clusters	883	883	883

Excluded category is individuals who communicated more frequently than weekly. Standard errors clustered at individual level in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Table 11: Predicted Friendship - Split Sample

Variable	(1)	(2)	(3)	(4)
	Dep. Variable: Received Agricultural Advice from j			
	2SLS			
	Friend_{ij} <"Close"	Friend_{ij} ="Close"	Pr. Friend_{ij} <Median	Pr. Friend_{ij} ≥Median
Religious institution co-membership (ζ_{ij})	0.282**	0.258	0.444*	0.169
	(0.124)	(0.394)	(0.240)	(0.224)
Male_i - Male_j	0.020	-0.082	-0.011	0.008
	(0.015)	(0.055)	(0.027)	(0.015)
Male_i - Female_j	-0.008	0.038	-0.015	-0.002
	(0.015)	(0.050)	(0.031)	(0.012)
Female_i - Male_j	0.006	-0.032	-0.006	0.009
	(0.010)	(0.040)	(0.024)	(0.010)
Known $j \geq 10$ years	0.066**	0.006	-0.022	0.042
	(0.030)	(0.051)	(0.094)	(0.078)
Immediate family of j	0.047	0.167	0.152***	0.243*
	(0.050)	(0.113)	(0.056)	(0.129)
Extended family of j	0.048***	0.023	0.021	0.149**
	(0.018)	(0.035)	(0.020)	(0.067)
Distance to j (kilometres)	-0.125***	-0.189	-0.225**	-0.114***
	(0.038)	(0.127)	(0.101)	(0.044)
Squared distance to j (kilometres)	0.061**	0.071	0.095	0.062**
	(0.024)	(0.092)	(0.064)	(0.026)
Fixed effects (enumerator, vill., surv. month)	√	√	√	√
X_{ij}, A_{ij}	√	√	√	√
Observations	7,263	1,442	4352	4353
Number of clusters	882/883	580/616	785/831	831/853
Kleibergen-Paap rk LM P-value	0.000	0.069	0.011	0.002
Kleibergen-Paap rk Wald F-stat	9.30	10.93	26.25	5.31
Hansen J Stat. P-value (H_0: IVs are valid)	0.241	0.377	0.200	0.156

Two-way clustered standard errors (at individual level) in parenthesis. 2SLS instruments are explained in text. ***p<0.01, **p<0.05, *p<0.1.

References

Asadullah, M.N. 2017. "Who Trusts Others? Community and Individual Determinants of Social Capital in a Low-Income Country." *Cambridge Journal of Economics* 41:515–544.

Association of Religion Data Archives. 2010. "Most Baha'i Nations." Last modified April 18, 2017. http://www.thearda.com/ql2010/QuickList˙40b.asp.

Banerjee, A.V., A.G. Chandrasekhar, E. Duflo, and M.O. Jackson. 2013. "The Diffusion of Microfinance." *Science* 341:363–370.

Beaman, L., A. Benyishay, J. Magruder, and A.M. Mobarak. 2015. "Can Network Theory-based Targeting Increase Technology Adoption?" Unpublished.

Bompani, B., and J. Smith. 2013. "Bananas and the Bible: Biotechnology, the Catholic Church, and Rural Development in Kenya." *International Journal of Religion and Society* 4:910–929.

Cameron, A.C., J.B. Gelbach, and D.L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business and Economic Statistics* 29:238–249.

Chuah, S.H., R. Hoffmann, B. Ramasamy, and J.H.W. Tan. 2014. "Religion, ethnicity and cooperation: An experimental study." *Journal of Economic Psychology* 45:33–43.

Conley, T.G., and C.R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *The American Economic Review* 100:35–69.

Deke, J. 2014. "Using the Linear Probability Model to Estimate Impacts on Binary Outcomes in Randomized Controlled Trials." Working paper No. December, Mathematica Policy Research.

Fafchamps, M. 2006. "Development and social capital." *Journal of Development Studies* 42:1180–1198.

Fafchamps, M., and F. Gubert. 2007. "The formation of risk sharing networks." *Journal of Development Economics* 83:326–350.

Feder, G., and R. Slade. 1984. "The Acquisition of Information and the Adoption of New Technology." *American Journal of Agricultural Economics* 66:312–320.

Foster, A.D., and M.R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103:1176–1209.

Granovetter, M.S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–1380.

Iannaccone, L.R. 1998. "Introduction to the Economics of Religion." *Journal of Economic Literature* 36:1465–1495.

—. 1992. "Sacrifice and Stigma: Reducing Free-riding in Cults, Communes, and Other Collectives." *Journal of Political Economy* 100:271–291.

Iyer, S. 2016. "The New Economics of Religion." *Journal of Economic Literature* 54:395–441.

Johansson-Stenman, O., M. Mahmud, and P. Martinsson. 2009. "Trust and Religion: Experimental Evidence from Rural Bangladesh." *Economica* 76:462–485.

Jones, B. 2005. "The Church in the Village, the Village in the Church: Pentecostalism in Teso, Uganda." *Cahiers d'Etudes Africaines* 45:497–517.

Kenyatta, J. 1938. *Facing Mount Kenya: The Traditional Life of the Gikuyu*. London, UK: Heinemann Educational Books Ltd.

Krishnan, P., and M. Patnam. 2014. "Neighbors and Extension Agents in Ethiopia: Who Matters More for Technology Adoption?" *American Journal of Agricultural Economics* 96:308–327.

Meyer, B. 2004. "Christianity in Africa: From African Independent to Pentecostal-Charismatic Churches." *Annual Review of Anthropology* 33:447–474.

Moran, P.A.P. 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika* 37:17–23.

Munshi, K. 2004. "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution." *Journal of Development Economics* 73:185–213.

Pew Research Center. 2011. "Christian Population in Numbers by Country." Last modified April 18, 2017. http://www.pewforum.org/2011/12/19/table-christian-population-in-numbers-by-country.

Putnam, R.D. 2000. *Bowling Alone*. New York: Simon and Schuster.

Rogers, E.M. 1995. *Diffusion of Innovations*, 4th ed. New York: Free Press.

Sahn, D.E., and D. Stifel. 2003. "Exploring alternative measures of welfare in the absence of expenditure data." *Review of Income and Wealth* 49:463–489.

Santos, P., and C.B. Barrett. 2008. "What Do We Learn About Social Networks When We Only Sample Individuals? Not Much." Unpublished.

UN. 2009. "UNData." Last modified April 18, 2017. http://data.un.org/Data.aspx?d=POP&f=tableCode%3A28.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

World Council of Churches. 2013. "Anglican Church of Kenya." Last modified April 18, 2017. https://www.oikoumene.org/en/member-churches/anglican-church-of-kenya.

Figure A1: Project area (Western Kenya)

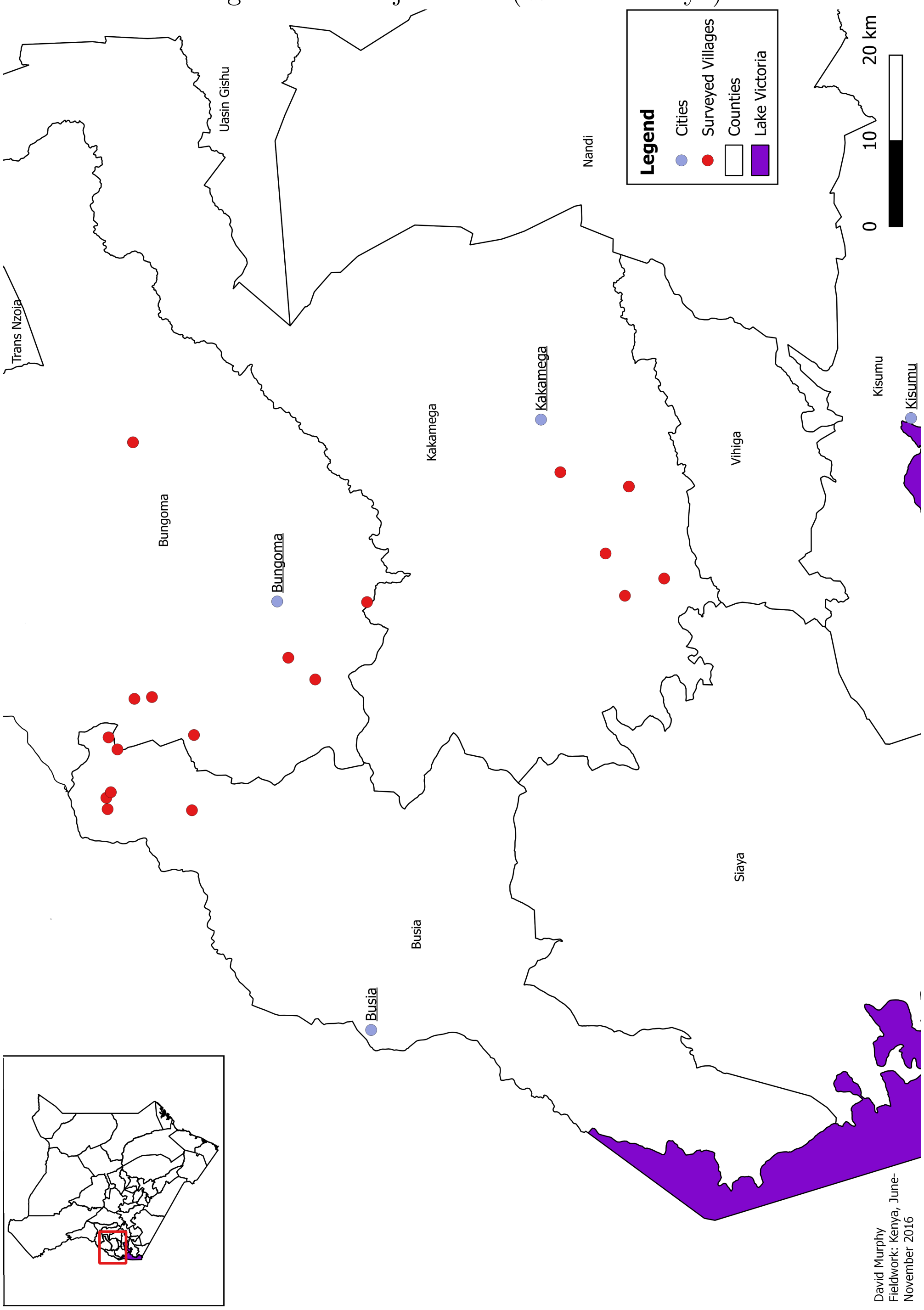


Figure A2: Sample Village and RIs

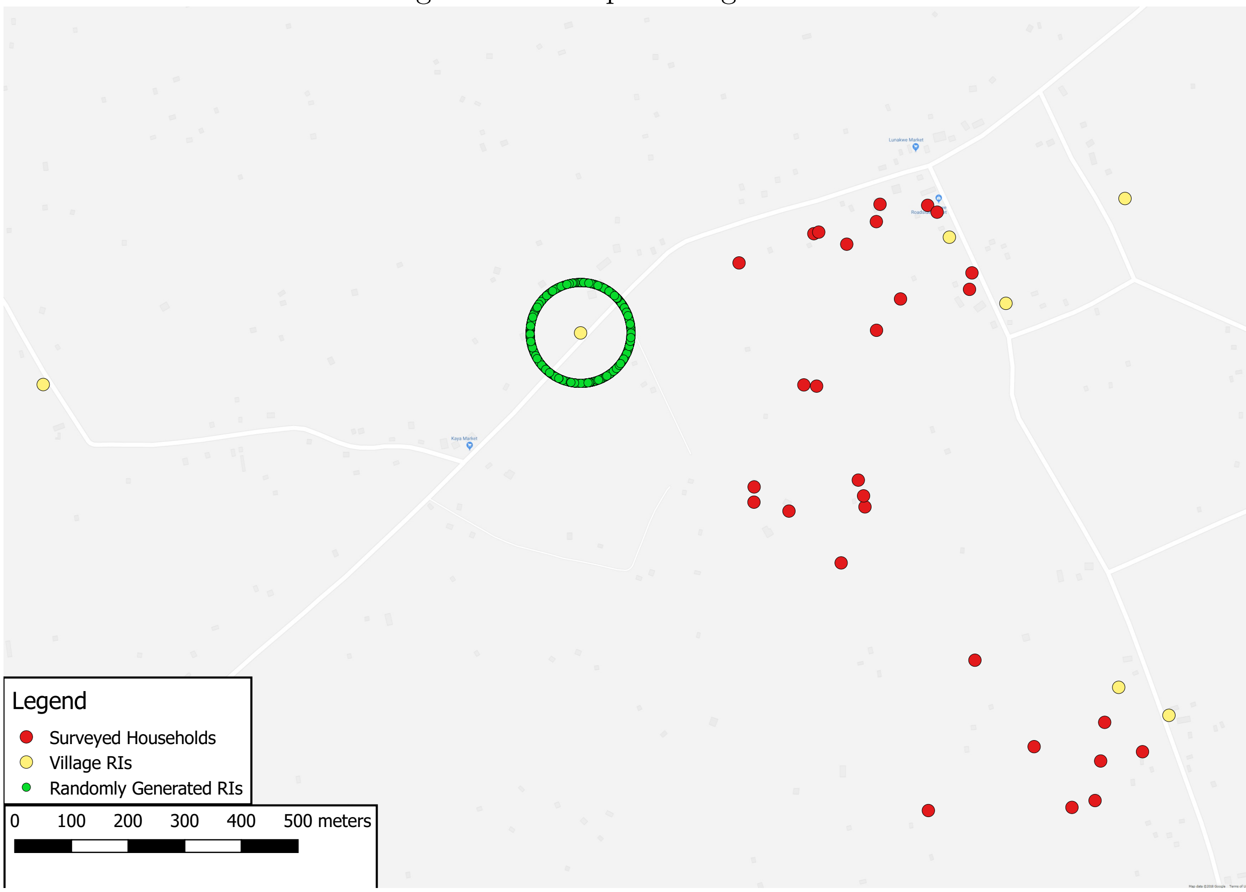


Table A1: Within Village RI Statistics

Village	N (RIs)	Dist. from Village Households (meters) (Mean)	(Median)	Year Est. at Location (Mean)	(Median)
1	5	2151	1656	1994	1998
2	6	1462	692	2004	2008
3	11	847	716	1998	2002
4	14	1824	1072	1999	2005
5	17	1307	1291	1983	1982
6	9	1325	1260	1997	2008
7	9	2855	1498	1987	1987
8	10	2384	1952	2001	2002
9	17	1884	1721	1996	1998
10	15	2211	1690	2001	2000
11	13	1492	1146	1999	2002
12	13	1626	1460	1978	1978
13	10	2273	1469	1981	1981
14	13	1826	721	1995	1998
15	11	1130	980	1998	2005
16	13	2382	1141	1985	1993
17	15	1125	683	1990	1996

Table A2: IV First Stage

Variable	Dependent Variable: Same Religious Institution (ζ_{ij}) (1)	(2)
Minimum ij RI distance (kilometers)	0.191***	0.160***
	(0.048)	(0.044)
Joint ij probability RI attendance	0.172	0.155
	(0.363)	(0.354)
Joint ij probability RI attendance (squared)	-0.151	-0.144
	(0.111)	(0.113)
Nearest ij RI matches	0.152***	0.141***
	(0.041)	(0.039)
Known $j \geq 10$ years	0.226***	0.227***
	(0.017)	(0.016)
Immediate family of j	0.191***	0.195***
	(0.038)	(0.037)
Extended family of j	0.084***	0.079***
	(0.020)	(0.019)
Distance to j (kilometres)	-0.168***	-0.179***
	(0.060)	(0.059)
Squared distance to j (kilometres)	0.104***	0.105***
	(0.036)	(0.034)
Male_i - Male_j		-0.031
		(0.020)
Male_i - Female_j		0.001
		(0.021)
Female_i - Male_j		-0.004
		(0.013)
Fixed effects (enumerator, village, survey month)	√	√
X_{ij}, A_{ij}	-	√
Observations	8,704	8,704
Number of clusters	883/883	883/883
R-squared	0.218	0.233
F-stat	46.33	19.03
Prob>0	0.00	0.00

Two-way clustered standard errors (at indvidual level) in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Table A3: Coefficient estimates for X_{ij} and A_{ij} (Sums)

Variable	Dependent Variable: Received Agricultural Advice (3)	(4)
	OLS	2SLS
Sums ($X_i + X_j$ and $A_i + A_j$)		
Polygamous household	0.007	0.008
	(0.009)	(0.010)
Repeated survey†	-0.020	-0.021
	(0.015)	(0.016)
Household head	0.003	0.005
	(0.008)	(0.009)
Age	0.002*	0.002*
	(0.001)	(0.001)
Age squared	-0.000**	-0.000**
	(0.000)	(0.000)
Years of education	0.003	0.003
	(0.003)	(0.003)
Years of education squared	-0.000	-0.000
	(0.000)	(0.000)
Household size	-0.001	-0.001
	(0.001)	(0.001)
Math ability‡	0.016*	0.011
	(0.008)	(0.011)
Asset index††	-0.005	-0.009
	(0.004)	(0.006)
TLU‡‡	-0.000	-0.002
	(0.002)	(0.002)
Widow	0.006	-0.001
	(0.010)	(0.013)
Total farm area (acres)	0.004	0.011
	(0.010)	(0.011)
Total farm area squared (acres)	0.002	0.001
	(0.002)	(0.002)
Bukusu subtribe	-0.019	-0.011
	(0.021)	(0.022)
Other Luhya tribe	-0.025	-0.015
	(0.021)	(0.023)
Other tribe	-0.037	-0.029
	(0.024)	(0.025)
Fixed effects (enumerator, village, survey month)	√	√
Observations	8,704	8,704
Number of clusters	883/883	883/883

† Controls for the few surveys that needed to be redone. ‡ Participant can do simple multiplication problem. †† Asset index compiled through factor analysis after Sahn and Stifel (2003). ‡‡ Tropical Livestock Units. Two-way clustered standard errors (at individual level) in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Table A4: Coefficient estimates for X_{ij} and A_{ij} (Differences)

Variable	Dependent Variable: Received Agricultural Advice (3)	(4)
	OLS	2SLS
Differences ($X_i - X_j$ and $A_i - A_j$)		
Polygamous household†	0.013	0.013
	(0.010)	(0.011)
Repeated survey	0.012	0.013
	(0.016)	(0.017)
Household head	0.003	0.002
	(0.007)	(0.008)
Age	-0.000	0.000
	(0.001)	(0.001)
Age squared	-0.000	-0.000
	(0.000)	(0.000)
Years of education	-0.004*	-0.003
	(0.003)	(0.003)
Years of education squared	0.000*	0.000
	(0.000)	(0.000)
Household size	0.001	0.001
	(0.001)	(0.001)
Math ability‡	0.007	0.006
	(0.010)	(0.011)
Asset index††	-0.012***	-0.012**
	(0.005)	(0.005)
TLU‡‡	0.003	0.003
	(0.002)	(0.002)
Widow	0.002	0.007
	(0.011)	(0.013)
Total farm area (acres)	-0.015	-0.015
	(0.009)	(0.010)
Total farm area squared (acres)	0.004*	0.004*
	(0.002)	(0.002)
Bukusu subtribe	0.012	0.017
	(0.020)	(0.023)
Other Luhya tribe	0.012	0.018
	(0.019)	(0.022)
Other tribe	0.016	0.029
	(0.021)	(0.024)
Fixed effects (enumerator, village, survey month)	√	√
Observations	8,704	8,704
Number of clusters	883/883	883/883

† Controls for the few surveys that needed to be redone. ‡ Participant can do simple multiplication problem. †† Asset index compiled through factor analysis after Sahn and Stifel (2003). ‡‡ Tropical Livestock Units. Two-way clustered standard errors (at individual level) in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Table A5: Observable Determinents of Friendship

Variable	Dep. Variable: Friendship level (i to j) (1)		
	Ordered Logit		
Known $j \geq 10$ years	3.463***	Constant cut1	0.395
	(0.109)		(0.427)
Immediate family of j	0.866***	Constant cut2	0.565
	(0.187)		(0.428)
Extended family of j	1.490***	Constant cut3	0.698
	(0.091)		(0.429)
Distance to j (kilometres)	-1.941***	Constant cut4	1.467***
	(0.241)		(0.432)
Squared distance to j (kilometres)	0.857***	Constant cut5	4.581***
	(0.142)		(0.440)
$Male_i$ - $Male_j$	0.066		
	(0.098)		
$Male_i$ - $Female_j$	-0.098		
	(0.099)		
$Female_i$ - $Male_j$	0.118*		
	(0.070)		
Sums ($X_i + X_j$ and $A_i + A_j$)		**Differences ($X_i - X_j$ and $A_i - A_j$)**	
Polygamous household	-0.060	Polygamous household	0.057
	(0.067)		(0.068)
Repeated survey	-0.199*	Repeated survey	-0.070
	(0.107)		(0.103)
Household head	0.029	Household head	-0.007
	(0.048)		(0.048)
Age	0.003	Age	-0.019***
	(0.007)		(0.007)
Age squared	-0.000	Age squared	0.000*
	(0.000)		(0.000)
Years of education	0.023	Years of education	0.012
	(0.016)		(0.016)
Years of education squared	-0.001*	Years of education squared	-0.001
	(0.001)		(0.001)
Household size	-0.001	Household size	-0.004
	(0.006)		(0.006)
Math ability	0.158***	Math ability	0.013
	(0.048)		(0.052)
Asset index	0.054**	Asset index	-0.025
	(0.026)		(0.027)
TLU	-0.000	TLU	-0.001
	(0.012)		(0.012)
Widow	0.055	Widow	-0.051
	(0.078)		(0.073)
Total farm area (acres)	0.098**	Total farm area (acres)	-0.061
	(0.047)		(0.045)
Total farm area squared (acres)	-0.016**	Total farm area squared (acres)	0.001
	(0.008)		(0.008)
Bukusu subtribe	-0.284**	Bukusu subtribe	0.169
	(0.117)		(0.106)
Other Luhya tribe	-0.273**	Other Luhya tribe	0.414***
	(0.138)		(0.129)
Other tribe	-0.171	Other tribe	0.247
	(0.164)		(0.155)
Fixed effects (enumerator, village, survey month)	$\checkmark$		
X_{ij}, A_{ij}	$\checkmark$		
Observations	8,705		

Standard errors clustered at the individual level in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Table A6: Spatial Autocorrelation P-values by Village

Village	Asset Index	Education	Anglican	Catholic	Pentecostal	Christian (Other)	Religion (Other)
1	0.000***	0.115	0.390	0.259	0.432		
2	0.304	0.365	0.391	0.254	0.200	0.311	
3	0.107	0.177	0.095*	0.228	0.289		
4	0.366	0.378	0.092*	0.171	0.137		0.130
5	0.496	0.480	0.000***	0.372	0.014**		
6	0.098*	0.114	0.268	0.425	0.198		
7	0.154	0.197		0.107	0.255	0.015**	
8	0.141	0.097*	0.053*	0.328	0.026**	0.177	0.013**
9	0.104	0.200		0.450	0.272	0.036**	0.137
10	0.146	0.496	0.407	0.121	0.240	0.198	
11	0.041**	0.382	0.491	0.058*	0.411	0.347	
12	0.223	0.202	0.283		0.428	0.146	0.056*
13	0.183	0.441	0.172		0.276	0.232	0.397
14	0.191	0.369		0.095*	0.262	0.215	0.085*
15	0.105	0.057*	0.185	0.458	0.420	0.233	
16	0.335	0.223		0.226	0.207		0.496
17	0.425	0.425			0.000***	0.000***	

Empty cells indicate that no members of that religion were sampled in that particular village. ***p<0.01, **p<0.05, *p<0.1.

Table A7: Autocorrelation Robustness Check

Variables	(1) Autocorrelated Villages Omitted: Assets (2 villages omitted)	(2) Autocorrelated Villages Omitted: Religions (3 villages omitted)
Religious institution co-membership (ζ_{ij})	0.328**	0.297*
	(0.146)	(0.154)
Known $j \geq 10$ years	0.081**	0.073*
	(0.034)	(0.042)
Immediate family of j	0.203***	0.185***
	(0.048)	(0.041)
Extended family of j	0.050**	0.073***
	(0.020)	(0.018)
Distance to j (kilometres)	-0.216***	-0.195***
	(0.045)	(0.048)
Squared distance to j (kilometres)	0.108***	0.097***
	(0.027)	(0.029)
$Male_i$ - $Male_j$	-0.006	0.004
	(0.017)	(0.018)
$Male_i$ - $Female_j$	-0.008	-0.000
	(0.018)	(0.018)
$Female_i$ - $Male_j$	0.002	0.007
	(0.013)	(0.014)
X_{ij}, A_{ij}	√	√
Observations	7,873	7,167
Number of clusters	792/792	727/727
R-squared	0.201	0.214

Two-way clustered standard errors (at individual level) in parenthesis. 2SLS instruments are explained in text. ***p<0.01, **p<0.05, *p<0.1.

Table A8: Family Relationships

Variable	Family (1)	Extended Family (2)
	2SLS	
RI co-membership (ζ_{ij})	-0.020	-0.111
	(0.066)	(0.184)
Known $j \geq 10$ years	0.046***	0.311***
	(0.018)	(0.052)
Distance to j (kilometres)	-0.415***	-0.529***
	(0.033)	(0.077)
Squared distance to j (kilometres)	0.210***	0.235***
	(0.018)	(0.042)
Male_i - Male_j	-0.001	-0.017
	(0.008)	(0.019)
Male_i - Female_j	0.000	-0.014
	(0.007)	(0.020)
Female_i - Male_j	-0.021***	-0.029**
	(0.005)	(0.014)
Fixed effects (enumerator, village, survey month)	√	√
X_{ij}, A_{ij}	√	√
Observations	8,704	8,704
Number of clusters	883/883	883/883

Two-way clustered standard errors (at individual level) in parenthesis. 2SLS instruments are explained in text. ***p<0.01, **p<0.05, *p<0.1.