Where Do Interviewers Go When They Do What They Do?
An Analysis of Interviewer Travel in Two Field Surveys

James Wagner[1,] Kristen Olson[2]
[1]University of Michigan, Institute for Social Research, 426 Thompson St. Room 4050, Ann Arbor, MI 48104
[2]University of Nebraska-Lincoln, Department of Sociology and Survey Research and Methodology Program, 703 Oldfather Hall, Lincoln, NE 68588

Abstract

Although interviewer-related variance and potential biases that arise when interviewers administer a questionnaire has long been studied, the role that interviewers play in obtaining contact and gaining cooperation is increasingly being explored. In this paper, we investigate the relationship between interviewer travel distance and contact rates, response rates, calls per complete, and hours per interview in two studies, the National Survey of Family Growth and the Health and Retirement Study. Using call record paradata that have been aggregated to interviewer-day levels, we examine the number of trips and number of call attempts interviewers make to sampled segments and the distance interviewers travel to segments.

Keywords: Interviewer travel, Survey Costs, Nonresponse

1. Introduction

Interviewers are a key source of variability in contact and cooperation rates in face to face surveys (O'Muircheartaigh and Campanelli, 1999; Durrant and Steele 2009). To make call attempts in face to face surveys, interviewers must travel from their homes to and among the sampled areas. If a sampled area is quite far from the interviewers' home, large amounts of time spent 'on the clock' is on travel, not interviewing or otherwise recruiting sampled units. Yet a surprisingly unexplored cause of variation in contact and cooperation rates across interviewers is the distance they travel to reach their sampled clusters from their homes and the distance traveled within and between sampled clusters. In this paper, we examine the relationship between the distance that interviewers travel during their work, in particular from their home to and among sampled segments, the number of segments visited during each trip, and the number of call attempts and successful outcomes, such as making contact and obtaining interviews.

Interviewers are trained to make call attempts at times during which they are likely to reach someone at home (e.g., Morton-Williams 1993). Yet not all persons who live in a sampled area will be at home at the same time. Interviewers are also trained to monitor (e.g, Mayer 1968) or limit the number of trips to sampled segments to keep costs under control (e.g., Morton-Williams 1993, p. 141; Campanelli, Sturgis and Purdon 1997, p. 3-20) Anecdotally, interviewers report visiting multiple cases once they have traveled to a particular sampled area (e.g., Peachman 1992). How many sampled cases interviewers visit when they travel to a sampled area and how these visits are constrained by the distance traveled, making contact, obtaining an interview, and other such factors is unexplored.

Interviewer travel has potentially great survey cost implications, although empirical examinations are few. Kalsbeek and colleagues (1994) showed that survey costs could be limited with fewer callback attempts in the NHIS by reducing interviewer mileage and salary related to travel to and from sampled segments, albeit at a decreasing rate with the largest cost reduction with substantial limitations on callbacks. Unfortunately, this approach also had substantial negative effects on the nonresponse bias properties of the estimates (Kalsbeek et al., 1994). In the late 1950s and early 1960s, interviewers were estimated to spend roughly as much as 40 percent of their time traveling to sampled segments and traveling within sampled segments (Sudman 1965-66). Assessments of travel costs during the early 1980s through the 1990s lead to estimates ranging from about 17 percent dedicated to mileage ($474,000/$4,725,000), with an additional 6 percent for between-PSU travel costs (Judkins, Waksberg, and Northrup 1990, Table 1), to costs for travel above 30 percent (Weeks, et al. 1983, Table 3). Although outdated, these few examinations show that interviewer travel can be a substantial part of the total interviewing budget, a trend that is likely to be reflected in today's data collection efforts.

Interviewer travel costs have been viewed as a constraint on sampling error (Sudman 1978; Hansen, Hurwitz, and Madow 1953, p. 274). These costs are assumed to be known and fixed for a given survey, although potentially varying for respondents and nonrespondents (Judkins, Waksberg and Northrup 1990). Yet only one examination of which we are aware has examined the relationship between interviewer travel and nonresponse probabilities, and this study used simulation methods rather than empirical data (Bienias, Sweet and Alexander, 1990).

Studying interviewer travel outcomes requires a record to be kept of travel itself. In most in person surveys, this can be obtained in two ways: (1) mileage reports from the interviewer when asking for travel reimbursement on their timesheets and (2) distances obtained from geocoding the locations of the sampled housing units as recorded in the call records. Both of these sources are likely to contain errors.

First, interviewers are unlikely to report mileage perfectly. We suspect that many interviewers do not keep track of the miles that they actually travel each time they visit a sampled unit, but instead calculate the mileage to a segment once and then use that mileage each time they visit that segment. Additionally, they may overlook cross-segment travel, especially for segments that are nearby, leading to lower reports of travel from the interviewers than actually traveled. Alternatively, interviewers may inflate their reported mileage slightly to increase their reimbursement. Finally, interviewers may simply report mileage as calculated by a mapping program, such as Google Maps or Mapquest, rather than actual mileage traveled.

Second, errors in the mileage calculated through geocoded travel information also are likely to be present. Furthermore, trips as reported in the call records may have errors in the order of visiting particular housing units. When this happens, then the geocoded data will reflect these recording errors. *We hypothesize that the mileage reported by interviewers and recorded through geocoding of sample segments will be highly, but not perfectly, associated.*

Largely unknown is the number of segments to which an interviewer travels on a given day. In this paper, we explore the number of segments that are visited each interviewer-day. We suspect that most interviewers will visit only one or two segments per day. We expect that as interviewers visit more segments, the mileage will also increase. Further, *we hypothesize that the relationship between the number of segments visited and total mileage traveled is not linear, but instead mileage increases at a decreasing rate as travel between multiple segments will require segments to be close together to be time effective.*

2. Hypothesized relationship between travel and field production outcomes

Given the limited research that exists on interviewer travel costs, we derive a number of predictions from the travel simulation by Bienias, Sweet and Alexander (1990) and our experience in the field. First, we assume that interviewers who spend more time travelling have less 'on the ground' time to make call attempts for a fixed amount of work time in a given day. *We hypothesize that as interviewers visit more segments, they make more call attempts, but at a declining rate, since travel takes an increasing amount of time during the day.*

Second, obtaining an interview will decrease an interviewer's time available to visit other sampled cases or other sampled segments assuming a fixed schedule for a given day. *Thus, we expect to see a negative relationship between number of cases visited per segment and cooperation rates. We also expect to see a negative relationship between the number of segments visited per day and cooperation rates.*

We also hypothesize that obtaining a contact without interview will decrease an interviewer's time available to visit other sampled cases, but it will not decrease the time as much as an interview. Similarly, higher numbers of successful contacts should decrease the amount of time that an interviewer has to visit other sampled segments. Alternatively, assuming a fixed schedule for a given day, visiting a sampled area (segment) and making few contacts may also increase the probability that the interviewer leaves the segment for other, hopefully more productive, areas. So, *we expect to see a negative relationship between the number of segments visited each day and contact rates.*

Unknown is how the distance between sampled segments (SSUs) affects the number of segments that an interviewer visits in a day and how they schedule their work among the segments. We will briefly examine the distance between segments and its relationship to the number of segments visited each day.

3. Data and Methods

3.1 Surveys
We examine two large-scale, national face to face surveys conducted by the Survey Research Center at the University of Michigan - the National Survey of Family Growth (NSFG) and the Health and Retirement Study (HRS). The two surveys differ in scope, target populations, and field periods. Additionally, the NSFG is a cross-sectional survey and the HRS is a longitudinal survey. The two surveys also manage field outcomes differently. In the NSFG, interviewers are instructed to visit particular cases whereas the HRS interviewers are instructed to visit particular segments.

3.2 National Survey of Family Growth
The 2006-2010 National Survey of Family Growth (NSFG) is carried out under a contract with the CDC's National Center for Health Statistics (contract #200-2000-07001). The data for this analysis come from Cycle 7 of the NSFG. The NSFG collects

information about fertility, childbearing, and sexual behaviors among women and men in the U.S. aged 18 to 45. Cycle 7 uses a continuous monitoring sample design, with a fixed field period and fresh sample released quarterly. In general, interviewers are assigned one primary sampling unit with three secondary sampling units. The assignment of interviewers to sampled segments is not random; in most PSUs there is only a single interviewer. They are assigned a random sample of segments from within the PSU. In those cases where there is more than one interviewer, the interviewers are assigned segments near their home location to try to minimize travel. Details about the NSFG sample design and management can be found in Lepkowski, et al. (2010). The data considered for this analysis come from the 2006-2010 NSFG data collections.

All of the NSFG field staff record call records and timesheets electronically. Travel information is recorded by the interviewer in their timesheets for purposes of mileage reimbursement in personal vehicles. The two systems, however, do not interface directly. That is, interviewers are not probed in their timesheets to record time specifically related to effort recorded in the call records.

3.3 Health and Retirement Survey

The Health and Retirement Survey (HRS) is a panel study of U.S. adults aged 50 and above, with initial data collection starting in 1992. Every six years, a new age-eligible sample is recruited to include the newly aged-in 51-56 year olds. In this paper, we examine the new cohort added to the 2004 data collection. The panel and the new cohort of the HRS are recruited using somewhat overlapping field staff at the beginning of data collection. However, we examine here the effort related to the interviewers who were assigned to contact the new cohort cases primarily. Among these interviewers, 91.7% of trips had no calls to the panel cases sample, and 5.2% had only one call to the panel cases. As such, although the HRS is a panel study, the data that we examine here are analogous to a cross-sectional study. In the HRS, multiple interviewers are (not randomly) assigned to a PSU, and each interviewer was (not randomly) assigned to a variable number of SSUs based on their geographic proximity to the SSUs. The field period for the HRS was about 10 months.

3.4 Level of analysis

Ideally, the level of analysis for each interviewer and sampled unit for this study is at the call level. Each sampled address could be geocoded and distance between the interviewer's house, the first sampled address, and among sampled addresses could be calculated. However, the timesheet data – the source of data for the HRS - are kept at the day level for each interviewer, not at the call level. Thus, we aggregate information for each interviewer for each day of the survey period, and analyze travel at an interviewer-day level.

3.5 Trips

As one measure of distance that can be calculated from call records, we examine the number of 'trips' that an interviewer makes to a segment. Trips are defined as a visit to a segment that involves travel between the interviewer's home and the segment, or between segments. For example, if an interviewer travels to segment A from their home, then goes to Segment B, and then returns to segment A, they have taken 3 trips.

3.6 Mileage Data

We do not currently have geocoded travel data from the HRS, and thus only have one source of data – their travel as reported in the timesheets. The NSFG mileage data come from two sources – geocoding and interviewer's reported travel. In the NSFG, we geocoded each interviewer's home address and the centroid of the sampled segment. We then calculated the distance in miles 'as the crow flies' from the interviewer's home address to their segments and among their segments. Although other options are available for calculating distance (such as 'best routes' calculated through Google Maps), we start with this approach for simplicity.

Discrepancies between the two sources of information can occur for a number of reasons. For example, interviewers can record work-related travel - such as travel to a training session - that is not related to field effort, interviewers who are flown into a sampled PSU will not record mileage because they do not need to be reimbursed for rental car mileage, and interviewers who work in major metropolitan areas may not use a car to travel among sampled units, instead using public transportation. Additionally, interviewers may fail to complete a call record for some types of travel, such as driving by a house and not seeing evidence of anyone at home (Wang and Biemer 2010), or may not complete their travel reports until the end of the day, potentially forgetting a trip or misremembering where they traveled. Interviewers may also enter travel information on the wrong date.

Geocoding has limitations, as well. In particular, geocoding distances from the centroid of the segment rather than from each individual housing unit leads to efficiencies in the geocoding task, but for large segments, will likely lead to an underestimate of distance relative to the actual distance traveled. Furthermore, geocoding requires an accurate recording of call attempts. To the extent that call attempts are not recorded accurately, the geocoded distances will not match actual travel, even if the travel reports are accurate.

Table 1 presents the distribution of missing mileage data in both surveys. As can be seen up to 24.6% of interviewer-days were missing geocoded mileage in the NSFG, 9.3% of interviewer-days in the NSFG were missing interviewer timesheet reports of mileage, and 3.8% of interviewer-days for the new cohort cases were missing mileage in the HRS. In both surveys, mileage was more likely to be missing on one segment trips.

Table 1. Distribution of missing data on mileage

	2006-2010 NSFG		2004 HRS New Cohort	
	n	% or Mean	n	% or Mean
Missing Data on Mileage				
Mileage not geocoded	449	24.62%	n/a	
Mileage not reported in timesheets	170	9.3%	498	3.80%
Missing interviewer timesheet mileage - all trips				
1 segment trips	997	12.84%	7040	5.75%
2 segment trips	399	5.01%	2808	1.92%
3 segment trips	225	5.78%	1605	1.25%
4 segment trips	93	5.38%	790	1.01%
5+ segment trips	110	3.64%	873	1.26%
Mileage available from time sheets	1654	90.68%	12618	96.20%
Overall	1824		13116	

4. Findings

4.1 Overall travel

We start by examining the overall distance traveled for each interviewer on each day of the field period. In the NSFG, there are a total of n=1824 interviewer days. NSFG interviewers travel an average of 87.71 miles (SD=94.20 miles), as calculated from the geocoded data, or 83.77 miles (SD=74.75 miles), as calculated from the interviewer-reported data, round trip from their home to the sampled segments and back. As hypothesized, the two sources of travel data are highly, but not perfectly correlated (r=0.780, p<.0001). In general, the interviewer-reported data tends to be higher than the geocoded data (Figure 1). HRS interviewers, in contrast, traveled an average of 53.4 miles each interviewer day.

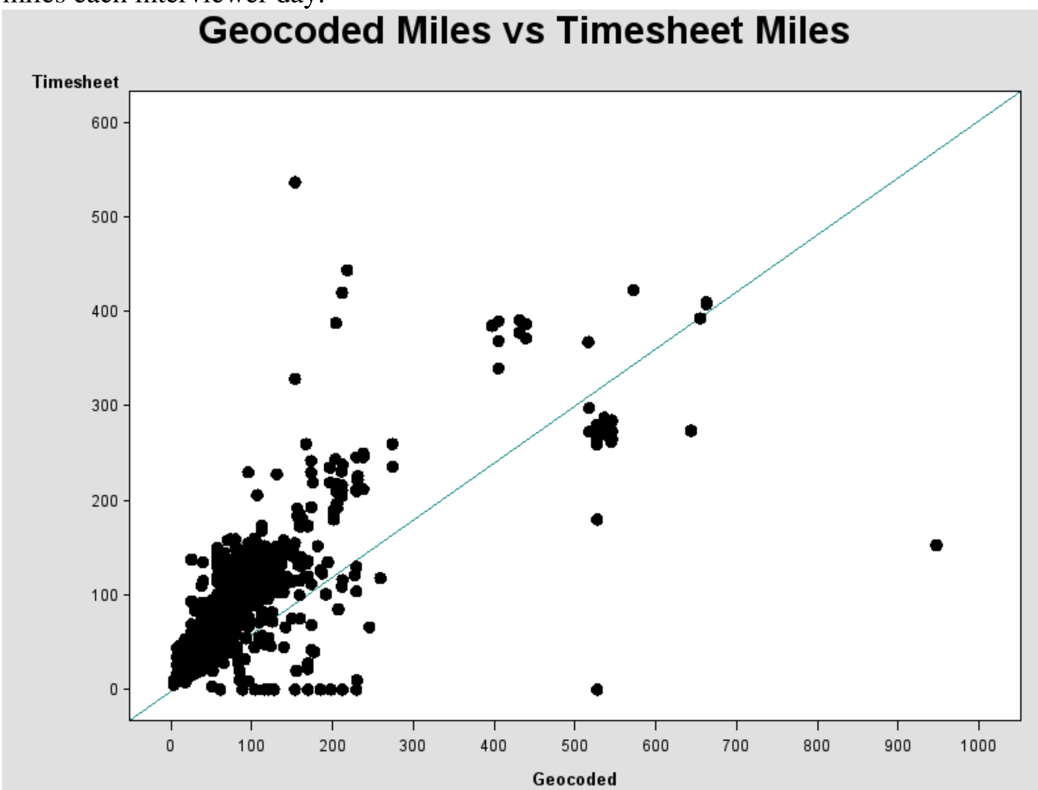

Figure 1. Scatterplot of geocoded miles and interviewer-reported miles in the NSFG

4.2 Miles by Number of Segments

Now, we examine how many segments interviewers are visiting each interviewer-day, and how far they travel when visiting those segments. As can be seen in Table 2, interviewers tend to visit one or two segments in a work day. The distribution of number of segments visited per interviewer-day is remarkably similar across the two surveys. In both surveys, roughly 54% of all interviewer-days are spent visiting one segment, and just under 22% of interviewer-days involve visits to two segments.

As interviewers visit more segments, their distance increases, but not as sharply as one might expect. On average, NSFG interviewers who visit only one segment each day travel an average of 70.3 miles, traveling an additional 14 miles to visit a second segment, for an average of 84.8 miles. NSFG Interviewers who visit 3 or 4 segments travel about 100 miles. In contrast, HRS interviewers travel only about 51 miles to visit one segment, with a marginal increase to 53 miles to visit two segments and 56 miles for three segments. These differences likely reflect the greater flexibility that HRS has in assigning SSUs to interviewers.

Table 2. Distribution of segments visited each interviewer-day, mean number of miles traveled by number of segments visited

	NSFG			**HRS**		
	n	% or Mean	SD	n	% or Mean	SD
Segments visited per interviewer-day	1631			13116		
1	865	54.66%		7040	53.67%	
2	381	21.88%		2808	21.41%	
3	254	12.34%		1605	12.24%	
4	131	5.01%		790	6.02%	
5+	110	6.01%		873	6.66%	
Mean number of miles traveled by number of segments visited						
1 segment	869	70.34	62.38	6635	51.17	47.42
2 segments	379	84.81	66.42	2754	53.26	47.46
3 segments	212	106.61	92.34	1585	56.20	55.83
4 segments	88	108.18	91.57	782	60.60	56.08
5+ segments	106	124.20	105.56	862	59.86	59.39

4.3 Relationship between Call Attempts and Travel

We separate calls for the NSFG into two groups – calls to screen the household for an eligible sample person and calls to complete the main interview. In the HRS, we look only at calls for the new cohort, but cannot disentangle screening calls from main interview calls. In both surveys, as the number of segments visited increases, the number of call attempts made during an interviewer day also increases. The rate of increase is slower in the NSFG than in the HRS, especially for screener calls.

Table 3. Mean number of call attempts by number of segments visited

	NSFG		HRS	
	Screener	Main	Screener	Main
Number of segments				
On 1 segment trips	9.34	2.97	7.90	0.97
On 2 segment trips	10.76	4.51	11.56	1.50
On 3 segment trips	12.18	5.94	14.66	1.95
On 4 segment trips	14.08	7.00	17.23	2.45
On 5+ segment trips	17.07	9.10	21.83	3.04

4.4 Number of segments and contact rates

We now examine the relationship between travel and contact rates. We expected that contact and cooperation rates would have a negative relationship with interviewer travel. As can be seen in Figure 2, there is a modest, but noticeable decline in contact rates in both surveys as an interviewers' travel load increased. In each survey, the rate of decline is relatively similar. .

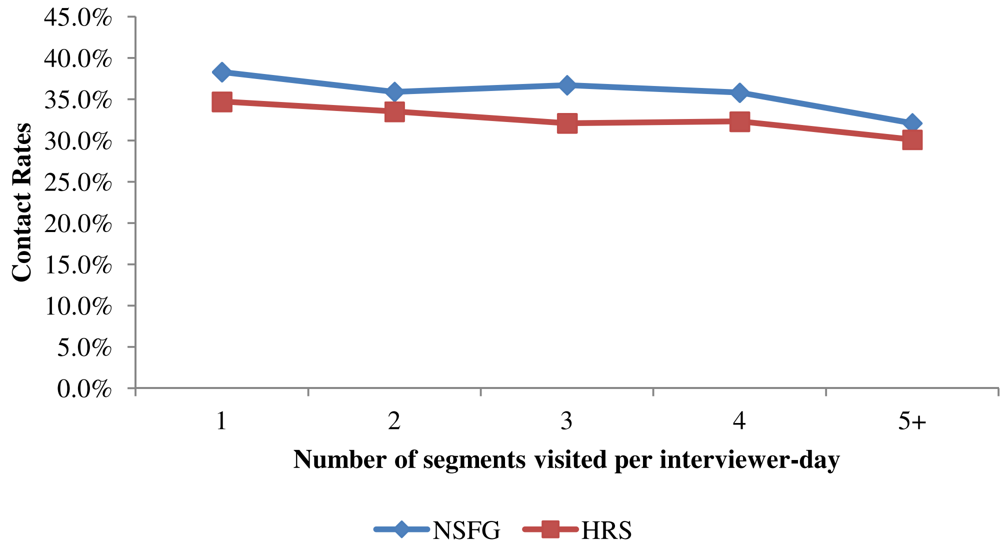

Figure 2: Contact rates by number of segments visited

4.5 Travel and Response Rates

Figures 3 and 4 show response rates for the NSFG (Figure 3) and HRS (Figure 4). As with contact rates, as interviewer travel increases, screener response rates and main interview completion days decrease.

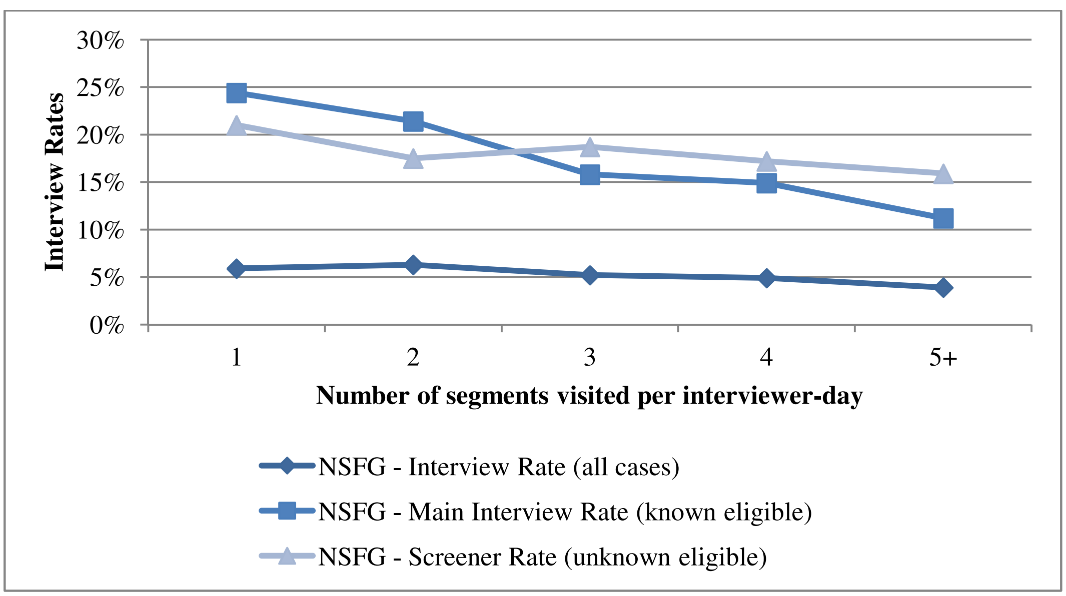

Figure 3. NSFG Response rates by number of segments visited

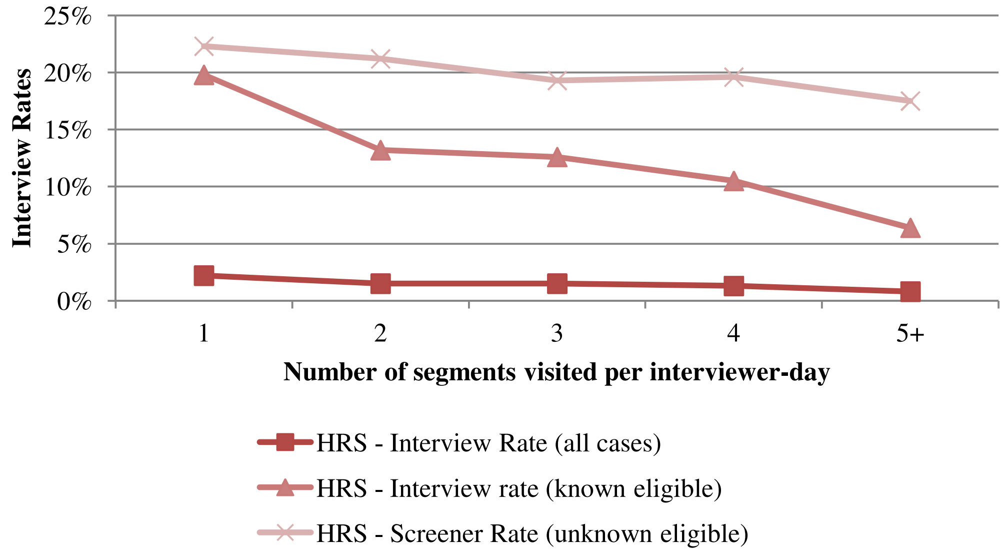

Figure 4. HRS Response rates by number of segments visited

4.6 Number of segments and HPI

In both surveys, hours per interview (HPI) fluctuates as the number of segments visited per interviewer-day increases. However, in the HRS, HPI tends to increase from 23.8 on one segment trips to about 26 hours per interview on 3 or 4 segment trips. In the NSFG, HPI is relatively similar for days with between 1 and 4 segments visited, but increases to 8.2 on longer trips.

Table 4. Hours per interview by number of segments visited

	NSFG	HRS
Number of segments visited		
On 1 segment trips	7.6	23.8
On 2 segment trips	6.6	28.3
On 3 segment trips	7.6	25.9
On 4 segment trips	7.2	26.4
On 5+ segment trips	8.2	38.7

5. Conclusions

This is an initial examination of the relationship between interviewer travel and survey field outcomes. We see a relationship between interviewer travel and survey outcomes. However, we cannot, from these data, disentangle causality – does more travel lead to worse field outcomes or do poor field outcomes result in more interviewer travel?. Clearly, more research is needed. Additionally, this analysis treated all interviewers and all sampled areas as the same, but it is plausible that characteristics of areas and interviewers affect the relationship between travel and field outcomes. For example, does the relationship between interviewer travel and survey field outcomes differ by urban versus rural areas? Does the relationship differ for interviewers who vary in experience levels? Furthermore, we operationalized travel as the number of segments visited per interviewer-day, but different stories may emerge when examining miles travelled, rather than segments. We will explore these questions in future research.

Additionally, more evaluations of the quality of travel data are needed. We have indications that the two sources – geocoded and interviewer-reported data – differ, but not an indication of which source is more accurate. Collection of real time travel data will help us understand the quality of both of these sources of data. Additionally, collection of real time travel data will help us examine the relationship between travel and field outcomes at the address level, rather than the interviewer-day level.

Future research should incorporate this information into explicit cost modeling and explore their implications for sample design. Although interviewer travel is often mentioned as a constraint on the number of clusters to select and the size of the clusters, we are unaware of incorporating contemporary empirical data for these measures into actual cost-error tradeoff models.

6. References

Bienias, J.L., Sweet, E.M., Alexander, C.H., (1990). A Model for Simulating Interviewer Travel Costs for Different Cluster Sizes, *Proceedings of the Survey Research Methods Section, American Statistical Association*, 20-27.

Campanelli, P., P. Sturgis and S. Purdon (1997). *Can you hear me knocking?: investigation into the impact of interviewers on survey response rates*, National Centre for Social Research.

Hansen, M. H., Hurwitz, W. N., and Madow, W. G. (1953), *Sample Survey Methods and Theory*, New York: Wiley.

Judkins, D., J. Waksberg, and D. Northrup. (1990). "Cost Functions for NHIS and Implications for Survey Design." *Proceedings of the American Statistical Association, Survey Research Methods Section*, 34-43.

Kalsbeek, W. D., S. L. Botman, J. T. Massey and P.-W. Liu (1994). "Cost-Efficiency and the Number of Allowable Call Attempts in the National Health Interview Survey." *Journal of Official Statistics* **10**(2), 133-152.

Mayer, C. S. (1968), "A Computer System for Controlling Interviewer Costs," *Journal of Marketing Research*, 5, 312-318.

Morton-Williams, J. (1993). *Interviewer Approaches*, Aldershot, Hants, England

O'Muircheartaigh, C., and Campanelli, P. (1999), "A Multilevel Exploration of the Role of Interviewers in Survey Non-Response," *Journal of the Royal Statistical Society, A*, 162, 437–446.

Peachman, J., (1992). Design and Methodology of 1991/92 Household Interview Survey, Australasian Transport Research Forum 17th Conference, 149-161.

Sudman, S. (1965-66). "Time Allocation in Survey Interviewing and in Other Field Occupations." *Public Opinion Quarterly* 29, 638-348.

Sudman, S. (1978). "Optimum Cluster Designs Within a Primary Unit Using Combined Telephone Screening and Face-to-Face Interviewing." *Journal of the American Statistical Association* 73, 300-304.

Weeks, M.F., R.A. Kulka, J.T. Lessler, and R.W. Whitmore. (1983). "Personal versus telephone surveys for collecting household health data at the local level." *American Journal of Public Health* 73, 1389-1394.