CHAPTER 5

PARADATA FOR COVERAGE RESEARCH

STEPHANIE ECKMAN
Institute for Employment Research (IAB), Nuremberg

5.1 INTRODUCTION

Coverage research involves studying the quality of the frames from which samples are selected, and the impacts of errors in frames on survey data. Coverage is an understudied area in the survey methodology literature, due in large part to the difficulty of obtaining the necessary data about errors on the frame. Fortunately, paradata can in many cases provide the missing data needed to study coverage. This chapter first provides an introduction to frame and coverage error and then explores how paradata can be used to study coverage in household surveys. It discusses several types of frames, and the studies related to each type that have made use of paradata. The chapter also suggests additional coverage research that could be done with paradata.

The list from which we select a sample for a survey is called a *frame*. The population we wish to reach and represent with the survey is called the *target population*. In the ideal situation, the frame contains one and only one entry for each of the units in the target population. For example, a professional organization might want to conduct a survey of its members, and it has a complete and up-to-date list of all of the members from which it can select a sample. The list of members is the frame and the members themselves are the target population for the survey. If all members are on the list and the list contains no one who is not a member, then the frame covers the target population perfectly.

Other types of frames do not contain the members of the target population itself, but instead consist of a list of clusters which relate to the members of the population. For example, a small town might have a list of all residential addresses inside the town boundaries, from which a sample of addresses could be selected. Of course, addresses cannot take part in the survey and are not members of the target population.

Improving Surveys with Paradata: Analytic Uses of Process Information, First Edition. Edited by Frauke Kreuter.

The target population is the people who live in those housing units. If all residents of the town have an address, and all addresses on the list are valid, then we can associate all members of the target population with the clusters on the frame and this frame also covers the target population perfectly.

In the real world, frames are almost never perfect. Lessler (1980) describes six types of errors that frames may have. *Undercoverage* is the exclusion of units in the target population from the frame. For example, a survey may have a target population of all adults in the United States, but choose to use a telephone number frame to select a sample. This frame will clearly not include adults who do not have a telephone. These members of the target population are not on the frame and are undercovered; they have no chance to be selected and no opportunity to participate in the survey. *Overcoverage* is the inclusion on the frame of units that are not members of the target population or are clusters which are not associated with any members of the target population. A telephone number frame is also a good example here. Telephone frames often contain numbers which are not in service or which reach non-adults (e.g., teenagers with mobile phones). Such cases are instances of overcoverage. These cases are on the frame, and may be selected, but they cannot or should not be interviewed.

Figure 5.1 depicts undercoverage and overcoverage in a Venn diagram. Those units that are in the target population, but not captured on the frame, shown in the figure with vertical stripes, are undercovered. Those units that are on the frame, but not part of the target population, shown in the figure with horizontal stripes, are overcovered.

While undercoverage and overcoverage receive most of the attention in the coverage literature, frames can have other types of errors as well. Multiplicity in a frame refers to the case when some members of the target population can be reached by more than one case on the frame. Such multiplicity is quite common with telephone

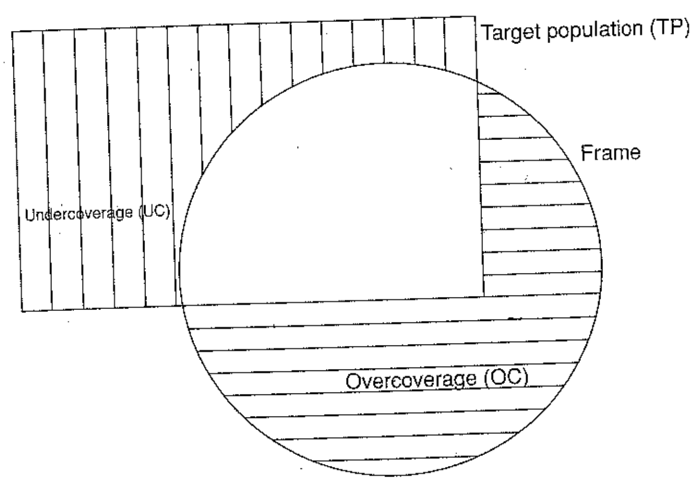
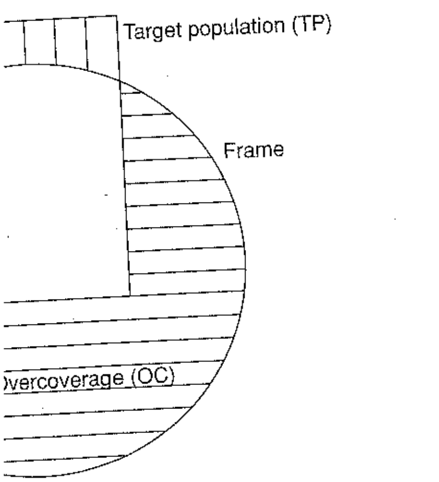


FIGURE 5.1 Overlap between target population and frame, showing undercoverage and overcoverage.

hose housing units. If all residents of
ıe list are valid, then we can associate
clusters on the frame and this frame

perfect. Lessler (1980) describes six
verage is the exclusion of units in the
a survey may have a target population
ıse a telephone number frame to select
adults who do not have a telephone.
ot on the frame and are undercovered;
ɔportunity to participate in the survey.
of units that are not members of the
ɔt associated with any members of the
is also a good example here. Telephone
in service or which reach non-adults
cases are instances of overcoverage.
lected, but they cannot or should not be

coverage in a Venn diagram. Those units
tured on the frame, shown in the figure
units that are on the frame, but not part
with horizontal stripes, are overcovered.
eceive most of the attention in the cover-
of errors as well. Multiplicity in a frame
the target population can be reached by
tiplicity is quite common with telephone



ılation and frame, showing undercoverage and

frames, as many people have more than one personal telephone number, such as a home phone and a mobile phone. Multiplicity on a frame is not itself an error: if all of the different ways each case could be selected are known and adjusted for, then there is no problem. In practice, it is often difficult to discover all the ways each case could have been selected. *Undetected multiplicity* on a frame is a concern. If the probabilities of selection of a sampled case do not adjust for multiplicity, the weights will be wrong and the survey estimates biased. Similarly, *undetected clustering* can lead to incorrect weights and inferences. This error arises when more than one member of the target population is reached via a single entry on the frame, but this fact is not discovered and adjusted for.

The final two errors that are possible on a frame are *incorrect auxiliary information* and *incomplete contact information*. The former includes bad data used as measures of size or as stratification variables in the selection process. When these data are incorrect, sample selection will be less efficient than it could have been, but estimates will still be unbiased. The latter refers to bad phone numbers, addresses or other contact data. If selected cases cannot be located or contacted, they are rather like undercovered cases in that they have no chance to participate in the survey. Unlike undercovered cases, however, such uncontacted cases are captured and reported in the denominator of the response rate.

Undercoverage and overcoverage are measured by coverage rates. The *undercoverage rate* is the number of undercovered units divided by the total size of the target population:

$$\text{undercoverage rate} \equiv \frac{N_{\text{uc}}}{N_{\text{tp}}} \tag{5.1}$$

The *overcoverage rate* is the proportion of units on the frame that do not belong to the target population:

$$\text{overcoverage rate} \equiv \frac{N_{\text{oc}}}{N_{\text{frame}}} \tag{5.2}$$

Some studies report instead the *net coverage rate* which combines undercoverage and overcoverage:

$$\text{net coverage rate} \equiv \frac{N_{\text{frame}}}{N_{\text{tp}}} \tag{5.3}$$

A net coverage rate near one does not indicate that a frame covers the target population well, only that it contains the right number of units: the frame may still contain a large amount of undercoverage and overcoverage. Oftentimes, however, net coverage is the only rate that can be estimated.

The coverage rates given in Equations 5.1, 5.2, and 5.3 measure the amount of undercoverage or overcoverage on a given frame. Coverage bias, on the other hand, captures the effects of these errors in the frame on estimates produced from survey data. Whenever a frame contains undercoverage or overcoverage, bias can arise. If the correctly covered units are different than the undercovered or overcovered units, then survey data will be biased. Undercoverage bias in the mean of a variable Y is a function of the undercoverage rate (Equation 5.1) and the difference between the undercovered and the correctly covered units on this variable.

$$bias_{\text{undercov}}(\bar{Y}) = \frac{N_{\text{uc}}}{N_{\text{frame}}}(\bar{Y}_c - \bar{Y}_{uc}) \tag{5.4}$$

This equation captures how wrong an estimate would be if only the covered cases were used to estimate the mean of Y. For example, data from the National Health Interview Study (NHIS), a national in-person survey in the United States, has shown that people who live in households that have no landline telephone are more likely to have no health insurance and also more likely to smoke (Blumberg and Luke, 2011). Because it is an in-person survey, the NHIS itself covers both those with and without telephones. However, many telephone surveys of the same target population cover only those with landline service.[1] Such surveys undercover these individuals and thus underestimate the prevalence of smoking and not having health insurance in the general population. These telephone surveys suffer from undercoverage bias.

The overcoverage bias in the estimate of the mean of Y is given in Equation 5.5.

$$bias_{\text{overcov}}(\bar{Y}) = \frac{N_{\text{oc}}}{N_{\text{tp}} + N_{\text{oc}}}(\bar{Y}_{\text{oc}} - \bar{Y}) \tag{5.5}$$

The example of overcoverage in a telephone frame discussed above included two different kinds of overcovered cases. The first were telephone numbers that were not in service; such cases do not contribute to overcoverage bias because no data about the variable Y exists for these cases. The second type of overcovered cases can contribute to overcoverage bias; in the example, these were cases that reached persons who were not eligible for the survey (age 18 and older in that example). Such persons can be mistakenly interviewed and thus can lead to bias in survey estimates.[2]

The formulas introduced above can help us understand frame error and coverage bias, but they are not easy to calculate. Most surveys cannot produce estimates of how many units in the target population are missing from the frame and thus cannot calculate the undercoverage rate in Equation 5.1. Even less often do surveys have any data about the undercovered cases with which to calculate undercoverage bias (Equation 5.4). Due to the difficulty in calculating all of the rates and biases given above, most surveys do not report any measures of frame quality. Paradata produced during frame construction, sampling, or data collection can help us better understand

[1] In many countries, dialing mobile telephones is more expensive than dialing landline numbers, and thus it is not uncommon for surveys to exclude mobile phones from the frame.

[2] See Wright and Tsao (1983) for a bias equation that combines undercoverage and overcoverage. For more background on frame error and coverage error, see Chapters 3, 4, and 5 of Lessler and Kalsbeek (1992).

the undercoverage and overcoverage in our frames, and also help us estimate coverage bias. (It is also true that paradata generated during frame construction can be useful in studying other error sources, such as nonresponse. I do not discuss this aspect here and instead focus on the use of paradata, from any survey process, to study coverage rates and coverage bias.)

This chapter is organized by the types of frames currently used for surveys of households: housing unit frames, telephone frames, rosters of household members to select one at random to participate in the survey, population registers from which persons can be selected directly, and frames for web surveys. In each section, I review briefly what is known about the coverage properties of the frame and then go into detail about how paradata has been used to estimate coverage rates and coverage bias, or the kinds of paradata that have been found to be related to coverage. I also draw attention to experiments and other studies which have not themselves used paradata but which identify the paradata that can and should be collected. Many of the studies cited below are proceeding papers or unpublished presentations. The published literature makes little use of paradata, and this area is a clear opportunity for future research. I also point out opportunities for additional coverage research using existing paradata and new types of paradata that could be collected and used to study coverage.

5.2 HOUSING UNIT FRAMES

Many household surveys begin with a frame of housing units from which a sample is selected. The selected cases are visited by an interviewer for a face-to-face survey, mailed a questionnaire, or matched to phone numbers for a telephone survey. Housing units frames can be derived from existing databases, such as postal delivery lists, or created in the field via listing. A hybrid method, called dependent listing, uses field staff to update existing lists in the field. In some countries, the random route (also called random walk) technique is common. This method combines frame creation and sampling into one step. Because they capture people at their homes, all of these methods undercover the homeless, and most exclude people in institutions such as hospitals, prisons, dormitories, etc. Each approach also has unique patterns of undercoverage and overcoverage, and researchers are only beginning to explore the many ways paradata can help us understand the quality of these frames.

5.2.1 Postal Delivery Databases

Countries often have centralized postal registers, which should include all residential addresses that receive mail, and thus most places where people live. These databases have obvious appeal as survey frames. Undercoverage could occur, however, if there are units which do not receive mail, if the database is slow to include new construction, or if residents have insisted that their addresses not be given to third parties. Overcoverage is also possible if the database includes addresses which no longer exist or are businesses. These frame problems could cause coverage bias, if the undercovered or overcovered units are different than those which are correctly covered.

Household surveys in the United Kingdom often use postal databases (Wilson and Elliot, 1987; Butcher, 1988; Lynn, 2009), and U.S. and Canadian researchers have also begun to explore the use of such databases in recent years (O'Muircheartaigh et al., 2002; Iannacchione et al., 2003; Thompson and Turmelle, 2004). Many studies have looked at the coverage of these databases for survey purposes. The coverage appears to be quite good in the United Kingdom (Wilson and Elliot, 1987; Foster, 1994) and Canada (Thompson and Turmelle, 2004; Turmelle et al., 2005). In the United States, the coverage for mailing purposes is high, but some addresses, especially in rural areas, are not valid for face-to-face surveys because they are only post office or rural route boxes, which interviewers cannot visit (O'Muircheartaigh et al., 2002; Dohrmann et al., 2006; O'Muircheartaigh et al., 2006; Dohrmann et al., 2007; Iannacchione et al., 2007; O'Muircheartaigh et al., 2007; English et al., 2010).

In the United States these databases are derived from the U.S. Postal Service's Delivery Sequence File. This database contains many additional fields used for mailing purposes: each address can be flagged as vacant, part of a college, or seasonal. From the survey researcher's perspective, such flags are paradata. Amaya et al. (2010) relate these flags on the frame to final case outcomes from the data collection process. They study the effects of excluding the addresses flagged as vacant, college, and seasonal on undercoverage and overcoverage rates. They find that the vacant flag is usually accurate at identifying vacant housing units: excluding all addresses flagged as vacant from a survey frame would reduce the overcoverage rate, and would increase the undercoverage rate slightly. In contrast, addresses flagged as college units are almost always private housing located off campus in their study and not dormitories or other institutionalized living quarters. This flag should in most cases not be used to exclude units from survey frames. They find that the available seasonal flag is harder to understand and warrants additional investigation. They do not explore the coverage bias that would result in survey data if these flagged cases were excluded.

One important piece of information for survey researchers which is often not included on the postal databases, is the location of each address. The only geographic variables available are postal geographies, such as ZIP or postal code. In the United Kingdom, survey researchers use the postal sectors as sampling units (Lynn, 2009), and additional information about the location of each housing unit is not needed at the frame creation stage. Such an approach has been proposed in the United States as well, but has not caught on because most U.S. area probability designs are based on Census geographies, about which many more demographic characteristics are available (Shook-Sa et al., 2010).

Because of the mismatch in the United States between the postal geographies on the database and the Census geographies used in sampling, very often one of the first steps in creating a frame from a postal database is to geocode the addresses. Geocoding assigns a latitude and longitude to each address and then translates that point into a Census area such as a block, tract, or county. Geocoding is performed by specialized software that contains a database of the location of all streets in the country and the range of addresses associated with each street segment. As an example, consider the

: coverage bias, if the undercovered
:h are correctly covered.
ən use postal databases (Wilson and
and Canadian researchers have also
:ent years (O'Muircheartaigh et al.,
Turmelle, 2004). Many studies have
vey purposes. The coverage appears
ıon and Elliot, 1987; Foster, 1994)
urmelle et al., 2005). In the United
gh, but some addresses, especially
urveys because they are only post
annot visit (O'Muircheartaigh et al.,
ı et al., 2006; Dohrmann et al., 2007;
al., 2007; English et al., 2010).
:ived from the U.S. Postal Service's
many additional fields used for mail-
acant, part of a college, or seasonal.
lags are paradata. Amaya et al. (2010)
tcomes from the data collection pro-
addresses flagged as vacant, college,
rage rates. They find that the vacant
ousing units: excluding all addresses
duce the overcoverage rate, and would
rast, addresses flagged as college units
campus in their study and not dormi-
. This flag should in most cases not be
y find that the available seasonal flag is
investigation. They do not explore the
if these flagged cases were excluded.
survey researchers which is often not
n of each address. The only geographic
ch as ZIP or postal code. In the United
ectors as sampling units (Lynn, 2009),
ı of each housing unit is not needed at
ıas been proposed in the United States
U.S. area probability designs are based
more demographic characteristics are

tates between the postal geographies on
d in sampling, very often one of the first
ıse is to geocode the addresses. Geocod-
ddress and then translates that point into
/. Geocoding is performed by specialized
ıtion of all streets in the country and the
et segment. As an example, consider the

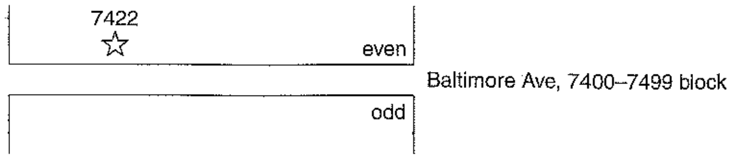


FIGURE 5.2 Assignment of geocode via interpolation.

address 7422 Baltimore Ave, College Park, MD 20740. The geocoding software first locates the appropriate section of Baltimore Ave in the associated city, state and ZIP code, as shown in Figure 5.2. It then interpolates the position of 7422 based on the known locations of the two ends of the 7400 block and the information that the even numbered addresses are on the north side of the street. 7422 will be assigned a coordinate that is 22% of the way down from the west end, on the north side. Having placed the address on the map and assigned a latitude and longitude, the software overlays a database of Census blocks to determine which block the address falls into. (For more details on all of these steps, see Zandbergen (2008) and Eckman et. al. (2012b)).

Only those addresses that are assigned to Census blocks that are selected for the survey become part of the frame. For this reason, accurate geocoding and correct block assignment is important to ensuring good coverage for frames derived from the databases (Morton et al., 2007; Shook-Sa et al., 2010; Eckman and English, 2012a). Fortunately, the geocoding process itself generates paradata with which we can study overcoverage and undercoverage. While assigning latitude and longitude, the software also reports a code which indicates how precisely each address could be geocoded. MapMarker geocoding software generates a 10-digit georesult code for each address that summarizes how the geocode was assigned (McMichael et al., 2008). The first two digits of this code are the most useful and are explained in Table 5.1. An S5 code indicates the highest level of precision, the type of geocoding demonstrated in Figure 5.2. Nationally, 83.3% of all addresses on the Delivery Sequence File geocode at this level (Eckman and English, 2012a).

Survey researchers can and should use this code to decide which addresses have geocoded precisely enough to use in the frame. Eckman and English (2012a)

TABLE 5.1 Meaning of Georesult Codes Returned by MapMarker Geocoding Software

First Digits	Meaning
S5	Single close match, point located at a street address position
S4	Single close match, point located at the center of street
S6, Z6[a]	ZIP Code centroid match for point ZIP
S3, Z3[a]	ZIP+4 centroid match (highest accuracy available)
S2, Z2[a]	ZIP+2 centroid match
S1, Z1[a]	ZIP Code centroid match
N	No close match

[a]There are minor differences between these two codes, see software documentation.

investigate the net coverage rates associated with different decision rules. They find that insisting upon using only addresses that geocode with the highest precision (S5) results in a national net coverage rate of 80%. Less stringent inclusion criteria increase the net coverage rate, but introduce overcoverage, as units are brought into the frame even when they do not really lie in the selected areas. The contribution of these inclusion criteria, which are based on geocoding paradata, to coverage bias has not been explored.

5.2.2 Housing Unit Listing

When such postal or other address databases are not available, or their coverage is not sufficient, surveys often use interviewers to create housing unit frames. Listing involves sending field staff to selected areas to make a frame, or list, of all residential addresses. These address lists are returned to the central office where they are entered into a database and become a frame for future sample selection. Major surveys using listed housing unit frames include: the General Social Survey (Harter et al., 2010); the Canadian Labour Force Survey (Statistics Canada, 2008); the Current Population Survey (U.S. Census Bureau, 2006); and several countries in recent rounds of the European Social Survey (Jowell and the Central Co-ordinating Team, 2003, 2005, 2007; Central Co-ordinating Team, 2010).

Two methods of listing are commonly used. In *traditional* or *scratch* listing, a lister is provided with a map which specifies the boundaries of the selected areas. See Figure 5.3 for an example listing map: the selected area is shaded and consists of five blocks. The lister starts with one block, say block 2004, at the small X shown on the map. She travels in the specified direction and records the address of every residential units she sees, without resident names (e.g., 104 Orchard St, Unit 201). In contrast, in *dependent* or *update* listing, a lister is provided with the map as well as an initial list of addresses, called the *input list*, which she updates in the field. Very often the input list comes from a postal database. The lister travels around each block, just as she does in traditional listing, but here she compares what she sees on the ground to the input list. She adds and deletes addresses as necessary. The goal of both types of listing is a full frame of housing units inside the selected area.

Dependent listing in particular produces rich paradata. As listers update the input list, they delete inappropriate units from the input list, add units which are not on the list, and confirm units that are already correctly listed. For each housing unit on the final frame, we know whether it came from the input list or was added by the lister. We also know which units the lister deleted from the input list. Pacer (2008) uses these paradata to explore the rates of addition, deletion, and confirmation of units, and the housing unit and area-level characteristics that correlate with these outcomes. Pacer finds that the units added by the dependent listing are likely to be in suburban or commercial neighborhoods. Those that listers delete from the input list are likely to be in multi-unit buildings and also in suburban areas.

This analysis could be extended to look at coverage bias, by comparing the responses of the units that the lister added and those on the input list that she confirmed. If the added units are no different, in terms of variables measured in the

different decision rules. They find
;ocode with the highest precision
%. Less stringent inclusion criteria
coverage, as units are brought into
selected areas. The contribution of
ding paradata, to coverage bias has

not available, or their coverage is
;reate housing unit frames. Listing
ıke a frame, or list, of all residential
;entral office where they are entered
nple selection. Major surveys using
Social Survey (Harter et al., 2010);
ıada, 2008); the Current Population
l countries in recent rounds of the
l Co-ordinating Team, 2003, 2005,

In *traditional* or *scratch* listing, a
ɔoundaries of the selected areas. See
;d area is shaded and consists of five
ς 2004, at the small *X* shown on the
ords the address of every residential
Orchard St, Unit 201). In contrast,
;d with the map as well as an initial
updates in the field. Very often the
er travels around each block, just as
ares what she sees on the ground to
necessary. The goal of both types of
selected area.
paradata. As listers update the input
ıt list, add units which are not on the
listed. For each housing unit on the
input list or was added by the lister.
ɔm the input list. Pacer (2008) uses
deletion, and confirmation of units,
;s that correlate with these outcomes.
ıt listing are likely to be in suburban
s delete from the input list are likely
n areas.
t coverage bias, by comparing the
nd those on the input list that she
n terms of variables measured in the

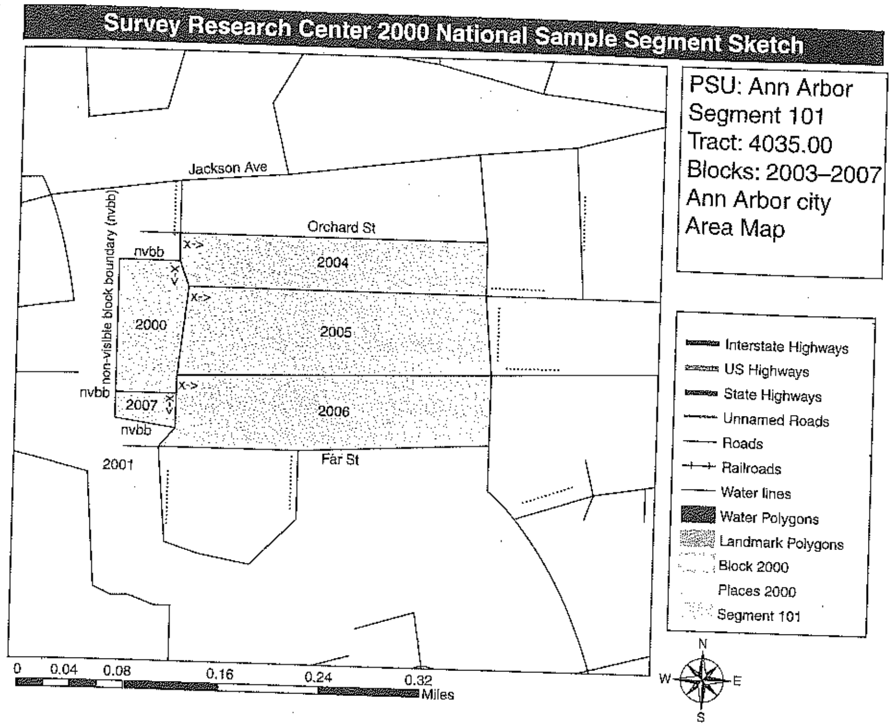


FIGURE 5.3 Example of listing map from Survey Research Center.

survey, then the dependent listing procedure did not affect the data. The undercoverage bias that we would see, if we did not do the dependent listing, would be zero. However, if the added units are quite different, then the dependent listing did change the survey results and undercoverage bias without the dependent listing would be large. Such changes to the survey data due to dependent listing should be compared to the cost of frame improvement. If it is expensive to travel field staff out to the selected area to update the frame, and such updates result in only minimal changes to the survey data, then dependent listing is not cost effective. To date, no studies have explored this cost/bias trade off.

Just as listers collect housing unit and neighborhood observations when they interview (see Chapter 2 of this book), they can also collect variables related to frame quality when they list. The Survey Research Center at the University of Michigan asks interviewers to record segment observations during their listing work. The variables they collect include: perception of safety in the neighborhood, road conditions, any locked buildings or communities, the language(s) spoken by the neighborhood residents, presence of commercial or industrial buildings in the area, and whether the lister drove or walked while listing. Eckman and Kreuter (2013) explore the relationships between many of these variables and undercoverage in traditional listing but find no significant relationships. Nonvisible block boundaries, that is, boundaries that do not correspond to a street or a body of water, as shown on the west side of the map in Figure 5.3, are also associated with undercoverage in traditional listing and so are errors on the listing maps. More analysis of the relationships between these sorts of paradata collected by listers or about their materials and the coverage of listed frames is warranted.

Other recent research has shown that listers can affect the frames they produce. Two different staff members, listing the same areas using the same method and materials, can produce different frames (Kwiat, 2009; Eckman, 2013). Inter-lister variability introduces not coverage bias but coverage variance (see Groves, 1989, p. 121 for a discussion of coverage variance). Paradata about the listing process, such as the weather at the time of listing, the time of day and of the year in which the listings were done may help to explain some of this variability. No studies have estimated coverage variance due to lister differences or explored the process variables that explain it.

Much more paradata about the listing process could be captured. When listing is performed with laptop- or tablet-based listing software, the time spent listing and keystroke data could be recorded and may relate to coverage. For example, a very fast listing might indicate that the field staff member rushed through the listing task and did not look carefully for hidden units such as basement apartments or those accessed via rear entrances. In contrast, a very slow listing might indicate that a lister had a hard time with the task, perhaps trouble finding the correct block or determining which units were inside the boundaries. Keystroke data that indicate many edits or rearranging of listed housing units can also be signs of problems. To my knowledge, such data has not been explored, or even captured.

The National Survey of Family Growth has recently started experimenting with the use of GPS-enabled cell phones during listing. Interviewers carry the phones so that they can easily contact their supervisors, but also activate GPS applications that record their heading, speed, and location (personal communication with J. Wagner of the University of Michigan). Market researchers use similar technology to track customer movements (Czaplicki, 2011). GPS recordings of lister travels could be used to capture the location of each listed unit and to study whether the lister included units outside the boundaries of the segment.

Despite the popularity of postal databases and population registers (discussed in Section 5.5) as survey frames, these are not always available. Listing is still commonly used in North American and European surveys. More investigation into the determinants of listing quality is needed, and paradata should play a central role in future research.

5.2.3 Random Route Sampling

A frame construction method somewhat similar to housing unit listing is random route (or random walk) sampling. This technique combines frame construction and sampling into one step. Interviewers start at a specified point and follow traveling

shown on the west side of the map
ıge in traditional listing and so are
elationships between these sorts of
ls and the coverage of listed frames

an affect the frames they produce.
ıreas using the same method and
2009; Eckman, 2013). Inter-lister
erage variance (see Groves, 1989,
Paradata about the listing process,
ne of day and of the year in which
of this variability. No studies have
es or explored the process variables

ıs could be captured. When listing
software, the time spent listing and
o coverage. For example, a very fast
rushed through the listing task and
ement apartments or those accessed
ing might indicate that a lister had
ıg the correct block or determining
oke data that indicate many edits or
gns of problems. To my knowledge,
d.
recently started experimenting with
ıg. Interviewers carry the phones so
t also activate GPS applications that
onal communication with J. Wagner
hers use similar technology to track
ordings of lister travels could be used
to study whether the lister included

and population registers (discussed
ot always available. Listing is still
ean surveys. More investigation into
d paradata should play a central role

lar to housing unit listing is random
que combines frame construction and
specified point and follow traveling

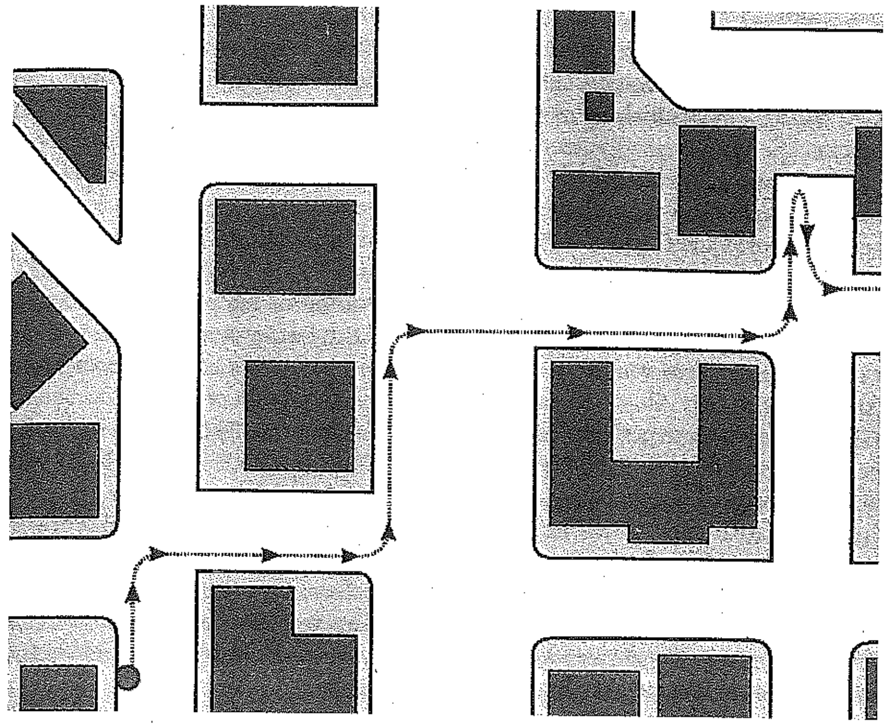


FIGURE 5.4 Example of implementation of route followed in random route sampling. From Bauer (2013).

rules for turning right and left at intersections and dead-end streets. There are several different sets of proposed traveling rules. Figure 5.4 gives an example of the implementation of one set of traveling rules, those by Noelle (1963), in a given neighborhood. The interviewer starts at the dot in the lower left corner, a randomly selected point, and proceeds north, takes the first available right, then the next left, then the next right, and so on. As she travels, she should select every kth housing unit that she passes into the sample and attempt the interview (Schnell et al., 2008; Diekmann, 2010). In the random route approach, there is no separation between the frame construction, sampling and interviewing stages. Because it eliminates the need for a separate trip to the selected area only for listing, the random route technique is less expensive and faster to implement than housing unit listing. Random route sampling is common in European surveys (Jowell and the Central Co-ordinating Team, 2003, 2005, 2007; Central Co-ordinating Team, 2010) and in developing countries (see Fink, 1963). However, the method is not used in major surveys in the United States or Canada due to concerns expressed more than 50 years ago that interviewers do not select representative samples (Manheimer and Hyman, 1949; Simmons, 1953; Boyd and Westfall, 1955).

Alt et al. (1991) used contact history paradata to explore this phenomenon more recently. Comparing the contact history records (paradata from the data collection process) for cases selected via random walk and via a population register, they show that random walk samples require fewer calls to complete the interview. They interpret these findings as evidence that interviewers preferentially select easier to complete cases when allowed to create their own sample. Because respondents who are easier to contact and interview can be different from those that are harder to contact (Groves and Couper, 1992), interviewers using random walk sampling could very well introduce undercoverage bias into the survey.

Because of its similarity to housing unit listing, all of the paradata that are or could be captured about the listing process should also be captured about the random walk process, such as which interviewer performed the selection, the time and date, neighborhood observations, duration, etc. Additional paradata of interest in random walk samples relate to performance of the routing instructions. Interviewers following the route sometimes find themselves caught in a loop or stuck in a dead-end with no new housing units to select (Bauer, 2013). Such breakdowns of the traveling algorithm should be recorded by the interviewer and relayed to the central office. A high number of such issues may indicate that an interviewer has trouble following the instructions, or that the instructions are somehow not appropriate for a given area. GPS data from mobile phones or other devices could also be useful here. Such paradata can reveal problems with the random route process that can spur additional research and improve our understanding of coverage.

5.2.4 Missed Unit Procedures

Because undercoverage is a risk in all housing unit frames, several techniques to find and select missed units have been proposed. The most commonly used technique is the half-open interval procedure, which involves checking for missed units associated with each selected case during interviewing (Kish, 1965, pp. 341–342; Groves, 1989, pp. 127–128). The Waksberg procedure (Judkins, 1997), and the recently proposed CHUM procedure (McMichael et al., 2008), ask interviewers to check a larger area, but have the same aim: to identify units missed during the original frame construction effort and give them a chance of selection.

The cases that are added by these techniques can be flagged in the final dataset, and this flag is paradata. Using such an indicator, Shore et al. (2010) demonstrate that units added by a missed unit procedure are less likely to be occupied residential units and, when they are occupied, more likely to be nonrespondents, than cases already on the frame. This analysis could be taken a step further to look at undercoverage bias had the missed unit procedure not been done, that is, by comparing results based on the full dataset to those excluding the cases added by the missed housing unit procedure. However, the procedures usually add so few cases to the respondent sample that such a study is not feasible.

Despite their theoretical ability to fix undercoverage, recent experiments have shown that interviewers do not perform the CHUM and half-open interval procedures correctly (McMichael et al., 2008; Eckman and O'Muircheartaigh, 2011).

:a to explore this phenomenon more
s (paradata from the data collection
and via a population register, they
alls to complete the interview. They
ewers preferentially select easier to
n sample. Because respondents who
erent from those that are harder to
s using random walk sampling could
survey.
ting, all of the paradata that are or
ld also be captured about the random
ned the selection, the time and date,
tional paradata of interest in random
g instructions. Interviewers following
a loop or stuck in a dead-end with
. Such breakdowns of the traveling
r and relayed to the central office. A
an interviewer has trouble following
mehow not appropriate for a given
vices could also be useful here. Such
oute process that can spur additional
erage.

nit frames, several techniques to find
ne most commonly used technique is
checking for missed units associated
sh, 1965, pp. 341–342; Groves, 1989,
ns, 1997), and the recently proposed
k interviewers to check a larger area,
luring the original frame construction

es can be flagged in the final dataset,
r, Shore et al. (2010) demonstrate that
likely to be occupied residential units
e nonrespondents, than cases already
tep further to look at undercoverage
done, that is, by comparing results
e cases added by the missed housing
ly add so few cases to the respondent

ercoverage, recent experiments have
CHUM and half-open interval proce-
kman and O'Muircheartaigh, 2011).

Future research into improving these procedures could make use of paradata. For example, if the reason for these findings is that listers are simply forgetting to do the procedure, a simple check list built into the interviewing software could help. The case management software could require interviewers to indicate with a check mark that they have performed the missed housing unit check and provide a space to describe any problems they encountered with the procedure. Such checklists have been found to improve outcomes in many fields (Hales et al., 2008; Haynes et al., 2009).

Housing unit frames are used throughout the world for important surveys, yet the quality of these frames is seldom studied. I hope that the discussion above about the helpful role that paradata can play in coverage research will instigate more work in this area.

5.3 TELEPHONE NUMBER FRAMES

Rather than selecting individuals for the survey through their addresses, some surveys use a frame of telephone numbers. Telephone surveys of households are less expensive to conduct than in-person surveys, but telephone frames tend to suffer from much more overcoverage and undercoverage than other types. Overcoverage occurs because randomly selected phone numbers often reach businesses, fax machines, modems, or nonworking numbers. Undetected multiplicity and undetected clustering are also problems, when several persons can be reached by several telephone numbers and these links are not discovered and adjusted for in the weights. Undercoverage is a concern if some households or persons have no phone number. Due to the expense of calling mobile phone numbers in many countries, some surveys exclude such numbers from their frames, and these surveys then undercover those who can be reached only via mobile phones.

The most common method of creating telephone frames in the United States uses the list-assisted methodology proposed by Casady and Lepkowski (1993). Rather than generating numbers entirely at random, banks of 100 consecutive numbers are assigned a score reflecting how many numbers in that bank also appear in the directory of listed phone numbers. The bank-level score, paradata from the frame creation process, is inversely correlated with overcoverage: telephone numbers in banks with zero or few listed numbers are less likely to be residential than those in banks with many listed numbers. Often, only banks with a score greater than or equal to some cutoff are used in the frame (many surveys use a cutoff of one). High cutoff scores lead to a lower overcoverage rate and thus more efficient samples: a high proportion of the numbers dialed will reach households. The downside of using a high cutoff score is undercoverage of residential numbers in the low-scoring banks. Using a low cutoff score protects against undercoverage, but leads to reduced efficiency (Casady and Lepkowski, 1993; Tucker et al., 2002; Fahimi et al., 2009). A similar technique, called the Gabler–Häder method, is used in Germany (Häder, 2000).

Two studies relate this bank score paradata to coverage bias. Brick et al. (1995) find mostly small and insignificant biases due to the exclusion of banks with zero listed numbers. Writing nearly 15 years later, Boyle et al. (2009) do find significant differences between the households reached in the zero banks and those reached in banks containing one or more listed numbers. Households with phone numbers in the banks with scores of zero tend to live in larger families with lower incomes and shorter tenure at their address.

Another method to reduce out-of-scope overcoverage in telephone sampling is to send all selected telephone numbers through a pre-screening service which uses databases of business and mobile phone numbers, and/or very short calls with an automated dialer, to identify out-of-scope numbers (Battaglia et al., 2005). This screening process generates paradata about each selected telephone number. Usually these paradata are used to exclude cases from normal data collection activities and thus to increase the efficiency of the dialing. The flagged cases are removed from the sample and not dialed. The cases are also not counted in the denominator of the response rate and are treated as if they had never been sampled: the flagged cases can be considered to have been removed from the frame.

These flags can substantially increase sample efficiency, by removing out-of-scope cases, but they can also introduce undercoverage when the cases do in fact reach residential phones. Battaglia et al. (2005) dialed these cases, which permits an assessment of how well the flags reduce overcoverage and whether they also introduce undercoverage. They find that 59% of the cases flagged as businesses are really business, and 29% are residential. Among those cases flagged as fax machines or modems, 64% are correctly identified and 14% are residential numbers. Those flagged as nonworking are correctly flagged in 88% of cases, and reach residential telephones in only 7% of cases. Thus removing the flagged cases reduces the overcoverage rate, but increases the undercoverage rate. A similar study by Merkle et al. (2009), focusing only on business flags, explores coverage bias due to removal of these flagged cases. They find that the respondents reached via the residential cases flagged as businesses are significantly older, more likely to be white, and more conservative in their political opinions, than those reached via telephone numbers that are not flagged as businesses. Some variables would be biased if the cases flagged as businesses were undercovered.

Another concern in telephone surveys, related to the high overcoverage rate, is the high proportion of cases that are unresolved at the end of the field period, even after many calls. These cases are often always busy, or are never answered at all. In the United States, approximately 9% of all sampled cases are unresolved, and this proportion has been increasing steadily over the last 15 years (Brick et al., 2002; Kennedy et al., 2008). Estimation of the proportion of these cases that are eligible for the survey, called e in the literature, is important in the calculation of the response rate. The estimate of e is used to calculate the denominator for this rate, the number of cases that were eligible to complete the survey. That is, $e\%$ of the unresolved cases are assumed to be eligible residential telephone numbers and are included in the denominator of the response rate, and $1 - e\%$ are assumed to be out-of-scope. (See Skalland (2011) for more details on the role of e in calculating response rates.) Calculating appropriate response rates is important for proper survey documentation, but it is not an issue of coverage. However, the determination of the share of unresolved cases that are in-scope (e) and out-of-scope ($1 - e$) for the survey

yle et al. (2009) do find significant
e zero banks and those reached in
ouseholds with phone numbers in
r families with lower incomes and

overage in telephone sampling is
. pre-screening service which uses
s, and/or very short calls with an
bers (Battaglia et al., 2005). This
elected telephone number. Usually
rmal data collection activities and
e flagged cases are removed from
counted in the denominator of the
een sampled: the flagged cases can
ne.

e efficiency, by removing out-of-
overage when the cases do in fact
aled these cases, which permits an
ge and whether they also introduce
flagged as businesses are really
cases flagged as fax machines or
residential numbers. Those flagged
es, and reach residential telephones
ases reduces the overcoverage rate,
ly by Merkle et al. (2009), focusing
to removal of these flagged cases.
dential cases flagged as businesses
more conservative in their political
s that are not flagged as businesses.
d as businesses were undercovered.
l to the high overcoverage rate, is
at the end of the field period, even
usy, or are never answered at all.
ampled cases are unresolved, and
er the last 15 years (Brick et al.,
proportion of these cases that are
is important in the calculation of
calculate the denominator for this
mplete the survey. That is, $e\%$ of
esidential telephone numbers and
ate, and $1 - e\%$ are assumed to be
ils on the role of e in calculating
ates is important for proper survey
However, the determination of the
out-of-scope $(1 - e)$ for the survey

is an issue of coverage, as out-of-scope cases are effectively treated as if they were not sampled. Paradata is often used in the estimation of e. The most common method uses the dispositions of the resolved cases to estimate eligibility for the unresolved cases. A newer method makes use of call history records to estimate survival models (Brick et al., 2002).

Survey researchers using telephone frames take the threat of undercoverage bias, and the inefficiencies of overcoverage, seriously. In the early years, users of telephone frames worried about undercoverage of households that had no telephone at all (Lavrakas, 1987, Chapter 1); these days, they worry about undercoverage of those with only mobile phones (Blumberg and Luke, 2011). Coverage studies in this area have made use of paradata since before such a term existed. However, many opportunities exist for paradata to play a larger role in telephone frame research.

5.4 HOUSEHOLD ROSTERS

When a survey wishes to select a random sample of individuals, but the unit on the frame is a household or an address, an additional stage of frame creation and selection is necessary. While various non-probability and quasi-probability methods are commonly used, such as next birthday or oldest male/youngest female, a true random selection requires taking a full roster of all members of a household and selecting one at random (Gaziano, 2005; Yan, 2009). Any errors of exclusion or inclusion at the roster stage are undercoverage and overcoverage of household members, which usually go undetected. Research has shown that when a survey's definition of a household does not match the definition used in a community, undercoverage and overcoverage can occur. College students tend to be overcovered in their parents' home and undercovered in their true place of residence; children in joint custody arrangements may be overcovered. Concerns about privacy can also lead to undercoverage (Valentine and Valentine, 1971; Hainer, 1987; Fein, 1990; Tourangeau et al., 1997; Martin, 1999, 2007).

Most of the literature on roster coverage focuses on errors made by respondents when they are asked who lives in their household. The available paradata, however, more often captures information about the behavior of interviewers during the rostering process. Thus the review below also focuses on the interviewer side of household roster errors, though interviewers are not the only or necessarily the largest source of error in these frames.

Paradata about the roster and selection process can help us to detect errors in household rosters. When the roster and selection is done on a computer, the software can record interviewers' keystrokes. Keystroke files are records of every key pressed during the course of the interview, and thus allow researchers to study in detail what happened during an interview. These paradata can indicate problems with the roster process. Dahlhammer (2009) found evidence in such keystroke files that interviewers taking rosters sometimes backed up after respondent selection was complete, to change the order of the roster or the number of household members, which he interprets as interviewers trying to force selection of another, more cooperative,

TABLE 5.2 Example of Kish Table for Respondent Selection from Household Roster

Household Size	Select Person on Row
1	1
2	1
3	3
4	2
5	1
6 or more	4

household member. Manipulation of the selection procedure to insure that one or more persons are not selected effectively gives those members no chance of selection. Thus they are undercovered on the frame of household members, even though they appear on the roster.

Interviewer manipulation of respondent selection is harder to detect when the roster is performed on paper. When taking a roster on paper, the interviewer should list the household members in a pre-specified order (e.g., from oldest to youngest). A Kish table then tells the interviewer which row of the roster holds the selected person. Table 5.2 gives an example of a Kish table. In a one-person household, obviously the only household member will be the selected respondent. In a two-person household, the two people have an equal chance of selection: in the example shown in Table 5.2, the person listed in the first row of the roster will be selected. In a three-person household in this example, the third person on the roster will be selected. These Kish tables are randomly generated for each case and printed on the roster materials or affixed as stickers (for more details, see Kish, 1965, Section 11.3B).

Although interviewers should list the household members in the specified order and apply the Kish table, they can ignore or circumvent these procedures by listing members out of order, or manipulating the sort variables to force a certain member to appear on the row they know will be selected.[3] The paper rosters and Kish tables are paradata in the selection process and should be carefully reviewed for these types of manipulations. If the interviewer lists household members in the incorrect order, or incorrectly applies the Kish table, these would be detectable with such paradata. However, manipulation of the sort variable, such as the age of each household member, would be harder to detect.

Although the literature on household rosters more often discuses the contributions of respondents to undercoverage and overcoverage, interviewers may also play a role. Several techniques, making use of paradata, can detect interviewer errors or intentional manipulations of the roster and respondent selection process.

[3]Eyerman et al. (2001) provide evidence that interviewers do engage in these sort of manipulations when using paper rosters, though that analysis relies on the gender of the selected individuals, not paradata.

Table for Respondent Selection

Select Person on Row
1
1
3
2
1
4

selection procedure to insure that one or
ives those members no chance of selection.
of household members, even though they

nt selection is harder to detect when the
; a roster on paper, the interviewer should
fied order (e.g., from oldest to youngest).
hich row of the roster holds the selected
. Kish table. In a one-person household,
be the selected respondent. In a two-person
ance of selection: in the example shown in
/ of the roster will be selected. In a three-
ird person on the roster will be selected.
d for each case and printed on the roster
etails, see Kish, 1965, Section 11.3B).
household members in the specified order
or circumvent these procedures by listing
e sort variables to force a certain member
lected.[3] The paper rosters and Kish tables
hould be carefully reviewed for these types
ousehold members in the incorrect order,
e would be detectable with such paradata.
such as the age of each household member,

sters more often discuses the contributions
rcoverage, interviewers may also play a
paradata, can detect interviewer errors or
respondent selection process.

nterviewers do engage in these sort of manip-
: analysis relies on the gender of the selected

5.5 POPULATION REGISTERS

An alternative to the selection of households through address or telephone frames, with household rosters to identify a selected individual, is direct selection of persons from population registers. Such registers are not available in all countries, but where they are available, they are often used as survey frames (Jowell and the Central Co-ordinating Team, 2003, 2005, 2007; Central Co-ordinating Team, 2010). Direct person selection removes the need for another stage of selection of persons within households and thus reduces design effects (Lynn and Taylor, 1995; Clark and Steel, 2002).

The patterns of overcoverage and undercoverage are unique to each country's register. In Sweden, the register overcovers people who have moved abroad but remain registered. Conversely, foreigners who move to Sweden and do not register are undercovered (Laitila et al., 2011). For centralized registers such as Sweden's, the available paradata may include information about when and by whom each record was updated. Such paradata could be analyzed in conjunction with paradata from the data collection process about out-of-scope cases to shed light on the quality of the register: perhaps some locales are prone to overcoverage, or some clerks have high data entry errors.

The German registers are somewhat unique in that they are decentralized. Each community keeps its own register and no central agency has access to the entire database. The registries contain the name, age, gender, and address of all registered residents. Although registration is mandatory, the databases do undercover immigrants and students, and overcover the elderly (Schnell, 2008). These registries are often used for sample selection, but because they are decentralized, survey researchers must contact each individual office to request a sample of eligible persons. Information from the process of contacting these offices and procuring frames is paradata that likely relates to coverage. If some offices require multiple prompts, or deliver incorrect samples, this paradata should be captured. These indicators may be signs of low quality registers, which could manifest as a high rate of overcoverage. In addition, some offices refuse to comply with requests for samples (Schnell, 2008). If there are some local offices that never respond to requests for samples, then the residents of those areas are undercovered by surveys selected from register frames, even though those people are registered. Survey researchers will never know about this type of undercoverage, unless they collect and share paradata about the process of gathering samples from the registry offices.

5.6 SUBPOPULATION FRAMES

For many surveys, the target population is not the entire population of the country, but only a subset. A survey about breast cancer screening, for example, will likely interview only women 18 years of age and older. Errors in identifying who is a member of the targeted subpopulation and who is not will lead to undercoverage and overcoverage in the subpopulation frame and may lead to both types of bias as well.

When selecting individuals from a population register, it may be possible to select only cases that meet the eligibility criteria. However, errors in the information on the register can lead to overcoverage and undercoverage. If a woman is 19, but incorrectly listed on the register as 17, she will be excluded from the frame for the breast cancer survey and will be undercovered. Conversely, if another woman is 16, but registered as 26 due to a typographical error, she will be overcovered. Ideally, such overcoverage would be detected during data collection by a survey question that verifies eligibility and reported back to the keepers of the registry so that the error can be corrected. If eligibility is not verified during the interview, overcoverage may go undetected and could lead to coverage bias (e.g., an underestimate of breast cancer screening rates). Paradata about registry quality, as discussed above, likely relates to the accuracy of subpopulation frames created from the registry.

When population registers are not available, or do not contain data about the target population for a given survey, a survey must select a general sample, contact each case, and screen for eligibility. Screening involves asking a few questions at the beginning of the interview to establish whether the respondent (or, for clustered frames, anyone in the household) is in the survey's subpopulation. Many surveys which screen for eligibility suffer from undercoverage, that is, an observed eligibility rate that is lower than expected. For example, in the National Longitudinal Study of Youth in the United States, which screened for youth 12–23 years old, the eligibility rate found during screening was 75% of the expected eligibility rate (Horrigan et al., 1999).

Research into undercoverage in screener surveys has so far made little use of paradata. Tourangeau et al. (2012) show that interviewers vary quite a bit in the eligibility rates that they achieve, which suggests that paradata about the screening process could explain undercoverage. Keystroke files and time stamps could shed light on undercoverage in screening, when screeners are conducted via computer.

5.7 WEB SURVEYS

Conducting surveys via the web has become quite popular in recent years, due to the low marginal costs and the ability to include multimedia materials. Respondents for such surveys are often recruited via mail, telephone, or in-person visits. Web surveys selected from such frames suffer from all of the coverage problems discussed above. However, web surveys also face an additional source of undercoverage: not all of the selected and recruited cases that wish to take part in the survey will be able to do so, due to the fact that they do not have web access or do not know how to use the web to complete the survey. To address this concern, several web panel surveys have begun to provide web access and training to those selected cases that wish to participate in the panel. These respondents are provided with a computer and internet access at their home, solving the problem of undercoverage. See Scherpenzeel and Das (2011) for details on how one such panel was recruited, the LISS panel in The Netherlands. Studies using similar techniques include the Knowledge Networks panel in the United States and the German Internet Panel, which will begin in 2013.

Researchers have begun to use the paradata indicating which cases were provided with internet to explore whether the inclusion of such cases reduces undercoverage bias. Leenheer and Scherpenzeel (2012) study how the demographics of the survey respondents are affected by the inclusion of these cases in the LISS panel. They show that provision of internet access makes the sample more representative of the overall Dutch population. Those cases that would have been undercovered are more likely to be older, to live in single-person homes, and to have migration backgrounds. Eckman et. al. (2012) extend this analysis by comparing the results of substantive analyses with and without these respondents, finding little evidence of undercoverage bias had the non-internet cases been excluded, though this topic warrants further investigation.

5.8 CONCLUSION

This review of coverage research has demonstrated the contributions that paradata have already made in this area, and pointed to possibilities for future research. Paradata produced during frame construction, sampling, and data collection, can provide insights into undercoverage and overcoverage. Most of the above studies focus on estimating coverage rates, or the correlates of coverage rates. Fewer studies explore coverage bias: the effects of such problems with the frames on the collected survey data. Just as the nonresponse literature has in recent years moved away from a focus on nonresponse rates to explorations of nonresponse bias, the coverage literature should whenever possible look not only at rates but at bias. Paradata can and should play an important role in this transition.

One theme that runs through many of the frames discussed above is that when interviewers are involved in frame construction, they introduce interviewer effects. These effects are well known in the measurement error literature (Schnell and Kreuter, 2005) and have also been demonstrated in nonresponse error (O'Muircheartaigh and Campanelli, 1999; West and Olson, 2010). We can see from the above review that interviewers also influence frames when they do traditional or dependent listing, roster household members, or screen for eligibility. Future research in this area to determine the source of these interviewer effects using interviewer-level paradata should be a high priority for the field of coverage research.

Another important lesson from this chapter, and all of the chapters in this volume, is that paradata generated at every stage of the survey process should be captured, retained, and, as much as possible, made available to survey researchers. Only when we have such data are the innovative analyses discussed in this book possible.

ACKNOWLEDGMENTS

The author thanks the Survey Research Center at the University of Michigan and Johannes Bauer of the Ludwig Maximilian University for the use of their maps (Figures 5.3 and 5.4). The following colleagues provided valuable assistance in developing this chapter: Frauke Kreuter, Lars Lyberg, Peter Lundquist, Aleksa Möntmann-Hertz, Joseph Sakshaug, Annette Scherpenzeel, Jennifer Sinibaldi, James Wagner and Ting Yan.

REFERENCES

Alt, C., Bien, W., and Krebs, D. (1991). Wie Zuverlässig ist die Verwirklichung von Stichprobenverfahren? Random Route versus Einwohnermeldeamtsstichprobe. *ZUMA-Nachrichten*, 28:65–72.

Amaya, A., LeClere, F., Fiorio, L., and English, N. (2010). Improving the Utility of the DSF Address-Based Frame through Ancillary Information. *Presentation at the American Association for Public Opinion Research Conference*.

Battaglia, M., Ryan, M., and Cynamon, M. (2005). Purging Out-Of-Scope and Cellular Telephone Numbers from RDD Samples. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 3794–3801.

Bauer, J. (2013). Selection Errors of Random Route Samples. *Sociological Methods and Research*. Forthcoming.

Blumberg, S. and Luke, J. (2011). Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, January-June 2011. Technical report, National Center for Health Statistics.

Boyd, H. and Westfall, R. (1955). Interviewers as a Source of Error in Surveys. *Journal of Marketing*, 19(4):311–324.

Boyle, J., Bucuvalas, M., Piekarski, L., and Weiss, A. (2009). Zero Banks: Coverage Error and Bias in RDD Samples Based on Hundred Banks with Listed Numbers. *Public Opinion Quarterly*, 73(4):729–750.

Brick, J.M., Montaquila, J., and Scheuren, F. (2002). Estimating Residency Rates for Undetermined Telephone Numbers. *Public Opinion Quarterly*, 66(1):18–39.

Brick, J.M., Waksberg, J., Kulp, D., and Starer, A. (1995). Bias In List-Assisted Telephone Samples. *Public Opinion Quarterly*, 59(2):218–235.

Butcher, R. (1988). The Use of the Post-Code Address File as a Sampling Frame. *Journal of the Royal Statistical Society. Series D*, 37(1):15–24.

Casady, R.J. and Lepkowski, J.M. (1993). Stratified Telephone Survey Designs. *Survey Methodology*, 19:103–113.

Central Co-ordinating Team (2010). European Social Survey Round 4 2008/2009. Final Activity Report ESS4e03.0, City University London.

Clark, R. and Steel, D. (2002). The Effect of using Household as a Sampling Unit. *International Statistical Review*, 70(2):289–314.

Czaplicki, A. (2011). GPS in der Markt- und Sozialforschung - Herausforderungen und Chancen. In *Presented at the Jahrestagung of the Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V.*

Dahlhammer, J. (2009). Using Paradata to Assess and Monitor Data Quality in the National Health Interview Survey (NHIS). In *Presented at the Statistics Canada/U.S. Census Bureau Methodological Interchange, Ottawa, Ontario, Canada.*

Diekmann, A. (2010). *Empirische Sozialforschung. Grundlagen Methoden Anwendungen.* Rowohlt, Germany, Reinbeck.

Dohrmann, S., Han, D., and Mohadjer, L. (2006). Residential Address Lists vs. Traditional Listing: Enumerating Households and Group Quarters. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 2959–2964.

Dohrmann, S., Han, D., and Mohadjer, L. (2007). Improving Coverage of Residential Address Lists in Multistage Area Samples. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 3219–3226.

Wie Zuverlässig ist die Verwirklichung von
ersus Einwohnermeldeamtsstichprobe. *ZUMA-*

glish, N. (2010). Improving the Utility of the
lary Information. *Presentation at the American*
Conference.

2005). Purging Out-Of-Scope and Cellular Tele-
Proceedings of the Section on Survey Research

m Route Samples. *Sociological Methods and*

ubstitution: Early Release of Estimates From the
y-June 2011. Technical report, National Center

rs as a Source of Error in Surveys. *Journal of*

Weiss, A. (2009). Zero Banks: Coverage Error
ed Banks with Listed Numbers. *Public Opinion*

(2002). Estimating Residency Rates for Unde-
ion Quarterly, 66(1):18–39.

er, A. (1995). Bias In List-Assisted Telephone
:218–235.

Address File as a Sampling Frame. *Journal of*
1):15–24.

ified Telephone Survey Designs. *Survey Method-*

Social Survey Round 4 2008/2009. Final Activ-
don.

ing Household as a Sampling Unit. *International*

ozialforschung - Herausforderungen und Chan-
e Arbeitsgemeinschaft Sozialwissenschaftlicher

sess and Monitor Data Quality in the National
ted at the Statistics Canada/U.S. Census Bureau
rio, Canada.

schung. Grundlagen Methoden Anwendungen.

006). Residential Address Lists vs. Traditional
oup Quarters. In *Proceedings of the Section on*
9–2964.

7). Improving Coverage of Residential Address
oceedings of the Section on Survey Research

Eckman, S. (2013). Do Different Listers Make the Same Housing Unit Frame? Variability in Housing Unit Listing. *Journal of Official Statistics.* Forthcoming.

Eckman, S., Brendel, M. and Kreuter, F. (2012). Did the Inclusion of Non-Internet Households in the LISS Panel Reduce Coverage Bias? *Presentation at the MESS Workshop*, Amsterdam, NL.

Eckman, S. and English, N. (2012a). Creating Housing Unit Frames from Address Databases: Geocoding Precision and Net Coverage Rates. *Field Methods*, 24(4):399–408.

Eckman, S. and English, N. (2012b). Geocoding to Create Survey Frames. *Survey Practice*, 5(4):1–8.

Eckman, S. and Kreuter, F. (2013). Undercoverage Rates and Undercoverage Bias in Traditional Housing Unit Listing. *Sociological Methods and Research.* Forthcoming.

Eckman, S. and O'Muircheartaigh, C. (2011). Performance of the Half–Open Interval Missed Housing Unit Procedure. *Survey Research Methods*, 5(3):125–131.

English, N., O'Muircheartaigh, C., Dekker, K., and Fiorio, L. (2010). Qualities of Coverage: Who is Included or Excluded by Definitions of Frame Composition? In *Proceedings of the Section on Survey Research Methods, ASA*, pages 3313–3321.

Eyerman, J., Odom, D., and Chromy, J. (2001). Impact of Computerized Screening on Selection Probabilities and Response Rates in the 1999 NHSDA. In *Proceedings of the Section on Survey Research Methods, ASA.*

Fahimi, M., Kulp, D., and Brick, J.M. (2009). A Reassessment of List-Assisted RDD Methodology. *Public Opinion Quarterly*, 73(4):751–760.

Fein, D. (1990). Racial and Ethnic Differences in U.S. Census Omission Rates. *Demography*, 27(2):285–302.

Fink, R. (1963). Interviewer Training and Supervision in a Survey of Laos. *International Social Science Journal*, 15(1):21–34.

Foster, K. (1994). The Coverage of the Postcode Address File as a Sampling Frame. *Survey Methodology Bulletin*, 34:9–18. Office of Population Censuses and Surveys, London.

Gaziano, C. (2005). Comparative Analysis of Within-Household Respondent Selection Techniques. *Public Opinion Quarterly*, 69(1):124–157.

Groves, R.M. and Couper, M.P. (1992). *Nonresponse in Household Surveys.* Wiley and Sons, Inc.

Groves, R.M. (1989). *Survey Errors and Survey Costs.* Wiley and Sons, Inc., New York.

Häder, S. (2000). Telefonstichproben. Technical Report 6, Zentrum für Umfragen, Methoden und Analysen, Mannheim. ZUMA How-to-Reihe.

Hainer, P. (1987). A Brief and Qualitative Anthropological Study Exploring the Reasons for Census Coverage Error Among Low Income Black Households. *Report prepared under contract with the Census Bureau.*

Hales, B., Terblanche, M., Fowler, R., and Sibbald, W. (2008). Development of Medical Checklists for Improved Quality of Patient Care. *International Journal for Quality in Health Care*, 20(1):22–30.

Harter, R., Eckman, S., English, N., and O'Muircheartaigh, C. (2010). Applied Sampling for Large-Scale Multi-Stage Area Probability Designs. In Marsden, P.V. and Wright, J., editors, *Handbook of Survey Research*, 2nd edition, pages 169–197. Emerald.

Haynes, A.B., Weiser, T.G., Berry, W.R., Lipsitz, S.R., Breizat, A.-H.S., Dellinger, E.P., Herbosa, T., Joseph, S., Kibatala, P.L., Lapitan, M.C.M., Merry, A.F., Moorthy, K., Reznick, R.K., Taylor, B., and Gawande, A.A. (2009). A Surgical Safety Checklist to Reduce Morbidity and Mortality in a Global Population. *New England Journal of Medicine*, 360(5):491–499.

Horrigan, M., Moore, W., Pedlow, S., and Wolter, K. (1999). Undercoverage in a Large National Screening Survey for Youths. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 570–575.

Iannacchione, V.G., Morton, K., McMichael, J.P., Cunningham, D., Cajka, J., and Chromy, J. (2007). Comparing the Coverage of a Household Sampling Frame Based on Mailing Addresses to a Frame Based on Field Enumeration. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 3323–3332.

Iannacchione, V.G., Staab, J.M., and Redden, D.T. (2003). Evaluating the Use of Residential Mailing Addresses in a Metropolitan Household Survey. *Public Opinion Quarterly*, 67(2):202–210.

Jowell, R. and the Central Co-ordinating Team (2003). European Social Survey 2002/2003. Technical Report ESS1e06.1, City University.

Jowell, R. and the Central Co-ordinating Team (2005). European Social Survey Round 2 2004/2005. Technical Report ESS2e03.1, City University.

Jowell, R. and the Central Co-ordinating Team (2007). European Social Survey Round 3 2006/2007. Technical Report ESS3e03.2, City University.

Judkins, D. (1997). Listing Updating Techniques. Internal Westat Memorandum.

Kennedy, C., Keeter, S., and Dimock, M. (2008). A "Brute Force" Estimation of the Residency Rate for Undetermined Telephone Numbers in an RDD Survey. *Public Opinion Quarterly*, 72(1):28–39.

Kish, L. (1965). *Survey Sampling*. Wiley and Sons, Inc., New York.

Kwiat, A. (2009). Examining Blocks with Lister Error in Area Listing. In *Proceedings of the Section on Survey Research Methods, ASA*.

Laitila, T., Wallgren, A., and Wallgren, B. (2011). Quality Assessment of Administrative Data. Technical Report 2, Statistics Sweden. Research and Development Methodology Reports.

Lavrakas, P.J. (1987). *Telephone Survey Methods: Sampling, Selection, and Supervision*. Applied Social Research Methods Series. Sage Publications.

Leenheer, J. and Scherpenzeel, A. (2012). Loont het om Huishoudens zonder Internet op te Nemen in een Internetpanel? *Jaarboek MarktOnderzoek Associatie*, 5:75–88. Translation provided by A. Scherpenzeel.

Lessler, J.T. and Kalsbeek, W.D. (1992). *Nonsampling Error in Surveys*. Wiley and Sons, Inc., New York.

Lessler, J.T. (1980). Errors Associated with the Frame. In *Proceedings of the Section on Survey Research Methods, ASA*.

Lynn, P. (2009). Sample Design for Understanding Society. Understanding Society Working Paper Series. Number 2009-01.

Lynn, P. and Taylor, B. (1995). On the Bias and Variance of Samples of Individuals: A Comparison of the Electoral Registers and Postcode Address File as Sampling Frames. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 44(2): 173–194.

Manheimer, D. and Hyman, H. (1949). Interviewer Performance in Area Sampling. *Public Opinion Quarterly*, 13(1):83–92.

Martin, E. (1999). Who Knows Who Lives Here? Within-Household Disagreements as a Source of Survey Coverage Error. *Public Opinion Quarterly*, 63(2):220–236.

Martin, E. (2007). Strength of Attachment: Survey Coverage of People with Tenuous Ties to Residences. *Demography*, 44:427–440.

McMichael, J.P., Ridenhour, J.L., and Shook-Sa, B.E. (2008). A Robust Procedure to Supplement the Coverage of Address-Based Sampling Frames for Household Surveys. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 4251–4257.

Merkle, D.M., Langer, G., Cohen, J., Piekarski, L.B., Benford, R., and Lambert, D. (2009). Unintended Consequences: The Cost of Purging Business Numbers in RDD Surveys. *Public Opinion Quarterly*, 73(3):484–496.

Morton, K., Iannacchione, V., McMichael, J., Cajka, J., Curry, R., and Cunningham, D. (2007). Linking Mailing Addresses to a Household Sampling Frame Based on Census Geographies. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 3971–3974.

Noelle, E. (1963). *Umfragen in der Massengesellschaft. Einführung in die Methoden der Demoskopie*. Rohwolt.

O'Muircheartaigh, C., Eckman, S., and Weiss, C. (2002). Traditional and Enhanced Field Listing for Probability Sampling. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 2563–2567.

O'Muircheartaigh, C., English, N., and Eckman, S. (2007). Predicting the Relative Quality of Alternative Sampling Frames. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 551–574.

O'Muircheartaigh, C., English, N., Eckman, S., Upchurch, H., Garcia, E., and Lepkowski, J.M. (2006). Validating a Sampling Revolution: Benchmarking Address Lists against Traditional Listing. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 4189–4196.

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, Series A*, 162(3):437–446.

Pacer, J. (2008). A Comparison of Commercial Listing to Traditional Listing: Coverage Rates in the National Survey of Family Growth, Cycle 7. Unpublished Paper, Program in Survey Methodology, University of Michigan.

Scherpenzeel, A. and Das, M. (2011). "True" Internet Panels: Evidence from the Netherlands. In *Social Research and the Internet: Advances in Applied Methods and New Research Strategies*, pages 77–104. Taylor & Francis, Boca Raton, FL.

Schnell, R. (2008). Avoiding Problems of Traditional Sampling Strategies for Household Surveys in Germany: Some New Suggestions. *Data Documentation 33, DIW Berlin, German Institute for Economic Research.*

Schnell, R., Hill, P., and Esser, E. (2008). *Methoden der Empirischen Sozialforschung*. Wissenschaftsverlag, Germany, Oldenburg.

Schnell, R. and Kreuter, F. (2005). Separating Interviewer and Sampling-Point Effects. *Journal of Official Statistics*, 21(3):389–410.

Shook-Sa, B.E., McMichael, J.P., Ridenhour, J.L., and Iannacchione, V.G. (2010). The Implications of Geocoding Error on Address-Based Sampling. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 3303–3312.

Shore, S.P., Montaquila, J.M., and Hsu, V. (2010). The Presence and Characteristics of Households at Addresses Obtained by Traditional Field Listing and from USPS-Based Lists. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 1709–1718.

Simmons, W. (1953). Prelisting in Market or Media Surveys. *Journal of Marketing*, 18(1):6–17.

Skalland, B. (2011). An Alternative to the Response Rate for Measuring a Survey's Realization of the Target Population. *Public Opinion Quarterly*, 75(1):89–98.

Statistics Canada (2008). Methodology of the Canadian Labour Force Survey. Technical Report 71-526-X, Statistics Canada.

Thompson, G. and Turmelle, C. (2004). Classification of Address Register Coverage Rates: A Field Study. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 4477–4484.

Tourangeau, R., Kreuter, F., and Eckman, S. (2012). Motivated Underreporting in Screening Interviews. *Public Opinion Quarterly*. Forthcoming.

Tourangeau, R., Shapiro, G., Kearney, A., and Ernst, L. (1997). Who Lives Here? Survey Undercoverage and Household Roster Questions. *Journal of Official Statistics*, 13(1):1–18.

Tucker, C., Lepkowski, J.M., and Piekarski, L. (2002). The Current Efficiency of List-Assisted Telephone Sampling Designs. *Public Opinion Quarterly*, 66(3):321–338.

Turmelle, C., Rodrigue, J.-F., and Thompson, G. (2005). Using the Canadian Address Register in the Labour Force Survey Implementation, Results and Lessons Learned. In *Proceedings of the Conference of the Federal Committee on Statistical Methodology*.

U.S. Census Bureau (2006). *Design and Methodology, Current Population Survey*. Technical report 66.

Valentine, C. and Valentine, B. (1971). Missing Men: A Comparative Methodological Study of Underenumeration and Related Problems. *Report prepared under contract for the US Census Bureau. Washington, DC*, 23.

West, B.T. and Olson, K. (2010). How Much of Interviewer Variance is Really Nonresponse Error Variance? *Public Opinion Quarterly*, 74(5):1027–1045.

Wilson, P.R. and Elliot, D.J. (1987). An Evaluation of the Postcode Address File as a Sampling Frame and its Use within OPCS. *Journal of the Royal Statistical Society, Series A*, 150(3):230–240.

Wright, T. and Tsao, H. (1983). A Frame on Frames: An Annotated Bibliography. In Wright, T., editor, *Statistical Methods and the Improvement of Data Quality*. Academic Press.

Yan, T. (2009). A Meta-analysis of Within-Household Respondent Selection Methods. In *Proceedings of the Section on Survey Research Methods, ASA*, pages 6134–6147.

Zandbergen, P. (2008). A Comparison of Address Point, Parcel and Street Geocoding Techniques. *Computers, Environment and Urban Systems*, 32:214–232.

PART II

PARADAT