Report on the International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys

Tom W. Smith
NORC/University of Chicago

June, 2011

The International Nonresponse Workshop was sponsored and supported by the National Science Foundation, the World Association for Public Opinion Research, and the International Association of Survey Statisticians. It was held in Chicago on June 2-3, 2011. The agenda of the workshop, including the list of attendees, appears in Appendix 1.

This workshop report considers 1) how auxiliary data (AD) can be used to analyze nonresponse bias, 2) other benefits of using AD, 3) recommendations regarding using AD for nonresponse analysis relating to a) sample frames, b) linked databases, c) paradata, d) aggregate data, and e) AD in general, 4) propensity models, 5) resources for nonresponse analysis, and 6) nonresponse bias measures.

Using AD to Detect and Adjust for Nonresponse

Nonresponse is appreciable in most general-population surveys and has been increasing over time (de Leeuw and de Heer, 2002; Couper and de Leeuw, 2003; Kreuter, forthcoming; Stoop et al. 2010a). Nonresponse usually does not lead to data that are missing at random and thus it is not ignorable. So surveys need to take steps to detect and adjust for nonresponse bias. Increasingly, it is recognized that this can be done by using multi-level data from sample frames, auxiliary databases, paradata, and related sources. As Sarndal and Lundstrom (2005) have noted, "The key to successful weighting for nonresponse lies in the use of powerful auxiliary information." Likewise, Groves (2006) remarked, "Collecting auxiliary variables on respondents and nonrespondents to guide attempts to balance response rates across key subgroups is wise."

AD can be thought of as encompassing all information not obtained directly from the interview of a respondent or case-specific informant. First, it starts with information from the sample frame itself. What is available varies greatly. In many samples of special populations there are often sample frames not only rich in information, but having abundant data closely related to the subject of the survey being conducted. Examples would include a health survey of HMO members that had access for all sampled cases to the full enrollment and health-care-use data from the HMO records or an educational study of students that had access for all sampled students of detailed school records.

But when dealing with general-population samples, such as national household surveys, the available information from the sample frames is more limited and highly variable across countries and data collector (for the use of sample-frame information to analyze nonresponse see Appendix 2.A). In a number of European countries (e.g. the Netherlands and Scandinavian countries), population registers can be used to sample individuals and households (Bethlehem, Cobben, and Schouten, 2011; Sarndal and Lundstrom, 2005). They typically contain demographic information on individuals and households and in some cases can be linked to a wide variety of other government records. However, access to the population register and linked-government records is often restricted to government statistical agencies.

In the United States, address-based sampling (ABS) is generally based on a combination of sample areas (e.g. segments and primary sampling units) selected on the basis of the United States Decennial Census and/or American Community Survey data and addresses selected from either the postal listing of addresses or the postal listings augmented by additional listings (especially in rural areas with a low incidence of city-style addresses (i.e. number + street name). While the postal list has about 40 address-specific variables related to the address and postal delivery matters, the Census information has no address-specific information. But the Census can provide rich, aggregate, geographic information at the block-group, Census tract, and larger units of aggregation. In the United States, RDD-phone samples have very limited information from the sample frame of the sampled numbers and especially due to the increased use and portability of cell-phones, aggregated geographical data are difficult to append. Other sample designs, such as random-route samples, also have very limited AD from the sample frames.

Second, AD consist of information that can be linked to the sampled cases (for the use of linked data from ancillary databases see Appendix 2.B). This would include both case-level and aggregate-level information and data from governmental and commercial sources. What can be linked varies greatly from sample-to-sample, across countries, and by data collector. ABSs in the United States can be linked to many databases with a wide ranging and growing set of both case-level and aggregate-level variables (DiSogra, Dennis, and Fahimi, 2010; Smith and Kim, 2009). RDD and other telephone number samples can be linked to a more limited range of case-level variables and the possibility of obtaining aggregate-level information is becoming more problematic due to the expansion of cell phones and their portability (See Appendix 3).

Third, AD includes paradata that are produced during the conduct of the survey itself (Couper and Lyberg, 2005; Kreuter and Casas-Cordero, 2010; Laflamme, Maydan, and Miller, 2008; Lynn and Nicolaas, 2010; Olson, forthcoming; Stoop et al., 2011). There are two main types of paradata (for the use of paradata see Appendix 2.C, D, E). First of all, there are process data that are routinely collected as part of the normal process of conducting the survey. Examples include records-of-calls, listing reasons for refusals, the offering and acceptance/rejection of incentives, and interviewer characteristics (e.g. demographics, experience, training). These might be particularly useful since a number of them are clearly closely related to nonresponse and should be valuable in the calculation of propensity scores. One special form of process paradata is when the interviewer obtains more information on the identity of the sample unit, such as the name of the respondent at the sampled address. This information can then be used to carry out additional database searches which can facilitate both data-collection efforts and nonresponse analysis. A second special form of process paradata consists of extended data collection especially via the use of computer-assisted, recorded interviews (Conrad et al., 2010; Kreuter and Casas-Cordero, 2010; Smith and Sokolowski, 2011). But this is essentially restricted to respondents only. A third special form of process paradata is keystroke data from CATI or CAPI programs.

Second of all, there are additional paradata, especially observational information, that are mostly collected to assess nonresponse or for some other purpose beyond the goal of survey management that is the purpose of the process paradata. For example, NORC's General Social Survey (GSS) (http://www.norc.org/GSS+Website) and the Centre for Comparative Social Surveys' European Social Survey (ESS) (http://www.europeansocialsurvey.org/) collect information on dwelling type and the ESS also collects information on the observed demographics of the contact person (e.g. age and gender) and the presence of trash, vandalism, and related neighborhood conditions. Of course some paradata may span these two categories. For example, interviewer estimates of the demographics of the sample unit and/or of the targeted respondent may facilitate both data collection and the assessment of nonresponse bias.

Finally, there is information collected from either doorstep interviews with non-respondents or post-survey, follow-up surveys of non-respondents (Cobben and Schouten, 2007; Schouten, 2007; Stoop, et al., 2010a). While these both involve surveys, they are AD in that they involve additional information being collected outside of the main survey itself. Naturally, there is nonresponse to such surveys of non-respondents.

AD from the sample frame, linked databases, and paradata have the potential to be available for all sampled units whether or not they become respondents. This would most often be true for information from the sample frame itself. But some variables might be missing in the sample frame data as well as for other AD.[1] The level of missing data varies by source and variable. For example, from linked databases in the United States using ABS, information on whether there is a listed telephone number is very complete, reports of age and race usually exist for over 90% of households, but information for education in typically known for less than half of sample units (Smith and Kim, 2009). For paradata, observations on

[1] In sample frames and any other source of AD, there also may be errant data (measurement error). The degree and nature of measurement error needs to be evaluated when deciding whether and how to incorporate any AD.

the neighborhood are generally nearly complete, but observations on the characteristics of households and of known or imputed respondents are much less complete. Missing data are of course even more common with door-step and follow-up, nonresponse surveys.

Figure 1 shows an example of how AD might apply to one type of general population survey, an ABS in the United States based on the postal listing of addresses and Census statistics using a multistage, area probability design. Three sources of data are listed in the columns: 1) the sample frame itself, 2) linked, ancillary databases, and 3) paradata from the survey itself. Paradata are further sub-divided into routine/process data and added/observational data. The rows show five levels of aggregation starting with the respondent followed by the household/address, the neighborhood (e.g. Census tract), the community, and higher spatial units (e.g. county, metropolitan area, state).

Beginning with the upper left-hand cell under the first column covering the sample frame, there is no information on the respondent from the sample frame because in these types of samples the frame identifies the address/household and sampled respondents have to be determined after the residence is contacted (e.g. via a Kish table or last/next birthday procedure). Going on down the first column, there is information on the household/address from the postal listings (but nothing from the public Census files). There is some Census-based data for the neighborhood (block group and/or Census tract). There is more Census information at the community and higher levels.

Column 2, dealing with linked sources, shows that if a respondent has been identified that individual can be linked to various ancillary databases. That, however, would only occur after field attempts to contact the household and the successful identification of the respondent. But at higher levels of aggregation, AD are directly available for addresses (at least for those with city-style addresses) (see Smith and Kim, 2009 for a fuller discussion). For example, any GIS-coded database can be linked to ABS addresses as can other address-coded databases such as listings of voting and registration. At the neighborhood level any address that has been coded by zip code and/or Census tract can be linked. At the community level an even wider range of government administrative data plus much commercial data are available. At higher levels of aggregation still more information, such as county-level votes and media-subscription rates, become readily accessible.

Columns 3 and 4 indicate that process paradata are available for respondents when the respondents have been identified (e.g. in single-adult households or households with the respondent-selection protocol completed). This might include reasons for refusals and other respondent-interviewer interactions. At the household/address-level it would include such things as the records-of-calls and related information. For added, observational paradata, when the respondent is known, information would include observed demographics of the respondent. At the household/address level, it would include descriptions of the housing unit and estimates of the household composition. For levels above the household, process paradata could be aggregated to cover neighborhoods, communities, and other units. Examples are the Census-tract-level, mail-return rates based on the decennial Census. For added/observational paradata, descriptions of neighborhood are often collected (e.g. the appearance of litter; broken windows, abandoned cars). Collecting added/observational paradata at the community level or higher is less common, but figures could be aggregated from the neighborhood-level reports. In brief, as this example illustrates, much AD can be appended to in-person ABSs in the United States. Other sample designs in other countries will vary as to what can be readily added. For example, population-register samples will typically start out with much more information, especially at the case-level, from the sample frame. In addition, the substantive focus of a survey and/or its research objective will help to determine what AD should be collected and used for nonresponse analysis.

Multiple Benefits of Using AD

While important for the detection of and adjustment for nonresponse bias, AD can serve many other equally important functions. First, AD can be used for survey design in general and sample design in particular (Dennis, 2011; DiSogra, Dennis, and Fahimi, 2010; Link, 2010; Smith and Kim, 2009). For example, achieving oversamples and/or the sampling of rare or hard-to-find population groups can be facilitated by incorporating information from AD.

Second, AD can be used to improve data collection. AD can assist interviewers in making contact with and gaining cooperation from sampled units (Durrant, D'Arrigo, and Steele, 2011; Link, 2010; Smith and Kim, 2009). If interviewers have more information about the sample units, they can undertake more effective approaches to make contact and more persuasive appeals to gain cooperation. AD can also be used to help validate interviews (Smith and Kim, 2009). In addition, paradata are in particular designed to help manage and oversee data collection.

Third, as the workshop focused on, AD can be used to detect and adjust for nonresponse bias. The bulk of this report elaborates on this point.

Finally, AD in general can be used to augment the data records of respondents for substantive analysis. In some cases AD will produce information of superior accuracy to what can be collected by direct inquiry from respondents. For example, in most cases GPS-based calculations of distance of the residence to various facilities (e.g. schools, power plants, hospitals, superfund sites) will be more accurate than respondent reports. In other cases, AD may be of equal reliability and can be used to cross-check interview reports and/or fill in missing data. Of course, in other cases the AD may be of lower quality and not useful to supplement the survey-based data.

Also, aggregate-level information covering neighborhoods, communities, and/or other spatial units can be used for contextual analysis (Smith and Kim, 2009; Smith, 2010). Social-science research is increasingly documenting that one needs to understand the context in which sample cases live. For general population surveys, this would include geographic dimensions like neighborhood and community-level effects. For employees, students, or groups members, it would include characteristics of the organizations they belong to.

Recommendations Regarding Using AD for Nonresponse Analysis

Sample Frames

All information available from sample frames should be retained and utilized to detect and reduce nonresponse bias. This should include both case-level and aggregate-level data.

GPS and/or longitude/latitude (L/L) readings are essential both for calculating case-level spatial variables (e.g. distance of residence to specific facilities such as hospitals or police stations) and for linkage to most aggregate-level data (e.g. hospital beds per capita in the local community or crime levels in the neighborhood). ABSs in the United States have GPS and L/L associated with city-style addresses. For non-city style addresses, GPS readings in the field should be collected to add this information. Of course, public-use files will have to censor or eliminate the GPS information for confidentiality reasons.

Linked Databases

AD from numerous databases can and should be linked to sample cases. While this is occasionally possible at the respondent level (e.g. by statistical agencies in countries with population registers), it will more generally be possible at the address or household-level (and higher levels of aggregation). Much research is needed on the augmenting of sample-frame data with information from other databases and sources.

First, the accuracy of links needs to be assessed in general and in particular the difficulty of linking apartments needs to be closely evaluated. As for residences without city-style addresses, while these can and should be GPS coded in the field, this will not materially enhance case-level linkages across databases since matching on GPS and L/L is not readily possible across most databases and other sources. Obtaining AD from linked databases necessitates good linkage protocols (Amaya, Skalland, and Wooten, 2010; Davern et al., 2009; Dixon, 2004; Durrant, D'Arrigo, and Steele, 2011; Herzog, Scheuren, and Winkler, 2007; Jenkins et al., 2007). Record linkage is often a complicated process and one that varies according to the criteria being used to establish links (e.g. name, address, phone number, social security number, person identification number, etc. and some combination of these identifiers) and the specific records being linked (involving such matters as duplications, omissions, and errors across databases). The basis for establishing links needs to be clearly documented and research is needed on the accuracy of record linkage. In addition, since both many databases and privacy regulations are country-specific, what can be done will vary across nations.

Second, the accuracy of the information from the various databases needs to be assessed. AD may be wrong because of wrong linkages, because the information is out-of-date (e.g. reflecting a former resident), or because it was never correct. If the AD are to be added to the survey record and directly used in analysis, one would want it to be highly reliable and perhaps be at least as reliable as data collected in the survey itself. But if the AD are only to be used to assess and adjust for nonresponse bias, the threshold of usability will be lower. One can assess whether sample units with a specific attribute are under- or over-represented among respondents, even if some inaccuracy in their actual, current characteristics exists. One way of assessing the accuracy and reliability of databases is to cross-check reported characteristics across sources. While correlated error can certainly exist across different sources, it would be useful to know their level of consistency. Another way would be to compare information from the survey of respondents to information on the respondents from databases. This would of course only generalize to respondents and the survey data would not always be correct, but this should give some sense as to the reliability and accuracy of specific variables from particular databases for all sample cases (respondents and non-respondents). When disagreements exist both among databases and between databases and the survey, research should attempt to adjudicate the inconsistencies and determine what the best information is.

Clearly too little is typically known about the procedures and criteria used in compiling many databases. As one workshop participant remarked, many databases are "black boxes" that do not disclose how they are constructed and what rules are followed. As Smith and Kim (2009) note, "(M)ost databases are limited in the documentation that they supply about their data. The original source of information is often not indicated nor is its recency. Likewise, quality-control procedures are rarely detailed. Definitions, data-procurement procedures, and quality-control procedures are often not indicated. Even obtaining limited information usually involves considerable digging and/or special requests from the providers. However, database providers will usually clarify matters and often supply additional documentation upon request. Also, additional information comes from using the databases and becoming familiar with their features through application and comparison. The impediment was usually not that information was being withheld to cover up poor procedures or serious flaws, but that the information and documentation had never been compiled, had not been prepared for dissemination, or was restricted to protect proprietary interests."

Paradata

Standard process paradata should be routinely collected and utilized to help manage the data-collection effort (its intrinsic purpose), to reduce nonresponse, and to detect and adjust for nonresponse bias. The protocol used in collecting paradata should be well-documented and quality-control measures to ensure their reliability should be implemented and also well-documented. During data collection, paradata should be used to model or predict case-level and overall survey outcomes and to optimize survey effort to reduce both overall nonresponse and/or nonresponse bias such as through responsive-design procedures (Groves and Heeringa, 2006; Groves et al., 2008; Laflamme, Maydan, and Miller, 2008; Mohl and Laflamme, 2007; Wagner, 2008). Also, monitoring key outcome variables and how they are changing across call attempts can indicate if results have stabilized and signal whether further data-collection efforts will materially change the profile of respondents (Wagner, 2008).

Process paradata that are collected for operational reasons as part of data collection including characteristics of interviewers from employment files (e.g. demographics, experience, training, production figures) should be routinely collected, appended, and utilized for nonresponse analysis. For example, the ESS includes records-of-calls data and the GSS appends information on the characteristics of interviewers.

Extra/observational paradata should be collected as long as they do not materially interfere with the prime interviewer task of conducting interviews. Given the extra costs associated with non-process paradata, research needs to determine what is the best set of variables to collect. Research is also needed on the verification and standardization of observational paradata and how to maximize their reliability. Additionally, research is needed on the best way to use process data. Records-of-calls data are especially complex and more work is required to determine the best ways to record, categorize, and analyze these data.

Given that the state of research on using process paradata has been limited to date, it is desirable that the paradata be made available in detailed formats including variables for each call attempt, verbatims of reasons for refusing, and other specifics rather than reducing the paradata to a few summary measures such as number of attempts per case. The paradata should be routinely appended to survey data as the ESS has done.

Paradata have several possible advantages over other types of AD. First, they are contemporaneous with the survey, while some other AD may be quite dated (and even outdated). Second, many paradata components are clearly directly related to survey response/nonresponse and as such are likely to contribute to propensity models (Casas-Cordero, 2010). Third, observational and other non-process paradata can be customized to fit the content and design of specific surveys and as information on the utility of particular paradata accumulates, surveys can be designed to include the most useful paradata. However, while it is a definite advantage that paradata can be customized to fit the topic and design of a particular survey, it is desirable to have general paradata measures that can be consistently applied across surveys for comparability.

Of course using paradata has a number of limitations. First, there are various costs involved in collecting paradata. Only process data are essentially free and if extra effort is devoted to their collection for the purpose of studying nonresponse rather than to facilitate data collection, then they both add costs and potentially distract from obtaining interviews (Kennickell, Mulrow, and Scheuren, 2011; Kreuter and Kohler, 2009). Observational and other paradata that are not collected to assist data collection clearly have added costs and may be even more likely to distract interviewers from their prime mission of securing interviews. Second, while theoretically paradata can be complete and exist for all sample cases, there are often missing data and that tend to be greater for non-respondents than for respondents (Stoop et al., 2010b). Such missing data naturally hinders their use in nonresponse analysis. Third, paradata may have low reliability and/or inconsistent reliability across surveys. For example, the Survey of Consumer Finance ceased the collection of AD by interviewers "because the inter-interviewer consistency of coding was too low… (Kennickell, Mulrow, and Scheuren, 2011)." The ESS has also been concerned about the variable recording of interviewer observations both across interviewers and more broadly across surveys in different countries (Matsuo and Loosveldt, 2010; Staehli, 2011). Much work is needed on the quality of AD in general and of each of the types of AD (e.g. sample frame, linked databases, paradata) (West and Kreuter, forthcoming; Casas-Cordero, 2010). Fourth, the richness of paradata varies greatly by mode. Observational data are most accessible for in-person surveys, very restricted for telephone surveys, and essentially non-existent for self-completion survey (e.g. postal surveys). Even process data are usually quite limited for most self-completion modes.

Aggregate Data

In addition to case-level data, aggregate-level data should be routinely used for nonresponse analysis (Smith and Kim, 2009). In general population surveys, aggregate data usually refer to information representing geographic units larger that an individual residence. For example, public-health surveys have increasingly used neighborhood-level measures to supplement the respondent-level data (Kawachi and Berkman, 2003). For other populations the aggregate data might represent an organizational unit such as a factory, school, or NGO chapter.

More research is needed on determining what levels of aggregation are most useful for nonresponse analysis, but in general having the lowest level of aggregation possible is most desirable. In addition, it would be valuable to have the same or similar variables at various levels of aggregation. For example, nonresponse might be greater in poor neighborhoods located in poor communities. When using case-level and aggregate data, appropriate multi-level models need to be utilized.

Besides the obvious impediment that aggregate data only refer to some larger units and not to the characteristics of either individual respondents or households, there are other limitations. First, the information may be dated. For example, data from the United States Decennial Census could be up to 12 years old before new data are available. Data based on the American Community Survey are more current being available over a five-year span for Census tracts and for shorter periods for larger geographical units. Second, much information may be available at only higher levels of aggregation (e.g. county vs. Census tract). This often means that one has to both use less geographically-focused data than desirable and that different variables are measured at different levels of aggregation leading to both theoretical and measurement complications. Third, while the advance of GIS and the increased assignment of residences to city-style addresses have greatly facilitated the collection of spatial data and their linkage to aggregate information, a number of challenges remain and some difficulties have grown. For example, matching Census data to other data collected by zip code is complicated and involves some estimating and fitting procedures (Smith and Kim, 2009). Geographic matching with non-ABS samples is even more difficult. In particular, in the United States the expansion of cell phones and their portability makes current assignment of geographic data to RDD samples using both landlines and cell phones highly problematic (see Appendix 3).

Interviewer effects and area effects are often hard to disentangle since there is often a heavy overlap or confounding between the two and interpenetrating and other designs to separate the effects are usually not available (Durrant, D'Arrigo, and Steele, 2011; O'Muircheartaigh and Campanelli, 1999; Schnell and Kreuter, 2005; Sturgis, Brunton-Smith, and Williams, 2011). More research using appropriate experimental designs is needed to better sort out these different effects.

Information on both geographic area and on interviewers should be routinely appendable to case-level survey data and not only added via special extra effort. Automatically having this information available means that costs and delay are minimized and, as a result, their utilization would greatly increase. Building these data into the standing system would lower the cost of using them in any particular survey, make their use simple and standard, and, by encouraging their use, advance research on nonresponse bias. Similarly, paradata collected during the data-collection effort should be routinely added to the survey file. Paradata should not be an extra that can be added through special effort, but information that is processed and appended as a standard product of conducting a survey.

AD in General

Adding more AD increases the likelihood of deductive disclosure and thus potentially undermines confidentiality. Files with extensive AD would either have to be restricted, sensitive-data files and/or the AD would have to be partly masked to eliminate the possibility of deductive disclosure. Data files can be thought of as falling along a sensitivity continuum from public-use files that have been designed to ensure that respondent identity cannot be broached via deductive disclosure or other means to in-house files with explicit respondent identifiers. In between these are restricted-access data files which contain more detailed information than the public files, but lack explicit identifiers. These may be accessible to researchers via sensitive-data access agreements and/or via data enclaves.

The increased use of AD will greatly improve our understanding of nonresponse bias and also promote ancillary advances in such areas as sample design and contextual analysis. This advancement will necessitate expanding survey-research expertise into several new areas. For example, this would involve such matters as gaining more familiarity with many auxiliary databases, learning about using GPS and GIS, better mastery of multi-level, geographical data, and improving record linkage procedures. In brief, it means that the repertoire of skills and expertise necessary for conducting survey research needs to expand.

Propensity Models

To detect and adjust for nonresponse bias one wants to identify propensity variables that predict a sample case's likelihood of becoming a respondent (Groves, 2006; Groves, Wagner, and Peytcheva, 2007; Olson, 2006). These may consist of general variables that tend to contribute to nonresponse across surveys in general (e.g. mistrust of surveys, dislike of intrusions, urbanicity) or survey-specific variables that are related to the content of a particular survey (e.g. lack of interest in politics for a pre-election survey; reluctance to discuss religion for a survey on religion), to survey mode (hard of hearing in a telephone survey, illiteracy in a postal survey), or to some other survey-specific attribute. So far the development of optimal propensity models has been limited. AD variables that both strongly predict propensity and relate to substantive variables of interest have been hard to identify (Kreuter et al., 2010; Kreuter and Olson, 2011). As Patrick Sturgis remarked, "We really lack good evidence of the extent to which geo-coded and paradata drive response propensities and are related to the sorts of things we are interested in measuring in surveys….I think the field is also in greater need of clear conceptual thinking regarding the causal system we are interested in, at the moment the approach is rather too much of the 'kitchen sink' variety…" A meta-analysis of the factors in general and AD variables in particular that are related to nonresponse bias would help to consolidate existing findings and best direct future research.

Clearly, the outcome of call attempts and the reason for non-cooperation in general and refusals in particular relate to nonresponse. If these factors are not related to nonresponse on variables of key, substantive interest, then this may indicate that the factors causing nonresponse are not related to these variables and thus that there is no meaningful nonresponse bias. But alternatively since the paradata measures will not produce complete models of nonresponse, it may be that nonresponse bias is occurring but the propensity models are too limited to detect it. Process paradata are much more likely to be related to the general causes of nonresponse and much less likely to be related to specific substantive variables. Their value is more in predicting nonresponse in general.

Generally, low correlations between available variables and survey nonresponse mean that propensity scores would predict little of the variance in responding. This would often be interpreted to mean that we lack strong variables to predict nonresponse and, as a result, we do not have a good propensity model. However, if responding was largely a quasi-random process (depending on transitory decisions and other idiosyncratic factors), then there would be few, if any, strong predictors and no propensity model would explain much of nonresponse. Showing that variables with a good theoretical basis for predicting nonresponse do not actually correlate with nonresponse might be seen as indicating that nonresponse is largely a quasi-random process and that there would therefore be relatively little nonresponse bias for most variables. Alternatively, if one had AD variables that were closely related to the substantive variables and showed that they did not vary across respondents and non-respondents, this would also be an indicator that nonresponse bias might be either limited in magnitude or at least not related to the target variables.

The use of different types of AD from different sources in multivariate models would help to develop general models of nonresponse and identify which approaches are most robust and generalizable. If different types of AD (e.g. frame, linked databases, paradata, etc.) or different measures from the various sources produce consistent results, we will gain a general understanding of the biases resulting from nonresponse. Since different variables will tend to come from different sources, multivariate models will also tend to be inter-source models. For example, with multi-stage, area-probability, ABS samples in the United States, much aggregate data and most demographics will come from the Census component of the sample frame, case-level data (respondent and/or household) will come from the postal listings component of the sample frame, linked databases, and paradata. Among the case-level data, measures closely related to survey-/respondent-related behavior will typically come from paradata while demographic data will tend to come from either paradata observations (including "partial" cases with household enumeration, but no interview) or from linked databases. Other types of information such as electoral registration, voting records, and political contributions) will come exclusively from linked databases (Smith and Kim, 2009).

In selecting AD one should consider both theoretical and empirical reasons for expecting the measures to be correlated with propensity scores and/or the key dependent variables. If a survey has a limited number of key variables (e.g. such as the unemployment rate for the Current Population Survey), the most relevant AD would relate to that key measure. But in surveys with no particular focus (or equal focus across many different variables, e.g. the GSS or ESS), then examining a limited set of substantively-related variables is neither possible nor useful. As Bethlehem, Cobben, and Schouten (2011) note, "when the nonresponse adjustment focused on single target variables, completely different choices can be made."

Resources for Nonresponse Analysis

One particularly useful aggregate bit of information is the Census-tract information based on the previous decennial Census and now the ACS that creates hard-to-count scores (HTCS) for each area. A wide range of Census data have been analyzed to predict what mail return rates will be by area. In 2000, the mail-return rates differed by 26.5 percentage points across Census tracts from the top and bottom deciles on HTCS (Bruce and Robinson, n.d.; US Census Bureau, 2009). Information is also available on the Census-tract-level, mail-return rate itself from the 2000 Census (US Census Bureau, 2000) and it is expected that similar figures will be available for the 2010 Census tracts.[2] In addition, some of this information is available at the block-group level. Related Census-tract level, nonresponse data may become available from recent ACSs. While relating primarily to mail-return rates and other Census related outcomes, these may well identify areas in which nonresponse to surveys in general are higher than average.

Creating case-specific information from prior survey efforts for specific addresses and phone numbers would facilitate the study of nonresponse. This might be done by starting with very large data collections (e.g. the National Immunization Survey or the American Community Survey) and then by getting professional and trade associations (e.g. the American Association for Public Opinion Research and/or the Council of American Survey Research Organizations) to encourage collaboration amongst members.

Existing aggregate data on propensity of areas to response should be used to target the more challenging areas perhaps by assigning the best interviewers, giving interviewers smaller assignments, having field supervisors more closely monitor production on these areas, etc.

Nonresponse Bias Measures

Besides being used to estimate nonresponse bias on particular variables, AD can be used to calculate general measures of nonresponse bias or representativity (R). These can be measures at the survey level or specific measures at the variable or estimate level (Groves, Brick et al., 2008; Kreuter, 2009 & forthcoming). The survey-level indicators include various proposed R-indexes or indicators (Cobben and Schouten, 2008; Groves, Brick et al., 2008; Schouten, 2007). Since applications of the various representativeness measures have been rather limited so far, more research is need to assess their various statistical and empirical merits in general and also to evaluate how to combine individual measures into a general, summary measure of representativeness. Currently, the most concentrated

[2] See interactive map of 2010 mail-return rates at http://2010.census.gov/2010census/take10map/

research effort along this line is the RISQ project supported by the European Union (http://www.risq-project.eu/).

Summary

Nonresponse is a serious problem in surveys and nonresponse bias is an important component in total survey error. Fortunately, multi-level AD from sample frames, linked databases, paradata, and related sources can be utilized to both detect and adjust for nonresponse bias. By fully exploiting such available data at both the case- and aggregate-levels, a better understanding of and adjustment for nonresponse bias can be achieved. Theoretical insight into the causes of nonresponse, empirical mastery of the various AD sources and variables, and statistical and methodological advances in the use of multi-level and multi-source models are needed to achieve optimal improvement. The recommendations of this workshop should facilitate that effort.

Figure 1: AD Available in General Population Surveys

Example (United States/ABS: Postal List + Census Statistics)

	Sample Frame	Ancillary/Linked Sources	Paradata Routine/Process	Paradata Added/Observed
Respondent	None	If R known	Refusal Reason	Race/Gender
Household/Address	Postal Info	GIS/Voting	ROC	Housing/HU Composition
Neighborhood	BG/Tract	Zip Code	If Aggregated	Trash/Broken Windows
Community	Census	Government	If Aggregated	If aggregated
County/Metro/etc.	Census	Vote Results	If Aggregated	if Aggregated

Table Notes: BG=Block Groups; ROC=Records of Calls; GIS=Geographic Information System; HU=Housing Unit

References

Abraham, Katherine G.; Maitland, Aaron; and Bianchi, Suzanne M., “Nonresponse in the American Time Use Survey: Who Is Missing from the Data and How Much Does it Matter? Public Opinion Quarterly, 70 (2006), 676-703.

Amaya, Ashley; Skalland, Ben; and Wooten, Karen, “What’s in a Match,” Survey Practice, December, 2010.

Bethlehem, Jelke G., “Weighting Nonresponse Adjustments Based on Auxiliary Information,” in Survey Nonresponse, edited by Robert M. Groves et al. New York: John Wiley & Sons, 2002.

Bethlehem, Jelke; Cobben, Fannie; and Schouten, Barry, Handbook of Nonresponse in Household Surveys. New York: John Wiley & Sons, 2011.

Biemer, Paul and Peytchev, Andy, “Nonresponse Boas Correction in Telephone Surveys Using Census Geocoding: An Evaluation of Error Properties,” Paper presented to the International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys, Chicago, June, 2011.

Billiet, Jaak, “Non-response Bias in Cross-National Surveys: Design for Detection and Adjustment in the ESS,” ASK: Research and Methods, 18 (2009), 3-43.

Blom, Annelies G. “Nonresponse Bias Adjustments: What Can Process data Contribute?” ISER Working Paper 2009-21, University of Essex, 2009.

Bruce, Antonio and Robinson, J. Gregory, “Tract Level Planning Database with Census 2000 Data,” Washington, DC: Census Bureau, n.d.

Cantor, David and Cunningham, Patricia, “Methods for Obtaining High Response Rates in Telephone Surveys,” in Studies of Welfare Populations: Data Collection and Research Issues, edited by C. Citro et al. Washington, DC: National Academy Press, 2002.

Casas-Cordero, Carolina, “Neighborhood Characteristics and Participation in Household Surveys,” Unpublished Ph.D. dissertation, University of Maryland, 2010.

Cho, Young Ik, “Using Community-level Correlates to Evaluate Nonresponse Effects in a Telephone Surveys,” Paper presented to the C_AAPOR Workshop on Nonresponse Bias in Household Surveys,” March, 2007.

Christian, Leah; Dimock, Michael; and Keeter, Scott, “Accurately Locating Where Wireless respondents Live Requires More than a Phone Number,” Pew Research Center Publication, July, 2009.

Cobben, Fannie, Nonresponse in Sample Surveys: Methods for Analysis and Adjustment. The Hague: Statistics Netherlands, 2009.

Cobben, Fannie and Schouten, Barry, “An Empirical Validation of R-Indicators,” Discussion Paper 08006, Statistics Netherlands, 2008.

Cobben, Fannie and Schouten, Barry, “A Follow-up with Basic Questions of Nonrespondents to the Dutch Labour Force Survey,” Discussion Paper 07011. Statistics Netherlands, 2007

Conrad, Frederick et al., "To Agree or Not to Agree? Impact of Interviewer Speech on Survey Participation Decisions," Paper presented to the American Association for Public Opinion Research, Chicago, May, 2010.

Couper, Mick P. and de Leeuw, Edith, "Nonresponse un Cross-Cultural and Cross-National Surveys," in Cross-Cultural Survey Methods, edited by Janet A. Harkness, Fons J.R. van de Vijver, and Peter Ph. Mohler. Hoboken, NJ: John Wiley & Sons, 2003.

Couper, Mick P. and Lyberg, Lars E., "The Use of Paradata in Survey Research," Proceedings of the International Statistical Institute, Sydney, April, 2005.

D'Arrigo, Julia; Durrant, Gabriele B.; and Steele, Fiona, "Analysing the Process Leading to Cooperation or Refusal Using Call Record Data: A Multilevel Multinominal Modelling Approach," unpublished paper, University of Southampton, 2011.

Davern, Michael. "Evaluating Linked Survey and Administrative Data for Policy Research," Paper presented to the Federal Committee on Statistical Methodology, Washington, DC, November, 2006.

Davern, Michael et al., "A Partially Corrected Estimate of Medicaid Enrollment and Uninsurance: Results from an Imputational Model Developed off Linked Survey and Administrative Data," Journal of Economic and Social Measurement, 34 (2009), 219-240.

De Leeuw, Edith and de Heer, W., "Trends in Household Nonresponse: A Longitudinal and International Comparison," in Survey Nonresponse, edited by Robert M. Groves et al. New York: John Wiley & Sons, 2002.

Dennis, J. Michael, "KnowledgePanel: Processes and Procedures Contributing to Sample Representativeness and Tests of Self-Selection Bias," Research Note, Knowledge Networks, 2010.

Dennis, J. Michael, "Using Ancillary Data Available for Address-Based Sampling to Measure Self-Selection Bias" Paper presented to the International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys, Chicago, June, 2011.

DiSogra, Charles, "The Promise of Address-based Sampling Using Ancillary Information," Knowledge Networks, 2010.

DiSogra, Charles; Dennis, J. Michael; and Fahimi, Mansour, "On Quality of Ancillary Data Available for Address-Based Sampling," Paper presented to the American Statistical Association, Vancouver, August, 2010.

Dixon, John, "Using Census Match Data to Evaluate Models of Survey Nonresponse," Paper presented to the American Statistical Association, Toronto, August, 2004.

Durrant, Gabriele B. et al., "Effects of Interviewer Attitudes and Behaviors in Household Surveys," Public Opinion Quarterly, 74 (2010), 1-36.

Durrant, Gabriele B.; D'Arrigo, Julia, and Steele, Fiona, "Using Field Process Data to Predict Best Times of Contact Conditioning on Household and Interviewer Influences," Journal of the Royal Statistical Society, Series A, (2011), forthcoming.

Durrant, Gabriele B. and Steele, Fiona, "Multivariate Modelling of Refusal and Non-Contact in Household Surveys: Evidence from Six UK Government Surveys," Journal of the Royal Statistical Society, A, 1972 (2009), 361-381.

English, Ned; Schwartz, Lisa; and Goble, Lisbeth., "Where, Exactly, Are the RDD Non-respondents: A GIS-Based Approach to Quantifying Potential Bias," Paper presented to the American Association for Public Opinion Research, Miami Beach, May, 2010.

Fone, David L.; Lloyd, Keith; and Dunstan, Frank D., "Measuring the Neighbourhood Using UK Benefits Data: A Multivariate Analysis of Mental Health Status," BMC Public Health, 7 (2007), 69ff.

Foster, Kate, "Evaluating Non-response on Household Surveys: Report of a Study Linked to the 1991 Census," Government Statistical Service Methodology Series No. 8, Office of National Statistics, 1998.

Fraboni, Romina; Rosina, Alessandro; Orsini, Sante; and Baldazzi, Barbara, "A Micro Analysis of the Macros Differences in Refusal Risk among Metropolitan Areas: The Case of the Italian Multipurpose Survey," Statistical Methods & Applications, 14 (2005), 283-295.

Freeth, Stephanie; White, Amanda; and Martin, Jean, "Using Matched Census-Survey Records to Evaluate the Quality of Survey Data," Proceedings of Statistics Canada Symposium, 2001.

Garrett, Joe; Dennis, J. Michael; and DiSogra, Charles, "Non-response Bias: Recent Findings from Addressed-based Panel Recruitment," Knowledge Networks report, 2010.

Gfroerer, Joseph; Lessler, Judith; and Parsley, Teresa, "Studies of Nonresponse and Measurement Error in the National Household Survey of Drug Abuse," In L. Harrison & A. Hughes (Eds.), The validity of self-reported drug use: Improving the accuracy of survey estimates (NIH Publication No. 97-4147, NIDA Research Monograph 167, pp. 273-295). Rockville, MD: National Institute on Drug Abuse, 1997.

Goodman, Anna and Gatward, Rebecca, "Who Are We Missing?" Area Deprivation and Survey Participation," European Journal of Epidemiology, 23 (2008), 379-387.

Groves, Robert M., "Nonresponse Rates and Nonresponse Bias in Household Surveys," Public Opinion Quarterly, 70 (2006), 646-675.

Groves, Robert M., J. M. Brick, M. Couper, W. Kalsbeek, B. Harris-Kojetin, F. Kreuter, B. Pennell, T. Raghunathan, B. Schouten, T. Smith, R. Tourangeau, A. Bowers, M. Jans, C. Kennedy, R. Levenstein, K. Olson, E. Peytcheva, S. Ziniel, and J. Wagner, "Alternative Practical Measures of Representativeness of Survey Respondent Pools," Survey Practice, 1 (October, 2008), at http://surveypractice.files.wordpress.com/2008/11/survey-practice-october-2008.pdf

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

Groves, Robert M. et al., "Responsive Design for Household Surveys: Illustration of Management Interventions Based on Survey Paradata," Unpublished report, University of

Michigan, 2008.

Groves, Robert M. and Heeringa, Steven G., "Responsive Design for Household Surveys: Tools for Actively Controlling Survey Nonresponse and Costs," Journal of the Royal Statistical Society: Series A (Statistics in Society), 169 (2006), 439–457.

Groves, Robert; Wagner, James; and Peytcheva, Emilia, "Use of Interviewer Judgments about Attributes of Selected Respondents in Post-Survey Adjustments for Unit Nonresponse: An Illustration with the National Survey of Family Growth," Proceedings of the Section on Surveys Research Methods, American Statistical Association. Alexandria, VA: ASA. 2007.

Guterbock, Thomas M.; Hubbard, Ryan A.; and Holian, Laura M., "Community Attachment as a Predictors of Survey Non-Response," University of Virginia, unpublished report, 2006.

Hansen, Sue Ellen, "CATI Sample Management Systems," in Advances in Telephone Survey Methodology, edited by James M. Lepkowski et al. Hoboken, NJ: John Wiley & Sons, 2008.

Herzog, Thomas N.; Scheuren, Fritz; and Winkler, William E., Data Quality and Record Linkage Techniques. New York: Springer, 2007.

Jenkins, Stephen P.; Lynn, Peter; Jaekle, Annette; and Sala, Emanuela, "The Feasibility of Linking Household Survey and Administrative Record Data: New Evidence from Britain," International Journal of Social Research Methodology, 11 (2007), 29-43.

Johnson, Paul, "Comparing ABS vs. Landline RSS Sampling Frames on the Phone Mode," Survey Practice, June, 2010.

Johnson, Timothy P.; Lee, Geon, and Cho, Young Ik, "Examining the Association between Cultural Environments and Survey Nonresponse," Survey Practice, June, 2010.

Johnson, Timothy P.; Cho, Young Ik; Campbell, Richard T.; and Holbrook, Allyson L., "Using Community-Level Correlates to Evaluate Nonresponse Effects in a Telephone Survey," Public Opinion Quarterly, 70 (2006), 704-719.

Johnson, Timothy P., Holbrook Allyson L., Cho, Young Ik, and Bossarte, Robert M. "Nonresponse error in injury-risk surveys," American Journal of Preventive Medicine, *31* (2006), 427–36.

Joye, Dominique et al., "A Geography of Non-response?" Paper presented to the International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys, Chicago, June, 2011.

Kawachi, Ichiro and Berkman, Lisa F., eds., Neighborhoods and Health. New York: Oxford University Press, 2003.

Kennickell, Arthur; Mulrow, Ed; and Scheuren, Fritz, "Paradata or Process Modeling for Inference," Paper presented to the International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys, Chicago, June, 2011.

King, John and Foster, Kate, "Understanding Household Survey Non-response through Geo-Demographic Coding Schemes," Survey Methodology Bulletin, No. 43 (July, 1998), 28-35.

Kreuter, Frauke, "The Nonresponse Challenge," Annals of the American Academy of Political and Social Science, forthcoming

Kreuter, Frauke, "Survey Methodology: International Developments," RatSWD Working Paper No. 59, German Council for Social and Economic Data, 2009.

Kreuter, Frauke and Casas-Cordero, Carolina, "Paradata," RatSWD Working Paper No. 136, German Council for Social and Economic Data, 2010.

Kreuter, Frauke and Kohler, U., "Analyzing Contact Sequences in Call Record Data: Potential and Limitations of Sequence Indicators for Nonresponse Adjustments in the European Social Survey," Journal of Official Statistics, 25 (2009), 203-226.

Kreuter, Frauke, Lemay, Michael, and Casas-Cordero, Carolina, "Using Proxy Measures of Survey Outcomes in Post-Survey Adjustments: Examples from the European Social Survey (ESS), Paper presented to the American Statistical Association, Salt Lake City, August, 2007.

Kreuter, Frauke et al., "Using Proxy Measures and Other Correlates of Survey Outcomes to Adjust for Non-response: Examples from Multiple Surveys," Journal of the Royal Statistical Society, A, 173 (2010), 389-407.

Kreuter, Frauke; Mueller, Gerrit; and Trappmann, Mark, "Nonresponse and Measurement Error in Employment Research: Making Use of Administrative Data," Public Opinion Quarterly, 74 (2010), 880-906.

Kreuter, Frauke and Olson, Kristen, "Multiple Auxiliary Variables in Nonresponse Adjustment," Sociological Methods & Research, 40 (2011), 311-332

Laflamme, Francois; Maydan, Mike; and Miller, Andrew, "Using Paradata to Actively Manage Data Collection Survey Process," Proceedings of the Survey Research Methods Section, American Statistical Association (2008). Alexandria, VA: ASA, 2008.

Lee, Sunghee et al., "Exploring Nonresponse Bias in a Health Survey Using Neighborhood Characteristics," American Journal of Public Health, 99 (2009), 1811-1817.

Linn, I-Fen and Schaeffer, Nora Cate, "Using Survey Participants to Estimate the Impact of Nonparticipants," Public Opinion Quarterly, 59 (1995), 236-258.

Link, Michael W., "Address Based Sampling: What Do We Know So Far?" American Statistical Association Webinar, November, 2010.

Lynn, Peter, "PEDAKSI: Methodology for Collecting Data about Survey Non-respondents," Working Paper for the Institute for Social and Economic Research, 2002-05, University of Essex, 2002.

Lynn, Peter at al., "The Effect of Extended Interviewer Efforts on Nonresponse Bias," in Survey Nonresponse, edited by Robert M. Groves et al., New York: John Wiley & Sons, 2002.

Lynn, Peter and Nicolaas, Gerry, "Making Good Use of Survey Paradata," Survey Practice, 3 (2010), http://surveypractice.org/2010/04/14/making-good-use-of-survey-paradata/

Maitland, Aaron; Casas Cordero, Carolina; and Kreuter, Frauke, "An Exploration into the Use of Para Data for Nonresponse Adjustment in a Health Survey," Paper presented to the American Statistical Association, Denver, August, 2008.

Martikainen, Pekka et al., "Does Survey Non-response Bias the Association between Occupational Social Class and Health?" Scandinavian Journal of Public Health, 23 (2007), 212-215.

Matsuo, Hideko and Loosveldt, Geert, "ESS Round 5: Contact Forms & Interviewer Training for Observable Data," Paper presented to the ESS National Coordinator's Meeting, the Hague, May, 2010.

Matsuo, Hideko et al., "Measurement and Adjustment of Non-response Bias Based on Non-response Surveys: The Case of Belgium and Norway in the European Social Survey Round 3," Survey Research Methods, 4 (2010a), 165-178.

Matsuo, Hideko et al., "Respondent-based Quality Assessment of ESS Round 4: Results for 30 Countries Based on Contact Files," Final report, Centrum voor Sociologisch Onderzoek, 2010b.

Mohl, Chris and Laflamme, Francois, "Research and Responsive Design Options for Survey Da Collection at Statistics Canada," Proceedings of the Survey Research Methods Section , American Statistical Association (2007), Alexandria, VA: ASA, 2007.

Mulrow, Ed and English, Ned, "Combining Survey and Administrative Information," Paper presented to the NORC Open House, April, 2011, Chicago.

Nicoletti, Cheti and Peracchi, Franco, "Survey Response and Survey Characteristics: Microlevel Evidence from the European Community Household Panel," Journal of the Royal Statistical Society, Series A, 168 (2005), 763-781.

Olson, Kristen, "An Investigation of the Nonresponse-Measurement Error Nexus," Unpublished Ph.D. dissertation, University of Michigan, 2007.

Olson, Kristen, "Paradata in Nonresponse Adjustment," Annals of the American Academy of Political and Social Science, forthcoming.

Olson, Kristen, "Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias," Public Opinion Quarterly, 70 (2006), 737-758

Olson, Kristen and Lindsey Witt. 2011, "Are we keeping the people who used to stay? Changes in Correlates of panel survey attrition over time." Social Science Research. Published online March 8, 2011, doi:10.1016/j.ssresearch.2011.03.001

Olson, Kristen and. Groves, Robert M., "An Examination of Within-Person Variation in Response Propensity over the Data Collection Field Period." Journal of Official Statistics. Paper accepted with minor revisions June 2010.

Olson, Kristen; Wood, Heather; and Smyth, Jolene, "Does Providing Respondents with Their Preferred Survey Mode Really Increase Participation?" Paper presented at the American Association for Public Opinion Research annual meeting, Chicago, IL, May 2010.

O'Muircheartaigh, Colm and Campenelli, Pamela, "A Multilevel Exploration of the Role of Interviewers in Survey Non- response," Journal of the Royal Statistical Society, Series A, 162 (1999), 437-446.

Paul, Jamie and Grady, Sarah, "Summary of Address Based Sampling Papers from 2009 AAPOR Conference," Survey Practice, 3 (2010), http://surveypractice.org/2010/04/14/summary-of-address-based-sampling-papers/

Peytchev, Andrew and Olson, Kristen, "Using Interviewer Observations to Improve Nonresponse Adjustments: NES 2004," Paper Presented to the American Statistical Association, Salt Lake City, August, 2007.

Peytchev, Andy et al., "Reduction of Nonresponse Bias in Surveys through Case Prioritization," Survey Research Methods, 4 (2010), 21-29.

Safir, Adam et al., "Effects on Survey Estimates from Reducing Nonresponse," Proceedings of the American Statistical Association, 2001. Washington, DC: ASA, 2001.

Salvo, Joseph J. and Lobo, Arun Peter, "The Effect of Neighborhood Characteristics on Nonresponse in the Bronx Test Site of the American Community Survey," Paper presented to the American Statistical Association, San Francisco, August, 2003.

Salvo, Joseph J. and Lobo, Arun Peter, "Moving from a Decennial Census to a Continuous Measurement Survey: Factors Affecting Nonresponse at the Neighborhood Level," Population Research Policy Review, 25 (2006), 225-241.

Sarndal, Carl-Erik and Lundstrom, Sixten, "Accessing Auxiliary Vectors for Control of Nonresponse Bias in the Calibration Estimator," Journal of Official Statistics, 24 (2008), 167-191.

Sarndal, Carl-Erik and Lundstrom, Sixten, Estimation in Surveys with Nonresponse. Chichester: John Wiley & Sons, 2005.

Sastry, Narayan and Pebley, Anne R., "Non-response in the Los Angeles Family and Neighborhood Survey," Labor and Population Program Working Paper Series 03-1, Rand, 2003.

Schnell, Rainer and Kreuter, Frauke, "Separating Interviewer and Sample-Point Effects," Journal of Official Statistics, 21 (2005), 389-410.

Schouten, Barry, "A Follow-up of Nonresponse in the Dutch Labour Force Survey," Discussion Paper 07004, Statistics Netherlands, 2007.

Schouten, Barry; Cobben, Fannie; and Bethlehem, Jelke, "Indicators for the Representativeness of Survey Response," Survey Methodology, 35 (2009), 101-113.

Shen, Min et al., "Census and Geographic Differences between Respondents and Nonrespondents in a Case-Control Study of Non-Hodgkin Lymphoma, "American Journal of Epidemiology, 167 (2008), 350-361.

Skalland, Benjamin J. and Blumberg, Stephen J., "Nonresponse Bias Analysis: 2005-06 National Survey of Children with Special Health Care Needs," NORC report, 2008.

Skalland, Benjamin J. and Blumberg, Stephen J., "Nonresponse in the National Survey of Children's

Health, 2007," Vital and Health Statistics, Series 2, forthcoming.

Smith, Tom W., "Contextualizing Social-Science Data," SBE 2020 White Paper, National Science Foundation, September, 2010.

Smith, Tom W., "Survey Non-Response Procedures in Cross-national Perspective: The 2005 ISSP Non-response Survey," Survey Research Methods, 1 (2007), 45-54.

Smith, Tom W. and Kim, Jibum, "An Assessment of the Multi-level Integrated Database Approach," GSS Methodological Report No. 116. Chicago: NORC, 2009.

Smith, Tom W. and Sokolowski, John, "Using Audio-Visuals in Surveys," in The Handbook of Emergent Technologies in Social Research, edited by Sharlene Hesse-Biber. Oxford: Oxford University Press, 2011.

Staehli, Michele Ernst, "Examples and Experiences from the Swiss Interviewer Training on Observable data (NEIGHBORHOOD CHARACTERISTICS) for ESS 2010 (R5)," Paper presented to the National Coordinators Meeting, Mannheim, April/May, 2011.

Steeh, Charlotte and Piekarski, Linda, "Accommodating New Technologies: Mobile and VoIP Communication," in Advances in Telephone Survey Methodology, edited by James M. Lepkowski et al. Hoboken, NJ: John Wiley & Sons, 2008.

Steele, Fiona and Durrant, Gabriele B., "Alternative Approaches to Multilevel Modelling of Survey Noncontact and Refusal," Working Paper M09/15, University of Southampton, 2009.

Stoop, Ineke A. L ., The Hunt for the Last Respondent: Nonresponse in Sample Surveys. The Hague: Social and Cultural Planning Office of the Netherlands. 2005.

Stoop, Ineke et al., Improving Survey Response: Lessons Learned from the European Social Survey. Chichester: John Wiley & Sons, 2010a.

Stoop Ineke et al., "Paradata in the European Social Survey: Studying Nonresponse and Adjusting for Bias," Paper presented to the American Statistical Association, August, 2010b, Vancouver.

Stoop, Ineke et al., "The Use of Paradata (and Other Auxiliary Data) for Nonresponse Adjustment in the European Social Survey (ESS)," Paper presented to the International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys, Chicago, June, 2011.

Stoop, Ineke; Billiet, Jaak; and Vehovar, Vasja, "Nonresponse Bias in a Cross-National Study," Paper presented to the International Statistical Institute, August, 2009, Durban, South Africa.

Sturgis, Patrick, "Nonresponse, Paradata, and Causal Inference," Paper presented to the International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys, Chicago, June, 2011.

Sturgis, Patrick; Brunton-Smith, Ian; and Williams, Joel, "Is Skill in Obtaining Contact and Cooperation Correlated with the Magnitude of Interviewer Variance?" Paper presented to the International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys, Chicago, June, 2011.

Trappmann, Mark and Mueller, Gerrit, "Paradata and Linked Administrative Data of the PASS Panel," Institute for Employment Research, University of Southampton, 2010.

Triplett, Timothy and Abi-Habib, Natalie, "Socio-Demographic Study of Telephone Nonrespondents," Paper presented to the FCSM Conference, November, 2005, Arlington. VA.

Tucker, Clyde and Lepkowski, James M., "Telephone Survey Methods: Adapting to Change," in Advances in Telephone Survey Methodology, edited by James M. Lepkowski et al. Hoboken, NJ: John Wiley & Sons, 2008.

Turrell, Gavin et al., "The Socio-economic Pattering of Survey Participation and Non-response Error in a Multilevel Study of Food Purchasing Behavior: Area- and Individual-level Characteristics," Public Health Nutrition, 6 (2003), 181-189.

US Census Bureau, "2010 Enhanced Planning Database," unpublished memo, October 20, 2009.

US Census Bureau, "Census 2000 Final Response Rates," 2000 at www.Census.gov/dmd/www/response/2000response.html

Van Goor, Henk; Jansma, Folbert; and Veenstra, Rene, "Differences in Undercoverage and Nonresponse between City Neighborhoods in a Telephone Survey," Psychological Reports, 96 (2005), 867-878.

Van Ingen, Erik; Stoop, Ineke; and Breedveld, Koen, "Nonresponse in the Dutch Time Use Survey: Strategies for Response Enhancement and Bias Reduction," Field Methods, 21 (2009), 69-90.

Voogt, Robert J.J. and van Kempen, Hetty, "Nonresponse Bias and Stimulus Effects in the Dutch National Election Study," Quality & Quantity, 36 (2002), 325-345.

Vehovar, Vasja and Zaletel, Metka, "The Matching Project in Slovenia: Who Are the Nonrespondents?" Tutkimuksia, 219 (1996), 193-203.

Wagner, James R., "Adaptive Survey Design to Reduce Nonresponse Bias," unpublished Ph.D. dissertation, University of Michigan, 2008.

West, Brady T. and Kreuter, Frauke, "Observational Strategies Associated with Increased Accuracy of Interviewer Observations: Evidence from the National Survey of Family Growth," Unpublished Paper, University of Michigan, 2011.

Yan, Ting and Raghunathan, Trivellore , "Using Proxy Measures of the Survey Variables in Post-Survey Adjustment in a Transportation Survey," Paper presented to the American Statistical Association, Salt Lake City, August, 2007.

Yu, Chong Ho et al., "A Data Visualization and Data Mining Approach to Response and Non-response Analysis in Survey Research," Practical Assessment, Research & Evaluation, 12 (2007), http://pareonline.net

West, Brady T. and Olson, Kristen, "How Much of Interviewer Variance Is Really Nonresponse Error Variance?" Public Opinion Quarterly, 74 (2010), 1004-1025.

Zaletal, Metka and Vehovar, Vasja, "Nonresponse and Socio-Demographic Characteristics of

Enumeration Areas," in Developments in Survey Methodology, edited by Anuska and Andrej Mrvar, Ljubljana: Metodoloski Zvezki, 2000.

Appendix 1: Workshop Agenda

An International Workshop on Using Multi-level Data from Sample Frames, Auxiliary Databases, Paradata and Related Sources to Detect and Adjust for Nonresponse Bias in Surveys

The workshop will held at NORC's Loop offices at 55 East Monroe St. Go to the 30th floor/room 3034. The NORC receptionist will direct you when you get off the elevator. If any issues come up, you can call Jibum Kim at 773-xxx-xxxx or email Tom W. Smith at smitht@norc.uchicago.edu.

Thursday June 2nd

9:00AM: Introduction and Welcome by Tom W. Smith

National Sample Frames in the US and Europe

Use of Aggregate/Geographical Data
- Using Addresses
- Using Phone numbers

Tom W. Smith and Jibum Kim, "Aggregate and Geographical Data for US Surveys"

Paul Biemer and Andy Peytchev, "Nonresponse Bias Correction in Telephone Surveys Using Census Geocoding: An Evaluation of Error Properties"

Dominique Joye, "A Geography of Non-response in Switzerland"

12:00 -1:00PM Lunch
1:00- 5:00PM Discussion Continues

Auxiliary/Linked Data
- Public/Administrative Data
- Commercial Databases

J. Michael Dennis, "Using Ancillary Data Available for Address-Based Sampling to Measure Self-Selection Bias"

Paradata
- Process/Survey Operations measures
- Observational/Extra Variables

Ineke Stoop Hideko Matsuo, Jaak Billiet, and Achim Koch, "The Use of Paradata (and Other Auxiliary Data) for Nonresponse Adjustment in the European Social Survey"

Arthur Kennickell, Edward Mulrow, and Fritz Scheuren, "Paradata or Process Modeling for Inference"

Patrick Sturgis, "Nonresponse, Paradata, and Causal Inference"

5:00PM – Workshop ends for day.
6:00PM – Group dinner
.

Friday June 3rd

9:00AM: Workshop continues: Focus on research agenda for the future
12:00 Noon: Workshop ends
12:00-1:00PM Lunch

Attendees

Paul Biemer, University of North Carolina
J. Michael Dennis, Knowledge Networks
Cheryl Eavey, National Science Foundation
Robert Groves, US Bureau of the Census
Timothy Johnson, University of Illinois, Chicago
Dominique Joye, University of Lausanne
Jibum Kim, NORC
Frauke Kreuter, University of Maryland & Institute for Employment Research
Kristen Olson, University of Nebraska
Colm O'Muircheartaigh, University of Chicago
Fritz Scheuren, NORC
Benjamin J. Skalland, NORC
Tom W. Smith, NORC (**Organizer**)
Ineke Stoop, Social and Cultural Planbureau
Patrick Sturgis, University of Southampton
Roger Tourangeau, University of Maryland

Appendix 2: AD Literature

A. Sample-Frame Data

Abraham, Maitland, and Bianchi, 2006 (CPS)
Amaya and Skalland, 2010 (postal)
Bethlehem, Cobben, and Schouten, 2011 (Population Register)
Blom, 2009 (Population Register)
Fraboni et al., 2005 (Census/local administrative registers)
Kreuter, Mueller, and Trappmann, 2010 (benefit recipients; address list)
Lin and Schaffer, 1995 (court records)
Link, 2010 (postal)
Lynn, 2003 (ACORN; Census)
Martkainen et al., 2007 (City records)
Sarndal and Lundstrom, 2008 (Population Register)
Schouten, Cobben, and Bethlehem, 2009 (Population Register)
Skalland and Blumberg, 2008 (phone listing)
Skalland and Blumberg, forthcoming (phone listing)
Smith and Kim, 2009 (Census; postal)
Stoop, Billiet, and Vehovar, 2009 (variable by country)
Van Ingen, Stoop, and Breedveld, 2009 (variable by country)
West and Olson, 2010 (government records)

B. Linked, Auxiliary Data

Amaya and Skalland, 2010 (Phone/MSG, Accurint)
Bethlehem, Cobben, and Schouten, 2011 (Social Statistical Database)
Cobben and Schouten, 2007 (various databases)
Disogra, Dennis, and Fahimi, 2010 (MSG)
Dixon, 2004 (Census)
Durrant, D'Arrigo, and Steele, 2011 (Census)
Durrant and Steele, 2009 (Census)
Fone, Lloyd, and Dunstan, 2007 (DWP benefits data)
Foster, 1998 (Census)
Fraboni et al., 2005 (family certificates)
Garrett, Dennis, and Disogra, 2010 (various, Experian, InfoUSA)
Groves, 2006 (general)
Groves and Couper, 1998 (Census)
Johnson, 2010 (phone)
Kreuter et al., 2010 (traffic database)
Kreuter, Mueller, and Trappmann, 2010 (government employment records)
Link, 2010 (phone)
Bethlehem, Cobben, and Schouten, 2011 (Population Register)Schouten, Cobben, and Bethlehem, 2009 (registry data/welfare)
Shen et al., 2008 (Census; water records)
Skalland and Blumberg, forthcoming (Phone/Census)

Smith and Kim, 2009 (various databases)
Stoop et al., 2010 (Population Register)
Vehovar and Zaletal, 1996 (various databases)

C. **Paradata (Process/Operational/Routine**)

Abraham, Maitland, and Bianchi, 2006
Bethlehem, Cobben, and Schouten, 2011
Blom, 2009
D'Arrigo, Durrant, and Steele, 2011
Durrant, D'Arrigo, and Steele, 2011
Durrant and Steele, 2009
Fraboni et al., 2005
Groves et al., 2008
Groves and Heeringa, 2006
Kreuter and Casas-Cordero, 2010
Kreuter, Couper, and Lyberg, 2010
Kreuter and Kohler, 2009
Kreuter et al., 2010
Laflamme, Maydan, and Miller, 2008
Lin and Schaffer, 1995
Maitland et al., 2008
Matsuo et al., 2010
Matsuo and Loosveldt, 2010
Mohl and Laflamme, 2006
Nicoletti and Peracchi, 2005
O'Muirchaertaigh and Campanelli, 1999
Peytchev and Olson, 2007
Peytchen et al., 2010
Trappmann and Gerrit, 2010
Stoop et al., 2011
Sturgis, 2011
Sturgis, Brunton-Smith, and Williams, 2011
West and Olson, 2010

D. Paradata (Extended)

Conrad et al., 2010 (audio recording)
Kreuter and Casas-Cordero, 2010 (audio recording/timing)
Smith and Sokolowski, 2011 (audio recording)

E. Paradata (Observational/Supplemental)

Billiet et al., 2009
Blom, 2009
Casas-Cordero, 2010

Durrant, D'Arrigo, and Steele, 2011
Kreuter and Koler, 2009
Kreuter, Lemay, and Casas-Cordero, 2007
Lynn, 2003 & 2010
Matsuo et al., 2010a; 2010b
Matsuo and Loosveldt, 2010
Peytchev and Olson, 2007
Staehli, 2011
Stoop et al., 2010
Stoop et al., 2011
West and Kreuter, 2011

F. Aggregate Data

Biemer and Peytchev, 2011 (phone)
Casas-Cordero, 2010 (observational paradata; Census tract)
Cho, 2007 (phone)
Cobben and Schouten, 2008 (urbanicity; region)
Durrant, D'Arrigo, and Steele, 2011; Durrant et al., 2010; Durrant and Steele, 2009 (Census)
England, Schwartz, and Goble, 2010 (Census tract)
Fone, Lloyd, and Dunstan, 2007 (government records)
Garrett, Dennis, and DiSogra, 2010 (various databases; Experian; Info USA)
Goodman and Gatward, 2008 (Census)
Guterbock, Hubbard, and Holian, 2006 (Census)
Johnson, 2006a; 2006b (phone)
Johnson. Lee, and Cho, 2010 (phone)
Joye et al., 2011 (communes)
King and Foster, 1998 (Mosaic, ACORN)
Lee et al., 2009 (phone; Census tract/county)
Lynn, 2003 (ACORN; postal)
Salvo and Lobo, 2006 (Census; government records)
Sastyr and Pebley, 2003 (Census tract)
Schnell and Kreuter, 2005 (election districts)
Shen et al., 2008 (Census; water records)
Skalland and Blumberg, forthcoming (phone/Census)
Skalland and Blumberg, 2008 (phone/Census)
Steele and Durant, 2009 (Census)
Stoop, Billiet, and Vehovar, 2009 (community-level data)
Trappmann and Mueller, 2010 (Mosaic; government statistics)
Triplett and Abi-Habib, 2003 (block group)
Turrell et al,, 2003 (Census)
Van Goor, Jansma, and Venestra, 2005 (government data)
Vehovar and Zaletal, 1996 (Census/urbanization)
Zatetal and Vehovar, 2000 (Census/CEA, Population Register, phone, government departments)

Appendix 3: Appending Geographic Codes to Telephone Number Samples

1. Even traditionally matching telephone numbers to geographic areas has been difficult because the smallest area defined by telephone number, the exchange, rarely closely maps to small geographic areas like Census tract and zip code. One study found that zip code imputed based on telephone number and that reported by respondents matched only 60% of the time (Johnson et al., 2006a). Even matches to larger units such as community and county are often not exact. Despite these difficulties, researchers often found it useful to link phone numbers to geographic information (English, Schwartz, and Goble, 2010; Johnson et al., 2006a & 2006b; Lee et al., 2009; Triplett and Abi-Habib, 2005).

2. For phone numbers listed in directories, addresses can be obtained and for those with city-style addresses, their Census tract, zip code, and other location information (e.g. longitude/latitude, GPS coordinates) can be determined (Guterbock, Hubbard, and Holian, 2006; Lee et al., 2009; Triplett and Abi-Habib, 2005). Unfortunately, many landlines and almost all cell phones are not listed in directories and thus no address is available. However, databases other than telephone directories do include additional landline and cell phone addresses (Smith and Kim, 2009). But directories and other databases may not capture recent moves.

3. Telephone portability and the expansion of cell phones have greatly increased the problem (Hansen, 2008; Johnson et al., 2006a & b; Steeh and Piekarski, 2008; Tucker and Lepkowski, 2008).

4. Landlines are not supposed to be ported outside the local exchange, but that may not always be the case. Even within local exchanges moves can be over appreciable distances and/or between quite different local neighborhoods.

5. Cell phone numbers often do not match residence even initially and can be transported across both exchanges and area codes.

6. In 2009, the Pew Research Center (Christian, Dimock, and Seeter, 2009) found that exchange-based and respondent-reported zip code agreed on county for 92% of landlines and 39% of cell phones. Also in 2009, NORC found that 94% of landlines agreed on county and 35% of cell phones did so.

References

Christian, Leah; Dimock, Michael; and Keeter, Scott, "Accurately Locating Where Wireless respondents Live Requires More than a Phone Number," Pew research Center Publication, July, 2009.

English, Ned; Schwartz, Lisa; and Goble, Lisbeth., "Where, Exactly, Are the RD respondents: A GIS-Based Approach to Quantifying Potential Bias," Paper presented to the American Association for Public Opinion Research, Miami Beach, May, 2010.

Hansen, Sue Ellen, "CATI Sample Management Systems," in Advances in Telephone Survey Methodology, edited by James M. Lepkowski et al. Hoboken, NJ: John Wiley &Sons, 2008.

Johnson, Timothy P.; Cho, Young Ik; Campbell, Richard T.; and Holbrook, Allyson L., "Using Community-Level Correlates to Evaluate Nonresponse Effects in a Telephone Survey," Public Opinion Quarterly, 70 (2006a), 704-719.

Johnson, Timothy P., Holbrook Allyson L., Cho, Young Ik, and Bossarte, Robert M. "Nonresponse error in injury-risk surveys," American Journal of Preventive Medicine, 31 (2006b), 427–36.

Lee, Sunghee et al., "Exploring Nonresponse Bias in a Health Survey Using Neighborhood Characteristics," American Journal of Public Health, 99 (2009), 1811-1817.

Steeh, Charlotte and Piekarski, Linda, "Accommodating New Technologies: Mobile and VoIP Communication," in Advances in Telephone Survey Methodology, edited by James M. Lepkowski et al. Hoboken, NJ: John Wiley & Sons, 2008.

Triplett, Timothy and Abi-Habib, Natalie, "Socio-Demographic Study of Telephone Nonrespondents," Paper presented to the FCSM Conference, November, 2005, Arlington. VA.

Tucker, Clyde and Lepkowski, James M., "Telephone Survey Methods: Adapting to Change," in Advances in Telephone Survey Methodology, edited by James M. Lepkowski et al. Hoboken, NJ: John Wiley & Sons, 2008.