Respondent Cooperation and Requests for Contacts in Longitudinal Research

By

Colter Mitchell

Department of Sociology and
The Institute for Social Research
University of Michigan

and

Vaughn R. A. Call

Department of Sociology
Brigham Young University

Paper prepared for the Annual Meeting of the Population Association of America, March 29-31, 2007, New York, New York

ABSTRACT

Researchers contemplating panel research designs regularly face the problem of panel attrition. A vital tool in reducing panel attrition is first wave information, which can be used to locate respondents who move. We analyze data from the 1988 National Survey of Families and Households (N = 13,007) to examine the extent of respondent cooperation to a request made during the initial interview for the names and addresses of friends and relatives who might help researchers locate panel members who move. About 90 percent of the respondents to the 1988 survey provided at least one contact. Compared to respondents who provide 3 contacts, respondents who provide zero contacts have 60% lower odds of being found the next year (26% lower odds for 1 contact). The paper explores theories for predictors of providing contacts: survey interest, social isolation, social exchange, and family and social networks. Utilizing censored Poisson and sequential logistic regression we find that respondents with greater interest in the survey and larger family and social networks supply more contacts. Conversely, respondents from racial minority groups and those who live in the East provide fewer contacts. Employing this information in preparation for and execution of a longitudinal study should alleviate some panel attrition.

Respondent Cooperation and Requests for

Contacts in Longitudinal Research

The number of major national and regional panel studies has increased greatly during the last three decades (Phelps, Furstenberg, and Colby 2002). There are several published reports on procedures used to track respondents and to reduce panel attrition (Coen, Patrick, and Shern 1996; Ribisl et al 1996; Cohen et al 1993; Gregory, Lohr, and Gilchrist 1992; Dodds, Furlong, and Croxford 1989; Ellickson, Bianca, and Schoeff 1988; Booth and Johnson 1985; Call, Otto, and Spenner 1982; Thornton, Freedman, and Camburn 1982; Freedman, Thornton, and Camburn 1980; Clarridge, Sheehy, and Hauser 1978; Temme 1975; Crider, Willits, and Bealer 1971), and the methodological problems associated with panel research designs are well known. Refusal to continue in the study, death, and residential mobility are significant attrition problems that plague longitudinal research (Kish, 1987). Of these problems, failure to locate panel members who move is one of the largest potential sources of panel attrition. A common theme in the attrition literature is the importance of direct access to the family and social networks of panel members to facilitate finding respondents who move and thus reduce tracking costs.

Access to contact names is a key mechanism to facilitate tracking mobile respondents and lowering the cost of longitudinal studies (Call, Otto and Spenner 1982). In fact, using National Survey of Families and Households (NSFH) data we found that, compared to respondents who provided three contacts, those who gave no contacts had 60 percent lower odds of being found one year later. Those respondents who supplied only one contact still had 26 percent lower odds of being found, compared to respondents who gave three contacts[1]. This suggests that having any contacts is helpful, but having multiple contacts is best for tracking and later follow-up

interviews.

Despite the value of asking for contacts, there is little information on the degree to which panel members cooperate with this request. More surprising is that studies into possible antecedents of such respondent cooperation are nonexistent. Understanding predictors for the number of contacts provided by respondents should help in preparing tracking designs that adjust for participants who provide fewer contacts and are more likely to attrite from the study. As well, panel studies of populations that are particularly prone to provide fewer contacts, thus being more difficult to track, may need to allocate greater resources to tracking. A better understanding of the correlates of the number of contacts provided will aid researchers in maintaining a more representative longitudinal sample and more efficiently organizing limited resources. To begin to fill this critical information void, we use data from the NSFH to examine the extent and nature of respondent cooperation in identifying people who will always know where the respondents live.

BACKGROUND

Residential mobility is the major source of potential panel attrition in most longitudinal studies. In 2003, 14 percent of the United States population moved, and this was one of the lowest rates of residential mobility in the last 50 years (Schachter 2004). More typically, about a fifth of the U.S. population moves each year (U.S. Bureau of the Census 1989). The magnitude of potential attrition risk from residential mobility varies considerably by age (Schachter 2004). For example, researchers contemplating a five-year longitudinal study of high school seniors face the prospect of almost all panel members moving to a new address. Conversely, a five-year longitudinal study of 55-year-olds would find a large proportion of panel members still living in the same residence. In longitudinal studies of the general population such as the NSFH, one-third of the panel members moved within three 3 years of the initial interview (Call 1991). People who are difficult to locate during follow-ups to panel studies also differ from other panel members

with respect to marital status, race, age, level of education, residence (urban or rural), and geographic location (Kandel, Raveis, and Logan 1983). Whether the risk of panel attrition becomes a reality, however, depends in large part on the researcher's ability to develop a comprehensive tracking strategy that minimizes the effects of residential mobility (Call et al. 1982). A key parameter of this tracking strategy is the amount and type of information the researcher has on panel members' family and social networks.

In order to compensate for contacts who move, researchers frequently obtain the names of at least two or three friends or relatives who could easily locate a respondent who changes residence. This may appear to be an easy task. However, some evidence suggests otherwise. For example, in one study, a request for the name of a contact at the end of a telephone interview resulted in only 72 percent of respondents providing a contact (Booth and Johnson 1985). Even though they completed the interview, over a fourth of the respondents in this study refused to provide a contact's name.

Obtaining names of contacts from respondents during the first wave of a panel study can reduce panel attrition in four ways. First, if a respondent moves, it permits the researcher to quickly identify someone who can provide the respondent's new address. Second, it gives legitimacy to a request for a new address. Instead of asking potential informants if they are related to Jimmy Smith and if they know where Jimmy Smith lives, the researcher can state that Jimmy Smith provided his aunt's name, address, and phone number so that he could be contacted if he moved. Offering such information, rather than questions that raise suspicions about the researcher's motives, should greatly enhance the probability of obtaining a new address from the informant. Third, the researcher can obtain a new address for Jimmy Smith without making an additional contact with him. This reduces the opportunity for a respondent to refuse further participation in the study without first receiving letters and information brochures that legitimize

the continuing study and increase the probability that the respondent will continue to participate (Groves and Lyberg 1988). Fourth, it permits the researcher to identify difficult tracking cases and the geographic and social characteristics of panel members who may be difficult to locate prior to a study's implementation. This allows researchers to tailor the tracking design to adequately address the extent and type of potentially difficult cases prior to initiating tracking rather than belatedly trying to adjust procedures to handle unexpectedly high numbers of "lost" panel members.

THEORY

We argue that there are at least four explanations for the number of contacts respondents supply at the end of a survey. That is, we propose that providing contacts is associated with: interest in the survey topic, social isolation, social exchange, and limited social and family networks. Our first three theories, although they have not been applied to providing contacts, are commonly used in survey research literature when dealing with survey cooperation (Groves and Couper 1998), and thus our discussion of them will be limited. We will discuss the theory of social and family networks in more detail.

Survey Interest

One explanation for a refusal to provide contacts centers on the respondent's interest in the survey content. Interest in a survey's topic is frequently linked to higher response rates (Groves 1989; Couper 1997; Groves, Presser, and Dipko 2004). In cross-sectional surveys a respondent agrees to participate based on vague attributions about the survey content and a potential level of interest in the survey topic. In panel studies, however, respondents experience the interview content and the level of burden the survey invokes. Since a request for contacts is usually made at the end of an interview, respondents assess their interest in the survey and determine if it is worth spending the time to do the survey again. Some respondents may refuse to

provide contacts simply because they are not interested in the survey and do not wish to repeat the experience.

Social Isolation

We expect that people who feel socially isolated will not be guided by the norms of society, or in this case will not comply with, to some extent, the request to provide contacts. Typically, feelings of social isolation occur in people who feel alienated from the larger society because of inequitable social policies and distribution of resources, as well as cultural differences (Groves and Couper 1998). That is, if people feel that they (or the groups with which they affiliate) have been treated unfairly they may not desire to follow the larger society's norms. Assuming the norm in a survey is for respondents to be available for re-interview and to provide contacts to make that feasible, those who are more socially isolated might be expected to provide fewer contacts. For example, one hypothesis is that racial and ethnic minorities are more socially isolated and thus less cooperative when it comes to providing contacts (Groves and Couper 1998).

Another form of social isolation stems not from people feeling alienated due to inequities, but rather because social norms dictate that they will be less connected. For example, males may be less connected to society, including family and friends, because social norms suggest that women maintain social and family relations (Auriat 1993). Similarly, some groups expect the elderly to slowly become more disconnected with society as they age (Quadagno 1999). In sum, because of social isolation, racial minorities, the elderly, and males are expected to have less cohesion with society and therefore provide fewer contacts.

Social Exchange

Social exchange hypotheses are nothing new to survey research (Dillman 1978, 2000). In essence, social exchange predicts that people, after weighing several costs and benefits (e.g.,

time, money, mental strain, incentives, etc.), will only choose to exchange if it is beneficial to them. In our case, a respondent may weigh the costs of providing contacts (e.g., giving out family and friends' information) versus the benefits (e.g., helping in scientific research), and if the cost is too high the respondent may not provide the contact information. One specific hypothesis using social exchange theory is that because surveys are seen as information-gathering devices and higher educated people have benefited from similar information gatherings, the higher educated may be more willing to exchange information (i.e., respond or provide contact information) (Groves and Couper 1998).

Family, Social, and Organizational Networks

People who do not provide contacts or provide only one contact may have smaller family and social networks that limit the number of names they can provide. Asking for someone who would know the respondent's whereabouts in five years might limit the choices even further. Thus, we do not know the extent to which a nonresponse or a limited response to a request for names and addresses of friends and relatives is a refusal to provide contact information or simply a reflection of the respondent's inability to provide contact information. Respondents may be willing to respond to a request for contacts but unable to comply.

All panel members have a network of family, social, and organizational linkages that tie them to the social fabric of the community in which they live. The pattern of ties, however, varies significantly. Many respondents have extensive ties to family members as well as numerous social and organizational ties. Some may be disconnected from their family but have numerous friends and organizational affiliations. The mix of family, social, and organizational links may determine, to some extent, how hard it will be to locate respondents who move.

Family Relationships

We expect that family relationships are the best source of information for quickly locating

a panel member who moves. Unless a person cuts all ties, family members usually know the location and activities of other family members. For example, over 87 percent of NSFH participants reported at least one living sibling. Almost 90 percent of those with a sibling had either seen or received a letter or telephone call from a brother or sister within the last month. Of respondents with a living mother, two-thirds had talked with her or received a letter from her within the last week. Even though a person may move thousands of miles away, family relationships are usually maintained.

The biggest difficulty in using family relationships to locate respondents is obtaining access to the respondent's family network. Aside from the strategies of phoning people with the same last name and contacting neighbors, there are few ways to identify family relationships other than getting the information directly from the respondent during the initial interview.

The number of family members and relatives a person can name as contacts varies considerably. The number of potential contacts is limited by the number of living parents, siblings, adult children, and relatives a respondent has. Thus, respondents with smaller kinship networks have fewer people they can list. The type and number of kin that can be mentioned also varies by the respondent's age. Younger respondents usually can name parents, grandparents, aunts, uncles, and other relatives. Among older respondents, most parents, older relatives, and many siblings may have died. Adult children often are the only family members older respondents can name as someone who would know their whereabouts in subsequent years. Marriage increases the number of relatives a respondent can name, while a subsequent divorce may eliminate some or all of these in-laws from consideration as a contact person.

Previous tracking experience suggests that siblings may not be as good a contact source as parents and children (Call et al. 1982). Siblings often decline to provide address information for a brother or sister. Instead, they will refer the researcher to their parent for the address. This

deferral by siblings appears to occur most frequently when the sought brother or sister is in some financial, marital, or legal difficulty. People who have two or fewer siblings are more likely to list no contacts or just one contact. This difference persists even when age is taken into account.

Social Relationships

In addition to family members and in-laws, researchers often ask for the names and addresses of friends to supplement the number of contacts available for each respondent. Social relationships are supplemental because friendships tend to be transitory relationships that are largely dependent on residence and employment. When respondents move or change jobs, their friendship networks may change substantially. Also, even though the distance moved may not be great, many moves are associated with major life-course events such as home-leaving, marriage, divorce, or employment changes that disrupt previous social networks. These transitions often result in new friendship networks and a gradual loss of contact with previous friends and neighbors. Young people and those who have never married are especially difficult to locate because they move more frequently and generally have limited social networks (Bright 1967). Once the annual exchange of Christmas cards ceases and multiple moves occur, former friends and neighbors lose contact with respondents. While they may still know valuable information about the respondent, they often do not know the respondent's whereabouts. While it is easy to locate the names and addresses of neighbors through city directories and list-marketing services, it is difficult to identify respondents' friends without the respondents providing friends' names or information about organizational affiliations.

Organizational Affiliations

People with extensive social networks in community-based organizations make numerous acquaintances while participating in organization activities. Some of these acquaintances may become close friends, while others only know the respondent through interactions at organization

meetings and activities. People who actively participate in community-based organizations are easy to locate through their acquaintances (Crider and Willits 1973). A residential move may not mean a change in church or club affiliations or employer. If a change is made, former pastors, club members, coworkers, and employers can often provide valuable information about the respondent's whereabouts or the whereabouts of another family member. Like social relationships, however, knowledge of the respondent's whereabouts quickly fades with time and multiple moves. Nonetheless, increased participation in community organizations provides respondents with a wider range of friends who may know where the respondent has moved.

Urbanicity and Population Density

An interesting variant of organizational affiliations is a small town. People living in rural areas are easy to locate given residents' greater individual visibility, personal social networks, knowledge of friendship and kinship networks, and the tendency to follow the activities of former residents (Crider and Willits 1973; Crider et al. 1971; Groves and Couper 1998). Also, as the size and population density of communities increase, the number of people a person knows in the community decreases (Goudy 1990).

It is important to note that although we discuss four separate theories for why people provide differing numbers of contacts, all four are highly interrelated. For example, people who have suffered inequitable social exchanges may also be more socially isolated (Groves and Couper 1998). Similarly, people who are very interested in a survey may count being re-interviewed as a large benefit for giving contacts. Also, those who have limited social networks are probably more socially isolated. And those who are involved in many organizations may be more readily interested in survey topics. Thus, we suggest that these theories in fact blend together, but for the sake of this paper we have separated them out to better understand factors associated with giving contacts.

In summary, the number of family, social, and organizational ties to a community provides the upper boundary to the number of contacts a respondent can list. In addition, the respondent's social isolation, survey interest, and amount of social exchange will predict the number of contacts provided. The National Survey of Families and Households also provides the opportunity to examine the hypotheses that the extensiveness of family, social, and organizational ties and the respondent's interest in the survey increase the number of contacts provided.

Our formal hypotheses are:

H1: The more positive the respondent's interest in the survey, the more contacts the respondent will give at the end of the survey.

H2: The greater the social isolation of the respondent, the fewer contacts the respondent will give.

H3: A greater benefit (and lower cost) for providing contacts will lead to more contacts being supplied

H4: The larger the respondent's family and social network, the more contacts the respondent will give.

DATA AND METHODS

Sample

The 1988 National Survey of Families and Households (NSFH) consists of interviews in 1987–88 with a nationally representative sample of 13,007 respondents. The sample design includes a main sample of 9,643 males and females aged 19 and over and an oversampling of some smaller populations of interest. Seventy-five percent of the eligible respondents completed the face-to-face interview. The average interview took an hour and 40 minutes to complete. A detailed explanation of the content and design of the NSFH is reported in Sweet, Bumpass, and

Call (1988).

Variables

—Table 1 about here—

Contacts

The research design for the NSFH contained provisions for a five-year follow-up with the 13,007 panel members. At the end of the initial interview, interviewers requested the names and addresses of friends and relatives who would know the panel member's whereabouts if he/she moved:

> *This is an ongoing research study. In about five years we may wish to contact you again to see how things are going. These last three questions are for our records only, so that we can get in touch with you if you move. Remember, everything you say is completely confidential.*
>
> *Think of three relatives who, five years from now, would know where you have moved. This could be your (or your husband's/wife's) parents, a brother or sister, an adult child, or a favorite relative you keep in touch with. Who are the three relatives who will know where you are?*

The interviewer recorded the name, address, telephone number, spouse's name, and relationship to the respondent for each name mentioned. If a respondent could not name a relative, interviewers probed for friends or someone who would know where they had moved. For older respondents, interviewers asked for the names and addresses of children or younger siblings.

During the data entry process we did not count listed contacts who resided outside the

United States, contacts who lived at the same address as the respondent (unless they had a different phone number), or situations where the respondent listed a contact's name but did not provide an address for that contact. These restrictions affected only 156 respondents.

About 90 percent of respondents to the NSFH provided the name and address of at least one contact. About a fifth provided one contact, a fourth provided two contacts, and just under half of all respondents provided three contacts. Of the respondents who did not provide any contacts, 35 percent overtly refused to provide references. An additional four percent could not think of anyone to list or insisted that there was no need for references. This latter reason was often given by older respondents who stated that they would "either be dead or still living here." Interviewers did not record any comment for the remaining 61 percent of the people without any listed contacts. These people may have refused or could not provide any names. Or these blanks may represent interviewer error; that is, the interviewers may have failed to ask for contacts.

Survey Interest

After the interviewer left the respondent's residence, the interviewer assessed the respondent's interest in the survey on a 7-point scale that ranged from 1 (*not interested*; 1% of the respondents) to 7 (*interested*; 61% of the respondents) with a mean of 6.3.

Social Isolation

Under this grouping we include the respondents' age, sex, and race. All three measures come from self-reported questions on the NSFH. Approximately 47 percent of the respondents were male. The ages of the respondents ranged from 19 to 95, with the average age being 43 and a median age of 38. Finally, our weighted sample reported the following racial and ethnic breakdown: 80 % *White*, 11% *Black*, 7% *Hispanic*, and 2% *other* (of which about 70% were Asian and 28% were Native American).

Social Exchange

Education is measured in years of total schooling (0 to 20), with a high school graduation or GED coded as *12,* two years of college or an associate's degree coded as *14,* a bachelor's degree recorded as *16,* a master's degree as *18,* and a professional degree or PhD as *20*. The average number of years of schooling for the sample respondents was 12.6 and the median was 12.

Family and Social Networks

The NSFH contains considerable detail on family relationships between children, siblings, and parents. This permits a count of the number of living parents, siblings, and children age 19 and older. Since a partner's parents and siblings are important contacts for most couples, the number of living partner's parents and siblings is included with the number of living biological parents and siblings. The count of children age 19 and above includes both biological and stepchildren. The total number of living parents, siblings, and children should determine the level of difficulty a respondent would have in complying with a request to name three close relatives.

The number of organizational affiliations is an indicator of the extensiveness of friendship networks in the community. The NSFH contains respondent reports on participation in 15 different kinds of organizations. Religious groups, sports groups, and school-related groups had the highest levels of participation. The organizational affiliation indicator is a sum of listed organizations that the respondent participated in at least several times a year.

Considering both family and social networks, we can determine the size of a respondent's "pool" of possible contacts. Only 42 percent of people age 19 and older reported that both their parents were still alive. For a third of the population, neither parent was still alive. When parent in-laws are taken into account, over a fourth of the sample still had no parent alive. Almost 20 percent had only one living parent, and approximately 13 percent of the population was married

or cohabiting with both sets of parents still living. Thirty-six percent of respondents reported one or more adult children. Over 93 percent of respondents had at least one living brother, sister, brother-in-law, or sister-in-law. Two-thirds of all respondents indicated that they participated in at least one type of organizational activity. When all factors were combined, less than 1 percent of respondents had no family or organizational ties, with most respondents having several ties to either family or social organizations.

Also included are the urbanicity and the region where the respondent lived. Approximately 75 percent of respondents lived in an urban setting, 16 percent lived in suburbs, and approximately 8 percent lived in a rural setting. We measure region broadly, with 25 percent of respondents living in the North Central United States, 22 percent living in the East, 20 percent living in the West, and 34 percent living in the South.

Analytic Strategy

Modeling the number of contacts provides an interesting statistical problem. Although the dependent variable is a count of the number of contacts given, it is unlike a standard count variable because certain values are truncated. That is, although the probability that someone would give more than three contacts may be low, because only three were requested (and recorded), we cannot know if indeed anyone would have provided any more than three.

If we did have a standard count variable we might have used a standard Poisson or negative binomial regression (Hoffmann 2004). However, in order to account for the truncation (or censoring) of those cases reporting three contacts (i.e., most would have still only given three, but some may have given more), a censored Poisson regression[2] is required (Hilbe and Judson 1999). The censored Poisson regression model can be extended to handle left- and right-censored data.

The censored Poisson regression correctly estimates regression coefficients, while

adjusting the standard errors and properly weighting the data. This allows us to use more complex survey sampling weights, which in the end means more generalizable and correct population estimates. This is the first substantive application of censored Poisson regression using survey sampling weights.

Hilbe and Judson (1999:187–88) document the likelihood function as

$$L(u, X) = \prod_{i=1}^{N} f(x_i, u)^{I(p_i=1)} \left(\sum_{j=0}^{x_i} f(j, u) \right)^{I(p_i=0)} \left(1 - \sum_{j=0}^{x_i} f(j, u) \right)^{I(p_i=-1)}$$

where

- N is the number of cases;
- $p_i = 1$ if the i^{th} observation is not censored, 0 if left censored, or -1 if right censored;
- $I(p_i)$ is the indicator function, taking the value 1 when the statement in parentheses is true, otherwise taking the value 0;
- f is the probability density of a Poisson random variable with parameter u;
- $u = \exp(X\beta)$;
- $1 - \sum_{j=0}^{x_i} f(j; u)$ is the probability of observing x_i, or more events when $E(Y) = u$;
- $\sum_{j=0}^{x_i} f(j; u)$ is the probability of observing x_i, or fewer events when $E(Y) = u$.

The model works assuming the count is non-negative; the distribution of that count can be reasonably acceptable as a Poisson distribution, and not all of the cases are censored (Hilbe and Judson 1999). A common problem in Poisson regression is that when the mean is not equal

to the variance, the estimates of coefficients are consistent, but the standard errors are incorrect. For this reason robust standard errors are used when reporting significance.

As mentioned, an important aspect of this model is its ability to correctly adjust for various types of weighted data. NSFH provides a case weight that is the product of the basic sampling weight, a screening nonresponse adjustment, an interview nonresponse adjustment, and a post-stratification adjustment (a detailed explanation of the weights can be found in Appendix L of the 1988 National Survey of Families and Households Codebook). This weight is the inverse of the individual respondent's probability of selection. In order to correctly estimate the model we use these probability weights throughout our analysis.

Although the censored Poisson regression models the number of contacts, two other models are also of interest: comparing those who give at least one contact versus those who don't give any and those who give multiple contacts versus those who only give one. These two models use binary logistic regression (Hoffmann 2004; Powers and Xie 2000), and the two dependent variables are coded as (a) *0* (those providing no contacts) or *1* (those providing at least one contact) and (b) *0* (those providing only one contact) versus *1* (those providing two or three contacts). Thus, these last two models are actually sequential models where the sample of the second model is contingent on being a "success" in the first model (Powers and Xie 2000). When compared to each other these models will provide evidence as to whether the model for measuring any contacts (whether or not respondents provided any contacts) is different from the model that measures the number of contacts respondents provided.

The intent of these final two models is to provide information as to whether there are two distinct processes (one regarding whether any contacts were provided and another concerning the number of contacts provided). This is similar—in concept—to zero-inflated Poisson and negative binomial models (Long and Freese 2003), neither of which will estimate correctly due to the very

low (or deflated) number of zeros.

RESULTS

—Tables 2 and 3 about here—

Interest in Surveys

Respondents who were interested in the survey provided slightly above average numbers of contacts. By looking at Model 1 in Table 2, we see that for every one-unit increase in interest, there is a 6%–7% increase in the number of contacts given. This is even the case when controlling for the other independent variables (see Model 5). When we consider the high overall interest in the survey (an average interest of 6.3 out of 7), this speaks more to those not interested than to those who are interested. That is, someone who is particularly not interested is more likely to provide few or no contacts, which is consistent with our expectation.

By examining Table 3, we see that survey interest has a similar effect on whether respondents provided any contacts and whether they provided multiple contacts (conditional on providing one contact). Also, both models suggest that for every unit increase in survey interest there is a positive effect on the odds of providing contacts. More specifically, as interest in the survey increases there is an expected 37 percent increase in the odds of providing at least one contact and an expected 20 percent increase in the odds of providing multiple contacts (conditional on providing at least one contact)[3]. Considering the significant effect of survey interest in all the models, there is evidence that the effect of survey interest on number of contacts provided is independent of many personal characteristics as well as family and social networks. Also, there is some limited evidence from Table 3 that survey interest plays a larger role in getting respondents to give any contacts over getting them to give multiple contacts.

Social Isolation

Under our theory for social isolation we expect that since certain groups of people may

feel isolated (through various mechanisms) from mainstream society they will tend to provide fewer contacts. We expect that males, isolated due to norms expecting women to maintain family relationship roles, would provide fewer contacts (Auriat 1993). And although the models in Table 2 estimate that males provide fewer contacts, the difference between males and females is not significant. In Table 3 we find a similar story, although now the effects are significant, with males having 15 percent lower odds of providing any contacts and 13 percent lower odds of providing multiple contacts when compared to women.

Assuming the elderly are more disconnected from society (Quadagno 1999), they should also provide fewer contacts. In Model 2 of Table 2, we do find that every 10 years of increase in age is associated with 2 percent fewer contacts. However, after controlling for other factors (see Model 5), there appears to be no significant difference for age. Similarly, Table 3 suggests that after controlling for other factors, age has little or no effect on providing contacts.

Although minorities may often have extensive social networks within the community, suspicions regarding a researcher's motive for contacting them is often cited as a reason for lower minority response rates (Lewis 1972). This suspicion may also apply to requests for the names of friends and relatives. We find that race has the single largest effect on providing contacts. Looking at Table 2, Model 2, compared to whites, blacks had about 10 percent fewer contacts, Hispanics provided over 14 percent fewer contacts, and other races provided 26 percent fewer contacts. Interestingly, these effects change very little even after controlling for other variables (see Model 5). By looking at Table 3 we see the powerful effects race has on providing contacts: compared to whites, blacks have 42 percent lower odds, Hispanics have 51 percent lower odds, and other races have almost 70 percent lower odds of providing any contact at all. Similarly, compared to whites, the other races have significantly lower odds of providing multiple contacts (32% lower for Blacks, 66% lower for Hispanics, and 47% lower for others)—

conditional on providing one contact.

Some of the race effect may be stem from recent immigration to the United States and the decision not to count contacts listed from other countries. For example, many of the foreign addresses in the NSFH were in Mexico. Also, illegal alien status may have discouraged some respondents from listing any contacts in the United States. Whatever the reason, minority reluctance to cooperate or inability to provide the names of contacts makes re-establishing contact even more difficult, which in turn makes later studies more prone to racially selective attrition.

Social Exchange

In Table 2, Model 3 provides some evidence that higher educated respondents may be more willing to exchange contact information with researchers. That is, every one-year increase in education is associated with a 0.8 percent increase in the number of contacts provided. However, after controlling for other factors in Model 5, that effect is attenuated. Further investigation shows that higher educated people had greater interest in the survey and that education's effect on providing contacts operated through survey interest's effect on providing contacts. Therefore, it still seems plausible that the higher educated are more likely to see a greater benefit to participating in the survey's request for contact information. Not surprisingly, because we are controlling for survey interest, education is not significant in either of the models in Table 3.

Family and Social Organizational Ties

Family Networks

Returning to Table 2, the fourth model examines the multiple indicators of family and social ties, or in another sense, the possible number of contacts. The first three indicators directly test the effects of the number of parents, children, and siblings. All three are estimated to have

positive effects on providing contacts, but only the number of parents has an effect that is statistically significant. More specifically, for every parent alive, the number of contacts increases by about 3 percent. After controlling for other variables, however, the effect of parents decreases and is no longer significant. Interestingly, more investigation showed that those with fewer parents had lower interest in the survey and that the effect of parents operates through survey interest. Another finding is that after controlling for the other variables in Model 5, the effect of siblings became significant, and every increase in a sibling resulted in 0.6 percent more contacts.

A more indirect method of estimating the effect of family network is examining respondent characteristics such as marital status, number of times married, and whether the person was raised in a two-parent home. Surprisingly, we find few significant differences. Marital status adds information only in that being widowed appears to significantly increase the number of contacts. In the full model, being currently widowed is associated with an increase of over 13 percent in the number of contacts. By comparison, that is twice the size of the effect of survey interest. This result is not too surprising considering that widows/widowers are probably more closely linked to family networks since they have lost a spouse. Although a similar argument could be made for divorcees, widows often seem to receive a great deal of (and various types of) support whether needed or not, whereas divorcees may not (Miller, Smerglia, and Bouchet 2004).

Turning to Table 3 again, we can examine the process of providing contacts in two parts. Here we find effects very similar to the Poisson model, except that the estimates are larger and more likely to be significant. We note that the number of living parents, siblings, and children are all significantly (and positively) related to providing at least one contact. However, the number of family members seems to have an even stronger effect on providing multiple contacts. Thus, it

may be possible that a family network has little to do with deciding to give any contacts but a great deal to do with how many contacts are given. Again, we find a positive correlation with being widowed, although it is only significant in deciding to give any contacts, not whether one provides multiple contacts. An interesting new result is that cohabiters are less likely than married respondents to provide multiple contacts. This may be due to cohabiters listing only their partner as a contact. Nevertheless, in general we see that the Poisson model tends to fit for both providing a contact and providing multiple contacts.

Social Organizations

The fourth model in Table 2 also tests the effect that the number of social organizations belonged to has on the number of contacts provided. As with the number of family members, we see that the effects are quite small, but nevertheless they are significant. We can see that the number of contacts provided by the respondent increases by approximately 1.4 percent for every social organization to which the respondent belongs. This effect reduces to about 1.0 percent for each social organization when controlling for the other variables in Model 5. We suggest that this is an indicator of the number of friends the person may have to list as a contact. It is also interesting to note that in the models in Table 3 we see that the effect of social organizations is slightly, although not significantly, larger for predicting whether any contacts are given than whether multiple contacts are provided.

Urbanicity and Region

Looking at Table 2, although people living in nonmetropolitan areas appear to provide more contacts than people living in metropolitan areas, these differences are not significant. Compared to people living in the North Central United States, people living in the West and South do not appear to provide a significantly different number of contacts. However, people

living in the Eastern region appear to provide about 10 percent fewer contacts than do people in the North Central region. Moving to Table 3 we can see that region affects the decision to give any contacts, though not the number of contacts. More specifically respondents in the East have 56 percent lower odds and those in the West have 32 percent lower odds of providing a contact than respondents in the North Central region. Also, note that evidence suggests that urbanicity may have a larger impact on the decision to give any contacts. Suburban and rural respondents, respectively, have 46 percent and 71 percent higher odds of providing at least one contact and 24 percent and 32 percent higher odds of providing multiple contacts compared to respondents in urban areas.

CONCLUSIONS

A key mechanism for reducing panel attrition in longitudinal research is to obtain, at the time of the initial interview, the names and addresses of each respondent's primary family (and social) contacts. When asked for the names and addresses of three relatives, about 90 percent of all respondents will provide at least one name as a contact. About half provide all three. The more contacts a respondent provides, the easier it can be to locate that respondent in the future.

Relative to other variables in the model, the extent of cooperation with a request for names and addresses of family and friends is primarily dependent upon the respondent's interest in the survey, the respondent's race, and the number of primary contacts available to a respondent. Of the social background and residence variables included in the model, race had the largest impact on the average number of contacts listed. Minorities provide a lower than average number of contacts.

As predicted, the number of primary family members and relatives impacts the average number of contacts listed. Net of other variables in the model, people with fewer adult living siblings provided fewer contacts. Increased participation in organizations followed a similar

pattern. Given the emphasis on relatives and the weak measurement of friendship networks, this study may not provide a good indication of the value of asking for friends as contacts. Nonetheless, participation in more organizations does increase the average number of contacts a person will provide.

Net of background and number of relatives, interest in the survey is a major predictor of the average number of contacts provided. It is hardly surprising that people who are not interested in the survey are reluctant to provide contacts to help locate them for a subsequent interview. What is surprising is the extent to which the number of contacts provided increased as interest in the survey increased by just one response category on the seven-point scale.

These findings are applicable to requests for contacts made in a face-to-face interview. The degree to which respondents will cooperate with a request for contacts during a telephone interview or on a mail questionnaire remains unanswered. Previous work using a mail-back questionnaire that included a request for two contact names at the end of follow-up interviews with 30-year-olds resulted in a 96 percent positive response to the request for contacts (Otto, Call, and Spenner 1981). By contrast, only 72 percent of the participants complied with a request for a contact name and address at the end of a telephone interview that was initiated using random-digit dialing procedures (Booth and Johnson 1985). This limited evidence suggests that there may be response differences among survey modes.

Locating 95 to 98 percent of panel members after a period of one year, and especially after 10 or 15 years since the first contact, does not happen by chance. A comprehensive, multimethod tracking plan must be developed to reduce panel attrition (Call et al. 1982). A major aspect of this tracking plan is to maintain contact with the respondent at least every 11 months between the waves. By using the U.S. postal forwarding service, which lasts for a year after the address change, researchers can reduce sample attrition due to mobility (Call et al. 1982).

Most respondents will provide the names and addresses of contacts if the researcher adequately explains why the contacts are needed and clearly defines the type of people who should be listed. While we did not code the presence of comments, respondents who listed fewer than the requested three contacts frequently were apologetic that they could not think of any other relative who would know where they were. A special sequence of probes for a forgotten relative or for a very close friend may be needed to obtain additional contacts. Rather than a specific probe, the NSFH only provided a set of general interviewer instructions. In retrospect, more contacts might have been obtained if a probe for close friends was included in the interview schedule.

The above findings reinforce the need for researchers to train interviewers to focus particular attention on obtaining contacts from minority respondents and respondents in the East. These respondents are less likely to provide contacts even when the number of potential contacts and their interest in the survey is taken into account. Thus, any survey or oversample of these groups is particularly prone to panel attrition. If known beforehand, special training and design may encourage more contacts and thus better long-term tracking.

Interviewer error is another source of missing contacts. Inspection of the tracking forms revealed that a few interviewers left the entire tracking page blank or filled in only a portion of the page. In some instances interviewers may have skipped the request-for-contacts page. Others may have obtained one contact and failed to ask for more. More interviewer training and closer inspection by supervisors should reduce many of these contact omission errors. This additional attention to training and supervision to ensure that contact information is available for each respondent may cause modest increases in the cost of the initial interview, but it could preclude substantial tracking expenses to locate hard-to-find panel members during the subsequent re-interview.

Although the NSFH requested three contacts, other surveys often request fewer. Little research has been conducted on this issue. Of those respondents who gave no contacts, only 87 percent were found five years later. On the other hand, of those providing one, two, and three contacts, 92%, 96%, and 97%, respectively, were found. Further analysis revealed that these differences in location rates are independent of other factors. There is a clear pattern between increased probability of being located later in the panel study and the number of contacts provided. Although there are significant differences between no contacts, one contact, and multiple contacts, there is no difference between two and three contacts. However, we suggest that a key part of any tracking strategy should be acquiring names and addresses for at least three relatives. The three contacts provide researchers with a sufficient number of primary family members who are the most likely to know a panel member's whereabouts. This is probably even more important for longitudinal studies that span several years.

In sum, researchers should not be reticent to ask respondents for contacts. By training interviewers to examine the respondent's social characteristics, the number of potential family contacts, and the respondent's level of interest in the survey and by providing strategies for coping with factors that reduce the number of listed contacts, researchers can minimize the cost of future follow-ups and substantially decrease panel attrition.

REFERENCES

Auriat, Nadia. 1993. "'My Wife Knows Best': A Comparison of Event Dating Accuracy Between the Wife, the Husband, and the Belgium Population Register." *Public Opinion Quarterly* 57:165–90.

Booth, Alan, and David R. Johnson. 1985. "Tracking Respondents in a Telephone Interview Panel Selected by Random Digit Dialing." *Sociological Methods and Research* 14:53–64.

Bright, Margaret. 1967. "A Follow-up Study of the Commission on Chronic Illness Morbidity Survey in Baltimore: I. Tracing a Large Population Sample over Time." *Journal of Chronic Disabilities* 20:707–16.

Call, Vaughn R. A. 1991. "Keeping Track of Panel Members in Longitudinal Research: The Experience of the National Survey of Families and Households." University of Wisconsin-Madison: NSFH Working Paper #38.

Call, Vaughn R. A., Luther B. Otto, and Kenneth I. Spenner. 1982. *Tracking Respondents: A Multi-Method Approach*. Lexington, MA: Lexington Books.

Clarridge, Brian R., Linda L. Sheehy, and Taissa S. Hauser. 1978 "Tracing Members of a Panel: a 17-Year Follow-up." Pp. 185-203 in *Sociological Methodology,* edited by Karl F. Schuessler. San Francisco: Jossey-Bass.

Coen, Anita S., Diane C. Patrick, and David L. Shern. 1996. "Minimizing Attrition in Longitudinal Studies of Special Populations: An Integrated Management Approach." *Evaluation and Program Planning* 19(4):309–19.

Cohen, Evan H., Carol T. Mowbray, Deborah Bybee, Susan Yeich, Kurt Ribisl, and Paul P. Freddolino. 1993. "Tracking and Follow-Up Methods for Research on Homelessness." *Evaluation Review* 17(3):331–52.

Couper, Mick P. 1997. “Survey Introductions and Data Quality.” *Public Opinion Quarterly* 61(2):317–38.

Crider, Donald M., and Fern K. Willits. 1973. “Respondent Retrieval Bias in a Longitudinal Survey.” *Sociology and Social Research* 58:56–65.

Crider, Donald M., Fern K. Willits, and Robert C. Bealer. 1971. “Tracking respondents in longitudinal surveys.” *Public Opinion Quarterly* 35:613–20.

Dillman, Don A. 1978. *Mail and telephone surveys: The total design method.* New York: Wiley-Interscience.

Dillman, Don A. 2000. *Mail and internet surveys: The total design method.* New York: Wiley & Sons.

Dodds, Suzanne, Andy Furlong, and Linda Croxford. 1989. “Quality and Quantity: Tackling Non-Contact Attrition in a Longitudinal Survey.” *Sociology* 23(2):275–84.

Ellickson, Phyllis L., Domenica Bianca, and Diane C. Schoeff. 1988. “Containing Attrition in School-Based Research: An Innovative Approach.” *Evaluation Review* 12(4):331–51.

Freedman, Deborah S., Arland Thornton, and Donald Camburn. 1980. “Maintaining Response Rates in Longitudinal Studies.” *Sociological Methods and Research* 9:87–98.

Goudy, Willis J. 1990. “Community Attachment in a Rural Region.” *Rural Sociology* 55:178–98.

Gregory, Marilyn M., Mary J. Lohr, and Lewayne D. Gilchrist. 1992. “Methods for Tracking Pregnant and Parenting Adolescents.” *Evaluation Review* 16(1):69–81.

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: John Wiley & Sons.

Groves, Robert M., and Lars E. Lyberg. 1988. “An Overview of Nonresponse Issues in Telephone Surveys.” Pp. 191–211 in *Telephone Survey Methodology,* edited by Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nichols II, and Joseph

Waksberg. New York: John Wiley & Sons.

Groves, Robert M., and Mick P. Couper. 1998. *Nonreponse in Household Interview Surveys* New York: John Wiley & Sons.

Groves, Robert M., Stanley Presser, and Sarah Dipko. 2004. "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly* 68:2–31.

Hilbe, Joseph, and Dean H. Judson. 1999. "Right, Left, and Uncensored Poisson Regression." P. 94 in *Stata Technical Bulletin Reprint*, vol.8. College Station, TX: Stata Press.

Hoffmann, John P. 2004. *Generalized Linear Models: An Applied Approach*. Boston: Allyn and Bacon.

Kandel, Denise, Victoria Raveis, and John Logan. 1983. "Sex Differences in the Characteristics of Members Lost to a Longitudinal Panel: A Speculative Research Note." *Public Opinion Quarterly* 47:567–75.

Kish, Leslie. 1987. *Statistical Design for Research*. New York: John Wiley & Sons.

Lewis, Morgan V. 1972. "Finding the Hard-to-locate: A Review of Best Practices." Pp. 165–73 in *Evaluating the Impact of Manpower Programs,* edited by Michael E. Borus. Lexington, MA: Lexington Books.

Long, J. Scott, and Jeremy Freese. 2003. *Regression Models for Categorical Dependent Variables Using Stata*. College Station, TX: Stata Press.

Miller, Nancy B., Virginia L. Smerglia, and Nicole Bouchet. 2004. "Women's Adjustment to Widowhood: Does Social Support Really Matter." *Journal of Women and Ageing* 16(3/4):149–66.

Otto, Luther B., Vaughn R. A. Call, and Kenneth I. Spenner. 1981. *Design for a Study of Entry into Careers.* Lexington, MA: Lexington Books.

Phelps, Erin, Frank F. Furstenburg, and Anne Colby (Eds.). 2002. *Looking at Lives: American Longitudinal Studies of the 20th Century*. New York: Russell Sage Foundation.

Powers, Daniel A., and Yu Xie. 2000. *Statistical Methods for Categorical Data Analysis*. Boston: Academic Press.

Quadagno, Jill. 1999. *Aging and the Life Course*. New York: McGraw Hill.

Ribisl, Kurt M., Maureen A. Walton, Carol T. Mowbray, Douglas A. Luke, William S. Davidson II, and Bonnie J. Bootsmiller. 1996. "Minimizing Participant Attrition in Panel Studies through the Use of Effective Retention and Tracking Strategies: Review and Recommendations." *Evaluation and Program Planning* 19(1):1–25.

Schachter, Jason P. 2004. "Geographic Mobility 2002 to 2003" Current Population Reports P20-594 Washington D.C.: U.S. Census Bureau.

Sweet, James A., Larry L. Bumpass, and Vaughn R. A. Call. 1988. "The Design and Content of the National Survey of Families and Households." University of Wisconsin-Madison: NSFH Working Paper #1.

Temme, Lloyd V. 1975. "The History and Methodology of the 'Adolescent Society' Follow-up Study." Washington, D.C.: Bureau of Social Science Research.

Thornton, Arland, Deborah S. Freedman, and Donald Camburn. 1982. "Obtaining Respondent Cooperation in Family Panel Studies." *Sociological Methods and Research* 11:33–51.

United States Bureau of the Census. 1989. Statistical Abstracts of the United States. Washington D.C.: United States Government Printing Office.

Table 1 Means, Standard Deviations, Medians, and Count of Valid Responses

Variables	Statistics			
	Range	Weighted Mean	Median	N (of 13,007)
Contacts	0–3	2.05	2	13,007
Interest in Survey	1–7	6.26	7	12,479
Social Isolation				
Sex (Male=1)	0–1	0.47	0	13,007
Age	19–95	43.37	38	13,003
Race	0–1		NA	12,982
White		0.80		
Black		0.11		
Hispanic		0.07		
Other		0.02		
Social Exchange				
Educational Attainment	0–20	12.56	12	12, 952
Family/Social Network				
Familial Ties				
# Living Parents	0–4	1.67	2	13,007
# Living Children	0–12	1.17	0	12,996
# Living Siblings	0–30	4.49	4	12,981
# Times Married	0–7	0.99	1	13,001
From an Intact Family	0–1	0.69	NA	13,005
Marital Status	0–1		NA	13,006
Married		0.61		
Cohabitating		0.04		
Separated/Divorced		0.09		
Widow(er)		0.07		
Single		0.18		
# Social Organizational Ties	0–15	1.90	1	13,007
Urbanicity	0–1		NA	13,007
Urban		0.75		
Suburban		0.16		
Rural		.008		
Region	0–1		NA	13,007
North Central		0.25		
East		0.22		
West		0.20		
South		0.34		

Table 2 Effects of Interest in Survey and Family and Organizational Ties on Number of Contacts
(Censored Poisson Regression, effects are exponentiated—e ^{coefficient}) N=12,387

	Model				
	1	2	3	4	5
Interest in Survey	1.069**				1.060**
Social Isolation					
Sex (Male=1)		0.986			0.992
Age		0.998**			0.999
Race					
White		—			—
Black		0.900**			0.907**
Hispanic		0.858**			0.871**
Other		0.740**			0.737**
Social Exchange					
Educational Attainment			1.008**		1.001
Family/Social Network					
Familial Ties					
# Living Parents				1.030**	1.014
# Living Children				1.009	1.010
# Living Siblings				1.002	1.006*
# Times Married				0.982	0.980
From an Intact Family				0.986	0.976
Marital Status					
Married				—	—
Cohabitating				1.024	1.037
Separated/Divorced				0.989	1.004
Widow(er)				1.106**	1.131**
Single				1.041	1.040
# Social Organizational				1.014**	1.010*
Urbanicity					
Urban				—	—
Suburban				1.022	1.008
Rural				1.049	1.023
Region					
North Central				—	—
East				0.900**	0.905**
West				0.969	0.994
South				0.988	1.006
G^2(df)	42.13 (1)**	31.54 (5)**	13.8 (1)**	35.39 (15)*	77.12 (22)**

*p<.05 **p<.001

Table 3 Logistic Regression for Providing at Least One Contact and Providing One versus Multiple Contacts

(Coefficients are Odds Ratios)

	Provided at least one contact (0 = 0 and 1 = 1, 2 & 3)	Provided Multiple contacts (0 = 1 and 1 = 2 & 3)
Interest in Survey	1.367**	1.202**
Social Isolation		
Sex (Male=1)	0.847*	0.869*
Age	0.994	1.002
Race		
White	—	—
Black	0.580**	0.683**
Hispanic	0.491**	0.339**
Other	0.305**	0.531*
Social Exchange		
Educational Attainment	1.093	0.995
Family/Social Network		
Familial Ties		
# Living Parents	1.131*	1.246**
# Living Children	1.093*	1.138**
# Living Siblings	1.072**	1.057**
# Times Married	0.958	0.920
From an Intact Family	0.811*	0.981
Marital Status		
Married	—	—
Cohabitating	1.429	0.712*
Separated/Divorced	0.893	1.150
Widow(er)	1.461*	1.236
Single	0.740	0.905
# Social Organizational Ties	1.089**	1.058**
Urbanicity		
Urban	—	—
Suburban	1.462*	1.236*
Rural	1.712*	1.321*
Region		
North Central	—	—
East	0.435**	1.122
West	0.679*	0.984
South	0.883	0.994
N	12, 387	11,230
G^2(df)	480.95 (22)**	380.12 (22)**

*p<.05 **p<.001

[1] The odds of finding respondents who provided two contacts was not significantly different than the odds of locating respondents that provided three contacts. These results take into account controls for age, sex, were similar to the chances of finding those who provided three race, education, urbanicity, and region of the country.

[2] We use the STATA command *Cepois*, which allows for weighted regression. We thank Joe Hilbe for modifying his *Cenpois* command to allow for weights.

[3] When comparing the censored Poisson and logistic regression models, the coefficients in the censored Poisson model are expected to be more conservative because of the censoring and the robust standard errors, therefore making statistical significance more difficult to achieve. Thus, we feel the censored Poisson estimates are the more reliable estimates, but the comparison does provide some useful information.