



Demographic Research a free, expedited, online journal
of peer-reviewed research and commentary
in the population sciences published by the
Max Planck Institute for Demographic Research
Konrad-Zuse Str. 1, D-18057 Rostock · GERMANY
www.demographic-research.org

DEMOGRAPHIC RESEARCH

SPECIAL COLLECTION 1, ARTICLE 4

PUBLISHED 19 September 2003, PAGES 109-142

www.demographic-research.org/special/1/4/

DOI: 10.4054/DemRes.2003.S1.4

Research Article

**How do we know we need to control
for selectivity?**

Susan C. Watkins

Ina Warriner

This special collection is edited by Susan Watkins, Eliya M. Zulu, Hans-Peter Kohler and Jere Behrman. The papers in this special collection were presented at the conference "Research on Demographic Aspects of HIV/AIDS in Rural Africa", held at the Population Studies Center, University of Pennsylvania, October 28, 2002.

© 2003 Max-Planck-Gesellschaft.

Table of Contents

1	Introduction	110
2	Data and context	113
2.1	Data	113
2.2	Context	116
3	Homogeneity and homophily: Do birds of a feather flock together?	117
3.1	Contexts and characteristics of informal conversations about family planning	117
3.2	Characteristics of network partners	119
4	Strategic selectivity on family planning	124
4.1	Evidence for strategic selectivity with respect to family planning	125
4.2	Qualitative evidence	126
4.3	Comparison of network characteristics across different topics of conversation	127
5	Conclusions	129
6	Acknowledgements	130
	Notes	132
	References	136x

How do we know we need to control for selectivity?

Susan Cotts Watkins ¹

Ina Warriner ²

Abstract

In the previous two decades there has been considerable progress in recognizing biases due to selectivity that are associated with the use of observational data to make causal inferences and in developing models to control for these biases statistically. Often there is a difference between estimates produced by models that attempt to control for selectivity and those that do not. Since a difference alone does not persuasively argue for one model over another, analysts typically rely on their *a priori* expectations of selectivity based on theory or intuition. Here we suggest that the analyst's judgement about the appropriate analytical model may be informed by simple descriptive statistics and qualitative data. We use data on social networks collected in rural Kenya, since the analysis of networks is likely to raise questions of selectivity, and simple examples. Although we do not provide general rules for assessing when models that control for selectivity should be used, we conclude by recommending that analysts inform their judgement rather than rely on theory and intuition to justify controlling for selectivity. Although our data are particular, the implications of our approach are general, since *a priori* evaluations of the credibility of assumptions on which analytic models are based can be made in other settings and for other research questions.

¹ Corresponding author: Susan Cotts Watkins, Tel. 215-898-4258, Fax 215-898-2124.

E-mail: swatkins@pop.upenn.edu

² Scientist, Department of Reproductive Health and Research, World Health Organization.

E-mail: warrineri@who.int

1. Introduction

There is currently interest among social scientists in the role of social networks in influencing attitudes and behavior. For example, demographers have found that the attitudes and behaviors of an individual are strongly related to the presence of these same attitudes and behaviors among those with whom she interacts (e.g. Montgomery and Casterline 1993; Entwisle et al 1996; Valente et al 1997; Munshi and Myaux 1997; Montgomery and Chung 1999; Entwisle and Godley 1998; Kohler, Behrman and Watkins 2001; Behrman, Kohler and Watkins 2002; Bühler and Kohler 2003), and similar analyses have been conducted with respect to more general behaviors (e.g. Brock and Durlauf 2001). It is thus tempting to conclude that individuals are influenced by others, i.e. that there is a “treatment” effect, where the treatment is the effect of an individual’s network partners. Establishing causal relations using observational data has, however, been a challenge for social scientists. When it is not possible to conduct experiments that assign individuals to “treatment” or “control” groups, it is typically assumed that rational actors can and do calculate for themselves whether they would be better off with a “treatment” (or not), and then select a treatment (or not). Although some of the bases for these choices may be observable in the data, others are likely to be unobserved.

In the previous two decades, there has been considerable progress in recognizing biases due to selectivity that are associated with the use of observational data to make causal inferences and in developing models to control for these biases statistically (Note 1). All of these, however, face important limitations (Manski 1992; Stolzenberg and Relles 1997; Winship and Morgan 1999; Freedman 1999; Winship and Sobel forthcoming). It might be thought that using statistical models to control for selectivity based on unobserved characteristics is justified by the outcome: if the introduction of controls have the effects that the analyst’s *a priori* theories or intuitions lead her or him to expect, the controls are justified. In this paper, we consider the effect of social networks on contraceptive use in rural Kenya. It is possible to tell plausible selectivity stories. For example, because small families motivated by deliberate choice and achieved by modern methods of family planning are new in this context, intuition suggests that it is reasonable to expect that actors would want to discuss this new model of reproduction with others, perhaps to evaluate the advantages and disadvantages of having fewer children or of using the new methods, perhaps to assess the normative acceptability of this new model of reproduction. It is also reasonable to assume that the choice of specific network partners with whom to discuss these issues is not random: actors can be expected to choose partners whose attributes are similar to their own and whose judgements are thus most pertinent, or partners who have had experience with the new methods, consistent with theories of learning and rational decision-making

under uncertainty. If we were to proceed directly to controlling for the expected selectivity, we would produce the results shown in Column 1 of Table 1 (other variables in the equation are not shown, and the data will be described in Section 2).

Table 1: *Linear probability models for women’s current contraceptive use. Respondent’s contraceptive use is measured at Kenya 1, Kenya 2 and Kenya 3.*

Method	OLS	Fixed Effects	GMM IV-FE
# of nwps using family planning, time t-	0.0703 (0.0085)**	0.033 (0.009)**	0.028 (0.012)*
Dummy for not married, time t	-0.0699 (0.0343)*	-0.003 (0.013)	-0.011 (0.024)
Children ever born, time t	0.0043 (0.0048)	-0.022 (0.049)	-0.034 (0.046)
Respondent has radio, time t	0.0291 (0.0210)	0.037 (0.028)	0.027 (0.026)
Respondent has metal roof, time t	0.0245 (0.0273)	-0.042 (0.045)	-0.036 (0.043)
Respondent has at least primary	0.0717 (0.0269)**		
Respondent has secondary education	0.0828 (0.0396)*		
Age	0.0375 (0.0093)**		
(Age/10) squared	-0.0522 (0.0125)**		
Dummy for survey wave Kenya 2	0.0167 (0.0207)	0.032 (0.023)	0.033 (0.023)
Dummy for survey wave Kenya 3	0.0411 (0.0230)+	0.036 (0.022)+	0.033 (0.025)
Constant	-0.6800 (0.1608)**		
N (number of women, each observed at three surveys)	497	497	497

Notes:

p-values: + < .1; * < .05; ** < .01

Robust standard errors are used to account for potential heteroscedasticity (especially in the linear probability model). For OLS estimates these robust standard errors additionally account for the correlation of residuals for the same individual across time periods.

Column 1 shows the estimates on the respondent's contraceptive use of having family planning users in her networks when the analytic model, OLS, assumes that unobserved characteristics of the respondents such as her preferences for certain kinds of network partners do not influence both the selection of these partners and her contraceptive use (the data will be described later). Suppose, however, that the analyst's theory or intuition leads to the expectation that women who are likely to choose to use contraception because they dislike the noise that children make are likely to select as network partners other women who are likely to use contraception because they dislike noisy children, and that our survey data do not include a measure of "likes tranquility". In Column 2 we thus control for unobserved fixed factors that determine contraceptive use and the selection of network partners, using longitudinal data from all three rounds of the Nyanza survey, as in Column 1. Going even further, in Column 3 we use a model that combines individual fixed effects and instrumental variables that take the change in social network measures over time into account (for details of the models, see Kohler, Behrman and Watkins 2002). The estimated effects of network partners on the respondent's contraceptive use are clearly different in magnitude. Analysts convinced by their theory or intuitions are likely to conclude that the OLS estimates are upwardly biased. Such a conclusion may not, however, be warranted.

Using models with strong assumptions risks the possibility that the cure may be worse than the disease. If the models are of "doubtful validity, the 'corrections' they suggest may make matters worse rather than better" (Winship and Morgan 1999; see also Bound, Jaeger and Baker 1995). A similar critique can be made of the instrumental variable techniques: they require important assumptions that are "sometimes unbelievable" and "often unreasonable" (Winship and Morgan 1999). More critically for the interpretation of the results, the validity of models that control for selectivity is difficult to evaluate, since the basic assumptions and predictions of these models are rarely challenged or tested (Freedman 1999) (Note 2).

If differences among estimates do not provide persuasive evidence for the choice of one model over another, how do we choose? Some critics suggest that judgement is central to the choice of a statistical model (Freedman 1999). Charles Manski and Daniel Ragin (1998), in an article comparing treatment effects with alternative models of how judges sentence convicted juvenile offenders, suggest how our judgement can be more informed:

"Our work clearly points to the need for further research analyzing how judges actually make sentencing decisions. Although the outcome optimization model and the skimming model differ from one another, they share critical simplifying assumptions--namely, that judges are concerned only with recidivism and that these judges have rational expectations about recidivism. It may be that these

shared assumptions are unrealistic. To make further progress in the analysis of recidivism using observational data, we need to better understand how sentencing decisions are actually made” (Manski and Ragin 1998:132).

In this paper, we follow the spirit of this recommendation. Our familiarity with the specific context leads us to ask about two motivations for the selection of network partners, one based on similar attributes (homophily), the other on a strategic search for information from those with relevant experience. Our goal is to illustrate strategies that may assist analysts in making judgements in other cases.

In Section 2, we describe the data on informal social networks that we use and evaluate its quality. We also describe the local context, since our detailed knowledge of this context informs the analyses we conduct and their interpretation. In Section 3, we consider homophily. Although the composition of the social networks in rural Kenya reflects the limited diversity in the pool from which network partners could be selected, we also consider whether preferences for network partners who are similar on a variety of dimensions further reduce network diversity. To do this, we use qualitative data to provide insights into women’s preferences and survey data to draw inferences from comparisons between male and female networks and between relatives and non-relatives. In Section 4, we focus on the extent to which network partners appear to be strategically selected precisely because they are believed to use family planning (or not). We use qualitative data that provides information on motivations that are not observable in the survey data in order to assess the magnitude of strategic selectivity (Note 3). We then use survey data to compare the composition and structure of conversational networks on four quite different topics that in this context are perceived quite differently. If the networks are different across topics, this would suggest that networks are chosen specifically to search for information about family planning or to provide support for a difficult choice about its use, and vice versa. Our conclusions summarize our findings and consider their implications for data collection and the selection of statistical models.

2. Data and context

2.1. Data

The data consist of qualitative interviews, focus groups and a longitudinal household survey, as well as less systematic observations by the research team (Note 4). The data were collected in rural villages in four sublocations in South Nyanza District, Nyanza Province, in the western part of Kenya and near Lake Victoria (Note 5). The first phase,

in June-July 1994, consisted of semi-structured interviews with 40 women and 40 men of reproductive age, as well as nine focus groups with women. Although the number of interviews and focus groups is small, they are valuable because they provide access to some of the motivations for selecting network partners that are not available in the household survey data. The respondents for the semi-structured interviews were selected systematically, such that they would be evenly distributed around the main roads or paths where much social interaction takes place (Note 6). Respondents and focus group participants were asked about their informal conversations on family size and family planning, on wealth flows (support of the elderly) and on *rariu*, a local female reproductive health ailment. The semi-structured interviews and focus groups were conducted in the local language, Luo, and were taped, transcribed and translated into English. They were then coded and analyzed using Ethnograph (Note 7).

The first wave of the longitudinal household survey, Kenya 1, was conducted in December 1994 and January 1995, with a sample of 926 married women and 859 husbands. The sampling frame was a list of villages in each of the four sites. From this list, enough villages were randomly selected to provide the desired sample size, consisting of all married women of reproductive age who were present, and their husbands if they were living at home (there is much male temporary migration for work, and sometimes wives accompany their husband to the city). Two years later the second wave of the survey, Kenya 2, re-interviewed these women and men. (A third wave was conducted in 2000, but we do not use this here). Some respondents were present at the time of both surveys, some only at one (Note 8). Although below we rely primarily Kenya 1, we use Kenya 2 for some questions that were only asked then (Note 9). Since for simplicity most of the analyses that follow use the women's data, we focus on that here.

In the two waves of the survey interviewers asked about partners in conversations on four different topics. In Kenya 1 half the women were asked with whom they chatted about wealth flows and half about the local illness *rariu*, and all were asked with whom they chatted about family planning; in Kenya 2 all women were asked about family planning networks and AIDS networks. The survey questions used the word "chat" to indicate that we were interested not in lectures or counseling sessions at the clinics, but rather in informal interactions. Respondents were first asked with how many others they had chatted about the specific topic; the number of network partners ranged from 0 to 21 (Note 10). Then detailed questions were asked about the characteristics of a maximum of four of these partners. Respondents were asked to place the network partner on a social scale according to economic status, educational level, and age, all of which are often found to be associated with the use of family planning. Network partners were also identified along several dimensions used in network studies to characterize the links between the respondent and her network

partners, such as the nature of the relationship (e.g. confidant, friend, or acquaintance), and the type of relationship (e.g. kin or non-kin). This format produced sets of respondent, or ego-centered, networks with one to four network partners (Note 11). There are a total of 1,987 network partners in Kenya 1 and 1,869 network partners in Kenya 2.

The household survey included socio-demographic questions taken from the 1993 Kenya Demographic and Health Survey (KDHS) to permit evaluation of the quality of the data by comparison with the KDHS. The comparisons in Table 2 are with the KDHS data for rural South Nyanza District and with the larger KDHS sample for rural Nyanza Province.

Table 2: *Comparison of the Kenya 1 Data to Ever Married Sample from 1993 KDHS Data for Rural South Nyanza. All estimates are percentages unless otherwise stated.*

Characteristics	Rural South Nyanza, Kenya 1	1993 DHS
<u>General Socioeconomic Characteristics</u>		
Mean Age	28.9	29.7
<u>Level of Education</u>		
Primary	67.7	66.1
Secondary	13.6	9.5
University	0.0	0.5
<u>Roof Material of Dwelling</u>		
Thatch	81.2	68.8
Metal	18.8	30.3
Household Owns a Bicycle	27.2 ^a	31.7
<u>Fertility and Family Planning</u>		
Mean Number of Children	4.5	4.5
Mean Number of Children who Died	1.2	1.0
Ever Used Family planning	22.7	33.9
Currently Using Family planning	11.9	11.8
N (Respondents)	926	256

Notes:

NB: Some categories do not sum to 100% percent because (a) 'other' and 'don't know' categories are excluded and (b) figures are rounded.

^a This question was only asked in Kenya 2 (1996-1997).

The close comparison with the KDHS in Table 2 is reassuring, since the DHS surveys are considered to provide the gold standard for survey data in developing countries. Interviewer effects were also calculated for these questions for the KDICP and the KDHS, and were similar in magnitude.

The use of family planning may be over-reported in these data as well as the KDHS, an issue that is important in assessing whether network partners are strategically selected with respect to their use of family planning. In rural Nyanza, family planning is associated with the Kenyan government, which has supported a relatively aggressive family planning program since the early 1980s (Watkins and Hodgson 1998; Ajayi and Kekevole 1998). Moreover, family planning is associated with *wazungu* (white foreigners), who are believed to have invented family planning and to encourage its use in Kenya (Watkins 2001). Despite the interviewer's introduction of their questions as university research, many respondents believed that our presence would "bring development to the community", either through the government or foreign donors, or from the team's own pockets. These expectations are likely to lead to over-reporting of family planning use by respondents, in order to present themselves as worthy of being rewarded (see also Christiaensen, Hoddinott and Bergeron n.d., for Mali). This may partly explain the even higher reports of ever-use of family planning in the KDHS, which is an explicit collaboration with the government, than in our survey.

2.2. Context

In addition to an active family planning program, other relevant features of the context are the settings in which social interaction takes place and the constraints on the formation of social networks. One of the sites is on an island in Lake Victoria, yet neither it nor the other sites are socially isolated. Across all sites, one half of the respondents reported owning a radio and 37% of women and 61% of men spent six months or more after marriage outside their village, usually in a city. Women leave their villages to visit their natal relatives elsewhere in Nyanza or a husband who has migrated for work (Note 12). Funerals, which are frequent due to the high level of AIDS, bring relatives and friends to the area for a few days. Nonetheless, because transportation is irregular and expensive and telephones very few, frequent interaction is largely restricted to members of the local community.

Within these areas, there is relatively little economic and social stratification compared to developed countries where most network studies have been done (see, for example, a recent review of homophily in networks by McPherson et al 2001, in which all citations appear to be to studies in developed countries). Most residents engage primarily in subsistence agriculture, supplemented by small-scale business, some wage

labor, and occasional remittances from urban relatives (Note 13). Almost all lived in mud huts with thatched roofs, although some have more costly and prestigious metal roofs. Those who complete secondary school look for work in the cities; those who do not find jobs and return to the rural areas are engaged in much the same activities as those who have never been to school. Only 14% of the 926 women interviewed in Kenya 1 had been to school beyond the primary grades.

Most of the respondents are Luo, a group characterized by exogamous marriage and patrilocal residence. Norms about appropriate network partners abound. Luo folk tales warn married couples not to involve their friends in family affairs (Ocholla-Ayayo 1976) and admonish women not to discuss domestic affairs in public (Oboler 1985). Community norms discourage informal interaction between the sexes (Ayodo 1994). Some topics, including family planning, are not to be discussed with unmarried women, and recently married women with no children told interviewers that they would be considered out of line if they indicated an interest in family planning.

3. Homogeneity and homophily: Do birds of a feather flock together?

That individuals form networks that are more homogeneous than had the networks been formed randomly might be due to constraints on the pool available for social interaction: poor means of communication with those outside the area, a relative lack of economic and social stratification, and normative rules may restrict opportunities for building diverse networks. Homogeneity in networks might also, however, be the result of preferences. The principal of homophily, in which people prefer to make connections with others who are most like them, is a persistent feature of social structure (Berelson et al 1954; Katz and Lazarsfeld 1955; Fischer 1982; Blau 1994; Mark 1998; Marsden 1988; Friedkin, 1993; Erikson 1996; McPherson et al 2001). This section examines homogeneity and homophily. It begins with a description of the contexts and characteristics of informal conversations about family planning and the characteristics of those the network partners, and then considers constraints and preferences.

3.1. Contexts and characteristics of informal conversations about family planning

The overall picture presented by both the qualitative data and the survey data is of conversations that occur frequently and casually, as women go about their daily activities on foot. The compounds are close to each other--from almost every compound we visited we could easily see neighboring compounds--and the area is criss-crossed by

paths, making it likely that women could easily meet by happenstance or by arrangement. In the semi-structured interviews, women described conversations that often occurred when more than two women were together: someone would mention that she thinks so-and-so is using family planning, provoking a discussion of the advantages and disadvantages of having fewer children or of using modern methods, as in the following excerpt from a focus group of women ages 20 to 29 (FGW identifies the speakers) (Note 14).

FGW: Sometimes we talk when going to fetch water or going to the market.

FGW: The one who uses the pills is the one who will tell you how that pill is affecting her.

FGW: Sometimes we've gone for a meeting, we talk before the meeting starts.

FGW: Some people say pills are not good with you and it makes you so thin. Even thinner than what you were before.

FGW: Some people also say you would give birth to a child with lots of disabilities.

...

FGW: There was a woman who's been using injections and she's been having monthly periods twice in a month. This woman really had trouble until she changed to pills. Now she took the medicine and by the time she was stopping to use pills she got pregnant and now she's pregnant.

Moderator: How did you get to know about this?

FGW: We were sitting down and she was telling how that thing happened to her.

This excerpt does not suggest a strategic search for information or opinions. Rather, it appears that even if a woman were not interested in hearing about family planning, she would be present when others were talking. In the semi-structured interviews and focus group, women said they typically chat about family planning as they are walking or sitting together, with one woman telling the others about her own experience or gossiping about the experiences of others she has heard about.

This picture of informal conversations among several women is supported by data from the household survey, shown in Table 3. Few of the conversations occurred in the

more formal settings of church and clinic, and the qualitative data suggest that even these were likely to be informal, for example while the women were waiting for a meeting to start or for the clinic nurse. Just over half of the conversations occurred when more than just the respondent and the network partner were present. Conversations that involved only two people, or those that occurred “at home” or “visiting” might have been more private, but need not have been, as most compounds are home to several women. The conversations also seem to be relatively frequent, perhaps because modern family planning is new: it is hard to imagine that nearly half the women living in contexts where the use of family planning is taken for granted would talk about it within the last month. Both the qualitative data and the survey data, then, lead to a view of family planning as “in the air”, such that even a young married woman for whom family limitation is not yet salient would learn much about the new methods of family planning.

Table 3: *Characteristics of Family Planning Conversations with Network Partners.*

Characteristics	Percent of network partners	Characteristics	Percent of network partners
Location of conversation		When conversation took place (K1)	
At home	27	This week	23
Visiting	31	This month	25
Market	8	This year	38
Walking	5	More than a year ago	14
Getting water	5	How many others were present? (K2)	
other (<i>posho</i> mill, <i>shamba</i> , clinic, women’s group meetings, church, Chief’s meeting and miscellaneous)	23	2 people	49
		3 to 4 people	31
		5 or more people	19

Note: Data are from Kenya 1 (N=1,987 network partners) and Kenya 2 (N=1,869 network partners). Numbers may not add up to 100 due to rounding.

3.2. Characteristics of network partners

The context and circumstances of the conversations described above appear to shape the characteristics of network partners, seen in Table 4.

Table 4: *Characteristics of the Family Planning Network Partners and the Relationship Between the Respondent and Her Family Planning Network Partners.*

Characteristics	Percent of network partners	Characteristics	Percent of network partners
Residence of network partner (K1)		Frequency of contact (K1)	
Same compound	10	Daily	34
Same village	37	Weekly	21
Same sublocation	12	Monthly	23
Elsewhere in Nyanza Province	24	Yearly	22
Outside Nyanza Province	16	Strength of relationship (K1)	
Don't know	1	Confidant	49
Relationship to network partner (K1)		Friend	39
Sister-in-law/ co-wife	32	Acquaintance	12
Friend	21	Length of relationship (K2)	
Mother-in-law	6	1 to 2 months	3
Mother	4	Less than 1 year	5
Sister	13	1 to 5 years	24
Brother/ Brother-in-law	7	More than 5 years	68
Miscellaneous ^a	17		

Notes:

Data are from Kenya 1 (N=1,987 network partners) and Kenya 2 (N=1,869 network partners). Numbers may not add up to 100 due to rounding.

^a Category includes fathers, other in-laws, aunts, uncles, grand-parents, co-workers, and others.

As the table indicates, most network partners are geographically and socially close: nearly one half of the respondents live in the same compound or same village and at least 62% are close relatives (Note 15). For one site the distances between compounds was mapped: the median number of compounds between respondents and network partners is three, the mean five and a half (White 1999). Across all sites, if a network partner does not live nearby, she is usually a maternal relative, the consequence exogamous marriage patterns and patrilocal residence (Warriner 1999). The majority of respondents and their network partners have known each other at least five years; over one half of the relationships of less than five years are between women who moved to the area recently (Warriner 1999). The social closeness is heightened by the density of the networks: most network partners are confidantes or friends of the other network partners as well as of the respondent (Note 16). In other analyses of these data, some of these characteristics have been found to be related to the use of family planning (Buehler and Kohler 2003; Behrman et al 2002; Kohler et al 2001). Although it might be reasonable to expect that those considering family planning would turn to “opinion

leaders” defined by their relatively high formal status, that is not the case here. In Kenya 1, a question, "Have you ever talked with a _____" listed these categories of higher status. Only 5 % had chatted with the chief's wife about family planning, 7 % had talked with a clergyman's wife, and 10 % had talked with a teacher or a teacher's wife; some respondents explained to the interviewers that they would only talk with a woman in these categories if she were a friend. The selection of network partners is also guided by normative rules. There are few male network partners, and few women talk with their mother-in-law about family planning-- both because of norms that limit topics that can be discussed with a mother-in-law and because women perceive that mothers-in-law want as many grandchildren as possible (Rutenberg and Watkins 1997).

Constraints and preferences

The homogeneity of networks described above is consistent with the relatively limited diversity of the available pool of network partners. It is also consistent with a high degree of homophily. To evaluate the extent to which homophily appears to operate, we compare men's networks to women's, and relatives and non-relatives. In South Nyanza the available pool of potential network partners is less constrained for men than for women; thus, men have more scope for personal preferences in selecting non-relatives than relatives for chats about family planning (see also Marsden 1987; Moore 1990). A finding that men's networks are also homophilous or that the characteristics of relatives and non-relatives in the networks are similar would suggest the operation of preferences, over and above the constraints imposed by the context.

The male population in rural Nyanza is more diverse than the female. Men have more education than women, and because of the highly gendered job market men are more likely to earn a regular salary than women (Watkins, Rutenberg and Wilkinson 1997). In addition, men are more likely to have lived outside Nyanza for six months or more—often in the large cities of Nairobi and Mombasa, where there is much more socioeconomic diversity. Table 5 compares the networks of men and women using the mean of network partner characteristics for all respondents. For example, a respondent with only two network partners, one who is older than she is and one who is younger, will have 50 % older and 50% younger network partners. For the sample as a whole (as shown in Table 5), 55% of the average respondent's network partners were characterized by the respondent as older than herself. Network partner characteristics need not be symmetric: some network partners live outside of the sampled areas and some women had more than the four network partners about whom specific questions were asked.

Table 5: Means for All Respondents of Family Planning Network Partner Characteristics Relative to Respondent by Gender

Characteristics	Network Partners		Characteristics	Network Partners	
	Female	Male		Female	Male
Relative Age			Relative wealth		
Older	.51	.39*	Number of durable goods ^a		
Same	.28	.27	Less	.35	.27*
Younger	.26	.35*	Same	.23	.25
			More	.42	.45
Education					
Secondary	.25	.40*	Relative location		
Primary	.54	.50	Same compound	.10	.03
None	.09	.05	same village	.48	.47
Relative education			Same sublocation	.08	.11
More educated	.13	.17	Other location	.06	.09
Same amount	.47	.53*	Nyanza	.20	.20
Less educated	.28	.24	Nair/Mom/Other	.08	.08
N (Respondents)	602	421		602	421

Notes:

* p < .05

^a Durable goods: radio, bike, sofa, lamp, plough.

Data are from Kenya 2. Figures do not add up to 100. All standard errors are all less than 0.02.

Despite the greater opportunities for men to select diverse network partners, the similarities between the male and female networks for both topics of conversation are notable. For example, although more men than women (61% vs 37%) have lived in Nairobi, Mombasa or other cities outside Nyanza, the proportion of male and female network partners reporting that they talked about family planning with geographically distant network partners is identical. With the exception of age—the tendency to select older network partners is more pronounced for women—there is little difference. We interpret this comparison of male and female networks as suggesting that although there are indeed constraints placed on the choice of networks by socioeconomic and cultural aspects of the context, preferences for homophily also operate.

There is less scope for choice in the selection of relatives than unrelated friends and neighbors. Women of necessity interact frequently and over a long duration with relatives in their compound and in their husband's village, which in principle is composed of male descendants of a single ancestor (e.g. a grandfather or great-grandfather) and their wives. Distinctions between relatives and others are clearly made. For example, trust appears to be greater among relatives: 51% of males and

42 % of females had lent money to a relative, compared to 38 % and 26 %, respectively, to non-relatives. To examine whether homophily is greater among women’s network partners who are relatives than friends, we begin with objective attributes, since most studies deal with objective homophily-heterophily (e.g. Rogers and Bhowmik 1971). Yule’s Q, which ranges from -1 to +1, is used to compare the degree of association between the respondent’s socioeconomic characteristics and the network partners’ socioeconomic characteristics, as reported by the respondent (Note 17). In Table 6, positive values show associations between people who are similar to one another with respect to ownership of durable goods and education, negative values indicate that members of a group become network partners with those who do not belong to the group, and values close to zero show a random association.

Table 6: *Measure of Isolation, Yule’s Q, for Network Partners Who Are Relatives and Non-Relatives of Respondents.*

Variables	Relatives	Non-Relatives
Metal roof	.04	.24
Radio	.11	.35
Bike	.24	.33
Lantern	.03	.15
Plough	.44	.58
Education	-.08	-.55
N (Network Partners)	1,500	369

Note:

Data are network partners from Kenya 2. The isolation measure, Yule’s Q, is based on a 2 by 2 contingency table and is calculated as $(ad-bc)/(ad+bc)$, where the upper left cell (a) contains the number of network partners and respondents who, e.g., both own the household good, the lower right cell (d) contains the cases where neither the network partner nor the respondent owns the good and the diagonal cells (b and c) contain the number of cases where either the respondent or the network partner own the good (see also endnote 12).

As the table indicates, there is a high level of clustering along socioeconomic dimensions for both relatives and non-relatives, further supporting the arguments for homophily. The values are greater for network partners who are non-relatives, suggesting that when choice is greater, similar socioeconomic characteristics play a greater role in the selection of network partners (Note 18).

The principle of homophily is no doubt broader than is captured by observed measures such as ownership of durable goods. For example, the use of the phrase “a woman like you” in the focus group excerpt below of women ages 20 to 29 as well as in many of the semi-structured interviews suggests that the women selected as network partners are similar along attributes such as their chattiness or their ability to keep a

secret, attributes that are unlikely to be observed in a data set. When the moderator asked who women talk with about their problems, the focus group responded:

FGW: It is only your husband you can talk to.

FGW: Aah. You can also tell a woman like you. When you are really annoyed you can go and tell a woman like you.

Moderator: What about you, Jennifer?

FGW: When I am annoyed sometimes I do not tell my co-wife [the Luo word for co-wife and sister-in-law is the same]. I go and tell my friend who is a woman like me and who will also tell me about their problems.....

FGW: These stories, we normally tell about them with young women just like us.

FGW: And more so to those who keep it a secret, not the talkative ones.

In summary, geographic, socioeconomic and cultural constraints play a role in the construction of networks that are more homogeneous than random chance would predict, and these constraints are undoubtedly more important in rural Nyanza than in more socially and economically differentiated societies with telephones and better roads. Here as elsewhere, homogeneity is heightened by personal preferences (Fischer 1982). One sort of preference, and one not measured in the survey data, is indicated by the woman who said that they talked with "those who keep it a secret, not the talkative ones", a characteristic particularly important if a woman is considering using, or using, family planning without her husband's knowledge, as a substantial proportion do (Watkins, Rutenberg and Wilkinson 1997). The most general preference, however, is for "a woman like myself."

4. Strategic selectivity on family planning

Although features of the local context constrain the pool of available network partners to a rather homogeneous group, that pool is differentiated by the use of family planning. Uncertainty about desirable family sizes and the use of family planning might lead women to turn to others to evaluate the information from the clinics and the media, to learn from the experiences of others, or to assess the degree of social support for, or opposition to, smaller families or the use of family planning. Thus, it is a reasonable

assumption that women already interested in using family planning to space or limit their childbearing would select from the available pool those who have had personal experience with modern methods, or who are believed to approve of them, for example to confirm their intention or to provide information about methods. The converse is also reasonable: that women who are not interested in spacing or limiting, or who have grave doubts about the methods, would turn to others who they believe share their disinterest or disapproval.

In what follows, we first present the most direct evidence that network partners are selected with respect to family planning. We then consider whether this is modified by other analyses. We begin with qualitative data, distinguishing between conversations that might be strategically motivated by a search for information or social support regarding family planning use and those that do not appear to be strategic. We then turn to what we think is a particularly useful perspective on selectivity with respect to family planning: the extent to which women appear to use different criteria for selecting conversational partners with whom to chat about quite different topics of conversation.

4.1. Evidence for strategic selectivity with respect to family planning

The strongest evidence to support an assumption that friends, relatives or neighbors are deliberately selected with respect to family planning comes from the disproportionate use of family planning by network partners. Although 31 % of respondents in Kenya 2 reported that they themselves used family planning, they reported that 56 % of their network partners used; a similar discrepancy was found in an analysis of the networks of urban Cameroonian women (Valente et al 1997) (Note 19). This certainly suggests that respondents select those who are believed to use family planning for conversations about family planning. This conclusion is supported by the higher levels of family planning network partners than AIDS network partners who were said to approve of family planning.

It is likely, however, that respondents are not fully accurate in reporting their network partners characteristics, as found in other studies (e.g. Lauman 1973; Bernard and Killworth 1977), and biases in the reporting of family planning use due to the promotion of family planning by the government and donors might account for at least some of these discrepancies. In order to assess the accuracy of respondents' reports about their network partners, in one site we linked the questionnaires of the women interviewed in Kenya 1 to those of their network partners in the sample. The reports of the respondent and her network partners agree most closely on the things that are clearly observable (White and Watkins 2000). There is agreement on possession of household goods such as a sofa set and a radio (81% to 78 % agreement). These figures

compare well with the 87% agreement on whether the alter's home has a metal roof, which was observed by the interviewer. The high level of agreement on possessions is matched by a high level of agreement on educational level (78 % agreement). In contrast, agreement on family planning use is only 55% to 57 %, due primarily to respondents' overestimates of family planning use by their network partners (for more discussion, see White and Watkins 2000). It is interesting to compare agreement between respondent and network partner on family planning use with the agreement between monogamously married husbands and wives, who should be in close agreement. In Kenya 1, only 77.6% of monogamously married husbands and wives in all four sites agreed that they were using, or not using, family planning, a figure similar to the agreement between monogamously married husbands and wives in a similar study in rural Malawi; analysis of husband-wife discrepancies in the Malawi and Kenya DHS are of similar magnitude (Reynar 1998; Miller et al 2001). The discrepancies between the respondents' reports and those of her network partners do not invalidate the possibility that the respondents' beliefs, however inaccurate, motivate her choice of network partners. Even if respondents are not accurate in reporting the family planning use of their network partners, a respondent's belief that a woman is using family planning may be relevant for her choice of this person as a network partner (Note 20). We thus turn to other evidence.

4.2. Qualitative evidence

In order to assess the extent of strategic selectivity by perceived use of family planning in the choice of family planning network partners, all conversations about family planning reported by the women in the 40 semi-structured interviews were analyzed and sorted into three categories. The first category includes conversations in which the topic of family planning came up incidentally in the flow of conversation. In the second are conversations about family planning in which a network partner was deliberately sought out for information (e.g. because they had some experience with family planning, or knew something about it). In the third category are conversations that could not be categorized. Over half (56%) of the respondents talked to others about family planning in the course of general conversations in which the topic of family planning was brought up by others who described their own experiences or gossiped about the experiences of others. About one third (36%) of the conversations could be labeled strategic and 8% are unclassifiable. Thus, the more detailed information from qualitative interviews suggests that strategic selectivity does not characterize the majority of conversations.

4.3. Comparison of network characteristics across different topics of conversation

The most compelling evidence that the degree of strategic selectivity is modest is a comparison of observed network characteristics across different topics of conversation. The two waves of the survey provide data on four conversational networks; family planning, a local reproductive illness (*rariu*), AIDS and wealth flows; in each wave women were asked about two conversational networks. Approximately 15% of the named family planning network partners in the first network are also named in the second, so for these partners there is no selectivity by topic (Note 21). For the other network partners-- a clear majority-- there may be strategic selectivity.

Although three of the topics concern issues that are perceived as similar by western advocates of programs to improve women's reproductive health (Hodgson and Watkins 1997), familiarity with this context provides reasons to expect that the different topics would elicit different network partners. For example, women go to family planning clinics for family planning methods, but those who have *rariu* are advised by their friends to go not to the clinics but to women who know the Luo herbs (Luke et al 2001; Luke 2000). These differences have implications for the characteristics of the network partners. Family planning network partners should be better educated, wealthier, and less likely to live in the same compound or village than the network partners with whom the respondents talked about *rariu*. In contrast to *rariu*, both family planning and AIDS are locally associated with modern life and perceived to come from outside the local area: family planning, however, is seen as a family matter, whereas AIDS is associated with promiscuity (Watkins and Zulu 1997).

The analysis in Table 7 is limited to women who were asked about a total of three networks: i.e. the woman was asked about family planning and either wealth flows or *rariu* in Kenya 1, and AIDS as well as family planning in Kenya 2. The analysis uses characteristics of the network partners that would be likely to be guiding women's choice of network partners if they were selecting them for characteristics associated with family planning. In the qualitative data women described family planning users as better educated and wealthier, and analyses of the household survey data show that these are among the characteristics of family planning users (Behrman, Kohler and Watkins 2002). Table 7 also considers characteristics of links among network members that network theorists expect to be associated with the diffusion of an innovation (Granovetter 1973). The proportion of weak ties, here defined as acquaintances, is relevant if respondents are searching for new information about family planning, as are those who are seen infrequently, and who live outside the area.

Table 7: Means for All Respondents of Family Planning Network Partner Characteristics and Three Other Topics of Conversation.

	I Family PI (K1)	II Rariu (K1)	III AIDS (K2)	IV Wealth (K1)
Variable				
Female	.96	.98	.74*	.88*
Age relative to respondent				
Older	.55	.59	.53	.63*
Same age	.21	.16	.21	.17
Younger	.22	.24	.26	.20
Education relative to respondent				
More education	.24	.17*	.31*	.20
Same education	.50	.45	.47	.53
Less education	.17	.26*	.14	.20
Economic status relative to respondent				
Wealthier	.43	.39	.42 ^a	.18*
Equally well off	.36	.32	.23*	.38
Less well off alters	.13	.21*	.35*	.39*
Frequency of contact				
Daily contact	.36	.37	.38	.44*
Weekly contact	.21	.22	.27*	.20
Monthly contact	.21	.19	.14	.19
Less than monthly contact	.25	.21	.21	.17
Strength of ties				
Confidants	.50	.50	.40	.58
Friends	.37	.37	.44*	.35
Acquaintances	.13	.13	.16	.07*
Dispersion of network partners				
Same compound	.12	.15	.09	.15
Same village	.36	.37	.44	.41*
Same sublocation	.13	.12	.13	.10
In Nyanza	.26	.23	.24	.26
Elsewhere	.11	.11	.07	.08
N (Respondents)	330	379	201	354

Notes:

^a The wealth variables for AIDS network partners was based on an objective comparison of durable goods (radio, bike, sofa, lamp, and plough) since respondents were not asked to assess the relative wealth of their network partners in Kenya 2.

* p < 0.05; t-test compares values for each network to the family planning network.

Standard errors are all less than 0.02.

Examining first the characteristics of the family planning network partners in Column I, respondents do appear to be seeking out network partners whose observable characteristics are associated with family planning use. When age, education and wealth are not the same as the respondents, the network partners are likely to be older, better educated and wealthier, supporting the view that family planning network partners are selective with respect to characteristics that women in these areas associate with the use of modern family planning. The comparison with the other networks is revealing, however. A particularly pertinent comparison is between family planning and *rariu*: both are considered largely women's issues, but the former is an innovation in these areas whereas the latter is not. Family planning and *rariu* network partners are similar on most characteristics; the striking difference is in terms of education (more for the family planning network partners, less for the *rariu* partners), which is what we would expect if strategic selectivity were operative. To the extent that there are statistically significant differences, these distinguish the AIDS and wealth flows networks from the other two. These analyses were repeated for males for the two networks in Kenya 2, family planning and AIDS (not shown). Again, the network characteristics of males and females are similar, and again the topic of conversation does not distinguish the characteristics of the networks (Note 22).

5. Conclusions

Writing about bias due to selectivity, Ross Stolzenberg and David Relles (1997) say that “researchers often have little more than their intuition to guide them” and they conclude that “intuition, informed judgement, simulation, experimentation and statistical methods are necessary to understand and manage inevitable problems in data.” We suggest that intuition and judgement can be informed by simple analytic strategies that do not require the assumptions demanded by more complex analytic techniques.

We demonstrated using an example where selectivity seemed *a priori* likely: the effect of social networks on contraceptive use. It is not difficult to tell a plausible story in which respondents choose network partners according to attributes that might also influence the respondents' use of contraception, thus raising questions about the direction of causality; indeed, a fertile imagination can no doubt produce a number of plausible stories. We looked for evidence of two types of selectivity. One is homophily. Given the ubiquitous finding in the network literature that networks are homophilous, we can expect to find it in rural Kenya. The second is more specific--that in a context such as rural Kenya where the use of modern family planning to control fertility is an innovation that has provoked considerable uncertainty, we can

expect women to strategically select as network partners those with experience of contraception in order to learn about their experiences, to evaluate the costs and benefits of contraceptive use or limiting family size, or to assess the social acceptability of contraception or small families. An analyst who then proceeded directly to model the data with a technique that controls for selectivity might be convinced by congruence between the stories and the results. In our initial comparison of statistical techniques, OLS produced distinctly higher estimates of the impact of social networks on contraceptive use than did models that controlled for selectivity based on unobserved characteristics. Yet surely it would be easier to justify the latter models as preferred were there other evidence of selectivity than a difference in estimates.

We used a variety of strategies to ask whether the selection of network partners appeared to be random or whether there was evidence of homophily on the basis of similar attributes and of strategic selection on the basis of the use of family planning. Some of these strategies exploited the survey data as if it were cross-sectional, others exploited qualitative data (semi-structured interviews and focus groups that had been systematically analyzed), and all took advantage of our familiarity with the specific research context. We concluded that in these data, there is indeed evidence of selectivity, primarily on the basis of homophily but with some strategic selectivity. Whether there is enough selectivity to justify the use of analytic techniques to control for it is a different question, one that we do not attempt to answer but that we think deserves attention by methodologists. For example: Would an analyst be justified in ignoring selectivity if it appeared that there was only “a little” selectivity? Could there be general rules for defining what is “a little” vs “a lot”? And what if the costs of a more complex technique were great, either in the expense of data collection or in credibility?

We thus conclude that along with continued efforts to improve statistical methods that control for selectivity biases and the collection of the data required to use these models in order to make causal inferences from observational data, careful albeit simple description of the survey data, the systematic collection and analysis of qualitative data, and a familiarity of the specific context in which all the data are collected should precede and inform judgements about the appropriate choice of analytic models.

6. Acknowledgements

An earlier version of this paper was presented at the annual meeting of the Population Association of America, New York, April 24-27, 1999, in a session on networks organized by Martina Morris. We are grateful for conversations about causal analysis with Paul Allison, Charles Manski and Herb Smith and with two of our collaborators on

the analysis of social networks in Kenya, Jere Behrman and Hans-Peter Kohler. We also received useful comments from Andrew Foster, David Freedman, Christopher Winship and from reviewers. The data were collected with financial support from grants from USAID's Evaluation Project to Watkins and Naomi Rutenberg, from the Rockefeller Foundation to Watkins and Eliya Zulu, and from NICHD to Watkins.

Notes

1. In this paper, we use the term “selectivity” as it is typically used in sociology. In the economics literature, the term is used to refer to the fact that some variables are observed only for a subset of a sample; what sociologists call selectivity is called endogeneity by economists.
2. Some have argued that it is not necessary for assumptions to be reasonable, only that the models are successful. Freedman, however, notes that “If regression models were generally successful in making causal inferences from associational data, that would be compelling evidence. In my experience, however, those who repeat [Milton] Friedman’s argument are seldom willing to engage in detailed discussions of the track record. Their reluctance is understandable” (Freedman 1999, 255).
3. For an excellent use of focus group data to determine the direction of causality between women’s work and childcare, see Short et al 2002.
4. The data are described in more detail in Watkins et al, this volume, and at www.pop.upenn.edu/networks. All the data are available from the senior author, including the field notes of Watkins, who was present throughout the data collection.
5. A sublocation is a small administrative unit; it was possible to walk (briskly) from one end of a sublocation to the other in about two hours. South Nyanza District was subsequently divided into three districts.
6. The focus group participants were meant to be selected randomly, but chiefs who were asked to call together all women between 20 and 40 with no more than primary education instead selected their brightest and best.
7. Initial coding was done independently by Watkins and Naomi Rutenberg; further coding was done by Warriner and Nancy Luke. In addition, two students, Caryl Feldacker (an undergraduate honors student at the University of Pennsylvania) and Sarah Varle, a student at the London School of Hygiene and Tropical Medicine, independently coded some of the same categories for their own research projects; there was considerable consistency.
8. The implications of attrition from the sample are analyzed in Alderman et al., 2000.
9. A third wave of data was collected in 2000, but these data are not used here.

10. Notes written on the questionnaires by interviewers as well as the qualitative data provide explanations for those who had no network partners. For family planning, this was likely to be a young woman who had no children.
11. The comparisons based on the survey data face several limitations. First, they only allow us to compare measurements of the networks and do not indicate the relative importance of the various discussions for individual women or provide information on the content of the conversations. In addition, we follow the practice common in ego-centered network analyses of asking specific questions about only a subset of network partners. Typically, the respondent is asked to report the number of people that form a particular network but is then asked detailed information about the first four or five respondents she names. Since it is unlikely that the nominating process is random, it is likely that some biases are embedded in the data. Kohler (1998) found bias in the estimating of density for truncated ego-centered network data. A second weakness of network surveys is that the measurements are based on recall and, as with any data reconstructed from memory, potential biases exist. Brewer and Webster (1999) and Brewer (2000) found that respondents tend to forget about 20% of their personal network partners when asked to list them from memory. These biases may affect the measurements of the composition of the networks, which therefore may not accurately represent the "true" pattern of discussions. Although selectivity in remembering and reporting network partners is potentially troublesome if biases and omissions in reporting are associated with selectivity toward family planning, we do not address it directly here. Consequently, measurements and results must be interpreted with some caution.
12. An analysis by Alan Ferguson of our data shows that most of the women were born in other areas of South Nyanza (unpublished).
13. When asked how they earned money, 40% of the women in the household survey sample said they "do nothing", another 40% said they had a small business (e. g. buying some bananas at a more distant market and selling them to neighbors), and 13% said they sell things from their *shamba*, a small farm plot, growing primarily maize but also crops such as millet and cassava.
14. FGW is a focus group participant. Unfortunately, in this exploratory and preliminary stage of the research we did not appreciate the importance of identifying each participant and thus cannot distinguish women who spoke more than once.
15. Because Luos are polygamous, the miscellaneous category includes a wide range of relatives, such as the respondent's husband's mother's co-wives, "cousin-sisters", and people who are termed kin out of respect or friendship.

16. A dense network is one in which all the network partners know each other; a sparse network is one in which none of the network partners know each other (they are connected only through being network partners of the respondent). The average density is .75 (a mean of 1 would mean that all the network partners know each other). This figure is relatively high compared to comparable calculations in developed countries. This characteristic of networks is of theoretical relevance in the study of the spread of information and the adoption of innovations, as the flow of new information and the adoption of new patterns of behavior are expected to be inhibited by when network members are closely connected (Granovetter 1973; Bott 1971; Marsden 1990; Valente 1995), as indeed we have found in other work (Kohler, Behrman and Watkins 2001).
17. Yule's Q is often used in studies of segregation by race and ethnicity (e.g. White 1994). For this table, a complementary data set was constructed in which the unit of analysis is the network partner and each network partner was assigned the values of their respondent. To calculate Yule's Q, we consider respondents as one cluster and network partners as another cluster using data from K2. Tabulations are weighted by frequency so that respondents who had more than one network partner are not over-represented. The isolation measure is based on a 2 by 2 contingency table in which, for example, those who own bikes are compared to those who don't own bikes, or those who have an education with those who don't. The upper left cell (a) contains the number of network partners and respondents who both own bikes. In the lower right hand corner of the table (d) are the cases where neither the network partner nor the respondent owns a bike. The diagonal cells (b and c) contain the number of cases where either the respondent or the network partner own a bike. For each contingency table, the measure of association Yule's Q is calculated as $(ad-bc)/(ad+bc)$. We do not present significance tests because the standard test assumes that the cases are independent of each other; since some alters are named by more than one respondent, this assumption does not hold here.
18. We do not have enough information to know whether respondents have a preference for those in the same economic category, as measured by durable goods, or whether the order in which the goods were purchased was influenced by the network partners. We think the latter unlikely, however, as these goods are highly desired in these communities; unlike the use of family planning, there is no uncertainty. In addition, because there are substantial differences in the cost of these items (lanterns are inexpensive, metal roofs costly), the order of their purchase is likely to be dictated by the availability of money.
19. This comparison is based on the ever-use of family planning rather than current use.

20. There are many reasons for a discrepancy between respondent's report of her network partner's family planning use and the network partner's own report: confusion about the reference period, secret use by the network partner that was confided to the respondent but not to us, projection, a tendency to report the typical behavior of network partners (c.f. Freeman et al 1987), or a desire to present oneself as someone who has socially and politically desirable friends. We found that the respondent's belief that a network partner is using is associated with the respondent's own use, but not associated with whether the network partner herself reports use (White and Watkins 2000). If this is the case, the apparent disproportionate use of family planning by network partners would be less than it appears.
21. Thirteen percent of *rariu* network partners, 18 % of wealth network partners, and 17 % of AIDS network partners are also named as family planning network partners.
22. It would be desirable to have information about the differences between those people with whom the respondent reported talking about each topic and those with whom she did not talk, as the latter might differ across the four topics. Because our network generator question was topics-specific, however, we do not have this information in the Kenya data. Our research in a similar context in rural Malawi, however, suggests that differences in network partners with whom a particular topic was not discussed are unlikely to be large. Following a household survey in 1998 which asked network questions about family planning and AIDS, in 1999 interviewers conducting semi-structured interviews were given the names of network partners that the respondents had reported in the 1998, and asked their respondent why they spoke with, for example, Teresa about family planning but not about AIDS, and Phoebe about AIDS but not about family planning. Strikingly, the respondents had great difficulty understanding the question, and almost invariably attributed the different network partners to quite specific circumstances. The typical response was "I talked with Teresa about family planning while we were walking to the clinic, and Phoebe didn't go that day." In this setting where, as in rural Kenya, women are close to their network partners both geographically and socially and both family planning and AIDS are "in the air", from the perspective of the respondents it seems that whether a specific conversation turned to family planning or to another topic is largely a matter of happenstance.

References

- Ajayi, Ayorinde and John Kekevole. 1998. "Kenya's Population Policy: From Apathy to Effectiveness" In Anrudh Jain, ed. Do Population Policies Matter? Fertility and Politics in Egypt, India, Kenya, and Mexico. New York: The Population Council 1998, pp, 133-156.
- Alderman, Harold, Jere R. Bherman, Hans Peter Kohler, John A. Maluccio and Susan Cotts Watkins. 2000. "Attrition in Longitudinal Household Survey Data: Some Tests from Three Developing Country Household Samples." The World Bank Group, Health and Population, [#2447](http://www.worldbank.org/research/working_papers).
- Ayodo, Awuor. 1994. "Definitions of the Self in Luo Women's Oral Literature," Research in African Literatures 25:121-29.
- Behrman, Jere R., Hans-Peter Kohler and Susan Cotts Watkins. 2000. "Empirical Assessments of Social Networks, Fertility and Family Planning Programs: Nonlinearities and their Implications." *Demographic Research* Vol.3. <http://www.demographic-research.org>.
- Behrman, Jere R., Hans-Peter Kohler and Susan Cotts Watkins. 2002. "Social Networks and Changes in Contraceptive Use over Time: Evidence from a Longitudinal Study in Rural Kenya." Demography 39(4):713-38..
- Berelson, Bernard, Paul F. Lazarsfeld, and William N. McPhee. 1954. Voting: A Study of Opinion Formation in a Presidential Campaign. Chicago: University of Chicago Press.
- Bernard, H. Russell and Peter D. Killworth. 1977. "Informant Accuracy in Social Network Data II," Human Communication Research 4: 3-18
- Blau, Peter M. 1994. Structural Contexts of Opportunities. Chicago: University of Chicago Press.
- Bott, Elizabeth. 1971. Family and Social Networks: Roles, Norms and Relationships in Ordinary Urban Families (2nd edition). London: Tavistock.
- Bound, John, David A. Jaeger and Regina M. Baker. 1995. "'Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak,'" Journal of the American Statistical Association 90:430 (June), 443-450.

- Brewer, Devon. 2000. "Forgetting in the Recall-Based Elicitation of Personal and Social Networks," *Social Networks* 22:29-43.
- Brewer, Devon and Cynthia Webster. 1999. "Forgetting of Friends and Its Effects on Measuring Friendship Networks," *Social Networks* 21:361-373.
- Brock, W.A. and S.N. Durlauf. 2001. "Discrete choice with social interactions. *Review of Economic Studies* 68(2): 235-260.
- Buehler, Christoph and Hans-Peter Kohler. 2003. "Talking about AIDS". *Demographic Research - Special Collection 1: "Social Interactions and HIV/AIDS in Rural Africa"*, edited by Susan Watkins, Eliya M. Zulu, Jere Behrman, and Hans-Peter Kohler. <http://www.demographic-research.org>.
- Christiaensen, Luc, John Hoddinott and Gilles Bergeron. n.d. "Comparing Village Characteristics Derived from Rapid Appraisals and Household Surveys: A Tale from Northern Mali," paper presented at a conference on "Imperfect Information and Fieldwork in Developing Countries", Yale University.
- Entwisle, Barbara and Jenny Godley. 1998. "Village Networks and Patterns of Contraceptive Choice," Paper presented at a meeting of the Population Committee of the National Academy of Sciences, January 29-30, Washington, D.C.
- Entwisle, Barbara, Ronald R. Rindfuss, David K. Guilkey, Apichat Chamrathirong, Sara R. Currin and Yothin Sawangdee. 1996. "Community and Contraceptive Choice in Thailand: A Case Study of Nan Rong." *Demography* 1996; 33(1)1-11.
- Erickson, Bonnie H. 1996. "The Structure of Ignorance," *Connections* 19:28-38.
- Fischer, Claude. 1982. *To Dwell Among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.
- Freedman, David. 1999. "From Association to Causation: Some Remarks on the History of Statistics," *Statistical Science* 14:243-258.
- Freeman, L., K. Romney, and S. Freeman. 1987. "Cognitive structure and informant accuracy," *American Anthropologist*. 89:310-325.
- Friedkin, Noah E. 1993. "Structural Bases of Interpersonal Influence in Groups: A Longitudinal Case Study" *American Sociological Review* 58:861-872.
- Granovetter, Mark. 1973. "The Strength of Weak Ties," *American Journal of Sociology* 78:1360-1380.

- Hodgson, Dennis and Susan Cotts Watkins. 1997. "Feminists and Neo-Malthusians: Past and Present Alliances." Population and Development Review 23 (3): 469-523.
- Katz, Elihu and Paul Lazarsfeld. 1955. Personal Influence: The Part Played by People in the Flow of Mass Communications. New York: The Free Press.
- Kohler, Hans-Peter. 1998. "Bias in the Estimation of Density on The Basis of Egocentric Networks with Truncated Size." Working Paper.
- Kohler, Hans-Peter, Jere R. Behrman and Susan Cotts Watkins. 2001. "The Density of Social Networks and Family Planning Decisions: Evidence from South Nyanza District, Kenya." Demography 38 (1): 43-58.
- Kohler, Hans-Peter, Jere R. Behrman and Susan Cotts Watkins. 2002. "Social Network Influences and AIDS Risk Perceptions: Tackling the Causality Problem." Paper presented at the annual meeting of the Population Association of America, Atlanta, GA, May 9-11.
- Laumann, Edward O. 1973. Bonds of Pluralism: The Form and Substance of Urban Social Networks. New York: Wiley Interscience.
- Luke, Nancy K. 2000. *Rariu and Luo Women: Illness as Resistance in Rural Kenya*. PhD dissertation, Graduate Group in Demography and Graduate Group in Sociology, University of Pennsylvania.
- Luke, Nancy K., Ina Warriner and Susan Cotts Watkins. 2001. "This *rariu*, it doesn't rhyme with Western medicine: Recognition and treatment of a reproductive illness in rural Kenya", in C. Maklout Obermeyer, ed., Cultural Perspectives on Reproductive Health. Oxford: Oxford University Press, pp. 58-83.
- Manski, Charles. 1992. Identification Problems in the Social Sciences Cambridge, MA: Harvard University Press.
- Manski, Charles and Daniel S. Ragin. 1998. "Bounding Disagreements About Treatment Effects: A Case Study of Sentencing and Recidivism." Sociological Methodology 28:99-137.
- Mark, Noah. 1998. "Beyond Individual Differences: Social Differentiation from First Principles," American Sociological Review 63:309-330.
- Marsden, Peter V. 1987. "Core Discussion Networks of Americans," American Sociological Review 52:122-131.

- Marsden, Peter V. 1988. "Homogeneity in Confiding Relationships." *Social Networks* 10: 57-76.
- Marsden, Peter V. 1990. "Network Data and Measurement." *Annual Review of Sociology* 16:435-463.
- McPherson, Miller, Lynn Smith-Lovin and James M. Cook. 2001. "Birds of a Feather Flock Together: Homophily in Social Networks." *Annual Review of Sociology* 21: 515-444.
- Miller, Kate, Eliya M. Zulu and Susan C. Watkins. 2001. "Husband-Wife Survey Responses in Malawi." *Studies in Family Planning* 30(2):161-174.
- Montgomery, Mark R. and John B. Casterline. 1996. "Social Learning, Social Influence, and New Models of Fertility," in John B. Casterline, Ronald D. Lee, and Karen A. Foote, eds. *Fertility in the United States: New Patterns, New Theories*. (Supplement to *Population and Development Review*, Vol. 22), pp.151-175. New York: The Population Council.
- Montgomery, Mark R. and John B. Casterline. 1993. "The Diffusion of Fertility Control in Taiwan: Evidence from Pooled Cross-Section, Time-Series Models." *Population Studies* 47(3): 457-479.
- Montgomery, Mark R. and Woojin Chung. 1999. "Social Networks and the Diffusion of Fertility Control: The Korean Case," in R. Leete, ed., *Values and Fertility Change*, Oxford: Oxford University Press
- Munshi, Kaivan and Jacques Myaux. 1997. "Social Effects in the Demographic Transition: Evidence from Matlab, Bangladesh," Boston: Boston University, mimeo.
- Moore, Gwen, 1990. "Structural Determinants of Men's and Women's Personal Networks," *American Sociological Review* 55:726-735.
- Oboler. R.S. 1985. *Women, Power, and Economic Change: the Nandi of Kenya*. Stanford: Stanford University Press.
- Ocholla-Ayayo. A.B. C. 1976. *Traditional Ideology and Ethics Among the Southern Luo*. Uppsala: The Scandinavian Institute of African Studies.
- Reynar, Angela 1998. "He Says, She Says, Who Gets Their Way? Marital Dynamics And Fertility In Kenya". Paper presented at the annual meeting of the Population Association of America, Chicago, IL, April 1-3.

- Rogers, Everett and Dilip Bhowmik. 1971. "Homophily-Heterophily: Relational Concepts for Communication Research," Public Opinion Quarterly 34:523-538.
- Rutenberg, Naomi and Susan Cotts Watkins. 1997. "The Buzz Outside the Clinics: Conversation and Contraception in Nyanza Province, Kenya." Studies in Family Planning 28(4): 290-307.
- Short, Susan, Fenian Chen, Barbara Entwisle and Zhai Fengying. 2002. "Maternal Work and Childcare in China: A Multi-Method Analysis." Population and Development Review 28(1): 31-57.
- Stolzenberg, Ross M. and Daniel A. Relles. 1997. "Assessing and Correcting Sample Selection Bias." American Sociological Review 62(3):494-507.
- Taylor, Charles. 1985. Philosophy and the Human Sciences: Philosophical Papers Vol. 2. Cambridge: Cambridge University Press.
- Valente, Thomas W. 1995. Network Models in the Diffusion of Innovations. Cresskill, NJ: Hampton Press.
- Valente, Thomas W., Susan Cotts Watkins, Miriam N. Jato, Ariane van der Straten and Louis-Philippe Tsitsol. 1997. "Social Network Associations with Contraceptive Use Among Cameroonian Women in Voluntary Associations," Social Science and Medicine 45:677-687.
- Warriner, Ina. 1999. Three Demographic Essays. PhD Dissertation, Graduate Group in Demography, University of Pennsylvania.
- Watkins, Susan Cotts. 2001. "Local and Foreign Models of Reproduction in Nyanza Province, Kenya." Population and Development Review 26(4):725-759.
- Watkins, Susan Cotts, Eliya M. Zulu, Hans-Peter Kohler and Jere R. Behrman. 2003. "Introduction". *Demographic Research - Special Collection 1: "Social Interactions and HIV/AIDS in Rural Africa"*, edited by Susan Watkins, Eliya M. Zulu, Jere Behrman, and Hans-Peter Kohler. <http://www.demographic-research.org>
- Watkins, Susan Cotts and Dennis Hodgson. 1998. "From Mercantilists to Neo-Malthusians The International Population Movement and the Transformation of Population Ideology in Kenya." Paper presented at a workshop on Social Processes Underlying Fertility change in Developing Countries, 29-30 January 1998, National Academy of Sciences 1998 Washington, D.C.
- Watkins, Susan Cotts, Naomi Rutenberg, and Steve Green. 1995. "Diffusion and Debate: Controversy About Reproductive Change in Nyanza Province, Kenya."

Paper presented at the 1995 annual meeting of the Population Association of America., San Francisco, CA., April 6-8.

- Watkins, Susan Cotts, Naomi Rutenberg, and David Wilkinson. 1997. "Orderly Theories, Disorderly Women," In W. Jones, R. M. Douglas, J. C. Caldwell and R. M. D'Souza, eds, The Continuing Demographic Transition. Oxford Clarendon Press, pp. 213-245.
- Watkins, Susan Cotts and Eliya M. Zulu. 1997. "Integration of HIV and Family Planning Services." Paper presented at the annual meeting of the Population Association of America, Chicago, IL, April 1-5.
- Weinreb, Alex. 2001. "First Politics, Then Culture: Accounting for Ethnic Differences in Demographic Behavior." Population and Development Review 37(3):437-467.
- White, Kevin. 1999. Essays on Developing Nation Mortality and Kenyan Contraception. PhD Dissertation, Graduate Group in Demography, University of Pennsylvania.
- White, Kevin, and Susan Cotts Watkins. 2000. "Accuracy, Stability and Reciprocity in Informal Conversational Networks in Rural Kenya." Social Networks 22: 337-355.
- White, Michael J., Robert F. Dymowski, and Shilian Wang. 1994. "Ethnic Neighbors and Ethnic Myths: An Examination of Residential Segregation in 1910," in After Ellis Island: Newcomers and Natives in the 1910 Census, Susan Cotts Watkins, ed. New York: Russell Sage Foundation.
- Winship, Christopher, and Stephen L. Morgan. 1999. "The Estimation of Causal Effects From Observational Data," Annual Review of Sociology 25:659-707.
- Winship, Christopher, and Michael Sobel. Forthcoming. "Causal Inferences in Sociological Studies", in Handbook of Data Analysis, Melissa Harding and Alan Bryman, eds. Sage Publications.

