Social Networks, Risk perception and changing attitudes towards HIV-AIDS: New Evidence from a Longitudinal Study Using Fixed-Effect Estimation

Stéphane Helleringer

Hans-Peter Kohler

In this paper we look at the effects of social networks on the adoption of strategies to prevent HIV infection in Sub-Saharan Africa. We try to assess the mechanisms through which social interactions shape these behaviors (social influence vs. social learning). Using longitudinal data collected in rural Malawi (MDICP), we implement fixed-effects models to allow for the possibility that important unobserved characteristics of the individuals, or of the communities they live in, determine not only the outcomes of interest but also the content and density of social interactions. We distinguish between social influence and social learning by analyzing data on the structure of the networks in which conversations about HIV prevention occur. Our analyses emphasize the need to control for the non-random selection of social networks. Standard estimates without such controls are likely to misrepresent the effects of social interactions on the adoption of different strategies to prevent HIV infection, and more broadly on all kinds of demographic behaviors.

The AIDS epidemic has reached crisis levels in Sub-Saharan Africa. Africans represent 70% of the world population living with HIV, and AIDS has become one of the major (if not the preeminent) causes of deaths in regions of Africa. The timing, pace and extent of behavioral responses to the epidemic is related to several factors: the knowledge of individuals about the disease, the availability of effective strategies to reduce the probability of infection, and the individuals' subjective perception of their own exposure to HIV. Recent surveys conducted in Sub-Saharan Africa (Demographic and health surveys) show that in these areas where HIV prevalence is high, almost everyone has the abstract knowledge to prevent infection. This paper thus focuses on the determinants of individual risk assessments, and their consequences for behavioral change.

Despite the emphasis on risk perceptions in many theories of AIDS-related behavior changes (UNAIDS, 1999), very little research has focused on the determinants of such perceptions in general, and on the role of social networks in particular. Indeed, recent economic theories suggest that assessments of risk about HIV infection, among other forms of opinions and judgments, are formed in interaction with others (McFadden, 1995, Manski, 2000), rather than in social isolation. Social networks, for example, can provide information about prevalence levels in a given community, or about what preventive strategies are available and/or acceptable, and thus modify one's perceptions of the environment.

Most of the evidence about the effects of social interactions on demographic behavior in developing countries comes from the literature on the diffusion of family planning. Indeed, family planning has certainly created a lot of "buzz" in all parts of Sub-Saharan Africa (Rutenberg and Watkins, 1997), and as a consequence, theoretical analyses of contraceptive choice have increasingly amended the traditional analytical frameworks relying solely on standard individual-centered variables, to include the effects of the context in which these decisions are taken. Numerous papers have thus shown that social interaction can influence and shape demographic change (Bongaarts and Watkins, 1996; Entwisle et al. 1996; Axinn and Yabiku, 2001; Kohler, 1997, 2000). In this paper, we take advantage of a unique dataset collected in rural areas of Malawi that provides information on conversations about HIV-AIDS, and build on the arguments developed in the family planning context, to assess how social interaction affects AIDS-related behavioral changes.

The literature on family planning has accumulated qualitative evidence and theoretical arguments for the relevance and importance of social networks in the analysis of demographic behavior. Quantitative evidence, however, has been harder to come by, and researchers engaging on that road have faced three major obstacles:

For one thing, data on social networks and conversations have rarely been collected: the view of individuals acting in isolation being vastly responsible for that. However, there has recently been a surge in interest for data on social networks.

On the other hand, quantitative analysis of the effects of social interaction is relatively difficult, often lacking an adequate theoretical framework. Indeed, one important question is: what is it exactly that networks do? The demographic literature has emphasized two mechanisms: social networks may work through social learning or social influence. Social learning implies that conversations and interactions help reduce the uncertainty associated with new behaviors and practices. In this context networks act as information conveyors. Social influence on the other hand implies that individuals make their decisions in a normative environment, and that social networks thus reinforce or alter the norms that prevail in a community.

Previous quantitative studies of the effects of social networks on demographic behavior have focused on network partner's behavior, and have typically reported that women whose network partners use family planning or worry about HIV infection are more likely to do so themselves. However, this specification doesn't provide much insight beyond the fact that conversational networks are relevant to the analysis of demographic behavior, and especially fails to determine the mechanism through which social interaction affects behavior.

Kohler et al. [2001] suggested an analytical framework to empirically distinguish between social learning and social influence: it is not only the content of interactions, i.e. the behavior of network partners, that matters, but also the structure of interactions, such as the size, the heterogeneity or the density of a network. Who people talk to, and in turn how are these people linked to each other defines the normative context in which interactions take place. According to the formal theory of social networks, social learning is maximized in sparse structures, in which network partners serve as mostly independent sources of information. On the other hand, dense structures, in which network partners are linked to each other, are expected to exert a stronger normative influence than isolated network partners.

Eventually, even when measures of social interaction are available and the theoretical framework of analysis is correctly defined, it is difficult to establish causal relations convincingly. The problem stems from the fact that conversational partners are not randomly assigned. Quite to the contrary individuals "select" themselves into different kinds of interactions depending on the goals they want to achieve, their own characteristics… Unobserved characteristics of individuals or their socio-economic environment may determine not only individual's demographic behavior or beliefs but also the content and structure of their interactions with others, thereby introducing bias in standard OLS estimates of such relationships. Most of the previous studies of social networks and demographic behavior maintained the assumption that network partners are somehow randomly assigned to the respondent, and have ignored this specification problem altogether (exceptions include Behrman et al., 2002).

In this paper, we are interested in the effects of social interactions on the formation of risk assessments about HIV-AIDS, and on the subsequent adoption of strategies to prevent infection. The following sections describe our strategy to address the theoretical and methodological issues mentioned above, before presenting estimates of the effects of social networks on HIV-related opinions and behaviors.

Data and Context

The analysis is based on data from the Malawi Diffusion and Ideational Change project (MDICP), whose aim is to evaluate and analyze the role of informal conversations in changing attitudes and behaviors about family planning and HIV/AIDS in rural Malawi. These are big topics of conversation in the surveyed communities: family planning has only been recently adopted, and the magnitude of the HIV/AIDS epidemic makes it a sensitive issue that men and women alike often discuss. The data consist of a longitudinal household survey, and a set of semi-structured interviews and focus groups that were collected in rural areas between 1997 and 2001. 1541 women of childbearing age and 1065 men were interviewed in 1998 in the Rumphi (north), Mchinji (Center) and Balaka (South) districts of Malawi. A follow-up survey was conducted in 2001. Of the 1541 women of the original sample, 1200 were re-interviewed successfully in 2001. In addition to this, field interviewers were asked to hold diaries, and relate some of the conversations about HIV-AIDS they had witnessed or had been involved in.

The regions covered by the survey are characterized by subsistence agriculture. Few men and women have attended school beyond the primary grades, even though most of them have had formal education. Education is highly valued as a route out of poverty, and those with higher education seek work in the neighboring cities of Blantyre, Lilongwe, or abroad. Remittances, wage labor or small-scale retailing provide the cash necessary to various kinds of expenses.

In one of the three survey sites (Rumphi district in the north) marriage is patrilinear, and residence of the spouses after marriage is patrilocal. This has implications for the dynamics of social networks formation since women have to form new networks at marriage, although they do sometime retain links with their natal families. In the other two sites, marriage is predominantly matrilocal, and it is the men rather than the women who have to form new networks at marriage.

Concerns about the risk of HIV infection are widespread in rural Malawi. 61 percent of the women surveyed expressed a high level of worry in 1998, and 47 percent in 2001. Knowledge about the epidemics is also important. In Malawi, most women knew that HIV/AIDS could be transmitted by sexual intercourse, but also knew about several other mechanisms (injections…). 87 percent of them knew of at least one recent death, which they suspected was caused by AIDS. Respondents are generally aware of different ways by which HIV/AIDS is transmitted, and of several ways of protection. Qualitative data collected in the various study sites showed that there also showed that there is a great deal of uncertainty about the desirability or even the possibility of reducing pre-marital or extra-marital sexual relations, as well as condoms in any kind of relationships.

The MDICP survey collected information on egocentric conversational networks. Respondents were asked about the total number of network partners they had ever chatted with about family planning and AIDS. AIDS is clearly a prominent topic in social interactions. 85% of the female respondents reported having discussed the probability of HIV infection with at least one network partner, and these percentages increased to 95% by the second round of data collection. In the different survey rounds, women talked with

4.3 to 5.7 network partners, suggesting an intensification of social interaction about HIV-AIDS[1]. These averages hide important geographical variations. In Balaka, women had 3.8 network partners in the first round, and 4.1 in the second round. The difference was not found to be statistically significant. In Rumphi, on the other hand, women reported having discussed HIV-AIDS with an average of 3.9 network partners in the first round, but this figure increased to 6.6 in the second round. Mchinji represents some kind of a middle ground, with averages of 5.2 and 6.3.

Additional information about network partners was obtained form the respondent for up to four partners, on attributes such as sex, residence, relationship to the respondent, education, risk perceptions about HIV/AIDS, and particularly relevant for our study, about the ties linking members of a given egocentric network.

Theoretical framework: the structure of social networks

Many social and demographic phenomena are embedded within networks of interdependencies. These networks have often been referred to as the "context" of these phenomena. Actors have to make decisions in an uncertain environment, and thus take into account the opinion and behaviors of others when establishing their own behaviors and attitudes. This process of information gathering and normative influence has often been referred to as "contagion", or "fads"… In the analysis of demographic events and behaviors, however, it has mostly been referred to as "diffusion".

The two most important mechanisms through which networks affect decisions are thought to be social learning and social influence. Social learning emphasizes the role of information in reducing uncertainty. Interlocutors are mainly conveyors of information, and the amount of normative pressure imposed by a network is reduced to a minimum. Learning through social networks about HIV-AIDS may entail learning the "history" and

[1] This increase could attributable to the reporting process, rather than to actual increases in the size of social networks. See Behrman et al. [2002] for a discussion of this point.

baggage of potential sexual partners, and may help evaluate the risk associated with a given partner… For example, the journals kept by the field interviewers in the framework of the MDICP, report many discussions, involving either males or females, that revolve around the previous partners somebody may have had, and how they may have infected him/her with HIV…

When social learning takes place, networks don't alter someone's preferences regarding family planning, or health behaviors… In the context of contraceptive adoption, women are thought as wanting to reduce their fertility, but lack information or fear the health consequences of modern methods of family planning. Informal conversations mainly ease their concerns, and social learning can help foster contraceptive change when there is an "unmet need" for family planning.

Social influence, on the other hand, extends beyond the pure process of learning, and allows for the possibility that preferences may be altered or even defined, by those with whom an actor interacts. Network partners may praise a woman believed to be using family planning because she is concerned about the health of her children, while they may blame her for doing so if they think she uses contraceptives because she cares about her looks... Alternatively they may express their disapproval of a man having sex with commercial sex workers, and thus exposing himself and his spouse to the risk of HIV infection.

Social influence thus occurs when an actor modifies his behavior/attitude according to the behavior/attitude of other actors in the social system, which then constitutes the frame of reference within which decisions are made.

Communication vs. Comparison

Two main processes are believed to lead to social influence: communication and comparison (Burt, 1987; Leenders, 2002). These processes differ mostly by the kind of reference group an individual uses when making decisions. When communication is the relevant process, the actors with whom an individual is directly tied constitute the relevant frame of reference. On the other hand, when comparison is the relevant mechanism, individuals observe the behavior and attitudes of actors they feel similar to (or aspire to be similar to) to make their decisions. For example, a woman may experience social influence through communication when deciding to adopt a given strategy to prevent HIV infection, if she trusts the judgments and arguments of her friends or her relatives. On the other hand, she may experience social influence through comparison when, for example in the context of family planning, she hears about the contraceptive behavior of women living in town, and perceives the imitation of such behavior as a way to achieve her specific goals (e.g. social promotion). It is worth noting that both kinds of processes are not exclusive, and may quite often coexist, or even conflict in a given community or at the individual level [Leenders, 2002].

In this paper, we chose to model social influence as a communication process, and we did so for two reasons:

First of all, analyses based on the comparison mechanism typically need "sociocentric" data on networks to assess whether actors share the same structural position in a network structure, and to identify relevant reference groups. This kind of data is typically collected for networks of small size, like in studies of organizations, or professions (Galaskiewicz and Wasserman, 1989) but is rarely collected in the case of population-based surveys of demographic behavior. Demographers have tried to capture imitation phenomena by including the prevalence of a behavior in a community as an independent variable in their models. For example, the probability that a woman adopts family planning depends not only on individual characteristics but also on the percentage of family planning users in her own community. Kravdal (2003) discussed the statistical problems implied by this practice.

We argue that this kind of approach builds on a poorly specified theory of social influence: is the "average" behavior the relevant reference in the decision-making process or is it just a statistical artifact that doesn't capture the process of comparison? In the case of studies focusing on rural parts of Africa, one implicit assumption that often seems to justify the use of the "average" behavior as a predictor is that communities and villages are fundamentally homogeneous. The anthropological literature has shown, however, that numerous stratification processes run through these communities. Here, we follow a different path use egocentric data (described above) to study social influence when it takes place through communication.

Qualitative descriptions of interactions about HIV-AIDS and descriptive statistics provide another reason to model social influence as a communication process. Indeed, contrary to family planning, most of the interaction about HIV-AIDS occurs within the household or the compound. Many respondents also report having discussed HIV prevention with one or both of their parents… So that it appears reasonable to assume that the opinions expressed by a respondent's network partners, and the behaviors he/she reports are relevant to the process of risk assessment.

The density of social networks

The common approach to the operationalization of a communication process in network analysis is via measures describing the structure of a network, often referred to as measures of centrality. These incorporate the number, length and strength of the paths connecting actors. Density is the formal concept we use here to define the degree of connectedness among network partners. Intuitively, the more the respondent's network partners are linked to each other, through links of friendship, family ties… the denser a network will be. Formally, the density of a network is the proportion of all possible social links that actually exist in a social network.

The influence exerted by a network is thus relatively strong if the network partners have not only a direct link to the respondent, but also indirect links through other network partners. On the other hand, social influence is relatively weak if the network partners are linked to the respondent only by a direct relation and have no indirect connections through other network partners.

The structure of social networks is important because networks not only provide opportunities for exchange of information, but also impose constraints on behavior for those who might otherwise wish to innovate. Social learning and Social influence can be distinguished on the basis of the above observations. If social learning dominates, we expect to find that both dense and sparse networks increase the probability that an individual will adopt a given behavior. Because sparse networks are more efficient conveyors of information, however, the effect of density should be negative. If social influence dominates, we expect a different pattern: variations in the prevalence of a given behavior in a network should exert only marginal effects on the individual's probability of adopting such a behavior in sparse networks, whereas these variations should be associated quite strongly with the likelihood of adopting this behavior in dense networks.

Empirical specification

Some evidence for the selectivity of social networks

We argued earlier that the hypothesis that network partners are randomly assigned to the respondent doesn't hold. The descriptive statistics reported in tables 1, 2 and 3 provide evidence for some aspects of selectivity that are important for the understanding of social interaction.

First of all, women who have interacted with at least two network partners about either family planning or HIV-AIDS are not a random sample of the population. Based on T-tests for equal means, we find that they are on average younger, more educated, and more likely to possess a radio. Furthermore, they use contraceptives and are worried about getting AIDS at a higher rate than the rest of the population. However, these women do not differ from the rest of the population on several other key components: indeed, they are roughly at the same parity, they are not more likely to have a metal roof on top of their house (which is an indicator of socioeconomic status), and they are not more likely to be engaged in a polygamous union.

Second, tables 2 and 3 show, by comparing the characteristics of the network partners with whom women discuss family planning and risks of HIV infection, that networks are not randomly assigned to them. Quite to the contrary, what constitutes an acceptable interlocutor for different topics of conversation seems to be socially defined, and family planning and HIV networks exhibit systematic patterns of differences. This runs contrary to what we would have expected, had network partners been assigned randomly.

Simple descriptive statistics thus show that women talk with their friends about family planning rather than about HIV/AIDS, but that they are more likely to discuss HIV risks with somebody living on their compound than family planning. Furthermore, we see that close to 90% of the interlocutors they mention having discussed risks of HIV infection with live in the same village as they do, against only 60% for family planning. Tables 2 and 3 also show some local variation in the characteristics of the "average" interlocutor, providing further evidence that the "choice" of network partners depends not only on individual characteristics (either observed or unobserved), but also on a series of characteristics of the social systems in which individuals are embedded. For example, women in Rumphi discuss family planning and HIV-AIDS with some of their female relatives a lot more than in Balaka or Mchinji. They also tend to discuss family planning with women from their own village a lot more in Balaka and Mchinji than in Rumphi. These aspects of selectivity motivate our empirical strategy.

Estimation strategies

Traditional models of network effects on demographic behavior are of the following form [Montgomery and Casterline, 1996]:

$$Y^{*}_{i,t} = X_{i,t}\beta + \delta_i \sum_{j \in N} \omega_{i,j} W_{j,t-1} + u_{i,t} \qquad (1)$$

where Y^* represents the propensity of an individual to adopt a given behavior, X is a set of both time-varying and time invariant covariates (including program effort). N is the set of individuals who belong to the social network of a given individual; W represents their respective behavior, the factors δ and ω represent weighting factors, and u is an error term, traditionally assumed to be "well-behaved". This equation has often been termed the "interdependent preferences" model (Pollak, 1976; Case, 1991), or the "autocorrelation" model (Anselin, 1982). Most of the previous studies of the effects of social networks have ignored the "weighting" matrix altogether and have focused on the information conveyed by social networks, thereby ignoring important dimensions of the process of norm formation. They have also overwhelmingly assumed that the disturbance term u is uncorrelated with both W, δ and ω, the network variables, and Y, the relevant demographic outcome.

Nonlinearities in the effects of social networks

As we mentioned earlier, one of the major findings of the literature on network effects is that networks not only provide opportunities for people, they also impose constraints on behavior. Sometimes they favor innovation while sometimes they discourage it.

Contemporary network theory has clearly defined how different structures of interaction may enhance or constrain behavioral innovation. For example, in his classical study of the labor market, Granovetter (1973) has shown that if the focus is on obtaining information about opportunities or new products, loosely tied networks are the most efficient social structure. On the other hand, information doesn't always flow quite as efficiently: in a network in which all network partners are linked to each other, most of the information gained form conversations is redundant, and social constraint rather than information is the relevant mechanism.

As we argued above, to empirically distinguish between social learning and social influence we modify (1) in order to take density and the interaction of density and content into account. The inclusion of nonlinearities in the effects of social networks adequately captures the different mechanisms through which social interaction affects behavior.
The logistic model to be estimated is of the following form, as in Kohler et al. (2001)

$$Y^*_{i,t} = X_{i,t}\beta + \delta_1 * (\%users) + \delta_2 * density + \delta_3 * density * (\%users) + \alpha_i + u_{i,t} \qquad (2)$$

where % *users* represents the content of social interaction, *density* represents its structure, and α_i represents all unmeasured variables (at the individual and local levels) that affect *Y*. When learning is the relevant mechanism through which networks affect behavior, we expect δ_1 (which measures the effects of the content of interaction in a network of density 0, the sparsest) to be quite large, positive and strongly significant. On the other hand, we expect δ_2 and δ_3 to be of little importance in the analysis (another hypothesis would be that δ_3 exerts a negative effect on behavior, because dense networks are less efficient sources information than sparser ones).

When influence is the relevant mechanism, we expect δ_1 to be of marginal importance, but δ_2 and δ_3 to be the coefficients of primary importance.

Endogeneity concerns

Previous studies have treated the α_i's as random variables and have ignored the fact that they could also be correlated with the variables representing social networks. However, some unobserved processes could jointly determine networks and demographic behaviors jointly, thereby biasing the estimates of network effects produced by standard OLS methods. The comparison with family planning networks has provided some evidence for such selectivity in the composition of conversational networks. This was true in the simple framework of social interaction being solely represented by its content, i.e. the prevalence of a given behavior among network partners. This is even more so when we adopt more complex and detailed representations of social interaction that also include the structure of social networks.

Indeed, just as women may choose their network partners based on their own preferences and inclinations, and on the kind of answer they expect, the structure of a woman's social network is likely to be determined by factors that also have an influence on various demographic outcomes (contraception, self-perceived risk of HIV infection…). For example, in rural Malawi, because of transportation and communication constraints, interaction occurs primarily with those who live in the same village, or in the villages nearby, thus constraining the networks toward higher homogeneity and density (see tables 2 and 3). Standard estimates neglecting this problem not only pick up the independent effects of social interaction (if any at all) but also of all those unobserved variables determining jointly treatment and outcome.

The panel structure of the data set allows us to use fixed-effects models to try to avoid this source of bias. Each individual serves as its own control and therefore provides an adequate specification if the confounding factors are believed to be time-invariant. More formally we have,

$$
\begin{aligned}
Y_{it}^{*} &= X_{i,t}\beta_1 + \delta_1 * (\%users) + \delta_2 * density + \delta_3 * density * (\%users) + f_i + \varepsilon_{i,t} \\
\%users &= X_{i,t}\beta_2 + hf_i + u_{i,t} \qquad (3) \\
density &= X_{i,t}\beta_3 + h' f_i + u'_{i,t}
\end{aligned}
$$

Where all the terms have the same meaning as in (2) except for f_i which represents all the time-invariant factors, which determine both outcome and networks.

What the fixed-effects specification does is that it takes the first-difference between two consecutive observations to yield consistent estimates of δ_1, δ_2 and δ_3. Indeed, when taking first-difference, the term f_i , which made the network variables endogenous and the coefficients of interest unidentified is being "swept out" of the estimated equation, thereby breaking the correlation of the between the endogenous network variables and the set of unobserved fixed effects subsumed in *f.* As long as $corr(\varepsilon, u)$ and $corr(\varepsilon, u')$ are 0, a fixed-effects specification yields consistent estimates of the parameters of interest. The correlation between *u* and *u'* however doesn't have to be 0.

Dealing with measurement error and attrition

One of the major challenges facing network surveys based on individual recall of conversation is that of measurement error. This error might be completely random, and bias coefficients toward zero in standard regressions. In fixed-effects analysis, the bias introduced by random measurement error is going to be even greater since the estimates are produced based on the deviations from within-individual averages.

Measurement error might also be systematic rather than random. For example, the way the survey is perceived might inflate the number of conversations reported by a certain factor, if the respondents somehow get to believe that interviewers are here to promote family planning… But to the extent that this systematic (under) over-reporting is the same across waves, it is adequately represented by a fixed effect, and is accounted for when we take the first-difference of equation (3). White and Watkins (2000) found empirical for such systematic misreporting using data collected in rural Kenya.

Another hurdle most longitudinal studies have to clear is that of attrition. Indeed, almost 20% of the original sample is lost to attrition between the 2 waves of the survey. Reasons for attrition include death, migration or refusal. The extent to which this loss of follow-up affects the estimation of the parameters of interest has been evaluated in the context of developed countries. Most studies concluded that attrition in panel data (most case studies considered the case of the long-standing PSID) even though of considerable extent, doesn't seriously undermine analysis (Fitzgerald et al., 1998). These analyses have also been replicated in the context of developing countries, and have reached the same conclusions (Alderman et al. 2001; Falaris, 2003)

However, one cannot simply rule out the importance of attrition, and the potential for bias stemming from loss of follow-up still depends on the particular model one tries to estimate, and its specification.

In the context of fixed-effects models, the prospects are rather encouraging. Indeed, to the extent that the mechanisms leading to attrition are due time-invariant characteristics of the individual, it is adequately "picked up" by taking the first-difference of equation (2), and the analysis conducted on the non-attritors (women who have been interviewed in Malawi 1 as well as Malawi 2) yields consistent estimates of the effects of social networks. Ziliak and Klisner (1998) shows that a simple fixed-effects specification does as good a job of dealing with attrition than the more complicated selection models "a la Heckman" in the context of the estimation of lifetime labor supply functions, thereby suggesting that time-invariant characteristics are at the origin of the loss of follow-up. On the other hand, if the mechanisms leading to attrition are wave-specific, attrition might be a bigger problem, and might require the use of a correction function or instrument variables.

When trying to assess the effects of social networks in rural Malawi, it is likely that most of the losses to follow-up originate in those time-invariant characteristics of the individual, and are therefore correctly represented by a fixed-effect. However, attrition due to death in the case of the analysis of the perceived risk for HIV infection is a mechanism that would lead to selective attrition: if the more worried people die at a higher rate than their less worried counterparts, the estimates are likely to be biased, for example.

Results:

Self-perceived risk of HIV infection

Our primary dependent variable is the respondent's own risk perception. The survey asked, "how worried are you that you might catch AIDS?" and respondents were given three choices: not at all, moderately and highly worried. The same questions were asked about one's network partners. We also look at another survey question, which asked respondents about the kind of strategies they would use to prevent HIV infection. More specifically, we're interested about knowledge and availability of effective strategies to reduce the risk of infection. Thus we look at whether a woman (man) would discuss HIV prevention with her (his) spouse and advise him (her) to take care.

Table 5 reports a set of coefficients for the fixed-effects model in relation (2). We also report random effects estimates, for comparison purposes. We initially represented our primary dependent variable as a continuous index from the recorded categorical values. In doing so, we adjusted for potential heteroscedasticity by using the Huber-White estimator of variance, and results are reported in an appendix table. Indeed, we found that respondents reporting being a little worried about HIV-AIDS were not statistically different from respondents reporting not to be worried about AIDS infection. Furthermore, network partners with medium level of worry had little effect on individual risk assessments (if any at all, results not reported). A reason for this may be that respondents answer "a little" to this question, just to please the interviewer, if they perceive that the survey team is also part of an effort to promote HIV prevention.

Nevertheless, this makes problematic treating the dependent variable as a continuous index: indeed this constrains the effects of network variables to be constant across all categories of the dependent variable, and biases the estimates toward zero. We thus pooled together the categories "a little worried" and "not worried at all" and estimated logit models of equation (2). Results are reported in table 5[2]. Also, to avoid the awkwardness of having to read and interpret multiple interaction terms in the analysis, we estimate our models with only the prevalence of highly worried network partners in one's network as an indicator of the content of social interactions.

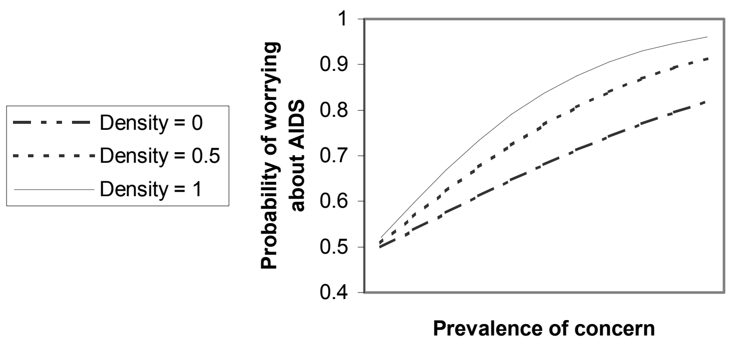
Estimates for Women

Table 6 shows, that when all the observations are pooled together, the coefficients δ_1 and δ_3 in equation (2) prove to be positive, large and highly significant, whereas the coefficient for density isn't statistically significant. As seen in figure 1, the probability of being worried about HIV infection is thus systematically higher in a denser network than in a sparse network. This would make for a situation typical of social influence, if it weren't for the slope of the curves. In a network of density 0, the effect of an additional "worried" network partner is almost linear, which is consistent with social influence. Every additional network partner worrying about HIV has the same effect on the probability of the respondent to worry. In a situation of social learning, we would have expected this effect to be nonlinear and the probability of being worried about getting AIDS to increase at a decreasing with each additional worried network partner.

However, in a network of density 1, the effects of an additional "worried" network partner are indeed nonlinear and the probability of being worried about getting AIDS does increase at a decreasing rate with the prevalence of worry in a network. In a situation typical of social influence, we would have expected them to increase.

[2] Qualitatively, the results based on a continuous representation of the dependent variable don't tell a different story than the results based on logit models. The point estimates, on the other hand, are systematically smaller (see appendix table).

Fig 1: The effect of the prevalence of concern in a network on individual risk assessments. Fixed effects estimates, all districts pooled.

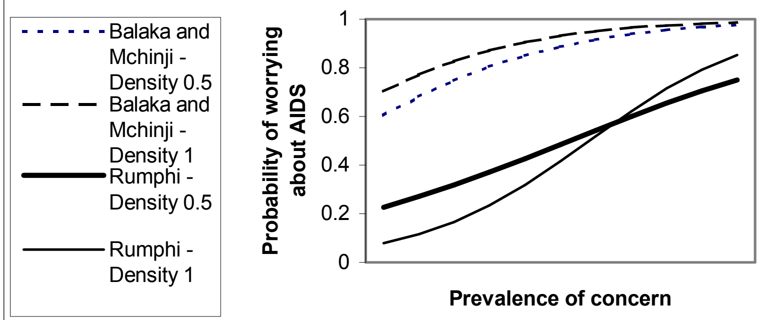


We argue that this unusual pattern is due to regional variation in the effects of social networks, and present estimates broken down at the district level (table 5). We found that Rumphi district in the north of Malawi, could not be pooled with the other districts. The coefficient δ_1 is not significant, and the effects of the network partners' risk assessments on one's own perception interact with the density of the network in which they are embedded. The pattern is clearly one of social influence, and shows that networks are an ambivalent institution in regard of risk assessments, sometimes constraining, sometimes facilitating. Indeed, these constraints operate against a pessimistic perception of the risk of infection when the prevalence of "worry" among network partners is low; however they do favor such a perception when the prevalence of worry is high.

In Balaka and Mchinji on the other hand, we have a pattern typical of social learning (figure 2). δ_1 is large, positive and highly significant, and even though the highest probabilities of being highly worried are found in a network of density 1, the important thing to consider here is the slope of the curves. It increases at a decreasing rate, implying that an additional worried network partner has a smaller and smaller effect.

Women in these districts thus merely use their networks as information conveyors rather than as a mean to evaluate appropriate attitudes and behaviors.

Fig. 2: The effects of the prevalence of concern about AIDS on risk assessments, Fixed effects estimates.



On figure 2, we also notice that at any given level of concern about AIDS in a network, the individual probability of forming a pessimistic assessment about HIV infection is higher in Balaka and Mchinji than in Rumphi.

Estimates for Males

The demographic literature on the effects of conversational networks has almost exclusively looked at women. Maybe the perception that women are more engaged in such networks, are more likely to "gossip" about various topics… is at the origin of such attention. However it seems worth looking at the effects of social interaction on males' risk assessments about HIV-AIDS. Indeed, the behavior of men is one major determinant of the speed and extent of the epidemic. Furthermore, casual observations in the sample villages indicate that men spend a great deal of time chatting about HIV, and the threat it poses to their health and their communities. Men report having discussed the risk of getting AIDS with about as many network partners as women, in all districts (see table

4). The time trends we identified for women (intensification of interaction in Rumphi…) hold for men as well. And just as for women, males' networks are highly gendered: on average 95% of the network partners are of the same gender.

Estimates of the effects of social networks on males' risk perceptions are reported in table 5. Here, the coefficient δ_1 proves to be large and highly significant, suggesting that interactions shape one's risk assessment through social learning. This finding is corroborated by numerous descriptions of conversations in the field journals held by the interviewers. Indeed, typical conversations among males about the probability of HIV infection evolve around the "history" of potential or past sexual partners, and try to evaluate the degree of risk involved in a given relationship/partnership…

Local variation in the effects of social interaction: the role of marriage patterns

Contrary to the patterns uncovered for women, however, estimates for males don't exhibit any sort of variation at the local level. This provides us with some hint to explain the different mechanisms through which social networks shape risk assessments in Rumphi and in Balaka and Mchinji.

As we mentioned earlier, marriage in Rumphi is patrilocal, i.e. that women leave their families once unions are formed, and start living in their husband's family. Marriage is clearly a turning point since, even if they keep contact with their origin family, they often have to form new networks. In Balaka and Mchinji, on the other hand, marriage is matrilocal, and networks are not altered by marriage in such a dramatic manner. As a consequence of this, we see that women in Rumphi engage in interaction with members of the family of their husband at a much higher rate than women in Balaka and Mchinji (Table 3). Especially, they report having discussed risks of HIV infections with one of their sisters-in-law much more frequently than in the other two districts: in Rumphi, on average, sisters-in-law represented more than 17% of the total number of network partners women reported, against only a little bit more than 7% in the other districts.

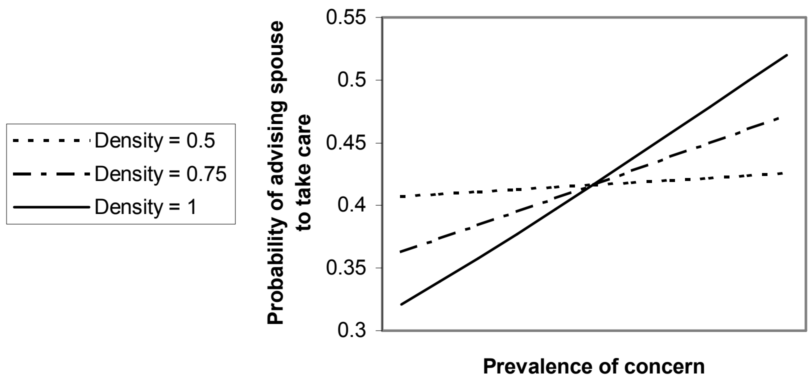
Reports of the type of conversations women have with their in-laws can be found in the field journals the MDICP survey gathered. They suggest that sisters-in-law represent an important vector of social pressure on married women in Rumphi. They appear to be frequently inquiring, or even “spying” about one woman’s actions and moves. Interaction about HIV-AIDS in Rumphi is thus a moment that is not free of normative pressure, quite to the contrary. It is fully embedded in the tension that runs within households and compounds, between a married woman and the family of her husband. Whereas in Balaka and Mchinji networks shape risk assessments by providing information about risky behaviors and preventive strategies, in Rumphi they do so by imposing judgments and evaluations on specific behaviors. We argued that this fundamental difference in the effects of social networks originates in the marriage rules that prevail in Rumphi, and their consequences for network dynamics. This may also provide some explanation for the pattern uncovered in figure 2, that at any given level of concern about AIDS in a network, the individual probability of forming a pessimistic assessment about HIV infection is higher in Balaka and Mchinji than in Rumphi.

Estimates for behavioral responses to HIV-AIDS

Spousal communication about preventing AIDS also constitutes an important determinant of behavioral responses to the risk of infection. Here, the dependent variable is a dummy variable taking value if the women advised her spouse to take care about HIV-AIDS. The specification of the right-hand side variables describing networks effects is analogous to the previous sections.

Our estimates indicate that network partners have a relevant and significant effect on spousal communication about HIV-AIDS. The pattern we unfold is typical of social influence, since the higher probabilities of advising one’s spouse to take care are found in networks of high prevalence of “worry” and of high density, and that the marginal effect of an additional network partner worried about getting AIDS increases with the density of a network. The estimates for males reveal the same pattern (table 6).

Fig. 3: The effects of the prevalence of worry in a network on the probability of adopting preventice behavior, Female sample. All districts pooled.



Some preliminary conclusions

Our results show that social environment is a major determinant of attitudes and judgments in the context of the AIDS epidemic, and has substantive implications for the adoption of preventive behavior. As previous studies of the role of social interaction in the context of family planning adoption have concluded, we found that the opinions and behaviors of others have significant effects on one's own attitudes toward HIV-AIDS. Furthermore, these effects are quite complex: if in many cases, respondents use their networks mostly as information conveyors when evaluating their own risk of infection (in Balaka and Mchinji for women, and in all districts for men), it also happens that networks impose constrains on the respondents, and help define what constitute a "socially acceptable" evaluation of one's risk of infection. We argued that marriage patterns and their effects on the dynamics of network formation were one major reason why social interaction influenced individual attitudes through a radically different mechanism in Rumphi.

Our analyses also emphasized the need to control for the unobserved processes that jointly determine the content and structure of social interactions and individuals' decisions regarding strategies to prevent HIV infection. The non-random selections of networks poses threats to the analyses of the role of social interactions on demographic behavior that few studies thus far have explicitly dealt with. As we have shown, standard estimates without such controls are likely to misrepresent the effects of social interactions and normative context on health behaviors, or other kinds of demographic outcomes (see appendix figures 1 & 2).

References

Alderman, H., J. R. Behrman, H.-P. Kohler, J. Maluccio, and S. C. Watkins (2001). "Attrition in longitudinal household survey data: Some tests for three developing country samples." *Demographic Research* 5(4), 79-123.

Axinn, W. G. and S. T. Yabiku (2001). "Social schange, the social organization of families,and fertility limitation". *American Journal of Sociology* 106(5), 1219-1261

Behrman, J. R., H.-P. Kohler, and S. C. Watkins (2002a). "Social networks and changes in contraceptive use over time: Evidence from a longitudinal study in rural Kenya" *Demography* 39(4), 713-736.

Bongaarts, J. and S. C. Watkins (1996). "Social interactions and contemporary fertility transitions". *Population and Development Review,* 22(4), 639-682.

Burt, R. S. (1987). "Social contagion and innovation: Cohesion versus structural equivalence" *American Journal of Sociology* 92(6), 1287-1335.

Case, A. C. (1991). "Spatial patterns in household demand". *Econometrica* 59(4), 953-965

Entwisle, B., R. D. Rindfuss, D. K. Guilkey, A. Chamratrithirong, S. R. Curran, and Y. Sawangdee (1996). “Community and contraceptive choice in rural Thailand: A case study of Nang Rong.” *Demography* 33(1), 1-11.

Galaskiewicz, J. Wasserman, S.S., (1989) “Mimetic processes within an organizational field: an empirical test”, *Administrative science quarterly,* 34, 454-479

Granovetter, M. S. (1973). “The strength of weak ties”. *American Journal of Sociology* 78(6), 1360-1380

Kohler, H.-P. (1997). “Learning in social networks and contraceptive choice”. *Demography* 34(3), 369-383

Kohler, H.-P. (2000). “Fertility decline as a coordination problem”. *Journal of Development Economics,* 63(2), 231-263.

Kohler, H.-P., J. R. Behrman, and S. C. Watkins (2001). “The density of social networks and fertility decisions: Evidence from South Nyanza District, Kenya” *Demography* 38(1), 43-58.

Kravdal, O. (2003) “The problematic estimation of ‘imitation effects’ in multilevel models”, *Demographic research, vol* 9

Leenders, R.Th. A.J., (2002), “Modeling social influence through network autocorrelation: constructing the weight matrix”, *Social networks,* 24, 21-47

Manski, C. F. (2000) “Economic analysis of social interaction”. *Journal of Economic Perspectives* 14(3), 115-136

McFadden, D. L. and K. E. Train (1995). “Consumer’s evaluation of new products: Learning from self and others”. *Journal of Political Economy* 104(4), 683-703.

Montgomery, M. R. and J. B. Casterline (1996). “Social learning, social influence, and new models of fertility”. *Population and Development Review* 22(Suppl.), 151-175.

Pollak, R. A. (1976). “Interdependent preferences”. *American Economic Review* 66(3), 309-320.

Rutenberg, N. and S. C. Watkins (1997). “The buzz outside the clinics: Conversations and contraception in Nyanza Province, Kenya”. *Studies in Family Planning,* 28(4), 290-307.

UNAIDS (1999). Sexual Behavioral Change for HIV: Where Have Theories Taken us? Geneva: United Nations.

White, K. and S. C. Watkins (2000). "Accuracy, stability and reciprocity in informal conversational networks in rural Kenya". *Social Networks* 22(4), 337-355.

Table 1: descriptive statistics of family planning network characteristics
districts of Balaka and Mchinji, and Rumphi. Women with 2 or more network partners in both waves.

	All Localities		Balaka and Mchinji		Rumphi	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
(Uncensored) size of network	4.49	6.18a	4.66b	5.77a,b	4.13	7.04a
Family planning Use by Network partners						
Average proportion of network Partners using Family planning	0.589	0.74a	0.598	0.73a	0.57	0.76a
Density						
Average density of network among network partners	0.77	0.898a	0.78	0.9035a	0.8	0.88a
Average proportions of Network Partners who are						
males	0.025	0.02	0.0266	0.022	0.02	0.0167
females	0.97	0.98	0.973	0.9778	0.98	0.9833
male relatives of respondent	0.007	0.008	0.007	0.0066	0.006	0.01
female relatives of respondent	0.33	0.35	0.3b	0.29b	0.37	0.46a
confidents	0.27	0.29	0.256b	0.27b	0.3	0.33
acquaintances	0.183	0.1a	0.18	0.106a	0.18	0.09a
friends	0.51	0.597a	0.51	0.6a,b	0.51	0.575a
Average proportion of Network partners who live						
in same household	0.006	0.007	0.006	0.01b	0.008	0.003
on the same compound	0.133	0.164a	0.113b	0.15a,b	0.1755	0.1864
in the same village	0.59	0.71a	0.62b	0.73a,b	0.565	0.674a
in the city	0.02	0.009a	0.02	0.006a	0.018	0.013
N	845	845	570	570	275	275

Results of two-tailed test for equal means:

a the difference between rounds is significant at the .05 level
b the difference between in Balaka/Mchinji and women in Rumphi is significant at the .05 level

Table 2: descriptive statistics of HIV network characteristics
districts of Balaka and Mchinji, and Rumphi. Women with 2 or more network partners in both waves.

	All Localities		Balaka and Mchinji		Rumphi	
	Round 1	Round 2	Round 1	Round 2	Round 1	Round 2
(Uncensored) size of network	5.62	6.57*	5.77	6.08[b]	5.32	7.58[a]
HIV Risk perceptions of network partners						
Average proportion of network Partners highly worried / HIV infection	0.32	0.254[a]	0.33[b]	0.26[a,b]	0.52	0.28[a]
Average proportion of network Partners not worried / HIV infection	0.166	0.29[a]	0.186	0.294[a,b]	0.18	0.357[a]
Density						
Average density of network among network partners	0.8	0.87[a]	0.78[b]	0.873[a]	0.79	0.86[a]
Average proportions of Network Partners who are						
males	0.05	0.0575	0.061	0.06	0.04	0.053
females	0.9446	0.9405	0.939	0.94	0.96	0.947
male relatives of respondent	0.015	0.027[a]	0.015	0.0231	0.014	0.035[a]
female relatives of respondent	0.33	0.29[a]	0.284[b]	0.25[b]	0.42	0.39
confidents	0.11	0.11	0.11	0.1[b]	0.12	0.147
acquaintances	0.32	0.36[a]	0.337	0.36	0.289	0.36[a]
friends	0.228	0.225	0.21[b]	0.22	0.258	0.23
Average proportion of Network partners who live						
in same household	0.12	0.126	0.11	0.12	0.138	0.137
on the same compound	0.57	0.64[a]	0.54[b]	0.63[a,b]	0.68	0.72
in the same village	0.87	0.9[a]	0.85[b]	0.9[a]	0.89	0.89
in the city	0.0058	0.0018[a]	0.006	0.0014[a]	0.005	0.0028
N	784	784	528	528	256	256

Results of two-tailed test for equal means:

[a] the difference between rounds is significant at the .05 level
[b] the difference between in Balaka/Mchinji and women in Rumphi is significant at the .05 level

Table 3: descriptive statistics of HIV networks, Males of 2+ network partners in both waves

	Balaka and Mchinji		Rumphi	
	Rd 1	Rd2	Rd1	Rd2
(Uncensored) size of network	7.44*	7.34*	6.02	7.95
Density	0.78*	0.86*	0.85	0.9
% Network Partners who are				
males	0.95	0.94	0.93	0.94
females	0.05	0.06	0.07	0.06
male relatives of respondent	0.167*	0.16*	0.27	0.28
female relatives of respondent	0.011*	0.012	0.023	0.011
confidents	0.12*	0.13	0.1	0.15
acquaintances	0.28	0.3	0.32	0.33
friends	0.18	0.22	0.18	0.22
~~%~~ Network partners who live				
in same household	0.14	0.13	0.16	0.1
on the same compound	0.56*	0.57	0.64	0.62
in the same village	0.82*	0.84	0.88	0.84
in the city	0.002	0.006	0.0008	0.006
N	466	466	182	182

* The difference between Rumphi and Balaka / Mchinji is significant at the .05 level

Table 4: Regression of AIDS worries on individual and network characteristics- Fixed effects and Random effects
Dependent variable: 1 if highly worried, 0 if not / Rounds 1 & 2

	Random effects			**Fixed effects**		
	All	Balaka + Mchinji	Rumphi	All	Balaka + Mchinji	Rumphi
Age	**	ns	*			
Parity	ns	ns	*	ns	ns	ns
Marital status	ns	ns	ns	ns	ns	ns
Metal Roof on the house	ns	ns	*	ns	ns	ns
Ever been to school?	ns	ns	ns			
Superior education	ns	*	*			
Has a job	*	ns	***	ns	ns	*
% nps with high aids worries	**2.32****	**3.08****	**0.6**	**1.51***	**3.07****	**0.45**
	(0.46)	***(0.58)***	***(0.8)***	***(0.72)***	***(1.01)***	***(1.2)***
Density	**-0.1**	**0.37**	**-1.24***	**0.038**	**0.86**	**-2.45****
	(0.3)	***(0.4)***	***(0.5)***	***(0.51)***	***(0.7)***	***(0.87)***
% nps * density	**1.04***	**0.52**	**2.48****	**1.61**	**0.42**	**3.76***
	(0.52)	***(0.7)***	***(0.89)***	***(0.85)***	***(1.11)***	***(1.52)***
Dummy for round=2001	ns	ns	**	ns	**	**
N	784	528	256			

*** p < .01 ** p<.05 * p < .1
Standard errors are adjusted to take into account the correlation of consecutive observations

Table 5: Risk perceptions of Males, individual and network variables, all districts pooled. Fixed effects and Random effects estimates

	Random effects	Fixed effects
Age	ns	
Metal Roof on the house	ns	**
Owns a radio	ns	ns
Ever been to school?	ns	
Superior education	*	ns
Has a job	***	ns
Program effort	ns	*
% nps with high aids worries	3.06*** *(0.45)*	3.88*** *(1.01)*
Density	-0.26 *(0.32)*	0.6 *(0.85)*
% nps * density	0.86 *(0.55)*	-0.21 *(1.1)*
Dummy for round 2	ns	ns
N	648	

*** p<.01 **p<.05 * p<.01

In the random effects specification, standard errors are adjusted to take into account the correlation of of consecutive observations

Table 6: probability of advising spouse to take care about HIV infection.
Random effects and Fixed effects specifications

	Females		Males	
	RE	FE	RE	FE
Age	ns		ns	
Parity	ns	*		
Marital status	***	**	ns	
Metal Roof on the house	ns	ns	ns	ns
Owns a radio	ns	ns	ns	ns
Ever been to school?	ns		**	
Superior education	ns	ns	ns	ns
Has a job	*	ns	*	*
% nps with high aids worries	**-0.78**** ***(0.33)***	-0.98 *(0.6)*	0.31 *(0.38)*	-0.68 *(0.64)*
Density	**-0.38*** ***(0.21)***	**-0.74*** ***(0.43)***	-0.11 *(0.29)*	**-0.83*** ***(0.49)***
% nps * density	**0.93**** ***(0.38)***	**1.51**** ***(0.66)***	-0.09 *(0.44)*	**1.13*** ***(0.69)***
Dummy for round=2001	-0.58***	-0.3**	ns	ns
N	784		648	

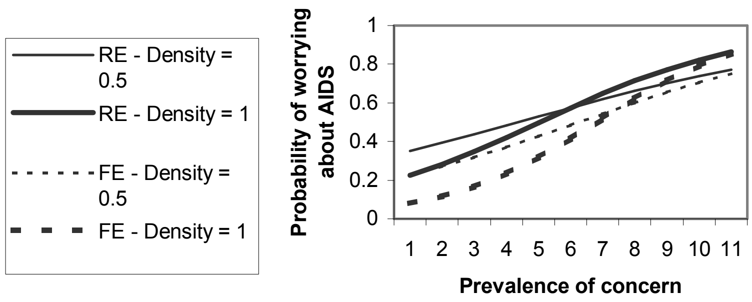
*** p<.01 **p<.05 * p<.1
In the random effects specification, standard errors are adjusted to take into account the correlation of consecutive observations

Appendix Table : Regression of AIDS worries on individual and network characteristics, Random effects and Fixed effects specification. Dependent variable treated as continuous.

	Random effects			**Fixed effects**		
	All districts	balaka+Mchinji	Rumphi	All districts	Balaka+Mchinji	Rumphi
Age	*	ns	ns			
Parity	ns	ns	ns	ns	ns	ns
Marital status	ns	**	ns	**	**	ns
Metal Roof on the house	ns	ns	ns	ns	ns	ns
Ever been to school?	**	ns	ns			
Superior education	ns	**	*			
Has a job	ns	ns	**	ns	ns	ns
% nps with high aids worries	**0.89****	**1.09****	**0.28**	**0.624****	**0.77****	**0.25**
	(0.13)	***(0.14)***	***(0.23)***	***(0.17)***	**(0.21)**	***(0.35)***
Density	**0.015**	**0.08**	**-0.34***	**-0.02**	**0.05**	**-0.26**
	(0.1)	***(0.12)***	***(0.17)***	***(0.13)***	***(0.15)***	***(0.22)***
% nps * density	**0.11**	**0.015**	**0.57***	**0.28**	**0.22**	**0.51**
	(0.14)	***(0.16)***	***(0.25)***	***(0.19)***	***(0.24)***	***(0.36)***
Dummy for round=2001	ns	ns	**	ns	ns	ns
N	784	528	256			

** p<.01 *p<.05
Standard errors are adjusted to take into account the correlation of consecutive observations,
and the potential heteroscedasticity emerging from clustering of observations at the village level

Appendix Fig. 1: Comparison of Random effects and Fixed effects estimates, Rumphi district.



Appendix Fig. 2: Comparison of Random effects and Fixed effects estimates. Probability of advising spouse to take care. Male sample.