

Address-based Sampling

AAPOR REPORT: ADDRESS-BASED SAMPLING

January 7, 2016

Prepared for AAPOR Council by the Task Force on Address-based Sampling, operating under the auspices of the AAPOR Standards Committee, with Task Force members including:

Rachel Harter, RTI International, Task Force Chair
Michael P. Battaglia, Battaglia Consulting Group, LLC
Trent D. Buskirk, Marketing Systems Group
Don A. Dillman, Washington State University
Ned English, NORC at the University of Chicago
Mansour Fahimi, GfK Custom Research, LLC
Martin R. Frankel, Baruch College
Timothy Kennel, U.S. Census Bureau
Joseph P. McMichael, RTI International
Cameron Brook McPhee, American Institutes for Research
Jill Montaquila, Westat
Tracie Yancey, Nielsen Company
Andrew L. Zukerberg, National Center for Education Statistics

The Task Force was supported by the following expert consultants:

Anne Connelly, Cigna, formerly of Valassis, Inc.
Philip Faulstich, Valassis, Inc.
David Malarek, Marketing Systems Group
Missy Mosher, SSI
Linda Piekarski, SSI
Bonnie Shook-Sa, RTI International

The task force gratefully acknowledges the helpful comments from J. Michael Brick, AAPOR council members, and their designees. The task force also gratefully acknowledges the editorial support provided by RTI International.

Contents

1. [Introduction](#)
 - 1.1 Why is Address-based Sampling Important?
 - 1.2 [What is Address-based Sampling?](#)
 - 1.3 [Goals of the Report](#)
 - 1.4 [Preview of the Report](#)

2. [Frame Creation: Sources, Filtering, Matching, and Deduplication](#)
 - 2.1 Vendors, Licenses, and the USPS
 - 2.1.1 Vendors, Licenses, and the USPS: Address Management System
 - 2.1.2 Vendors, Licenses, and the USPS: Computerized Delivery Sequence
 - 2.1.3 Vendors, Licenses, and the USPS: Delivery Sequence File Second Generation
 - 2.1.4 Vendors, Licenses, and the USPS: Considerations
 - 2.2 [Addresses as Sampling Units](#)
 - 2.2.1 Addresses as Sampling Units: Multiple Addresses per Household
 - 2.2.2 Addresses as Sampling Units: One Address for Multiple Households
 - 2.3 [Frame Creation Activities](#)
 - 2.3.1 Frame Creation Activities: Address Standardization
 - 2.3.2 Frame Creation Activities: Coverage Improvement
 - 2.3.3 Frame Creation Activities: Matching Files
 - 2.3.4 Frame Creation Activities: Conversions
 - 2.3.5 Frame Creation Activities: Auxiliary Data
 - 2.3.6 Frame Creation Activities: Deduplication
 - 2.3.7 Frame Creation Activities: Filtering
 - 2.3.8 Frame Creation Activities: Updating
 - 2.4 [Checklist of Questions to Ask Vendors](#)
 - 2.5 [Section Summary](#)
3. [Auxiliary Variables](#)
 - 3.1 Frame Supplementation With Auxiliary Variables
 - 3.1.1 Two Methods for Appending Auxiliary Data: Geocoding and Matching
 - 3.1.2 Geocoding
 - 3.1.3 Data Matching
 - 3.2 [Types and Sources of Auxiliary Variables](#)
 - 3.2.1 Phone Numbers
 - 3.2.2 Commercial Data
 - 3.2.3 Government Data
 - 3.3 [Auxiliary Data Quality](#)
 - 3.4 [Section Recommendations](#)
4. [Designing and Implementing ABS Surveys](#)
 - 4.1 Sampling Frame Considerations for ABS Surveys
 - 4.1.1 Drop Points
 - 4.1.2 Seasonal and Educational Addresses
 - 4.1.3 Vacant Addresses
 - 4.1.4 Addresses versus Households
 - 4.2 [Address Sample Selection](#)
 - 4.3 [Sampling Adults Within Households](#)
 - 4.4 [Sampling Subgroups](#)
 - 4.5 [Use of Auxiliary Variables in Sampling](#)
 - 4.6 [Sample Design Consideration for Local versus National Surveys](#)
 - 4.7 [ABS for Nonhousehold Surveys](#)

- 4.8 [Survey Designs for Single- and Mixed-mode Surveys](#)
 - 4.8.1 ABS for In-person Surveys Using Area Probability Samples
 - 4.8.2 ABS for Single-mode Mail Surveys
 - 4.8.3 ABS for Mixed-mode Surveys
 - 4.8.4 Contact Mode versus Interview Mode
- 4.9 [The Use of Incentives](#)
- 4.10 [Section Recommendations](#)
- 5. [Eligibility, Response Rates, and Weights](#)
 - 5.1 Eligibility Issues
 - 5.1.1 Determining the Eligibility of Each Sampled Address
 - 5.2 [Response Rates](#)
 - 5.2.1 Estimating e
 - 5.3 [Weights](#)
 - 5.3.1 Design Weights
 - 5.3.2 Multimodality Weighting Issues
 - 5.3.3 Adjusting for Unknown Eligibility
 - 5.3.4 Unit Nonresponse Adjustments
 - 5.3.5 Poststratification to Control Totals
 - 5.4 [Section Summary](#)
- 6. [Reporting Guidelines](#)
 - 6.1 Current Best Practices for Reporting
 - 6.2 [Section Summary](#)
- 7. [Quality and Cost Issues for Address-Based Samples](#)
 - 7.1 Frame Variable Quality
 - 7.2 [Determining Coverage](#)
 - 7.2.1 Effect of Sample Design on Coverage Error
 - 7.2.2 Coverage Evaluations
 - 7.2.3 Decisions to Make When Evaluating Coverage
 - 7.2.4 Impact of Geocoding Error
 - 7.3 [Methods for Improving Coverage](#)
 - 7.3.1 Linking Procedures
 - 7.3.2 Supplementary Frames
 - 7.4 [Cost and Quality Tradeoffs](#)
 - 7.4.1 Factors to Consider When Looking at Costs
 - 7.5 [Section Recommendations](#)
- 8. [Summary of the Current State of ABS](#)
 - 8.1 Strengths and Limitations of Address-based Sampling
 - 8.1.1 Sampling Frame versus Delivery Database
 - 8.1.2 Reliability of Auxiliary Variables
 - 8.1.3 Frame Multiplicity, Vacant Addresses, and Other Frame Issues
 - 8.1.4 Mode Effects
 - 8.1.5 Required Infrastructure, Time, and Costs
 - 8.2 [Report Summary](#)
 - 8.3 [Future Research](#)
- 9. [References](#)
- Appendix A. [Definitions](#)

SECTION 1 INTRODUCTION

1.1 Why is Address-based Sampling Important?

Arguably, address lists updated via the United States Postal Service (USPS) Computerized Delivery Sequence (CDS) file are the best possible frames for today's household surveys in the United States. National coverage estimates vary, but are very high overall and nearly 100% in many areas, and coverage continues to improve. In addition, many address lists are regularly updated with changes from the USPS CDS file, reducing the need for expensive field work by survey organizations. Historically, field-generated frames were the only option for in-person surveys, but the high cost was prohibitive for many important national surveys, not to mention other valuable research surveys at the state, region, or community level.

For many years, telephone surveys have been the low-cost alternative to in-person surveys with field-generated frames. However, the nature of telephony has shifted dramatically toward cellular technology (Blumberg and Luke 2014; Keeter et al. 2007). With more households switching from landline to mobile telephones, the coverage of landline-based random digit dialing (RDD) frames has dwindled (Blumberg and Luke 2014). Furthermore, because of legislation regarding how survey researchers may dial cell phones, and because of generally lower response rates for cell phone numbers, the cost of telephone surveys that seek coverage of cell-only households is increasing (AAPOR Cell Phone Task Force 2010). Address-based sampling (ABS) offers attractive solutions to these coverage and cost problems in the United States (Link et al. 2008).

The accessibility of address frames has reduced the cost of in-person surveys and brought about a resurgence of relatively inexpensive mail surveys. ABS is often used in multimode studies, where different modes may be used for contact versus response in data collection or to follow up with nonrespondents (Alexander and Wetrogan 2000; de Leeuw 2005). Alternatively, advance mailings can be used to direct selected households to web surveys, with the hope that doing so may dramatically reduce costs. Furthermore, the ability to append geocodes, phone numbers, demographics, and other data to the address frame, although imperfect, can provide deep stratification and aid in designing more cost-efficient studies.

Society is changing through the way people communicate. Letters and telephone calls are largely being replaced by texts, tweets, e-mails, and other electronic communications, although mail is still used for some formal and official communications. Surveys that push selected individuals to respond to surveys electronically (e.g., via the web) take advantage of today's prevalent modes of communication. Without general frames of electronic addresses, mail addresses provide excellent coverage of households. At the same time, initial contact by mail ensures that virtually every selected household can be reached, regardless of electronic capabilities. Creative use of ABS provides many options for reaching busy households and gaining cooperation.

The purpose of this report is to describe the nature of ABS and its uses for conducting surveys. Multiple specific goals of the report are presented in Section 1.3. The report discusses in detail technical aspects of constructing ABS frames and samples, and the technical aspects reveal both its strengths and limitations. These aspects are important for effective use of ABS in survey design and implementation, as described in the report.

1.2 What is Address-based Sampling?

ABS is sampling from address frames that are usually based, in part, on the USPS files. Although approaches for selecting samples of addresses for mail or in-person surveys have existed for decades, the survey's success depends heavily on the availability of a suitable address frame. For some geographies or some subpopulations, special purpose address lists may be available. Only in the prior decade, however, did survey organizations begin to use commercially available address lists updated at least in part by the USPS (Iannacchione, Staab, and Redden 2003). Quite simply, the address lists available today are the best frames available for national U.S. household surveys.

ABS is best understood in the context of its history and its potential. Before ABS frames were commercially available, survey designers sometimes built a frame of housing units within small geographical areas selected through a multistage design (Kish 1965). At the first stages, high-quality frames exist for geographical areas. Counties or metropolitan statistical areas are commonly used as primary sampling units, and smaller, well-defined areas such as census blocks might be secondary or tertiary sampling units. With a sample of manageably small geographies, field personnel can traverse the sampled areas and generate lists of addresses or housing units for frames at the housing unit level. Such designs are sometimes referred to as area probability designs, and the process of gathering listings in the field is called field enumeration, field listing, or traditional listing. This practice is still done for studies such as the National Survey of Drug Use and Health (Center for Behavioral Health Statistics and Quality 2014) and the National Health and Nutrition Examination Survey (Johnson et al. 2014). Although field listings were considered the best frames for many years, they have their drawbacks. First and foremost, collecting addresses in the field is extremely costly, and very few survey organizations are willing and able to generate an address or housing unit frame in this way. Second, field listing takes considerable time, and frames can become out of date rather quickly. Third, listers sometimes make mistakes (Eckman 2010). Field methods to correct for errors or omission are also error prone (Eckman and Kreuter 2011).

In truth, using a national address list for sampling is not new. The U.S. Census Bureau has sampled from national decennial address files for decades, but their files are confidential and not available outside the U.S. Census Bureau until 72 years have passed from the census. In 1994, Congress amended Section 412 of Title 39 in the U.S. Code (1994) to require the USPS to provide postal addresses to the Secretary of Commerce for surveys and censuses conducted by the U.S. Census Bureau, and this

action may have spurred interest in the USPS addresses elsewhere in the survey industry. However, Title 39 prohibits the release of USPS address lists to other parties, and Title 13 further protects the addresses the Census Bureau receives.

Even so, the USPS was already offering to “correct” existing mailing lists for mass mailers (updating the lists and putting addresses in proper format), provided the mailing lists covered at least 90% of addresses in a ZIP Code and did not exceed 110% of the addresses in a ZIP Code (U.S. Postal Service 2013a). (NOTE: In this context, covering between 90% and 110% of the addresses in a ZIP Code is referred to as “owning” a ZIP Code.) Marketing companies with address lists for mass mailings were among the first to take advantage of the service. Qualified licensees of the USPS’s CDS file (Section 2.1) can receive updates weekly or every 2 months for their proprietary address lists in the geography they own, which they can, in turn, license in part or in whole to survey organizations. A small number of marketing database companies receive CDS updates and make their proprietary address frames available to survey organizations. Also, a small number of survey organizations have licensed the entire set of U.S. addresses from these vendors as frames for their own sample selections.

More recently, organizations that historically provided telephone samples to survey organizations have also obtained address lists from vendors, and they can now select and provide samples of addresses to survey organizations. Many survey organizations have taken advantage of such samples without incurring the costs of developing or leasing their own national frame.

Both types of vendors, frame providers and sample providers, typically offer value-added data or services, such as the ability to append householder name or telephone number, indicators regarding the type of address, and demographic information on occupants of the households, where available. Survey organizations of all sizes can now obtain frames or samples from frames that nearly cover the entire population of U.S. mailing addresses. Sampling from address frames that are highly dependent on the USPS files has come to be known as address-based sampling or ABS. Iannacchione (2011) and Link (2010) provide helpful overviews of ABS.

Even though ABS evolved initially as a lower cost method of building frames of housing units for in-person surveys, its greater potential is as an alternative to RDD or part of a multimode mix. AAPOR members are more likely to encounter ABS in these applications. This report covers the basic ABS concepts to help the survey profession understand the possibilities, limitations, and options inherent in ABS.

1.3 Goals of the Report

Any major new development in the science of surveys deserves research and documentation of current best practices. Effective use of ABS requires knowledge of the files and their properties, including coverage and timeliness. Addresses are often geocoded for targeted areas, an enhancement that increases the utility of these files but introduces the need to deal with the implications of geocoding error. When combined

with auxiliary data such as census demographics for the geographies or marketing data for the households, the resulting designs can be more efficient in targeting specific subpopulations, but sometimes with additional risk of bias and variance. Methodologies are evolving with the technologies, and survey sponsors may be confused by the tradeoffs of one methodology versus another. Furthermore, responsible reporting of ABS surveys should disclose aspects of the design that currently are not reported consistently.

The overall goals of this report are (1) to educate the survey community about the current state of ABS; (2) to supplement AAPOR guidelines for ABS issues, consistent with existing standards for scientific surveys; (3) to encourage the survey industry to think creatively about the uses of ABS so that surveys can keep pace with technology and the changing ways in which information is communicated; and (4) to encourage research to fill the gaps in our knowledge of ABS. In particular, the goals of this report are to

- Standardize and clarify ABS terminology.
- Inform the survey industry about frame building and frame characteristics.
- Comment on frame enhancements and the benefits and risks.
- Discuss sample design possibilities with ABS.
- Discuss data collection methodology with ABS.
- Discuss special issues with case dispositions, especially with ABS mixed-mode surveys.
- Recommend methods for weighting and for computing response rates for ABS studies.
- Recommend standards for reporting ABS studies to help assess the quality of a study.
- Discuss quality and cost issues associated with ABS.
- Point out the limitations of ABS and opportunities for advancing the field.

The reader should also be aware of the limitations of this report. The current review is not meant to stand the test of time in the rapidly changing survey environment. Rather, this report represents the current state of the source files and methods at the time of this writing. Throughout the document, we have noted areas in need of research, and we expect that additional research will continue on aspects that have already been investigated in some fashion. In addition, we fully expect that aspects that are integral to this methodology (e.g., features of the files used in constructing ABS frames, and changes in the ways people communicate, including increased use of smartphones) will influence how ABS surveys are conducted in the future. Although every effort has been made to reflect the current best practices and considerations for ABS studies, we expect that there will be a need to revisit and revise this report periodically.

1.4 Preview of the Report

Various aspects of ABS are discussed in the remaining sections. To educate the reader adequately about the features and limitations of ABS, the material is presented with considerable technical detail. Although detailed, the material should be accessible to the AAPOR reader familiar with survey research designs and methods in practice. ABS terminology that may be unfamiliar to survey researchers is explained for the ABS context in Appendix A.

Key aspects of ABS are covered in the remaining sections of the report. The reader will note some repetition of technical details and general concepts across sections, but some repetition is necessary in discussing various aspects of ABS in sections structured as follows:

- There are various issues related to address frames that survey designers should discuss with their ABS frame or sample providers. The available frames, variables they typically include, and other related issues are discussed in Section 2.
- There are many ways in which an address frame can be enhanced by auxiliary variables. The varying coverage and accuracy of auxiliary variables appended to the frame, and the dizzying array of possibilities for sample design and data collection, add to the creativity of ABS. The rapidly changing options pose both opportunities and challenges. These aspects are explored in Section 3.
- ABS encompasses a very broad array of designs and methodologies. Section 4 describes how to apply ABS methods to designing survey frames and samples, and ways of using them for data collection. This section discusses both single- and mixed- mode survey designs.
- Tricky issues of eligibility and dispositions may be encountered in ABS studies, especially mixed-mode studies. These issues feed directly into computational methods for response rates and weights. Section 5 details these issues and prescribes methods for dealing with them.
- ABS studies should adhere to general reporting standards, but encompassed within these are some points and considerations that are specific to ABS. Section 6 reviews current best practices for reporting on studies as they apply to ABS studies.
- All survey design options involve tradeoffs. Section 7 discusses the cost and quality aspects of ABS. Issues such as coverage, accuracy, timeliness, and cost are discussed, along with methods that have been developed to improve coverage. Even so, our profession has not yet achieved a full understanding of cost and quality tradeoffs in the context of ABS designs, and this area is ripe for research.
- Although ABS has opened the door to a new collection of methods, and the use of existing methods in a new context, it is not without limitations and unknowns. Section 8 summarizes the limitations of ABS that are discussed in prior sections, and the strengths that make ABS so promising. This section summarizes ABS topics suggested for further research in its concluding remarks.

Although the field of ABS is rapidly evolving, this report attempts to summarize concepts, present research literature to date, and address current best practices for the scientific survey industry, while noting areas where additional research is desirable. Especially in the realm of multimode data collection, including web data collection, it is incumbent on us to thoroughly research the tools necessary for the future of our profession. Ideally, survey methodologists, statisticians, and practitioners will think seriously about the possibilities presented by address-based frames and move the survey industry into the future of gathering valuable information in our changing culture.

SECTION 2

FRAME CREATION: SOURCES, FILTERING, MATCHING, AND DEDUPLICATION

Typically, the ABS frame creation process is done by a data vendor, whose business rules, decisions, and procedures used to create ABS frames are often proprietary. However, these rules, decisions, and procedures can have a major impact on the quality and cost of a survey. For example, how a vendor deduplicates an address list and filters the list for erroneous inclusions can directly influence the cost, quality, and coverage of a survey. A frame with many out-of-scope units may result in a smaller number of completed interviews, which in turn may result in larger variances and costs per completed interview. On the other hand, frames with systematic omissions may result in biased estimates. Whether a frame or sample is purchased from a vendor, the techniques vendors use to create, process, and update address files can add or reduce errors and costs to the survey.

One of the primary goals of this chapter is to illuminate some of the processes that vendors employ when creating frames and illustrate how they may affect the cost and quality of the survey. A primary source of all ABS frames, the Address Management System (AMS), is protected by law, and the processes used to create, update, and maintain it are not generally known. A second goal of this chapter is to establish a common understanding of the AMS and highlight some known issues with the AMS that may impact the quality of ABS surveys.

The process of creating an ABS frame often involves minimizing both overcoverage and undercoverage, subject to cost functions. Undercovered units are in the target population, but erroneously excluded from the frame or omitted from it. Overcovered units are on the sampling frame, but are not in the target population. Duplicates and erroneous inclusions are examples of overcovered units. To illustrate the balancing of overcoverage and undercoverage errors, subject to cost, consider the situation in which a small number of residential addresses are misclassified as business addresses. Including all business addresses in a survey of households would result in a great amount of overcoverage for the sake of small reductions in undercoverage. Thus, it usually is advisable to exclude business addresses in household surveys. Similarly, sometimes addresses in Alaska and Hawaii are excluded from sampling frames because interviewing in those states is not within budget constraints. Although decisions about which addresses to include and exclude in the sampling frame are proprietary, in most cases vendors may be open to creating extracts that meet users' needs.

In this section, the different types of vendors are explained, the relationship between addresses and sample units is discussed, and details about typical ABS frame creation processes are given. Parts of this section are highly technical and detailed. Readers who are new to ABS and want some background prior to engaging a data vendor may want to skim to the checklist of questions to ask vendors and conclusions at the end of this section. On the other hand, survey methodologists, sampling statisticians, readers familiar with ABS frames, frame creators, and other researchers analyzing ABS frames may want to spend some time engaging with the technical material.

2.1 Vendors, Licenses, and the USPS

Sometimes called information resellers, data brokers, or direct mailers, vendors sell ABS frames and samples. Vendors differ in their relationship to the USPS, the source of their addresses, the services they provide, and their geographic coverage. Given the amount of variability among vendors, the process of selecting a vendor can be daunting. Furthermore, there are important technical distinctions between vendors and the quality of the frames they produce. Some of the types of vendors are discussed here. The choice of a vendor is very important to the total quality and cost of a survey. If a sample is being purchased, it is essential to have a trustworthy vendor because the survey researcher will not have access to the frame to verify or evaluate the frame or sampling process.

There are two types of vendors that provide ABS frames and samples. Primary vendors hold a Delivery Sequence File Second Generation (DSF2) or CDS license with the USPS, while secondary vendors do not. Before obtaining a CDS license, primary vendors must already have an address file that meets minimum quality thresholds (Dohrmann, Han, and Mohadjer 2007; U.S. Postal Service 2013a). Vendors with a DSF2 license do not need to meet the same rigorous standards, but the files they receive are generally complete and derived from the same source as that provided to the CDS vendors (U.S. Postal Service 2014a). Primary vendors rely on different sources of addresses, including local tax records, phone directories, or credit card databases. Some vendors enhance the USPS lists with additional addresses from these other sources, especially in areas with known USPS undercoverage of locatable addresses.

Because they do not have a relationship with the USPS, secondary vendors often purchase addresses from a primary vendor and add additional variables or addresses to the purchased list. Although primary vendors undergo a bit more scrutiny than secondary vendors and often have more technical knowledge and expertise with frame creation, high-quality frames and samples can be purchased from primary or secondary vendors.

Vendors also differ in the services they provide. Some vendors provide only samples, some provide only address lists, and some provide both. Some vendors append

auxiliary data such as geocodes, phone numbers, and person-level data, while others focus on providing only addresses.

Some vendors have national address lists, while other vendors focus on specific geographic areas. Furthermore, even the vendors that have national address lists may differ considerably at local levels because vendors may not hold DSF2 or CDS licenses for all ZIP Codes and address types in the United States (Dohrmann et al. 2007; Iannacchione 2011; U.S. Postal Service 2013a).

Much of this section outlines the various techniques that vendors use to create frames. Vendors approach these methods differently and employ different business rules and procedures to create sampling frames and samples. When selecting a vendor, one should make sure the vendor has the ability to meet the frame or sample requirements.

Although not a vendor, it should be noted for completeness that the U.S. Census Bureau has a special relationship with the USPS. All addresses from the USPS are delivered to the U.S. Census Bureau twice a year for statistical and frame creation purposes (Section 412 of Title 39 in the U.S. Code). The Census Bureau refers to these files as the raw DSF (equivalent of AMS database). There are no licensing agreements for the raw DSF, and the Census Bureau's address frame cannot be disseminated outside the Census Bureau.

Despite the variability among vendors, there are some forces that bring conformity among vendors. First, all vendors with a CDS or DSF2 license must be certified by the USPS and undergo a rigorous application process (2013a; 2014a). Second, all vendors with a CDS or DSF2 license clean, verify, and validate their address lists with information from the same source (U.S. Postal Service 2013a; U.S. Postal Service 2014a). Although vendors may enhance their frames with other sources, the size and frequency of updates from the USPS provides an incentive for vendors with a CDS or DSF2 license to become aligned with parts of the AMS. Additionally, secondary vendors that purchase lists from vendors with a CDS or DSF2 license may also sell frames that mirror the AMS because their primary input is aligned with the AMS.

In general, the frames created by primary vendors are updated more frequently than frames created by secondary vendors. Thus, the coverage of frames from primary vendors is expected to be slightly higher than that of secondary vendors, although there is no empirical research to support this claim. On the other hand, secondary vendors may have specialized lists that can be used to target segments of the population or enhance the address frame with phone numbers, person names, e-mail addresses, and other variables for customization.

Typically, ABS frames are created by vendors who have a license agreement with the USPS to update and verify their address lists. Depending on their licensing contracts, vendors can receive updates through the DSF2 or the CDS file (Dohrmann, Han, and Mohadjer 2006). The CDS updates are more comprehensive, so CDS vendors are

usually preferred to those with a DSF2 license (Kalton, Kali, and Sigman 2014). However, because of proprietary contracts, the coverage and address quality differences between vendors with DSF2 and CDS licenses have not been explicitly published in the literature. Certainly, there are some high-quality lists from vendors with DSF2 licenses as there are high-quality lists from vendors with CDS licenses. Both the DSF2 and CDS files are part of the Address Information System (AIS), which consists of extracts from the USPS AMS database.

2.1.1 Vendors, Licenses, and the USPS: Address Management System

The USPS maintains the AMS for sorting and sequencing of mail, which is why the primary extract is called the Delivery Sequence File. Addresses on letters are scanned and routed to post offices and further sorted based on the DSF. When most letter carriers receive their mail, it is usually sorted based on their assigned route. Thus, the letters delivered to the first house on a delivery route are sorted to the top of the pile and letters delivered to the last house on the carrier's route are sorted to the bottom. The DSF determines how mail is sorted.

Although the AMS serves as a major source of ABS frames and samples, it was not created as a sampling frame. For this reason, vendors and survey researchers must put significant energy into finding ways to transform the AMS into a sampling frame. To do so, it is useful to begin with some basic, albeit technical, information about the purpose, construction, and maintenance of the AMS.

The AMS contains more than 170 million addresses. Using proprietary business rules to filter out some units, the USPS creates a number of extract files through the AIS from the AMS database. These extracts form the CDS, DSF2, No-Stat Files, Delivery Statistics Files, City-State files, 5-digit ZIP product, ZIP+4® product, Zip4Change product, ZIPMove file, enhanced Line of Travel (eLOT®) product, Congressional District Code Files, and Carrier Route Files, which vendors use to update and verify their address lists. More information on the USPS files can be found in the USPS AIS Products Technical Guide (U.S. Postal Service 2015b) and the CDS User Guide (U.S. Postal Service 2013a).

Very little is publicly known about methods used to update and clean the AMS. Mail carriers provide many updates to addresses on their routes, but local governments, post offices, and some vendors also provide new addresses and address updates to the USPS. The quality and timeliness of these updates can vary considerably by geography. For example, one local government may provide the USPS with a list of addresses for a proposed new housing development several years before construction begins. Other local governments might not regulate new construction and addresses at all and thus might not provide new construction addresses to the USPS very far in advance of occupancy, if at all.

Addresses in the AMS are parsed into standard components or address fields(Section 2.3.1). In addition to these address fields, the AMS includes the following attributes

associated with each address (Iannacchione 2011), to be discussed in subsequent sections:

- Seasonal indicator
- Educational indicator
- Vacant indicator
- Delivery mode type indicator
- Residential indicator
- Business indicator
- Drop indicator
- Drop count
- Locatable Address Conversion System (LACS) indicator
- No-Stat(istics) indicator
- Address throwback indicator (U.S. Postal Service 2015c)

These attributes are appended to vendor lists for both CDS and DSF2 licenses. It is also important to note that the CDS and DSF2 licenses deal with mailing address information. Person names, phone numbers, e-mail addresses, global positioning system (GPS) coordinates, number of housing units, or persons living at an address are not provided by the USPS.

2.1.2 Vendors, Licenses, and the USPS: Computerized Delivery Sequence

Section 412 of Title 39 of the U.S. code prevents the USPS from disclosing specific names and addresses to any person or business, apart from the U.S. Census Bureau. Thus, vendors do not receive complete address files through the CDS program. Rather, the USPS provides services that are helpful in cleansing and updating the vendor files. To qualify for a CDS license, vendors initially send their address lists to the USPS, which sequences the vendor files, removes undeliverable addresses, and adds new addresses (U.S. Postal Service 2013a).

Before receiving updates through the CDS program, a vendor must be classified as a qualified mailer. To qualify, a vendor must already have a list containing between 90% and 110% of the USPS addresses in a ZIP Code, and if a vendor meets this qualification for a ZIP Code, the vendor is said to “own” the ZIP Code (Dohrmann et al. 2007; Iannacchione 2011; U.S. Postal Service 2013a). Vendors may qualify for an entire ZIP Code or for just certain address groups within a ZIP Code (U.S. Postal Service 2013a). There are six address groups:

- City Carrier Residence Only
- City Carrier Business
- City Carrier Combination Residence and Business

- Post Office™ Box
- Rural Route and Contract Delivery Service Route (U.S. Postal Service 2013b)
- Combined Delivery Type

Thus, one vendor may qualify to receive CDS updates only for City Carrier Residences in a ZIP Code, while another vendor may qualify for all addresses in the ZIP Code. The updating process is done separately for each ZIP Code or address group. Vendors with a CDS license do not necessarily own all ZIP Codes in the United States, and thus do not necessarily receive updates from the USPS for all ZIP Codes in the United States (U.S. Postal Service 2013a).

Vendors with a CDS license receive weekly or bimonthly updates from the USPS for the ZIP codes and address groups they own (U.S. Postal Service n.d.-a). Weekly updates include only the addresses that changed from one week to the next. In addition to these frequent updates, they also receive base files containing changed and unchanged addresses twice a year. An added advantage of the CDS license is that the updates include vacant and seasonal housing flags. The weekly change files specifically include “the addition of new addresses, changes in delivery sequence, addresses moving from one route to another, or changes to No-Stat or Vacancy coding” (U.S. Postal Service 2013a, p. 5).

Vendors with a CDS license may also purchase the No-Stat File, which contains planned addresses in new housing developments, vacant addresses on rural routes, addresses on rural routes where mail is forwarded to post office boxes, and addresses in some gated communities where mail is delivered to a central point. Most addresses in the No-Stat file do not receive mail; however, one analysis found that using a portion of the No-Stat File improved rural coverage by 2.2%, without adding too many erroneous inclusions (Shook-Sa et al. 2013). Addresses on the No-Stat File can be incorporated into the regular CDS files through the eLot file (U.S. Postal Service n.d.-a), which provides a link in the sequence between the two files.

2.1.3 Vendors, Licenses, and the USPS: Delivery Sequence File Second Generation

The DSF² license is available to a broader group of vendors who do not need to have existing address lists that meet the qualification guidelines under the CDS (Iannacchione 2011). Some vendors with a DSF² license could qualify for a CDS license, but choose not to. Others may not be able to meet the qualification standards. Vendors with the DSF² license send their addresses to the USPS and receive a file in return indicating if each address appears on the AMS; if the address is vacant, residential, business, or seasonal; if the address is delivered to a curb- side mailbox or to a door slot; if the address is on the No-Stat File; and if the address is a throwback address (U.S. Postal Service 2015c). Addresses must be coded through a ZIP+4 system prior to being sent for updates through the DSF² system. Unlike the updates provided to CDS licensees, change files with new addresses and other changes are not included with the DSF² license.

Although DSF² vendors do not receive update files from the USPS with new addresses, some DSF² vendors may get updates through supplemental files not originating with the USPS. Thus, the coverage of a DSF² vendor's address list is not necessarily less than the coverage of CDS vendor lists, but it is important for those considering a DSF² licensed vendor to obtain information about any such supplementation the vendor does.

2.1.4 Vendors, Licenses, and the USPS: Considerations

High-quality frames and samples can be purchased from primary or secondary vendors. When selecting a vendor, one should consider how much of the target population the vendor owns. For example, large undercoverage errors should be expected for a survey of rural Americans from files purchased from CDS vendors that do not own any Rural Route and Contract Delivery Service Route addresses. Secondary vendors should be ready to disclose if their primary vendor does not own any specific ZIP Codes or groups of addresses.

2.2 Addresses as Sampling Units

In many household surveys, the ultimate sampling unit is either a person or a household. Although there is usually a one-to-one correspondence between addresses and households, there are numerous exceptions where one household has multiple addresses or multiple households share the same address. Households that appear on the frame multiple times can bias survey measures because they are given a greater chance of being included in the sample. Furthermore, if multiple addresses for the same unit fall into sample, then one address is often excluded from the sample. This reduces the sample size and can inflate variances. On the other hand, households that share the same address are often sampled at a lower rate than intended. Weighting adjustments can correct for this type of error, but they tend to decrease the precision of estimates. Furthermore, if all households at the sample address are not enumerated at the time of interview, then this error can cause undercoverage bias.

Many frame errors are found at the time of interview, but sometimes frame creation activities can mitigate problems found in the field. When a sample is purchased, it is nearly impossible to detect coverage errors because the frame is not available to the survey designers. However, the inability to detect or correct for undercoverage and overcoverage does not imply that serious coverage errors do not exist. Section 2.2.1 describes some of the common situations where addresses and households may not have a one-to-one relationship. Section 2.2.2 reviews these same issues and the problems they cause in frame creation.

2.2.1 Addresses as Sampling Units: Multiple Addresses per Household

One housing unit, household, or person can be associated with multiple addresses. Some of the common situations are seasonal units, mergers, reconfigurations,

throwbacks, and P.O. Boxes. Although addresses are quite different from phone numbers, RDD surveys have developed procedures to deal with households that contain multiple phone numbers. This section provides basic situations that may need adjustments to account for address multiplicity.

When creating a sampling frame, vendors must decide to include or exclude seasonal addresses. Some people live in different housing units at different times of the year. For example, someone may live in the northeastern United States in the summer and in a warmer climate in the winter. If the same person inhabits both addresses at different times of the year, he or she may have two chances of selection on the frame. According to the USPS, the Seasonal Delivery Indicator “specifies whether a given address receives mail only during a specific season (e.g., a summer-only residence)” (U.S. Postal Service 2013a, p. 25). This indicator also includes educational delivery points. In 2010, there were about 1 million seasonal delivery addresses and about 200,000 educational delivery points in the AMS. About 38% of the seasonal delivery points were found to be occupied in the 2010 Census, and 40% of the educational delivery points were occupied (Kennel 2012). One option is to include a sample address as eligible only if a household currently resides there. Alternatively, one could remove all addresses flagged as seasonal. In many cases, it is necessary to ask respondents how many addresses they have. Because there is a lag in updating whether an address is seasonal or not, and because seasonal addresses represent potentially occupied units, surveys that wish to maximize coverage may want to include seasonal units in their sampling frames. Even though one stops receiving mail at an address, the USPS still maintains the address and includes it in regular CDS and DSF2 updates. Thus, it is important for the survey designer to explicitly inform the vendor if seasonal units should be included or excluded from the frame. If they are included, undercoverage will be reduced, but some overcoverage should be expected.

Mergers often occur in multiunit apartment buildings when one household buys an adjacent unit or developers consolidate units so that two or more units become one unit. For example, 101 Main Street Apartment A and 101 Main Street Apartment B could merge into 101 Main Street Apartment A/B. If 101 Main Street Apartment A and 101 Main Street Apartment B are not deleted and 101 Main Street Apartment A/B is not added to the address list, the merger could result in the single unit having two addresses on the ABS frame. It is also possible that 101 Main Street Apartment A and 101 Main Street Apartment B are not deleted, but 101 Main Street Apartment A/B is added to the frame. In this case, the same household would have three chances of selection. Survey researchers should anticipate this type of duplication and develop procedures to minimize this error, if possible.

When street names, street numbers, or within-structure identifiers are changed, an address is said to be reconfigured. For example, 101 Main Street Apartment A could be renovated and renamed 101 Main Street Apartment Oak, or First Avenue could be renamed Liberty Avenue. If a new address is added to the ABS frame, but the old address is not deleted, one housing unit could receive multiple chances of selection. For reconfigured units, the exact outcome is somewhat dependent on field procedures. If

the frame has both addresses, an effort should be made not to interview units with the old address. Often the USPS maintains both addresses on the AMS for a period of time so that mail sent to the old address can be redirected to the new address. Such changes should be recorded on the LACS file.

Throwback addresses are city-style addresses for which mail is redirected by the USPS to a P.O. Box (Iannacchione 2011). Persons who have a city-style address, but receive mail only at a P.O. Box, receive mail intended for either address at the post office, not at the throwback address. Otherwise, households with both a P.O. Box and a city-style address can receive mail at either address. Households with both a city-style address and a P.O. Box in the frame, whether or not a throwback, have no linkage noted between the two addresses. Furthermore, because P.O. Boxes are opened by individuals, a housing unit may have a physical address and multiple P.O. Boxes for multiple persons living in the same housing unit. Because the risk of duplication between mailing addresses and P.O. Boxes is high, and the chance of locating a housing unit on the basis of the P.O. Box is low, P.O. Box addresses are often excluded from sampling frames. For personal visit surveys, the city-style address is the appropriate address to use, anyway. For mail surveys, either address can be used. Without linkages between city-style and specific P.O. Box addresses, duplications are minimized when P.O. Boxes are removed from the frame, even for mail surveys.

There are exceptions, however. Some P.O. Box addresses have no corresponding city-style address and are not duplicates of other housing unit addresses on the frame. These addresses are not in normal delivery routes, and, in fact, constitute their own “route” in the CDS. Some vendors flag these P.O. Box addresses as OWGM (only way to get mail). Often these OWGM P.O. Box addresses are retained in the frame for mail surveys when other P.O. Boxes are removed.

Sometimes address files contain multiunit placeholders or header records for an apartment building. A multiunit placeholder is an address record for a multiunit building without a unit designation. For example, 101 Main Street may have two apartments: Apartment 1 and Apartment 2. If the frame has an address record for 101 Main Street without any unit designation in addition to the two addresses with unit designations, then the 101 Main Street address is considered a multiunit placeholder. Multiunit placeholders allow letters without unit designations to be sorted and delivered to the correct building, even though clerical sorting is needed to deliver the letter within the building. Also, some multiunit placeholders are used by the main business office of apartment buildings, although they may be shared by resident households as well. If the frame also contains addresses for one or both units, then the multiunit placeholder should be considered an erroneous inclusion. Otherwise, the multiunit placeholder should be treated as a drop point (see Section 2.3.2), even if it is not flagged as such.

Multiunit placeholders threaten the integrity of the sample design, and procedures should be developed to mitigate errors associated with them. For example, for in-person surveys, interviewers may be instructed not to interview at an address without a unit designation in a multiunit building unless the only record on the frame is the

undesigned unit. For surveys using other modes and organizations without access to the sampling frame, multiunit placeholders are difficult to detect. If mail to the multiunit placeholder is always given to the same unit, then some coverage bias may result. On the other hand, if mail to the multiunit placeholder is distributed randomly among the units, then there will be some increase in frame variance. For ABS surveys with a phone component, errors depend on whether the multiunit placeholder is matched to a phone number; some vendors may erroneously match the multiunit placeholder to the phone number of one of the units.

Finally, ABS frame files may contain multiple copies of the same address. In some apartment buildings, one can expect to see multiple undesigned units. For example, consider a building with four units. The AMS may contain four copies of the same address without any unit designation. In this case, four records look like duplicates of the same unit when they actually represent four different housing units. In 2010, the Census Bureau detected more than 3.5 million basic street addresses with multiple undesigned units. In fact, over 95% of the basic street addresses containing five or more units had at least two undesigned addresses on the AMS extracts delivered to the Census Bureau (Kennel 2012).

In summary, the presence of seasonal housing, P.O. Box addresses, mergers, reconfigured units, and multiunit placeholders on frames can inflate costs, variance, and sometimes bias (Amaya et al. 2014). Of course, even in the absence of frame issues, a person or household can be associated with multiple addresses (e.g., a household with a separate weekend home), so the issue is not unique to ABS. Survey researchers should anticipate situations where one housing unit or person will be associated with multiple addresses. When possible, duplicates should be removed from the sampling frame without removing any valid units. However, in many situations the frame does not contain information that can be used to determine duplicates. Without access to the full sampling frame, it is impossible to identify duplicate addresses in advance unless the main address and its duplicate both fall into sample. When duplicates cannot be removed prior to interviewing, survey researchers typically inflate the initial sample size to account for reduction in sample sizes caused by duplication, and residency rules are imposed for households with multiple addresses discovered during field work. For more information for dealing with overcoverage see Groves et al. (2009).

2.2.2 Addresses as Sampling Units: One Address for Multiple Households

The converse situation is multiple households sharing the same address, which can happen for a number of reasons. Multiple households may share the same address for convenience (e.g., shared residences) or privacy concerns, or there may be a lag in the USPS updating the AMS when housing units split. In some cases mail is delivered to a single drop point (Section 2.3.2) for subsequent distribution by someone at the drop point. Lastly, multiple units with the same street address may be represented by a single multiunit placeholder. Usually, in all these situations, field procedures are developed to make sure all households at the same address have a chance of selection. Each of these situations is defined and discussed in turn.

In an effort to maintain privacy, clandestine households may share an address with another household. For example, a household renting a converted pool house or garage may share an address with the household living in the rest of the house. Both households share the same mailbox. Individuals in the households divide the mail among the households. As their name implies, clandestine households are difficult to detect and are often a source of omissions. For in-person surveys, sometimes clandestine households can be detected and interviewed along with the main unit or subsampled. To detect clandestine households, some survey instruments begin the survey by asking about extra or additional housing units on the property that may share the same address.

A split occurs when one housing unit is divided into multiple units. Splits often occur when a single housing unit is converted into condominiums or apartments. Ideally, the sampling frame should have addresses for the separate split units and not for the single unit prior to the split. This requires that new units are added to the frame and the old unit is either removed from the frame or marked as ineligible. Sometimes the old unit is not removed from the sampling frame. Other times, the new units are not added. Thus, the sampling frame may have one address (just the old address), distinct multiple addresses (just the new addresses), or duplicate addresses (the new and old addresses).

Different field procedures should be implemented for these three situations, and documented for proper weighting. If only the old address is on the frame, then either all new addresses should be interviewed or one of the addresses should be subsampled for interview. If the old and new addresses are on the frame, then the old address should be classified as an erroneous inclusion, if it is selected. Of course, if one does not have access to the frame, then the errors cannot be identified and adjustments cannot be made to account for these errors.

Drop points are mail delivery points that serve multiple households or businesses. According to the USPS,

A drop point is a single delivery point or receptacle that services multiple businesses/families. Examples of drop sites include: single box shared by more than one business/family, boarding or fraternity houses, and gated community where mail for all homes is delivered to a gatehouse. A Commercial Mail Receiving Agency (CMRA) holds or forwards mail to an addressee. Each CMRA is registered with the Post Office responsible for delivery to the CMRA. Each CMRA is identified as a drop site, the carrier delivers the mail to one point, and the company makes final distribution (U.S. Postal Service 2013a, p. 22).

The AMS contains about 1 million delivery drop points that are flagged as such. Many of these drop points correspond to multiunits, and sometimes the AMS indicates how many units are associated with a drop point. More research into where drop points are located, whether they actually correspond to single or multi units, and how to handle them is needed. To reduce undercoverage, units with drop point addresses should be given a chance of selection. For in-person surveys, this can be done by interviewing all

units at a sample drop point or by listing the units at the drop point and subsampling the list. For web, mail, and phone modes, more research is needed to develop and test methods that improve coverage and give all units at a drop point address a chance of selection.

2.3 Frame Creation Activities

2.3.1 Frame Creation Activities: Address Standardization

At some point, vendors must clean their files. Cleaning activities include parsing addresses, standardizing addresses, and sometimes adding additional information to addresses. There are a number of software products and businesses that specialize in parsing, standardizing, and adding additional information to complete addresses. These processes are described in this section.

Parsing and standardizing addresses are necessary to receive discounts for bulk mail and to improve matching and deduplication processes. One of the primary goals of parsing and standardizing is to bring consistency between vendor files and the USPS. Consistency between vendors and the USPS adds efficiency to mailing and matching processes. The USPS maintains a list of postal addressing standards (U.S. Postal Service 2015b).

Parsing is the process of breaking down one line of an address into various components. According to USPS standards, the full street name should be parsed into the street direction prefix, street name, street type, and street direction suffix. For example, the street name in 101 North Main Street Northwest would be parsed into four parts: North, Main, Street, and Northwest.

The process of standardization involves comparing parsed address components to valid values and formatting constraints. For example, the state in an address might be "Michigan," "MI," "MI.," "Mich.," or "Mich.". Standardization replaces all of these values with "MI," the standard abbreviation for Michigan. The process of standardization corrects misspellings and internal variations in the way addresses are written by aligning them to the set of address standards used by the USPS. For example, 101 North Main Street Northwest would be standardized to 101 N MAIN ST NW.

After parsing and standardizing the vendor address file, additional data can be added to the addresses for completeness. For example, some address cleaning systems will assign ZIP+4 Codes to addresses. Section 3 summarizes the process of adding auxiliary data to addresses.

2.3.2 Frame Creation Activities: Coverage Improvement

Vendors often supplement their files with additional addresses to improve the coverage of their files. Supplementation may be desired for newly constructed housing units, areas lacking adequate coverage on the USPS files, and types of units commonly

omitted on the USPS files.

CDS license holders receive weekly or bimonthly updates from the USPS with new addresses for their “owned” ZIP Codes (U.S. Postal Service n.d.-a). The vendor, in turn, can update its address files on a schedule of its choice; some vendors update their files weekly, monthly, bimonthly, or semiannually. Naturally, the frequency of these updates can influence the frame coverage.

Some vendors may add other USPS addresses to enhance the coverage of their frames. For example, vendors may add certain types of addresses from the USPS No-Stat File. The USPS does not maintain a complete list of addresses in certain ZIP Codes, however. For example, addresses on simplified carrier routes and addresses on some Native American Reservations are not well covered in the USPS files. In both cases, local post offices sort the mail without using the DSF. Some vendors get street addresses in these ZIP Codes from other sources such as credit card data, consumer databases, public property tax files, or telephone directories. Vendors may also supplement coverage by matching addresses to local government files, or by adding addresses from marinas or hotels to improve coverage of persons without a standard residence elsewhere. Shook-Sa and Currivan (2011) found that using commercial files to improve the coverage of HUs with unlocatable mailing addresses on CDS frames was not as cost effective as using the USPS No-Stat File (Shook-Sa and Currivan 2011). Matching is discussed more in the next section.

2.3.3 Frame Creation Activities: Matching Files

Vendors often match addresses from multiple files when creating frames. Vendors may merge files to add auxiliary variables to the frame, to supplement the frame with new addresses, or to update the frame over time. The process of merging two files can introduce duplicates, erroneous inclusions, erroneous exclusions, sampling variance, and coverage variance, depending on the type and quality of the matching. Exact matching requires all components of an address in the merging files to be the same to be considered a match. Exact matching results in the least undercoverage, but may introduce a substantial amount of false nonmatches, if the address files and address components were not constructed in similar ways. A false nonmatch occurs when two addresses are essentially the same and represent the same housing unit, but are not matched because the matching rules are too strict. False nonmatches result in duplication with one housing unit having multiple addresses on the frame.

In probability matching, two addresses are considered the same if they meet some threshold of equivalence. A high threshold will match two addresses if they are very similar. A high threshold results in somewhat smaller decreases in undercoverage compared to exact matching, but may still increase duplicates considerably. If a low threshold for matching is used, then it is likely that truly different addresses will erroneously be called a match. Matched addresses that represent two different units are called false matches. False matches can lead to ambiguity in locating the sample unit or to undercoverage if one address is removed. On the other hand, two addresses that

represent the same housing unit but have slightly different forms may be matched, resulting in fewer false nonmatches than exact matching. For probability matching, the goal is to determine a threshold that balances the coverage/bias tradeoffs of false matches and false nonmatches.

Vendors should be able to provide information about whether they use exact matching or probability matching when combining data. If probability matching is used, then the vendor should know the threshold and have information about how the threshold was determined. Of course, some vendors consider this information proprietary, in which case the survey researcher must trust the vendor to match files in ways that minimize coverage errors.

2.3.4 Frame Creation Activities: Conversions

Sometimes addresses change, for a variety of reasons. When post offices split or merge, the ZIP Codes of addresses often change: “New ZIP Codes and major route adjustments are generally implemented on July 1st of each year” (U.S. Postal Service 2013a). In an effort to assist early responders to locate emergency 911 calls, rural addresses are assigned new “city-style” addresses with a house number and street name. Sometimes local governments change the names of streets or renumber houses. The USPS maintains the LACS, which provides a crosswalk for address conversions. Vendors with a CDS license receive address conversions through the LACS in owned ZIP Codes. Vendors may add the converted addresses to their files and mark the outdated addresses as invalid, or simply replace the old addresses with the new ones.

2.3.5 Frame Creation Activities: Auxiliary Data

During frame creation, auxiliary variables are often added to address files. Section 3 provides more detail about appending auxiliary data to ABS frames. Geographic codes, phone numbers, e-mail addresses, GPS coordinates, housing values, property tax assessments, and other auxiliary variables can be added to address lists by matching to other files. The quality of the added variables depends on the coverage and accuracy of the external files and the quality of the matching procedure.

2.3.6 Frame Creation Activities: Deduplication

As noted in Sections 1.2 and 2.2.2, vendors with a CDS license must contain no more than 110% of the CDS addresses in a ZIP Code to own the ZIP Code (Dohrmann et al. 2007; Iannacchione 2011; U.S. Postal Service 2013a). Thus, frames purchased from vendors with a CDS license usually have less than 10% overcoverage (relative to the mailable addresses in the AMS), although vendors that enhance CDS files with addresses from other sources may introduce some additional overcoverage in some areas. Much of that overcoverage could be duplicate addresses for the same housing units, perhaps caused by errors in matching and merging files. Sometimes duplicate addresses are exactly the same, but often they are slightly different. Data standardization reduces but does not eliminate the probability of these false

nonmatches.

Vendors may seek to remove duplicates from frames by removing addresses with a high probability of being the same. Like probabilistic matching, probabilistic deduplication can result in false duplicates and false nonduplicates. During probabilistic deduplication, one seeks to correctly identify duplicates without removing too many false duplicates. A false duplicate is a pair of addresses that are identified as duplicates, but actually represent two different housing units. In 2010, more than 3.5 million basic street addresses contained multiple undesignated addresses. Furthermore, there were fewer than 300,000 basic street addresses that contained duplicate addresses with a unit designation (Kennel 2012). Thus, one should anticipate undesignated duplicates, especially in large multiunit structures.

2.3.7 Frame Creation Activities: Filtering

A final important step to creating an ABS frame involves removing out-of-scope addresses, also known as erroneous inclusions. The goal is to remove as many erroneous inclusions without creating erroneous exclusions (i.e., addresses for in-scope housing units that were removed from the frame in error). Depending on the target population, survey budget, and survey rules, it may be desirable to remove vacant units, businesses, group quarters, seasonal housing, demolished units, and multiunit place holders from the frame extract. Filtering of addresses, also sometimes called extracting or screening, is the process of subsetting the address lists to those desired for the address frame (Loudermilk and Kennel 2005; Martin and Loudermilk 2008; Tomaszewski and Shaw 2013; Ying 2012). Removing addresses flagged as erroneous inclusions has cost benefits, because less money is spent fielding interviews to ineligible units. However, if in-scope units are removed from the frame because they are incorrectly flagged as ineligible, then the filtering process can result in increased undercoverage. Cost and coverage targets of the survey should be considered when determining filtering rules.

The USPS files contain variables and information that may be used in the process of extracting addresses into an ABS frame. For example, the Delivery Point Usage Code uniquely classifies all addresses as Residential, Business, Primarily Residential with some Business, Primarily Business with some Residential, and General Delivery. In 2010, the USPS address files contained more than 13 million business addresses, of which fewer than 2% contained residential households in the 2010 Census (Kennel 2012). Therefore, addresses classified as business addresses are often excluded from residential ABS sampling frames.

In the 2014 Fiscal Year, the AMS contained about 141.0 million residential delivery points and 12.9 million business delivery points (U.S. Postal Service 2014a). In general, business addresses are excluded in sampling frames for household surveys. There are numerous types of business addresses, some of which are more likely to contain residential households than others. For example, the Census Bureau found about 7% of the 1.5 million business curb-line delivery points (mail receptacle located at the curb)

and 2.3% of the 5.5 million business other delivery points to be occupied housing units in Census 2010. Misclassification rates for other types of business delivery points were smaller.

In 2010, there were about 16 million residential P.O. Box addresses on the AMS. P.O. Box addresses are often excluded from ABS frames because they pose a number of survey problems. First, persons with P.O. Boxes may live in housing units that also receive mail at their street addresses. Thus, P.O. Boxes may be a source of overcoverage. If P.O. Box addresses that are not OWGM are removed from the frame, no duplication is expected between the city-style addresses and the P.O. Box addresses. Second, P.O. Box addresses do not identify physical locations. For surveys conducted in person, it is unlikely that an interviewer will be able to locate a housing unit from its P.O. Box address. Furthermore, for phone and e-mail surveys, the frequency of finding phone numbers or e-mail addresses corresponding to P.O. Boxes is low. For mail mode, however, P.O. Box addresses that are a household's only way to get mail (flagged as OWGM in some address files, see Section 2.3.1) may be included to increase coverage.

The USPS delivers mail to fewer than a million active residential rural route delivery points. Rural route addresses are sometimes excluded from sampling frames, as well, especially for the problems posed to surveys conducted in person. Namely, housing units associated with rural route addresses can be difficult to locate because the link between a housing unit and a rural route delivery point is often poorly defined or ambiguous. Mailboxes often have the rural route identified on them, but housing units do not. If multiple mailboxes are at the end of the lane, it may not be clear which mailbox corresponds to which housing unit. Like P.O. Box addresses, matching rural route addresses to phone numbers and e-mail addresses is challenging; current methods are not successful at obtaining phone numbers and e-mail addresses for rural route addresses. Including rural route delivery points will improve coverage for mail mode surveys.

In 2010, there were fewer than 30,000 residential general delivery points (Kennel 2012). These delivery points are thought to be for transient living quarters and are often excluded from sampling frames. Little is known about the coverage and quality implications of excluding these delivery points, but they represent a very small proportion of the total population of residential addresses.

It should be noted that vendors do not receive simplified addresses with their CDS or DSF2 licenses. Simplified addresses are ones for which the only delivery information held by the USPS is the city, state, ZIP Code, and the number of postal customers receiving mail at that ZIP Code. Therefore, many address frames will undercover the simplified addresses unless they are supplemented from other sources.

The USPS maintains an address vacancy indicator. In 2010, there were about 13 million addresses indicated as vacant in the AMS (in the raw DSF delivered to the Census Bureau). However, recent studies have found between 38% and 41% of units classified

as vacant to be occupied (Amaya et al. 2014; Kalton et al. 2014). A housing unit that is vacant may quickly become occupied. According to USPS guidelines, an address must be unoccupied for 90 days to be classified as vacant (Iannacchione 2011). Given the difficulty in maintaining an accurate and timely vacancy indicator and the coverage concerns with using this variable in frame creation, caution should be exercised in removing addresses classified as vacant.

The USPS also maintains the Seasonal Delivery Indicator (Section 2.2.1). For similar concerns about the timeliness and accuracy of the Season Delivery Indicator, one should exercise caution before removing addresses classified as seasonal delivery. Including seasonal housing may improve coverage, but information about the seasonality may be necessary for appropriate weighting. Excluding the seasonal housing means that occupants' probabilities of selection are based solely on their main residences, which simplifies the calculation of the base weights but can lead to lower response rates when people are not present at their main address.

The majority of addresses on the No-Stat File do not receive mail delivery. Thus, the decision about which types of No-Stat addresses to include on an ABS frame is based on a tradeoff between coverage and cost. Research has found that including all but the drop points (duplicates of addresses already in the frame) in rural areas improves the coverage in those areas by about 4% (Shook-Sa et al. 2013), but with some inefficiencies from the inactive addresses similar to those of a field enumerated frame. However, when limited only to active addresses on the No-Stat file (rural throwback addresses), frame coverage for in-person surveys is improved by about 2.2% with no apparent loss of efficiency. Shook-Sa et al. (2013) provide some guidance on coverage/cost tradeoffs of the No-Stat address types.

2.3.8 Frame Creation Activities: Updating

Some CDS vendors receive weekly transaction files with address updates. Over time, the attributes of many addresses may change. For example, the vacancy status of an address may periodically change. Furthermore, addresses may move between the business and residential files or between the No-Stat and regular files from week to week.

Some vendors maintain a history of address changes; others keep only the most recent status of the address. If the vendor keeps a history of changes, the history can be used in filtering addresses to create a frame. For example, in an effort to improve coverage, one might want to include addresses on the residential file that appeared at least once in the residential file in the past 12 months. Or, one may want to include addresses that appeared on any update in the past year, even if it did not appear as an active delivery point (Martin and Loudermilk 2008).

To date, there has been very little research into update rules and their effect on the coverage, cost, and quality of frames. Most of that research has been done within the

Census Bureau (Loudermilk and Kennel 2005; Martin and Loudermilk 2008; Tomaszewski and Shaw 2013; Ying 2012). See Section 8.3 for research suggestions.

2.4 Checklist of Questions to Ask Vendors

As noted in the introduction, this chapter contains numerous details. Sorting through the many details and considerations can be daunting for those new to ABS. To help navigate the vendor selection process, we recommend discussing the following questions with vendors.

For all vendors:

- Does the vendor have a CDS or DSF2 license?
- How often does the vendor update its address database?
- Are any coverage improvements made? If so, what is the source of the addresses?
- Is the file unduplicated? If so, are there any processes in place to minimize removing false duplicates?
- Are the following addresses included or excluded: P.O. Box throwbacks, seasonal housing, vacants, units on the No-Stat file, mostly business with some residential, P.O. Boxes, rural routes and highway contracts, or general delivery?
- Are any edits or filters applied to multiunit placeholders?
- What kind of quality assurance techniques are in place to minimize erroneous inclusions?
- What kind of quality assurance techniques are in place to minimize omissions and erroneous exclusions?
- Are there any special edits done to undesignated addresses in multiunit structures? For example, are unit designations ever imputed?

For primary vendors:

- What was the source of the vendor's addresses prior to obtaining a license?
- How often does the vendor receive updates from the USPS?
- What ZIP codes are not "owned" by the vendor?
- Are there any address groups that are not "owned"?

For secondary vendors:

- What was the source of the primary vendor's addresses prior to obtaining a license?
- What ZIP Codes or address groups are not owned by the primary vendor?

2.5 Section Summary

The frame creation process affects the coverage, quality, and cost of a survey. On the one hand, the process of updating vendor lists through USPS license agreement brings great conformity among vendors. To qualify for a CDS license agreement and address corrections for a ZIP Code, vendors must already have the vast majority of the USPS addresses in that ZIP Code from an independent source. Also, once a primary vendor meets the required threshold, periodic address updates allow vendors to quickly refresh their databases to align with the coverage and standardization of the USPS. On the other hand, variations in vendors' frame creation processes, including screening/filtering address types to include in the frame, deduplicating, and supplementing with addresses from other sources can increase or decrease the cost, coverage, and variance of the survey. Furthermore, the type of agreement between the vendor and USPS (i.e., DSF2 or CDS) may affect frame coverage.

Regardless of the vendor and frame creation process, researchers should question potential vendors on the correspondence between addresses and households, with variations caused by types of addresses included and methods for processing addresses (Section 2.4). That is, researchers should be clear about the relationships among sampling units, data collection units, weighting units, and units of analysis. The researcher should seek to minimize errors and ambiguous situations through clear field procedures and, in cooperation with the vendor, through frame refinements. The vendor's methods of deduplication, matching, updating, and other file handling procedures all affect the quality of the frame.

Address frames have both undercoverage and overcoverage. Vendors may supplement their address lists to improve coverage for certain address types or in certain geographies. Overcoverage is partially addressed by deduplication and by filtering the address list to remove unwanted address types. In the context of household or person surveys, addresses that are exclusively or mostly business are excluded. Common frame refinements include removal of P.O. Boxes that are not OWGM to reduce duplication. For in-person surveys or mail surveys restricted to specific geographies, other unlocatable addresses such as OWGM P.O. Boxes or rural route boxes may be removed, as well. The removal of addresses flagged as vacant or seasonal is risky, with efficiency traded for coverage of potentially eligible addresses. Drop points, multiunit addresses, and other special situations require careful planning.

SECTION 3

AUXILIARY VARIABLES

3.1 Frame Supplementation With Auxiliary Variables

Although all ABS frames contain addresses, it is often highly desirable or necessary to have some auxiliary information about some or all addresses on the frame. In fact, one of the primary advantages of ABS designs is the potential availability of rich frame data. Such frame data could have numerous uses in an ABS environment. For example, ABS frames used in a mixed-mode design could benefit from landline phone numbers, cell phone numbers, or e-mail addresses in some combination. Furthermore, auxiliary frame

variables can be used to reduce sampling errors through both stratification and post-stratification. Auxiliary variables could theoretically be used in filtering the frame. For example, a survey of children might benefit from a variable indicating if children live at the address. Indeed, auxiliary variables can be used to reduce costs and sampling errors, especially for surveys of rare populations. However, the effectiveness of the auxiliary variables depends on the content, accuracy, and completeness of the auxiliary variables and on the quality of matching addresses to the auxiliary data.

Vendors often differ in the auxiliary variables they offer. Furthermore, the methods used to attach auxiliary variables also vary between vendors. Because the overall quality of any auxiliary data depends on the source and the programs used to geocode and match addresses, survey practitioners should discuss the specific frame sources, geocoding methods, and matching methods when working with vendors. Rather than comparing specific vendors or the many possible auxiliary variables, this section introduces two major components of appending auxiliary data: geocoding and matching. Then, the general sources of auxiliary variables are reviewed. The main drivers of quality of auxiliary data are noted and discussed here and also summarized in Section 7.

3.1.1 Two Methods for Appending Auxiliary Data: Geocoding and Matching

Adding auxiliary data to an address list usually involves geocoding and matching. Geocoding is the process of determining the geographic location (often in longitude and latitude coordinates) for an address, often to facilitate the appending of state, county, tract, block group, and block codes. After each decennial census, the Census Bureau parses the nation into blocks and tracts. These blocks and tracts are numbered, and census summary data are produced at the various geographic levels. Once an address has been geocoded, a wealth of summary-level geographic data can be appended at the block, block group, tract, county, and state levels. Analysts often desire geocodes for purposes of estimation. Even if geocodes are not directly used in the sample design, survey operations, or estimation, they are often added to ABS frames because they allow matching of hundreds of block- and tract-level counts and estimates from the American Community Survey (ACS) and decennial census.

Matching an address file to a supplementary data source involves linking the two files, usually on a common set of variables. Address files can be matched to person, household, address, housing unit, or geographic source files to gain even more information about addresses on the frame or sample. The following sections describe geocoding and matching in more detail and highlight some of the potential errors associated with both processes.

3.1.2 Geocoding

The AMS database includes state, county, and congressional district codes, but it does not contain lower levels of geography. Tract, block group, and block codes are neither defined nor maintained by the USPS. The U.S. Census Bureau draws tract, block

group, and block boundaries every 10 years, based on the physical and political geography of the nation and on the most recent decennial census. The basic building block of Census codes is the four-digit tabulation block code. All other boundaries are defined as collections of blocks. The block group is the first digit of the tabulation block code. Most vendors and programs provide block group– level geocodes. Some also assign the full four-digit tabulation block code and longitude and latitude coordinates.

ABS vendors often use software packages or secondary vendors to geocode addresses, although programs to geocode addresses are publicly available. There are numerous algorithms to geocode addresses.

First, addresses can be geocoded by directly observing the coordinates of an address. To do so, field representatives use a GPS device to capture the coordinates of the front door of the housing unit of the address. With the coordinates in hand, some Geographic Information System applications can reference Topographically Integrated Geographic Encoding and Referencing (TIGER) lines (or commercially enhanced versions of TIGER lines) to assign the coordinates to specific blocks. Directly assigning geocodes in the field is rarely done because it is expensive; the Census Bureau used direct field observations to capture the GPS coordinates of nearly all addresses in the nation prior to the 2010 census. Of course, GPS signals are not available everywhere, and handheld devices often have both random and systemic errors when determining coordinates (Alkire 2010; Bonner et al. 2003; Fiorio and Fu 2012; Ward et al. 2005).

Second, geocoding can be done clerically by overlaying either TIGER or TIGER-based lines or block boundary files over aerial photographs or satellite imagery to assign a block code to each housing unit. This method is also rare because of the labor cost of manually assigning blocks. Satellite imagery and manual processes are subject to random and systemic errors, as well.

Third, latitude and longitude coordinates can be assigned to an address by matching to a public TIGER line file or to an enhanced proprietary TIGER database. (There are a number of vendors that specialize in geocoding. The origination of such proprietary databases is usually the TIGER line files, which are enhanced and updated by the geocoding vendor.) To geocode an address by this method, it must have a house number, street name, and ZIP Code (Cayo and Talbot 2003). The address is matched to a TIGER line segment and its corresponding range of house numbers. Even if the address file is standardized and parsed appropriately, there are numerous opportunities for errors in matching. Probability matching and SOUNDEX algorithms are often used to increase match rates (Goldberg, Wilson, and Knoblock 2007). If probability matching is used, one should be aware of the probabilistic confidence level for matching. If address components are relaxed to improve matching, careful quality control procedures should be implemented to minimize false matches (Goldberg et al. 2007).

Once the address matches to a TIGER line segment, a specific coordinate can be assigned through interpolation. Interpolation is the process of assigning coordinates to an address from a TIGER line. Basic geocoding software usually makes two

assumptions: the parcel homogeneity assumption and the address range existence assumption. The parcel homogeneity assumption assumes that all house numbers along a line segment are uniformly distributed along the segment. The address range existence assumption assumes that all addresses in an address range exist. Some commercially enhanced TIGER files offer shorter ZIP+4 line segments and a correspondingly narrower range for more precise interpolation. Some more sophisticated geocoders use parcel and other data to more realistically interpolate addresses (Goldberg et al. 2007). If the exact coordinate is of interest, the interpolation method can be quite important. However, for sampling needs it is often enough to place the address in the correct block. In such situations, the quality of the interpolation method used to assign the exact coordinate is less salient because all addresses on the same side of a TIGER line segment usually fall in the same block. For more information on which types of geocoding error affect block assignment, see Eckman and English (2012b).

The underlying TIGER line files may be spatially inaccurate or out of date. Cayo and Talbot (2003) noted that some address ranges were incorrect or reversed in their study, and that coordinates in rural areas were less accurate than in urban areas. Addresses on new streets, P.O. Box addresses, and rural route addresses are the most difficult to geocode using address ranges (Eckman and English 2012a). Addresses on new streets cannot be geocoded until the TIGER lines are updated with the new streets and address ranges.

Fourth, the address can be geocoded based on the nine digit ZIP Code (ZIP and ZIP+4) or some fraction of the ZIP Code. The reference data often include the ZIP Code associated with each TIGER line. Thus, the address can be matched to one or more TIGER line segments using the ZIP code. If the address matches to multiple TIGER line segments, the best segment is selected. A coordinate is then assigned based on the address ranges along the line segment. Of course, there are various methods of interpolating coordinates from ZIP Codes. For example, all addresses with the same ZIP Code can be assigned the coordinates at the geographic center of the ZIP Code Tabulation Area ZCTA (Goldberg et al. 2007). ZIP Codes represent a collection of delivery routes, but do not define an area with boundaries. The U.S. Census Bureau has drawn polygons around the delivery routes of a single ZIP code. The resulting geographic areas are called ZCTAs. One can expect the majority of addresses in a ZCTA to have the same ZIP code, even though a ZCTA may contain addresses with different ZIP codes (U.S. Census Bureau n.d.- b). For assigning blocks, the accuracy of geocoding by matching on the ZIP Code is much better using the ZIP+4 Code than when fewer digits in the ZIP Code are used.

Fifth, addresses can be geocoded by matching to parcel data. Some local governments have publicly available parcel datasets containing the boundaries of each parcel and the coordinates of each parcel's centroid or the center of the structure on the parcel. Proprietary datasets also exist for many counties. Addresses can be matched to such datasets and coordinates assigned based on coordinates in the parcel data (see Cayo and Talbot 2003). Once the coordinates are assigned, one can use TIGER boundaries

to geocode the address.

Often, addresses that match to a TIGER line are geocoded using the third method. Then, the remaining addresses are geocoded using the ZIP+4, ZIP+2, and 5-digit ZIP Code methods. In one national frame, this method resulted in correctly geocoding 84% of all addresses to the correct block group (Kennel and Li 2009). As expected, the addresses that were geocoded based on their address range and ZIP+4 Codes resulted in much more accurate geocodes than those based on ZIP+2 or 5-digit ZIP Codes (Kennel and Li 2009). McElroy et al. (2003) describe one study that uses all three geocoding methods in a sequential fashion.

Vendors and software programs can provide information about their geocoding methods (depending on the ability to match the addresses to the TIGER reference file); see Kennel and Li (2009); Eckman and English (2012a; 2012b), Fiorio and Fu (2012), and Dohrmann et al. (2012) for further discussion of this. Overall, the accuracy of geocodes depends on the way addresses are matched, the interpolation algorithms, and the reference dataset (Goldberg et al. 2007). When choosing a vendor or software package, one should consider these three aspects of geocoding and determine whether the geocoding methodology will meet geocoding accuracy requirements. Although matching, interpolation, and reference file accuracy can have a large effect on the accuracy of the geocoded coordinates, matching and interpolation have less impact when assigning higher level areas than blocks, such as block groups and tracts.

3.1.3 Data Matching

The addition of supplementary data can enrich an address frame or sample tremendously. Broadly, there are two levels of auxiliary data: unit- or address-level variables and summary- or geographic-level variables. Unit- or address-level variables are usually assigned by matching addresses on the address frame to addresses on other auxiliary files, although sometimes address- level auxiliary variables are assigned by a model. Summary- or geographic-level variables are typically assigned based on Census block or tract codes or ZIP Code. Both types of variables, which are available in supplementary files, must be matched to the address frame.

When matching addresses to supplemental files, one should be aware of several potential errors. The same general matching errors discussed in Section 2 can occur. In addition, one should be cautious of one-to-many matches. For example, matching an address file to a person file (or a collapsed person file) may add variability or bias to a frame if every person at the address does not share the same characteristic. Also, one should be cautious of data obtained through matching based on part of the complete address. For example, sometimes frame addresses may not match on the exact address, but may match to a supplemental file if the unit identifier is missing from an address in a multiunit building. In this case, the auxiliary data may actually correspond to another unit in the building.

Although the addition of detailed and rich supplementary data can result in an extremely

rich frame, matching errors can threaten the utility of the auxiliary data. If auxiliary variables play a major role in the survey design, a careful understanding of the matching methods and the completeness and coverage of the supplemental files is essential. One should certainly ask the ABS vendor for information about the match rates of the files and the accuracy of the variables. Also, some vendors can tailor the matching process to increase match rates for a particular auxiliary variable, which is valuable to marketers who are trying to reach as many consumers with a particular characteristic as possible. It is very important to discuss the survey goals, coverage needs, and matching process with the ABS vendor.

In general, it is not recommended to use appended auxiliary data to filter the frame or sample. Doing so may result in serious undercoverage errors. For example, consider an address frame that is matched to a phone number list. If the persons who do not have a listed or otherwise matched phone number are removed from the frame, estimates from the survey may suffer from undercoverage bias because, depending on how the phone number list was created, persons with cell phones, persons on the do not call list, persons with unlisted numbers, and persons without phones will not be represented by the survey estimates. Similar coverage errors can occur for any variable that has a match rate to the frame of less than 100%. In general, it is far better to stratify by auxiliary variables than to filter the frame or sample based on matched auxiliary variables.

3.2 Types and Sources of Auxiliary Variables

Address-level variables appended to the ABS frame originate from a variety of sources, and the rates of completeness and accuracy of the variables vary considerably. Despite this variation in specific sources, phone numbers, commercial data, and government data are the most common types of auxiliary data attached to ABS frames.

3.2.1 Phone Numbers

Phone numbers can play an important part in survey operations. For surveys with a telephone component, phone numbers are needed to contact the sample unit. For in-person surveys, they can be used to help locate the unit and establish contact with the sample unit. The shift from landline phones to cell phones has added complexity to assigning one phone number to an address. Some addresses may be associated with multiple numbers and some addresses may be associated with no phone numbers. Furthermore, because cell numbers are more portable than landline numbers, cell numbers may not be associated with the sample address. Amaya, Skalland, and Wooten (2010) discuss issues related to the matching of telephone numbers to addresses.

Telephone match rates can vary by address type and geography. For example, obtaining phone numbers in rural areas can be difficult because matching to rural routes and P.O. Boxes is difficult. Furthermore, many of the matched telephone numbers may not be accurate. Amaya et al. (2010) found that 74% of addresses could be linked to a

phone number, but only 54% to a working phone number. Furthermore, addresses associated with multiunit buildings had very high match rates, but the high match rate was because of linking one phone number to multiple units at the same address.

The lack of accuracy of the phone numbers linked to the addresses can produce bias and variance increases. For example, when multiple phone numbers match to an address, if the lowest phone number is assigned, then the sample may skew toward people from certain states that traditionally have been assigned small area codes. On the other hand, if one of the matching phone numbers is selected at random, then there may be some additional frame variance because the frame would be dynamic. To reduce bias and variance, frame creators should make efforts to verify that each phone number corresponds to the address or unit sampled.

Most phone-to-address matching is based on listed landline numbers. As households continue to switch to cell phone service only, matching to landline numbers can cause differential undercoverage. According to Blumberg and Luke (2014, p. 3), "Adults living in the Midwest (40.6%), South (39.7%), and West (37.8%) were more likely than those living in the Northeast (23.6%) to be living in households with only wireless telephones." They also found that Hispanic adults, adults living in poverty, renters, adults aged 25 to 29, and men were more likely to be living in cell-only households than other demographic groups. Relying solely on matched landline phone numbers for a survey will likely produce coverage bias. McMichael and Roe (2012) and Harter and McMichael (2013) explored the feasibility of matching both cell phone numbers and landline numbers to address frames and found that both the match rate and accuracy rate of matched cell phone numbers were substantially less than those of landline phone numbers. Of course, one advantage of ABS is that the households that do not match to a phone number can be invited to participate in the survey by other modes, such as mail or a personal visit.

3.2.2 Commercial Data

Vendors often maintain person- and address-level proprietary datasets. The content, coverage, and quality of these datasets vary considerably from vendor to vendor and also within vendor. Variables range from whether someone at the address is interested in beauty products to credit scores to the number of persons living at the address to housing values to the age and race of the householder. Credit agencies and direct mailers often amass and compile personal and household information from thousands of sources to better engage or target customers. Commercial files include both proprietary and public files such as real estate files, property tax assessment files, and voter registration files. Data can be extracted from credit card purchases, magazine subscriptions, or even from warranty cards that have check boxes for socioeconomic and demographic information. Auxiliary data can also be modeled or imputed based on geocodes or covariates from supplementary data sources.

Buskirk, Malarek, and Bareham (2014) describe two sources of variables that can be appended to ABS frames and samples. First, proprietary data likely obtained through third-party data aggregators often include household- and person-level variables, such as shopping behaviors, purchase history, household composition, and voter registration. Second, sample vendors offer variables such as the housing type, ethnic surnames, telephone numbers, and owner/renter status.

For auxiliary data derived from commercial sources, information being matched to an address usually refers to at least one person who resides at that address. For this reason, it is completely possible that information appended to an address from one proprietary source may not match that from another source for a given address because the appended information refers to different reference persons who reside at that address (Buskirk et al. 2014; Dohrmann et al. 2014). This aspect of appended data can lead to undercoverage when used to identify members from a specific target subpopulation (Valliant et al. 2014). For example, a household with two adults, one with a PhD and one without, may be classified as not having a PhD because the householder on the commercial data file was the person without the PhD. If the households without a PhD are filtered out of the frame, then this household would be excluded. On the other hand, if the commercial data file has the person with a PhD as the householder, then this household would be covered. Thus, the process of selecting a householder or aggregating person data to the address level may result in significant errors, if not done appropriately.

Resident names are one special type of commercially available data. According to Fahimi and Kulp (2009), about 85% of sample addresses could be name-matched at that time, the hope being that adding a name to an address might increase the probability of the recipient opening the survey mailing. As noted with other auxiliary variables from commercial sources, however, the names will not be available for all addresses, will usually be based on the “primary householder,” and are likely to include some that are outdated or incorrect. Having a name appended is likely to increase response rates when correct, and increase the discard rate in households with an incorrect name. Dillman et al. (2014) found that response rates to mailed contacts can be improved through other design aspects. At the present time, matching names to addresses is not recommended for most purposes.

Buskirk et al. (2014) discuss the overall append rates for a collection of demographic and household variables and compare the distributions of appended data to national benchmarks. They note that append rates can vary depending on the underlying variable of interest; the distribution of the number of adults based on appended data tends to overestimate the number of single adult households and underestimates the number of two-adult households, for example. Some vendor variables represent a compilation or standardization of related variables obtained from proprietary sources. Others represent a conversion of continuous metrics into binary variables (e.g., converting a renter propensity score into a binary renter variable).

Because commercial data may be captured, coded, and defined in many ways, it is

important to have a basic understanding of the commercial variables appended to the address frame. For example, the variable has children might be operationalized as a binary 0/1 variable, where 0 indicates no children and 1 indicates the presence of children in the household; however, what is usually appended for such a variable is a 1 for those addresses with a child and 0 for those addresses without children and those for which information regarding the presence of children is missing or not available. There are many reasons information about households may not be available, including time lags from administrative database updates or underlying low prevalence rates for the given variable of interest in the population. Of course, even when data are available, the value for a particular address may not be accurate.

Pasek et al. (2014); DiSogra, Dennis, and Fahimi (2010); Roth, Han, and Montaquila (2013); Valliant, Dever, and Kreuter (2013); and Ridenhour, McMichael, Harter, and Dever (2014) discuss findings involving the completeness and accuracy of appended address-level variables. Pasek et al. (2014) compared survey responses to covariates from one vendor. They found numerous variables that disagreed with the survey responses and were missing values not at random. Because third-party sources and ABS vendor methods change over time, continued evaluations of commercial auxiliary data (similar to those cited here) are recommended. Until more research is done, commercial auxiliary data should be approached with a healthy dose of skepticism.

3.2.3 Government Data

The linkage between geocodes and census geographies allows a bevy of census-related variables available to be appended to the sampling frame or the ABS sample. The link is commonly done at the block group level, but can also be done at higher levels of geography, such as ZCTA, place, tract, county, and state. There are two primary sources of block group and tract level data, the ACS and the decennial census. Every year, the Census Bureau publishes a file with 5-year estimates for hundreds of variables at the block group and tract levels. Additionally, every 10 years, the Census Bureau publishes a summary file of block group and tract-level counts based on the most recent decennial census.

As long as an address has been geocoded, it can usually be matched to Census data. Thus, the match rate to Census data is usually about the same as the geocoding rate. However, there are a number of other errors to consider when matching to census data. First, ACS estimates at the block or tract level often have large standard errors, in addition to nonsampling errors. Second, ACS block and tract estimates represent 5-year estimates, which may not reflect the current characteristics of the area. Third, decennial census counts are subject to nonsampling errors and reflect counts at one point in time. Fourth, errors in the geocoding process could place an address in the wrong tract or block (Dohrmann et al. 2007; Kennel and Li 2009). Finally, because the census data are area-level variables, their association with characteristics of the particular household at the address may be limited (Biemer and Peytchev 2012; Biemer and Peytchev 2013).

3.3 Auxiliary Data Quality

In considering the utility of auxiliary variables, one should consider two related aspects: the match rate of appending data to an address and the accuracy of the data themselves. The two issues are distinct but related, with similar factors affecting each. We can define the match rate as the percentage of addresses from an ABS sample or frame that could have data appended when requested, while the accuracy of the data refers to the data values corresponding to the current truth.

The match rate and accuracy of auxiliary variables appended to ABS datasets will vary based on a number of key factors. First, the type of variable appended, whether it is an area-level, household-level, or individual-level characteristic will affect the expected match rate and accuracy. Certain variables are easier for vendors to compile, maintain, and match than others. For example, individual characteristics would be expected to have higher error rates than those at the area level, which are likely correct as long as the address has been placed in the right geography. Moreover, variables that change, such as age, or are modeled, such as income, are likely more problematic than others.

Second, the nature of the addresses chosen for augmentation will have an influence on match rate and accuracy. Single-family homes will have higher match rates than addresses in multiunit buildings, drop points, or non-city-style addresses. Also, because single family homes tend to have lower turnover rates than multiunit addresses, the accuracy may be better for single family homes. For this reason we can expect single family homes to have more accurate telephone numbers, for example, than multiunit buildings when matched for a multimode survey.

Third, neighborhood-level characteristics can influence the accuracy and match rate of auxiliary variables. Specifically, areas with populations that tend to be lower income, immigrant, or highly mobile will have lower match rates and accuracy than areas that are more stable.

English et al. (2014) evaluated three vendors of targeted lists to see how each could perform at identifying households with small children in specific underserved neighborhoods in Los Angeles County, CA. The authors found little variation among vendors, but all three were challenged by the hard-to-reach population. The most successful vendor was correct 39% of the time in identifying households with small children, which compares with a 20% incidence rate in the general population.

Pasek et al. (2014) examined one vendor to understand the accuracy and usefulness of matching homeownership status, household income, household size, marital status, education, and age to selected households. The authors explored whether auxiliary variables could be used as a replacement for missing survey data by comparing them to respondent-reported values. In so doing, they found inconsistencies between self-reported data and the consumer data files, and suggest that users proceed with caution when considering a similar approach. Specifically, they found data were missing in lower-income households more often than in others, and that variables such as household income, age, and marital status tended to be incorrect considerably more

often than they were correct.

In summary, users of ABS samples with auxiliary data appended should be aware of associated caveats and limitations. First, not all records will have associated data, and those with successful matches will be expected to have variable degrees of error. Second, some data will be less reliable than others, with income and other modeled quantities being the least accurate. In general, the usefulness of an auxiliary variable is defined by the match rate to the frame, the accuracy of the information, and the prevalence of the target population.

3.4 Section Recommendations

Researchers who intend to append auxiliary data to address files should be aware of the limitations and associated issues, as with any data. Specifically, one should be careful of match rates and expected accuracy associated with a particular appended auxiliary variable, whether it is related to geocoding or to individual-, household-, or area-level demographics. The risk is that coverage biases may be introduced depending on the data appended and the way they are used in a survey context. Moreover, different data are considerably easier for vendors to append than others. For example, we can expect that nearly all residential addresses from a given list could be geocoded in some manner, and thus have their location determined with reasonable certainty in order to append Census or ACS data. Fewer addresses would be expected to have household-level characteristics appended, and individual-level information such as e-mail addresses are likely to be the most limiting. Urban areas, and especially multiunit buildings, generally have the least successful household- and individual-level matches. We expect the least accurate data to be at the individual or household level, especially for those residing in multiunit buildings in unstable areas.

Consequently, practitioners should not exclude addresses on the basis of auxiliary variables such as names, telephone numbers, household demographics, or e-mail addresses unless they are willing to introduce the risk of coverage bias. Survey practitioners should also be cautious of the veracity of such matched data, especially in environments where the reliability is most questionable, such as telephone numbers in multiunit buildings

Auxiliary variables can be extremely useful to stratify frames for more efficient sampling of subgroups, even if the auxiliary variables are not entirely complete and accurate. The strata defined by auxiliary variables can be classified by density for the characteristic of interest. As long as all strata are sampled, albeit with a lower rate for a low-density stratum, there is no coverage error as there would be if the low density stratum were dropped entirely. The data collection for a relatively rare group can be much more efficient with such stratification, especially if the targeted subgroup is relatively rare. The tradeoff is larger design effects resulting from unequal weighting. The survey designer should consider the tradeoffs in the context of the survey's specific goals and requirements.

SECTION 4

DESIGNING AND IMPLEMENTING ABS SURVEYS

The design and implementation of surveys using ABS necessitates considering multiple aspects of survey design, including sample selection and data collection methods. This section discusses design and implementation considerations for both single-mode and multimode surveys.

Sample selection procedures include stratification techniques and within-household respondent selection methods. Although ABS is very closely related to list sampling (Kish 1965), we discuss special aspects of ABS of which AAPOR members need to be aware. The sampling of population subgroups (e.g., households with children) is also discussed, along with differences between national and local address-based samples.

Data collection methodologies are also covered, including single-mode versus mixed-mode survey designs, and single-phase versus multiphase data collection. Limitations of the various types of ABS data collection approaches are discussed.

4.1 Sampling Frame Considerations for ABS Surveys

Section 2 discussed the creation of the ABS frame. This section discusses considerations specific to sampling from an ABS frame. With a single- or mixed-mode survey, the addresses included in the frame can impact the coverage, cost, and quality of the survey. The USPS provides indicators with an address to identify certain types of addresses (e.g., drop point, educational, seasonal, or vacant). Dohrmann et al. (2014) provide detailed examples of these indicators. As discussed in Section 2, the inclusion or exclusion of these addresses in the frame can impact both coverage and survey design.

4.1.1 Drop Points

As described in Section 2, drop points are mail delivery points that serve multiple units (residences or businesses). The proportion of residences in drop point addresses is small, but the proportion varies by area, with the majority clustered in New York City, Chicago, and Boston (Amaya et al. 2014; Dekker et al. 2012). A local geographic area can contain a limited number of drop points with a large number of apartments/housing units in each drop point. Other local areas may contain a large number of drop points that typically have a small number of apartments in each drop point. Because drop points tend to be clustered in particular areas, excluding them from the sampling frame can result in coverage bias. When determining the sample design, there are a few options available when deciding to include drop points: (1) include the drop point address, (2) include the drop point address and use supplemental data frames to augment/identify specific units, or (3) send field enumerators to list all addresses associated with the drop point. The choice used will depend on the survey design and the size of the drop points in the sample. Augmentation was found to be useful in improving coverage in a mail survey (Kalton et al. 2014).

4.1.2 Seasonal and Educational Addresses

As described in Section 2.2.1, seasonal addresses, which include some educational addresses (but not all educational addresses), are occupied only during certain times of the year. Including seasonal addresses in the frame can result in multiple chances of selection for the part-time occupants. When including these types of addresses in the frame, an effort should be made to identify the primary residence to ensure equal probability of selection or at least account for the multiplicity in the weighting computations. In addition, many educational addresses may be student housing in group quarters, which may be excluded from “general population” surveys. There may, however, be some off-campus student housing such as a house that is rented by two or more students. If one is doing a sample of adults in households using a primary residence rule, then one may want to exclude those educational units from the sampling frame. On the other hand, if one is using a current residence rule, then they should be included. If they are included, one needs to add a question to the educational address questionnaire to determine whether the sample address is a housing unit. Although excluding seasonal addresses may increase the efficiency of the sample, it can result in bias.

4.1.3 Vacant Addresses

The USPS provides a flag for addresses that have been vacant 90 days, which, if it is reliable, can be used to determine if the address should be included in the frame for the data collection period, thus reducing this bias. However, as noted in Section 2.3.7, the vacancy flag is considered unreliable and time sensitive. As described in Section 2.3.7, addresses identified as vacant may, in fact, be occupied by the time of the study. Amaya et al. (2014), using a national sample of addresses, estimated that excluding vacant addresses would have reduced the frame sizes for their study by 5.1%, excluded about 2% of occupied housing units, and reduced the portion of identified vacant addresses by 29.3%. Because of changes in the occupancy of housing units, especially in urban areas, excluding vacant addresses can introduce undercoverage and coverage bias (Amaya et al. 2014).

4.1.4 Addresses versus Households

It is important to keep in mind that the selection of addresses using an address-based sample is just that, a sample of addresses. This does not imply a one-to-one correspondence with households. A single address may be associated with multiple households, as is the case with multiunit dwellings without individually identified addresses, known as drop points. In addition, individual households may be reached at multiple addresses. This is the case, for example, when a household or subset of household members receive mail at both a city-style residence and a P.O. Box or when some or all household members reside at multiple properties throughout the year, including seasonal and educational addresses. Because probability surveys require a known probability of selection to make valid statistical inferences about a population, the

lack of one-to-one correspondence between addresses and households must be considered when designing a study using ABS. See Section 2.3 for more details about these one-to-many and many-to-one situations.

4.2 Address Sample Selection

One approach for survey organizations to obtain an ABS sample is to obtain the sampling frame from a vendor and select the sample themselves; an alternative approach is to use an ABS sampling vendor to select a sample of residential addresses. This section discusses the latter approach. There are multiple options for working with ABS vendors to facilitate sample selection. It is crucial that survey practitioners communicate sampling requirements or specifications with the vendor providing the ABS sample and that practitioners understand the sample selection methods used.

ABS selection can be conducted using either unstratified or stratified designs and, under either design, the sampling frame can be sorted by address-level characteristics leading to the systematic and implicitly stratified selection of addresses (Kish 1965). For example, the address list can be sorted geographically (e.g., by ZIP Code) or by address characteristics such as address type (e.g., city-style, rural route, P.O. Box). The addresses on the frame may also be sorted by characteristics associated with census block group or tract-level demographics (e.g., percent of tract population below poverty). This sorting can be conducted for the entire sampling frame or within designated strata. If not specified, it is important for survey practitioners to request information from the vendor about how the address list is sorted prior to selection. It can also be useful to ask the vendor to provide a variable indicating the sample selection order with the sample.

The ABS frame can also be stratified to select addresses with certain characteristics with probabilities other than what would be obtained with a random sample or a systematic (proportional) random sample. There are two options for selecting a stratified sample. First, the entire ABS frame may be stratified by characteristics available for all addresses on the frame such as those noted above. For instance, a study may wish to stratify by address type and undersample P.O. Boxes to reduce the likelihood of duplication issues while mitigating coverage bias. Another study may wish to oversample addresses in census blocks or tracts with certain proportions of particular demographic characteristics (e.g., proportion black, proportion with individuals under 18, proportion who do not speak English). Sample selection can occur using specified rates within strata formed using the entire ABS frame.

The second option is to select a stratified sample in two phases (i.e., two-phase sampling or double sampling).¹ As discussed in Section 3, auxiliary variables (typically demographic variables) can be appended by the vendor to a selected sample to augment the information available about the sample. If it is not feasible to append this information to the full frame prior to stratification, practitioners can request a sample that is significantly larger than the final required sample size. The auxiliary demographic data can be appended to the large sample, and then the large sample can be stratified

by one or more of these auxiliary variables. A second phase of sampling can occur in which a stratified sample is selected from the larger vendor- provided sample. In situations in which this two-phase sampling design is used, it is important to keep in mind that many of the auxiliary demographic variables are unavailable for many addresses. If these variables are used for stratification purposes, it is often necessary to include addresses with missing data as their own stratum, because these addresses tend to differ demographically and socioeconomically from those addresses with available data (Buskirk et al. 2014).

¹ *Traditionally, two-phase referred to double sampling in which a sample is selected in phase 1, some relatively inexpensive data are collected, and the resulting set of data forms the frame for subsampling the units in phase 2. In recent literature, two-phase has been used to describe two data collection operations applied sequentially. In this report, two-phase sampling refers to the first definition, while a two-phase survey refers to the second definition.*

4.3 Sampling Adults Within Households

When adults comprise the target population, a variety of methods for within-household sampling may be considered, recognizing that virtually every household will contain at least one eligible adult. Methods include those for both single-phase survey (i.e., data collection) designs and two-phase survey designs. Both the single- and two-phase approaches have advantages and limitations. Although these approaches have been studied extensively for telephone and in- person modes, they have not been evaluated as comprehensively for use in the mail mode.

Battaglia et al. (2008) were the first to compare and evaluate within-household selection approaches in the context of a single-phase, mail-based ABS design where the ultimate sampling unit is the person. They compared three approaches: (1) selecting any adult in the household, (2) using the next-birthday selection approach, and (3) selecting all adults in the household. The any-adult method involves asking the household to determine (purposely) which adult should respond; thus, it is not a random (or even quasi-random) method. The next-birthday method, in which the adult with the next birthday is selected, is viewed as quasi-random; it relies on the household members to do the selection and thus is subject to implementation error. The all-adults approach involves no selection; instead, all adults in the household are asked to complete the survey.

Battaglia et al. (2008) found that all three methods resulted in comparable household-level response rates; however, the all-adults method leads to lower person-level rates when individual nonresponse among some household members is taken into consideration. The any- adult method did yield a set of adults with demographic distributions less comparable to population distributions than the other two methods. In addition, although the next-birthday method is a quasi-random selection method, Battaglia et al. (2008) found that, in reality, in households with two or more adults, the incorrect adult completed the questionnaire between 31% and 60% of the time. It should

also be noted that the all-adults method requires enclosing multiple survey forms (many of which will be extraneous for most households) in the questionnaire mailing. This could have cost and resources implications for the survey.

The methods tested by Battaglia et al. (2008) all assumed a single-phase design. In considering these findings, and the limitations of the three methods, Brick, Williams, and Montaquila (2011) and Montaquila et al. (2013) tested a two-phase survey design in which a screener is used to enumerate eligible persons within the household; the returned screener instrument is processed by the survey organization, and an eligible person is randomly selected. In the second phase,² the survey is administered to the randomly selected person. This two-phase survey approach allows survey practitioners to have the most control over within-household selection. A variety of mode options may be available for the second phase. If the second-phase survey is attempted by mail, the survey is mailed to the household with instructions for the randomly selected person to complete it. The two-phase approach was tested by Brick et al. (2011) in the context of a subpopulation survey (discussed in Section 4.5), but has subsequently been used in other general population surveys. Additional early testing of the two-phase mail survey approach includes Han et al. (2010) and Mathiowetz et al. (2010).

One key benefit of the two-phase approach is the assurance that the sampling is undertaken accurately. Additional benefits include the ability to personalize the questionnaire mailing based on information obtained on the screener form, easing respondents into the response process by starting with the shorter screener form before sending the larger survey questionnaire, and not having to waste resources mailing survey questionnaires to households or persons that end up being ineligible for the survey (Montaquila et al. 2013). However, there are also potential drawbacks associated with this approach. First, although the within-household selection is truly random, there is no guarantee that the individual selected will, in fact, be the person who completes the questionnaire. Furthermore, additional labor is needed to process the returned screener forms and conduct the within-household selection. Finally, using two phases in data collection can lead to the need for longer survey administration windows, which may affect the timeliness of survey results.

An additional consideration is the effect of a two-phase survey approach (vs. a single-phase approach) on response rates. Having a second phase of data collection allows for a second opportunity for nonresponse; however, several of the aspects listed above as benefits of the two-phase approach have the capacity to have a positive effect on response rates. At this point, for a given survey, it is unclear whether the two-phase approach or a single-phase approach is likely to yield higher response rates, because there has been no experimental testing to evaluate this.

With the two-phase approach, careful attention should be given to the design of the screening survey. Well-established design principles, as described in Dillman et al. (2014), should be followed. In addition, to facilitate within-household sampling, the screener must clearly communicate to the household the eligibility criteria to be applied when enumerating household members, and it must collect information required for

sampling and for identifying the sampled household member. Although a name (or initials or nickname) may be useful for identifying a household member, it may not be necessary, and in some situations it might be preferable not to ask for a name. In addition, considerations should be given whether to include “engaging” items in the screener, in an effort to increase the overall response rate to the detailed survey (Williams et al. 2014) or to use these items to adjust for nonresponse to the second-phase questionnaire.

² *Eligibility information is collected in the first phase, and the second phase sample is a subsample of the eligible addresses. In this usage, however, the second phase is combined with person selection, which is an additional stage of sampling in traditional terminology because different sampling units are introduced. But because the screener and interview are handled in two separate data collection efforts, it is called a two-phase survey in recent literature.*

4.4 Sampling Subgroups

For subpopulation surveys in which not every household contains a member of the target population, the two-phase approach has additional utility. The two-phase approach precludes the survey organization from having to mail questionnaires (possibly multiple questionnaires) to sizable proportions of sampled households that do not contain any eligible members. In addition, two phases eliminate the complication of the household having to apply both eligibility criteria and selection rules to determine who should be sampled for the survey. Also, for situations in which entire questionnaire modules would be either administered or skipped depending on characteristics of the person (e.g., a child’s school enrollment status), information obtained in the first phase can be used to determine the questionnaire (or modules) administered to the particular person, rather than having to rely on the respondent to follow skip instructions.

4.5 Use of Auxiliary Variables in Sampling

The auxiliary data available for ABS frames (Section 3) can be valuable for designing or weighting a sample or for a tailored recruitment protocol. There is a growing body of research regarding the uses of the appended data for all facets of the survey design and research process. Biemer and Peytchev (2012; 2013) used census-related appended variables for nonresponse bias evaluation. Burks and Buskirk (2012) and Buskirk, West, and Burks (2013) used a combination of census, proprietary, and vendor variables to predict response prior to fielding the survey and discussed how these variables and response propensities could be used as part of a responsive/tailored survey design. Link and Burks (2013) used combinations of the appended variables to tailor incentives and contact mode to sample units. Appended data can also be used as part of the sampling design to identify specific population subgroups to make household screening more efficient, or to define geographical, social, behavioral, economic, or managerial subgroupings of addresses as part of a stratified random sampling design.

The nature of these appended variables has implications for their use in sampling

designs (Roth et al. 2013). For example, suppose one needed to stratify the household population into two strata based on whether children are present. Based on the appended information, it may be possible to identify those households for which children are present (i.e., appended flag is 1), but because the flag is missing for all other households, the second and remaining stratum would comprise all households that do not have children or for which there is no information about the presence of children. If not having children is related to the primary outcome of interest, then the second stratum may in fact have more heterogeneity in the outcome as a result of combining the addresses with no children and those for which no information is available. This extra variability may have implications in the required sample sizes needed. Care should be taken by sample designers to understand the extent to which the appended variables match intended strata and when they represent combinations of strata (Kolenikov et al. 2013).

A second important point about using appended variables for stratification involves the structure of the ABS frames. Although some frames contain both the address and myriad information about each address, it is also possible for the frame of addresses to be housed separately from the proprietary sourced information that can be appended to the frame. This separation is in part because of licensing issues and restrictions for keeping proprietary information from multiple sources separate and distinct. As a result, the use of appended variables for stratification designs would occur as part of a two-phase sampling process (Roth et al. 2013). As described in Section 4.4, in the first phase, a sample is selected from the ABS frame, and prior to the selection of the final (second phase) sample of addresses, the information needed for stratification is appended to each of the first-phase sampled addresses.

Issues with appended telephone numbers are relevant for the design of contact and interview mode. As noted in Section 3, appended telephone numbers are more likely to be landline numbers than cell numbers, and the landline and cell populations differ in demographic and behavioral characteristics. Thus, ABS studies that plan to use telephone for the initial contact with the sampled household should have alternative procedures for contacting households for which no phone number match or an incorrect phone number match was obtained.

4.6 Sample Design Consideration for Local versus National Surveys

Local ABS surveys face different design considerations than national ABS surveys. Local surveys cover individual counties, cities, Metropolitan Statistical Areas, and even neighborhood areas defined by ZIP Codes, census tracts, block groups, etc. Most ABS surveys seek to sample households or to sample adults residing in households. Group quarters and adults residing in group quarters are often excluded from “general population” surveys.

The first major issue relates to the distribution of addresses in the ABS frame by address type for the local geographic area being surveyed. This distribution is generally available and should be examined prior to sample selection. As noted in Section 4.2.1,

the number of drop- point and educational addresses in the ABS frame can vary considerably among local areas. Survey researchers need to be cognizant of the distribution of these addresses in the geographic area they are studying and will need to implement specific sampling strategies for drop points or educational addresses when the local geographic area contains a large number of drop point units or student housing units.

Two useful sources of information for local area surveys are the most recent decennial census and the ACS. These data sources can provide a comprehensive picture of the local geographic area during the design stage, even if that local area is a neighborhood. Of particular interest is the relationship between the number of persons living in households and the resident population of the local geographic area. A large difference between these two population counts indicates that there is a substantial group quarters population in the local area. In this situation, one should carefully review the distribution of addresses in the local area by address type to see if any address types should be excluded from the sampling frame.

Another issue for local areas is the situation where one is sampling from a county or city but the sample is stratified by neighborhoods. The neighborhoods may be defined by census tracts, block groups, etc. The stratification of the sampling frame will use default assignment rules for assigning P.O. Box addresses to a geographic stratum, typically based on the location of the post office. In that situation, particularly if the true geographic location of the respondent within the local area is of analytic interest, it may be important to include a question on geographic location in the questionnaire for P.O. Box addresses to determine actual residence stratum location.

4.7 ABS for Nonhousehold Surveys

It is well known that the vendor address files contain residential addresses, but the USPS also delivers to nonresidential addresses (commercial, industrial, agricultural, etc.) included in the AMS, and therefore, it is, in theory, possible to use ABS for nonresidential addresses. In practice, we are aware of very few nonresidential surveys that have been conducted using USPS- based ABS frames.

Often commercial list frames are used to draw a sample of business establishments. Other business surveys collect information from nonresidential buildings. In this situation, multitenant (i.e., multiestablishment) buildings must be sampled rather than sampling the individual business establishments within the building. List sampling along with area probability sampling is most likely needed in this situation.

The Energy Information Agency (EIA) of the Department of Energy (O'Brien 2013) collects information from a national sample of commercial buildings. A pilot test using the business addresses in the vendor files used to construct ABS frames found that the relationship between the postal address and the unit of analysis in the survey was highly variable. Issues arose primarily for strip malls, business campuses, and

multitenant buildings. The experience of the EIA pilot study for buildings likely means that multiestablishment structures may cause problems for sampling business establishments (similar to the issues posed by drop-point addresses for household surveys) using ABS frames. ABS may hold greater potential for surveys of business establishments more likely to have a one-to-one relationship with nonresidential addresses.

It may be better to view the vendor files used to construct ABS frames as tools for supplementing commercial list sources of business establishments. This approach could improve coverage of small business establishments that are more likely to be missing from commercial list frames. For example, suppose the ABS frame has an address for a building that contained Restaurant XX, but that restaurant recently closed and Restaurant YY now occupies the building. Commercial list frames generally have a time lag for adding new small business establishments, but the address list would include this newly opened establishment. This use of ABS for business surveys qualifies as an area for future research.

4.8 Survey Designs for Single- and Mixed-mode Surveys

The choice of survey modality is a key aspect of survey design. The primary modes of data collection are in-person, mail, telephone, and web. For ABS, one can use a single-mode survey design or a mixed-mode survey design. Single-mode surveys are those that use only one contact mode and request a response by that same mode. Thus, in-person surveys that use ABS addresses to contact sampled addresses and an interviewer to conduct the survey are a form of single-mode survey. A second type of single-mode survey is the use of only postal contacts to request completion of a paper questionnaire. If contact is made by one mode (e.g., postal mail) and people are requested to respond by a different mode (in this case, by web or telephone, for example), by definition the study is mixed mode. In the latter case, people may be required to respond by a single mode other than the contact mode, or they may be given an opportunity to respond by more than one mode.

4.8.1 ABS for In-person Surveys Using Area Probability Samples

Although Section 4.8 focuses primarily on the use of ABS for modes of data collection other than in-person surveys, this section covers the use of ABS for in-person surveys because, as Iannacchione (2011) points out, ABS has become an integral component of the sampling frames of such well-established, in-person household surveys as the National Survey of Family Growth (Lepkowski et al. 2010), the American National Election Study (Lupia et al. 2009), and the General Social Survey (O'Muircheartaigh et al. 2009).

In-person surveys consist of national probability samples and state and substate probability samples and may collect information on the household, all household members, a randomly selected adult, or a randomly selected child. They all make use of area probability sampling to select a sample of households. Area probability samples

are typically clustered sample designs with two or more stages of sampling. For example, the first stage of sampling may be counties, the second stage census block groups, and the third stage individual housing units. Different types of geographical areas may be used in the early sampling stages.

Historically, the development of the sample of households for area probability surveys typically involved sending trained listers to the sampled geographic areas and creating a listing of all households (field enumeration) in each area. However, it is well known that the typical listing process resulted in households being missed by the listers (Eckman 2010), and is expensive to implement. ABS, therefore, was originally developed as an alternative to field enumeration (Iannacchione 2011). This is accomplished by using mailing lists of residential addresses from USPS-qualified vendors for each sampled geographic area. All locatable residential address types are generally included in the listings for each area, including vacant and seasonal residential addresses.

Although an ABS frame is less costly and time consuming than traditional field enumeration, there are some limitations related to undercoverage of households that need to be addressed when the survey is conducted. These coverage issues are discussed more extensively in Section 7. Many area probability surveys based on ABS frames include a method to give households not on the frame a chance of selection. The original missed housing unit procedure in area probability studies is the half-open interval (Kish 1965) in which a missed housing unit's chance of selection is tied to an address that is on the frame, relying on the sequence order of addresses. This method is difficult to apply with frames ordered in delivery sequence (McMichael et al. 2008a). One approach to add missed housing units to the sample, with some similarities to the half-open interval approach, is the Check for Housing Units Missed (CHUM) procedure (McMichael, Ridenhour, and Shook-Sa 2008b; McMichael et al. 2014). Kalton et al. (2014) and Dohrmann et al. (2012) describe an alternative Address Coverage Improvement (ACE) procedure (referred to in some of the earlier references as Coverage Enhancement Procedure, or CEP). Also, O'Muircheartaigh, Eckman, and Weiss (2003; 2002) describe an enhanced field listing procedure. Another option involves augmenting the CDS file with additional listings of simplified addresses (Section 2.2.1) that are sourced from third-party administrative data sources to improve coverage in primarily rural areas (Buskirk et al. 2014). ABS sampling vendors may be able to add addresses using other proprietary data sources. Finally, approximately 10 million records (Buskirk et al. 2014) from the USPS No-Stat file can be added to an ABS frame. Although many of the records are inactive and cannot receive mail, some are active addresses (Shook-Sa et al. 2013). The presence of No-Stat addresses in a frame could assist survey researchers with field enumerations or other listings required by an ABS design (Buskirk et al. 2014) and can lead to efficiencies in a hybrid frame design (Shook-Sa 2014). The address counts from an ABS frame can then be used for area probability samples to estimate the measure of size of secondary sampling units with reasonable effectiveness and at minimal cost (Dohrmann, Li, and Mohadjer 2011). This point is especially important if the alternative measures of size are based on decennial census data, especially for studies conducted late in the decade.

4.8.2 ABS for Single-mode Mail Surveys

Designing single-mode surveys that use a paper contact to obtain paper responses poses new challenges. It requires surveyors to contemplate in different ways the construction of survey questionnaires than has typically been done in the past, when in-person and telephone interviews dominated survey data collection. For in-person and telephone surveys, interviewers were relied upon, not only to interpret questions, but also to redirect respondents when they seemed not to understand or be responsive to particular questions. Interviewers also provided encouragement and direction so that “no answers” were few and far between. On paper, however, questionnaire respondents are not compelled to answer every question.

Another challenge with mail surveys is that branching questions lead to respondent error. For computer-assisted interviewing, branching questions increased in number to customize and shorten the questionnaire for the respondents based on their individual responses, but increased the complexity of the instrument. Branching questions are not always followed correctly when the respondent reads the directions in a paper questionnaire. Contacting postal addresses by mail and asking a household member to self-administer questionnaires demands significant adjustments to the design and administration of household surveys.

There are also advantages to using self-administered postal questionnaires compared to interviewer-administered modes. Research shows that, typically, interviewers obtain more socially desirable answers and greater acquiescence (agreement with items) than self-administered surveys by mail. In addition, perceptions of the interviewer have been found to influence respondent answers. These considerations were generally considered less significant than the above-mentioned problems associated with postal surveys, in part because of the lack of adequate household sampling frames for mail contact and response.

ABS, which now provides far better coverage than that provided by landline telephones, has made it desirable to reconsider the use of mail as a data collection strategy. Although challenges in getting compliance with branching instructions and collecting detailed information from open-ended questions may need to be addressed, research has shown that improved visual layout and design can improve the quality of answers to branching questions in most surveys (Redline et al. 2003). In addition, answers to open-ended questions can be improved significantly with improved visual layouts (Dillman et al. 2014; Smyth et al. 2010). Nonetheless, when questionnaire topics require a great amount of branching, requiring respondents to follow different routes through complicated sets of items, it is difficult to make mail work effectively.

Experimental tests of mail-only surveys using ABS have shown that it is possible to get response rates considerably higher than those now associated with many RDD telephone surveys. A series of 14 experiments used a mail-only approach to data collection in regional and state samples in which households were asked to complete a

12-page booklet questionnaire requiring 20–30 minutes to complete. These studies produced response rates from 38% to 71%, with a mean of 52% (Dillman et al. 2014; Edwards, Dillman, and Smyth 2014; Messer 2012; Messer and Dillman 2011; Smyth et al. 2010). Response rates for national RDD telephone surveys of adults vary depending on sponsorship, questionnaire topic, use of incentives, level of effort, etc. The Pew Research Center (2012) has documented the decline in RDD response rates over time. The response rate of a typical Pew Research Center telephone survey was 36% in 1997, and by 2012 it was only 9%.

Response rates to the 14 experimental studies were likely affected by the use of multiple design factors. Each study used four to five requests to respond, all of which were sent by mail. The questionnaires were printed in color and had a cover page that conveyed the general topic of the survey and graphical design that identified the state or region being surveyed. The topics varied from quality of life and economic well-being to future energy production and water management. The person asked to reply to the survey was “the household member with the most recent birthday” or the household member most knowledgeable about the topic (e.g., “how the 2008 economic decline had affected their household”). Personalized letters were addressed to “Dear (city) resident,” using the city named in the mailing address. In addition, cash incentives of \$4 to \$5 were included with the survey requests.

Some of the experiments involved surveying distant populations, which appeared to lower response rates. Specifically, questionnaires sent to Nebraska residents from Washington State University obtained lower response rates than the same questionnaire sent from the University of Nebraska. It was also true that responses from Washington households were lower when questionnaires originated from the University of Nebraska rather than Washington State University. The lowest response rate (38%) in the 10 studies was from Alabama, which has a significantly lower education level, but was sponsored from a distant state (Washington) (Messer 2012). Comparisons of the demographics with ACS data showed that, although many demographics were comparable, respondents were less likely to have higher education or higher incomes and more likely to be married (Messer and Dillman 2011).

Mail-only data collection has also been tested in a two-phase data collection sequence, in which an initial set of contacts seeks to identify households with a particular characteristic of interest (e.g., a child under age 18). Then a second data collection sequence was used to collect additional data from those households with children of the right age (see Section 4.5 for a description of this methodology). Comparisons of the two-phase, mail-only method to an approach that used telephone for nonresponse follow-up showed that the mail-only approach yielded response rates significantly higher than the approaches that included telephone at different stages of follow-up (Brick et al. 2011). Similar results were obtained in a study of salt- water anglers in North Carolina (Brick, Andrews, and Mathiowetz 2012).

In summary, mail-only data collection using ABS appears to have considerable possibilities for use in surveying U.S. households. However, it is important to continue to

do research on additional procedures and techniques that might improve response rates and reduce differences in characteristics between responding and nonresponding households.

4.8.3 ABS for Mixed-mode Surveys

Research has also been undertaken on mixing survey modes with ABS samples. In particular, much interest exists in being able to obtain responses over the Internet when contacting households by mail. Such studies could ask people to respond only over the Internet, or they could also be encouraged to respond by paper if they are unable or unwilling to respond over the Internet. Combining mail and Web in these mixed-mode designs is significantly more complicated than simply conducting a mail-only ABS survey.

It has been shown that providing households a choice of whether to respond by mail versus Internet does not improve response rates, and may lower response rates (Medway and Fulton 2012). This may occur because offering choice encourages potential respondents to defer a decision (Tversky and Shafir 1992), or it encourages them to think of other options such as not responding at all (Schwartz 2004).

In addition, research has shown that when given a choice of responding by mail versus Internet, people are far more likely to respond by mail (Smyth et al. 2010). This tendency defeats the much desired goal of reducing survey costs by eliminating return postage and processing costs.

An alternative strategy of data collection has emerged in an effort to encourage a greater portion of people to respond over the Internet, which may be described as Web-push. Under this strategy, households receive a mail contact asking them to respond over the Internet. Then, later in the data collection sequence, at the time of the third or fourth contact, nonrespondents are offered the possibility of responding by mail. These Web-push strategies were compared to the single-mode, mail-only response strategy mentioned earlier in this section. The mean response with the Web-push strategy ranged from 31% to 55%, with a mean of 44%.

It is clear from these results that offering Web first, followed later by mail, does not improve overall response rates, but instead decreases them by about 8 percentage points. The decrease was fairly consistent across all tests (Dillman et al. 2014). However, it was also learned that in the state of Washington the mail respondents were quite different, demographically, than the Web respondents, being significantly younger, having less education, and more likely to be married with children in the household (Messer and Dillman 2011). It was also learned that the demographic composition of the mail-only treatment groups was quite similar to the combined Web-plus-mail composition of the Web-push treatments.

Early in this sequence of experiments, a mail-push method was used, withholding the offer of Web until the third or fourth contact. It turned out to be ineffective, with only 1 to

3 percentage points of increase in response from individuals who responded over the Internet at that time (Dillman et al. 2014). It was concluded that adding Web as an alternative late in the data collection process was not worth the cost. It seemed clear that the types of individuals who preferred responding over the Internet would also generally have completed mail questionnaires, but not vice versa.

To put into context the potential uses of phone with mail and Web, it is useful to note another study conducted by Finamore and Dillman (2013). This survey was of college graduates whose names, addresses, and household phone numbers were drawn from early responses to the ACS. Three treatment groups were identified, a Web-first group, a mail-first group, and a telephone-first group. Attempts to survey each of these groups then switched to the other two modes sequentially, so that all three modes were used for each group. The final response rates were nearly the same, in the range of 75% to 77%. Most respondents answered by the original mode that was requested. It was also found that the telephone-first mode cost \$75/respondent, compared to \$66 for the mail-first mode and \$48 for the Web-first mode. In addition, respondent characteristics differed very little by mode of response. Thus, in contrast to the household survey situation where a Web-first methodology produces a lower response rate and is more costly, this study showed that Web-first methodology is more effective from a cost standpoint, and there is virtually no difference in response rates. It is unfortunate that we have not yet found a way to bring telephone into ABS designs effectively.

4.8.4 Contact Mode versus Interview Mode

The ability to link auxiliary variables to ABS frames or samples not only allows researchers to target or stratify the population of addresses in multiple ways, but it also provides the possibility for multiple contact strategies as part of the survey data collection process. For example, Nielsen's TV Diary Samples are selected from an ABS sampling frame. From there, telephone numbers are matched to sampled addresses and are used as part of a phone recruitment strategy (Bailey, Grabowski, and Link 2010). Addresses without appended telephone numbers are asked to complete the screener by phone, mail, or web. The Nielsen TV diary example shows how auxiliary information can be used to support recruitment.

A growing number of surveys based on ABS samples use the appended phone information not only as a recruitment tool, but also as a response mode option (Brick et al. 2012). In this design, respondents are mailed a paper version of the questionnaire, and those addresses with matched phone numbers receive follow-up phone calls in an effort to encourage response. Households at addresses that do not respond to the initial mailings, but for which a phone number is available, receive phone follow-ups. Although these calls may be used for reminder purposes only, other studies may offer households the option to complete the survey by telephone.

Olson and Buskirk (2015) explore the potential for nonresponse bias as a result of heavy reliance on appended phone numbers for phone follow-up. In some studies, it is completely possible that the addresses for which phone numbers are available will have

higher response rates simply because of the use of an alternative mode of contact to reach nonresponding households. Furthermore, Montaquila et al. (2013) found that addresses for which telephone numbers could be appended responded at a higher rate than addresses for which no telephone number could be appended, even when mail was the only mode used. As suggested by Buskirk et al. (2014) and illustrated by Olson and Buskirk (2015), the potential for higher responses to be obtained via phone from ABS sampled addresses with phone matches compared to those addresses with no matched phone number who can complete only by mail creates an increased potential for nonresponse bias if addresses with appended information (in this case, the phone number) tend to differ from the overall population in the survey outcomes of interest.

Overall response rates are connected to the potential for nonresponse bias. Response rate calculations for mixed-mode designs where some sampled addresses have more than one way to complete a survey may need to be revised. In this situation, it may be completely possible to confirm eligibility or ineligibility via a phone follow-up, but such confirmation from addresses with no phone number can only be made if those addresses respond to the mailing. Another related issue is the situation in which a respondent refuses via the phone but ends up completing the survey via mail, or conversely, an address for which a postal return code yields an ineligible category (such as deceased) but for which a survey was completed via the phone. The interleaving of response options and eligibility confirmation options across the two modes (mail versus phone) creates potential complications for using AAPOR response rate calculation formulae because a single sampled address may yield different information about eligibility from the two different sources (mail versus phone, in this case). A rule to reconcile these differences with respect to computing response rates should be decided upon for the study. This issue is discussed in more detail in Section 5.

4.9 The Use of Incentives

Research on the use of incentives in ABS surveys with mail as the mode of contact has generally found that incentives function in mail surveys much the same ways they were found to function in studies using other modes. Studies have shown that small monetary incentives increase response rates in both single-phase and two-phase mail surveys (Church 1993; Han, Montaquila, and Brick 2013; Lesser et al. 2001; Singer et al. 2014). In addition, prepaid incentives are almost always more effective than promised incentives that are contingent on response (Church 1993; Petrolia and Bhattacharjee 2009). However, findings are mixed on the ideal incentive size and whether the effects of incentives are linear (Cantor, O'Hare, and O'Connor 2007; Gelman, Stevens, and Chan 2003; Han et al. 2013; Yu and Cooper 1983). A meta-analysis by Singer and Ye (2013) found that there is no clear-cut evidence for how large an incentive should be; however, in general, response rates increase as the size of the incentive increases, but they do so at a declining rate.

Somewhat less is known about whether prepaid incentives should be sent to all sampled individuals with the first contact or used as a refusal conversion tactic for initial nonrespondents (Brick et al. 2005) or how incentives function in multiphase mail

studies. In a two-phase household survey of children's parents, Han et al. (2013) found that a larger incentive at the first phase (\$5 compared to \$2) increased response rates at the first phase, but did not necessarily affect the average number of sampled addresses needed to obtain a final complete case at the second phase. In addition, the authors found that a small incentive at the second phase significantly increased second-phase response rates, but larger incentives did not correlate with a measurable increase in response.

There are limitations associated with the use of incentives in ABS mail surveys, as well. Aside from the incentive cost, mail surveys offer a comparatively inexpensive survey mode and, therefore, are often used in sample surveys of significant size. Although strategies can be used to maximize response rates, return rates might not be high enough to justify the expense of sending prepaid monetary incentives to all sampled addresses. Furthermore, with the large amount of "junk" mail received by households, there is concern that the survey, and thus the incentive, may be thrown away without opening, rendering the prepaid incentive ineffective. Another limitation of incentives in mail surveys is that the targeted respondent may not be the person who opens the mail and finds the incentive, also potentially decreasing the effectiveness of the prepaid incentive.

Furthermore, although incentives generally increase response rates, the increase does not always guarantee a decrease in nonresponse bias in survey estimates. Although some studies have shown incentives to decrease nonresponse bias (Griffin et al. 2011; Groves et al. 2006; Petrolia and Bhattacharjee 2009), other studies have shown that incentives can bring in individuals more similar to those who would have responded without the incentive, potentially increasing nonresponse bias (Juster and Suzman 1995).

Research has also been done on the use of incentives in ABS mixed-mode surveys, including comparisons of incentives for mail-only versus multiple (Web-push) modes (Dillman et al. 2014). A comparison of including a \$5 incentive in the initial contact showed that it improved response to mail by only 13 percentage points; however, it increased response to the Web questionnaire used in the Web-push strategy by 18 percentage points, from 13% to 31%. This suggests that token cash incentives may be more important for Web-push surveys than for mail-only approaches. The inclusion of such incentives seems critical for getting people to go from the mail contact mode to providing a response over the Internet. The greater importance of an incentive for getting people to go to the Web is not surprising, inasmuch as it would seem to require more effort to go from the postal medium of delivery to the Web medium of response than is required for simply filling out a questionnaire enclosed in the same envelope as the incentive. To put it another way, the importance of the incentive is greater when the effort needed for responding is greater.

Use of an additional incentive in contacts subsequent to the initial contact was also tested. Normally, surveyors do not use second incentives for respondents, the rationale being that to do so would be to reward having not responded earlier. However, it was

reasoned in this sequence of experiments that when people were being asked to switch modes, and different arguments for responding were being included in the later contacts in an effort to encourage response from different kinds of respondents, inclusion of a second incentive could be worthwhile. The inclusion of a second \$5 incentive in the third mailing when individuals were being encouraged to respond by mail resulted in a nonsignificant increase of four percentage points, from 48% to 52%, with more responses being received by mail (18% to 15%) than when the incentive was not used.

However, the mail-only treatment group that received a second incentive of \$5 exhibited a significantly higher response rate of 68% compared to 59% for the group that did not receive a second incentive. Thus, the second incentive appears to have reinforced the previously offered incentivized request to respond in the same mode. This leads to wondering why use of a second incentive to encourage people who had not responded over the Web to instead respond by mail was less effective. One possibility is that by giving people an incentive to respond by mail, a mode that they had been denied earlier because of the Web-push strategy, the respondents had to overcome certain feelings of disgruntlement associated with not receiving this choice in the original contact.

The use of first and second incentives is an issue that remains relatively unexplored, and more research is clearly needed in thinking how incentives might be used most effectively in ABS mixed-mode studies.

Another key consideration for future research is the cost effects of Web-push strategies of data collection. In the designs tested by Messer and Dillman (2011), it cost more per respondent to use Web-push data collection, in part because of the lower response rates. Also, it is important to remember that the use of mail contacts and incentives to encourage response are similar for obtaining both paper-only and Web-push responses. However, over time, it seems likely that greater portions of the household populations will have access to and feel comfortable with responding over the Internet.

4.10 Section Recommendations

ABS is closely related to list sampling, and many list sampling techniques can be used in ABS designs. Sampling methods such as explicit stratification, equal probability sampling within strata, implicit stratification (i.e., list sorting), disproportionate sampling, and double sampling may be appropriate to a specific ABS design. List sampling considerations also deal with blank elements and foreign elements, and these concepts carry over to ABS designs.

It is important to do some initial design work related to the specific geographic area (e.g., a county) that will be surveyed. Unless group quarters are included in the target population, one should determine the size of the group quarters population in the geographic area and have a strategy for including only households in the survey. One should determine the number of drop point addresses, P.O. Box addresses, seasonal addresses, educational addresses, and vacant addresses. The design should explicitly determine which of these address types will be included. Assuming drop point

addresses are included, one also needs to determine if the geographic area contains any drop points with a large number of housing units and include a strategy for sampling these drop points and their housing units.

There are several data collection approaches that can be employed in ABS surveys. It is useful to think of the data collection strategy along two dimensions. The first relates to the mode of data collection. An ABS survey may employ a single mode of data collection, most typically a mail modality, or it can use two or more modes of data collection (i.e., a mixed-mode design). The second dimension relates to the number of phases of data collection. An ABS survey can use a single phase of data collection, or two or more phases can be used. The most common example of a two-phase survey design, using a single mode of data collection, is a mail screener at the first phase to obtain a roster of adults in the household. For the screeners that are returned, one adult is randomly selected from the household. In the second phase of data collection, a survey is mailed to the sampled adults. One can easily extend this design approach by including additional modes of data collection in subsequent phases.

A common alternative is to use mail contact, but ask respondents to provide responses over the Internet. Tests of both the mail-only approach and the push-to-Web approach with a mail questionnaire alternative being sent later in the implementation process have revealed that, when token cash incentives are used, response rates over 40% can be obtained in some situations, with the higher response rates typically obtained by the mail-only approach.

In developing an ABS data collection approach, one should take response rate considerations, length of field period, and expected cost per completed interview into account. It is also important to consider how an adult will be randomly selected from the household if the design calls for this.

SECTION 5

ELIGIBILITY, RESPONSE RATES, AND WEIGHTS

Considerations of eligibility issues, response rate computations, and weighting methodologies are certainly not unique to ABS studies, and the literature on these topics is wide ranging. However, ABS studies have their own particular considerations related to each of these issues, and in this section, we discuss these considerations. We begin with a discussion of eligibility issues in ABS studies in Section 5.1. In Section 5.2, considerations in the computation of response rates for ABS studies are presented, and Section 5.3 contains a description of weighting nuances for ABS studies.

5.1 Eligibility Issues

Address-based samples offer researchers flexibility in terms of auxiliary data that can be appended to samples and frames to facilitate identifying and recruiting specific subpopulations, stratifying the population, and applying tailored survey protocols. For example, it may be possible to append auxiliary information such as a phone number or

e-mail to addresses sampled from an ABS frame and then use this information to conduct initial contacts, screening, or nonresponse follow-ups or a combination of these. The sampling unit for ABS designs, however, is the address, and eligibility of these addresses for most general population studies centers around verifying that the sampled unit is, in fact, a household or that someone from a specific targeted subpopulation resides there.

When appending additional information to sampled addresses that open up additional modes of contact, researchers could obtain a much wider set of signals for the address that might indicate the presence of a household, but in many cases, these might not be definitive within the context of the ABS design. For example, an e-mail might bounce back from the e-mail provider as nondeliverable, and yet the address to which the e-mail was appended is still a household. Similarly, phone messages regarding a nonworking number or a “ring no answer” might imply ineligibility for an RDD study, but would not necessarily imply that the address is ineligible for the current study. Disregarding this address based on these dispositions may incorrectly inflate the response rates (because this address is counted as ineligible rather than as a nonrespondent, for example) or affect the accuracy of the calculation of an eligibility rate.

The eligibility classification of the address will also affect final sampling weights. Care must be taken to define clearly eligibility rules for ABS studies, including how signals from various modes of contact are used in resolving eligibility for sampled addresses. In this section, we broadly discuss the determination of eligibility for general population surveys and surveys intended to target specific subpopulations that are based on ABS designs. We will also discuss how additional information obtained from possible alternate modes of contact might help clarify the eligibility determination for each address in the sample.

5.1.1 Determining the Eligibility of Each Sampled Address

When considering the eligibility of each sampled address, there are two aspects of eligibility to take into consideration: (1) eligibility of the address itself and (2) eligibility of the household residing at the address (provided the address is eligible). We consider each of these aspects in turn.

Eligibility of the Address

Determining the eligibility of the address equates to determining whether the address corresponds to a household address. Although sampling frames for ABS studies are generally restricted to contain only addresses designated as being used for “residential only” or “residential, some business” purposes (see Section 2.3.7 for further discussion), it should not be assumed that all addresses in these frames actually correspond to household addresses. For example, some addresses will correspond to vacant units or P.O. Boxes not currently in use. Thus, for the purpose of computing weights and response rates, it is important to appropriately account for the eligibility of the sampled

addresses. The information available and approach for determining the eligibility of the address depend on the mode(s) of data collection.

For single-phase studies that use mail as the only mode of data collection, and for multiphase studies that use mail as the only mode for the first phase of data collection, the only information available to inform whether the address corresponds to a household address is (1) whether a questionnaire was returned by the household and (2) whether a postmaster return (or Post Office Return) was received. If a questionnaire was returned (completed, blank, or with annotations [e.g., indicating refusal to participate in the study]) by the household, then the address is generally assumed to correspond to a household address. If no questionnaire was returned and a postmaster return was received for at least one mailing, the address may be considered ineligible depending on the reason the mail was nondeliverable indicated on the postmaster return. The AAPOR Standard Definitions report (2015) is currently being revised to include guidelines for ABS studies; once revised, this report will provide further guidance regarding how the specific reasons for nondeliverability should be treated in determining address eligibility. If no questionnaire was returned and no postmaster return was received, then the address should be classified as having unknown eligibility; it is assumed that some proportion, e , of such addresses are associated with households. See Section 5.2.1 for discussion of approaches for estimating e .

If mail is used, an important consideration is the timeline for receiving returns (from both households and postmaster returns). Without fail, returns will continue to trickle in long after most have been received. Thus, it is necessary to establish a cut-off date for processing the information from the returns. There might be one cut-off date for completed questionnaires (which might be driven by schedules for other aspects of data collection and data editing) and a different cut-off date for other household returns and postmaster returns (which typically would affect case dispositions).

Another important consideration for mail returns is reconciling results from multiple mailings. The study should establish rules for handling:

1. questionnaires being returned from more than one mailing (e.g., a household screener is returned in response to the first mailing, and another household screener is returned in response to the second mailing; in such cases, it is necessary to determine which questionnaire will be used);
2. seemingly conflicting information about whether the address corresponds to a household (e.g., a form returned by the household indicating refusal to participate, and a form from a subsequent mailing returned as undeliverable); and (if applicable)
3. conflicting information about the household itself (e.g., three person-level questionnaires were completed from a given addresses, but a household respondent reported only two eligible people in the household).

For studies using modes other than mail (perhaps in conjunction with mail) in the initial (or only) phase of contact with the household, the other modes may provide additional information about whether the address corresponds to a household. If telephone, web,

or in- person data collection are used and the initial (or only) questionnaire is completed, then the address is assumed to be eligible, provided the respondent confirms the sampled address. Because of the potential for mismatches, it is very important to include this address confirmation in the questionnaire. If the respondent does not confirm the address, and the actual address (corresponding to the household) is sufficiently different from the sampled address, then the sampled case should be treated as if no contact was made (i.e., the completed questionnaire should not be processed, and the case disposition should be reassigned). The study should establish rules for determining whether the actual address is sufficiently different from the sampled address; for example, a ZIP Code change or a difference in the city alone does not indicate a sufficiently different address, but a complete difference in the street address does (e.g., the number, street name, and street type do not match). In most cases, with a few simple rules, it will be very apparent whether a difference indicates a sufficiently different address, but studies should plan for a review process that captures the relevant information and allows for a home office staff member to review questionable differences.

If multiple modes are used to attempt data collection for the same case in the initial (or only) phase, it will be necessary to reconcile the results from the different modes. In general, a completed questionnaire takes precedence and results in the case being treated as a respondent. In the absence of a completed questionnaire, if information has been received that indicates the address is a household address (e.g., a form is returned blank by the household, or a household member answers the phone and refuses to respond but does confirm the address), then the case should be treated as an eligible nonrespondent (with respect to address eligibility), regardless of information received from other modes. In the absence of a completed questionnaire or other information (from any mode) to indicate that the address is a household address, the address should be treated as having unknown eligibility.

A word of caution is warranted for studies using telephone as a mode for the initial (or only) phase. The classification of telephone results into the broad classes used for weighting and response rate computation (respondents, eligible nonrespondents, ineligible, and cases with unknown eligibility) is different for ABS studies than for RDD studies. It is important to remember that the goal of such classifications is to determine what information the result provides about whether there is an eligible household at the sampled address. For example, a case attempted only by telephone for which the result was a nonworking phone number should be classified as a case with unknown eligibility for an ABS study, in contrast to the ineligible classification it would have received in an RDD study.

Regardless of the mode(s), the ABS study should anticipate that partial completes may be present in the data, and, in a multimode study, that these partial completes may come from different modes. Procedures are needed for identifying, reviewing, and classifying (with final dispositions) partially completed questionnaires.

Eligibility of the Household at the Address

For general population surveys (i.e., household-level surveys or surveys for which virtually every address corresponding to a household is assumed to contain at least one eligible person, such as surveys of adults), the only eligibility consideration is the eligibility of the address. For subpopulation surveys, it is also necessary to take into account whether the address contains eligible members (given that the address corresponds to a household). There may be differences among modes in the information conveyed to verify eligibility of the household (or of members of the household). For example, a nonreturned mail questionnaire could be an indication of nonresponse (among eligible or ineligible households) or may be because of the household not containing any eligible persons (and if one objective of the study is to estimate the subpopulation's prevalence, this may lead to a phenomenon referred to in some contexts as avidity bias; see, for example, Brick, Andrews, and Mathiowetz (2012)). To facilitate estimating the response rate and aiming to reduce avidity bias, ABS mail studies might want to include instructions for the completion and return of forms even if the household does not contain any eligible household members.

Eligibility Considerations for Two-Phase Data Collection Efforts

Extra consideration must be given to determine eligibility for sampled addresses and households in the context of two-phase data collections. As described in Section 4, a two-phase collection typically consists of a screening survey, designed to collect information about the presence of eligible respondents in the household, followed by a longer topic-specific survey directed at one or more of the eligible household members. Under this two-phase design, the criteria for determining eligibility at the first phase may differ from those used at the second phase. As discussed previously, in multiphase studies in which mail is used as the only mode of collection in the first phase, households for which no questionnaire is returned (either completed or partially completed, refused, or returned as nondeliverable by the postmaster) should be classified as cases of unknown eligibility at the first phase. However, in the second phase, the survey is typically mailed to first-phase responding addresses with a household member who meets the eligibility criteria for the survey (e.g., the parent of a child of a certain age, an adult over age 18, a household member who engages in some specified activity). Once that specific household member is sampled and the second-phase survey is mailed to the address, a survey returned as nondeliverable should, in most circumstances, be classified as an eligible nonrespondent rather than an ineligible address, because the household and person were deemed eligible based on the first-phase response.³

It is also important that eligibility criteria be established based on a single point in time. Two-phase ABS designs that use mail as the contact mode for one or both phases likely include some time lag between the first phase and second phase data collection. For example, if survey eligibility is based on a household member's age, that eligibility criterion needs to be set either as the person's age at the time of the first phase, his or her age at the time of the second phase, or his or her age at a specified point in time (e.g., January 1); and the time point chosen may affect the specific questions asked at

the first-phase survey. Continuing the example, if eligibility is based on age at the time of the first-phase survey, it may be acceptable to ask only for a person's age in years. However, if eligibility is based on age at a point in time other than the time the question is being asked, it may be necessary to ask for a person's full or partial date of birth. Similarly, sampled household members may die or cease involvement in the activity upon which eligibility was initially determined between phases. The eligibility in these cases, then, will be determined based on eligibility at the predetermined point in time.

Finally, other sources of error such as measurement error can lead to collecting erroneous data affecting the determination of eligibility for a given sampled household. For instance, misreporting of a person's age at the first phase may lead to a household being classified as including an eligible individual who is, in fact, out of scope for the study. This would lead to the assignment of an "ineligible" code at the second phase for an address classified as "eligible" at the first phase. In summary, it is crucial for study designers using ABS frames to consider the intricacies and nuances of eligibility coding at each phase of the survey design separately.

³ *It is possible that an individual selected for the second phase of a mail survey could move away from the sampled address between the first and second phases of data collection and a forwarding address may be provided on mail returned from the post office. Depending on the resources available, a researcher may wish to mail materials to the sampled person's new address because the individual is now the unit of study. However, this adds expense and complexity. Therefore, if the sampled case is not followed to the new address, it should be coded as an eligible nonrespondent.*

5.2 Response Rates

The previous section laid the groundwork for the computation of response rates for ABS studies; the key first step is determining the eligibility of each sampled address, and classifying each as a respondent, a nonrespondent, an ineligible case, or a case with unknown eligibility. Once those classifications have been made, the computation of response rates is done using one of the response rate computations from the AAPOR Standard Definitions (2015), which, as noted above, is currently being revised to include guidelines for ABS studies. In this section, we discuss specific considerations that often pertain to ABS studies.

5.2.1 Estimating e

As in surveys using other types of sampling frames, calculating response rates for address-based sample surveys can be complicated by cases (addresses) of unknown eligibility. This is particularly true for ABS surveys because, like RDD surveys, there are often a moderately large portion of cases with unknown eligibility. AAPOR response Rates 3 and 4 defined in the AAPOR Standard Definitions (2015) include estimates of the proportion of cases of unknown eligibility that are actually eligible for the survey (e). AAPOR standards mandate that this estimate is "guided by the best available scientific information on what share eligible cases make up among the unknown cases" (AAPOR,

2015, p. 20). The best available information for a particular study's estimate of e may come from the study's data or from another survey such as the American Household Survey.

Smith (2009) reviews the current methods that can be used to estimate eligibility rates among cases of unknown eligibility. These methods include minimum and maximum allocation, proportional allocation (Council of American Survey Research Organizations n.d.), and allocation based on disposition codes, survival methods, and linking to other records. However, the methods outlined by Smith are primarily discussed in the context of telephone and in-person surveys and, therefore, may not map clearly onto the disposition codes associated with ABS surveys, particularly when mail is the data collection mode. In addition, the assumptions made about the proportions of eligible cases among cases with unknown eligibility may differ for ABS surveys compared to telephone or in-person surveys. For example, it is believed that in RDD surveys, the proportional allocation method for computing e overestimates the actual proportion of eligible telephone numbers because the "real" ineligible numbers are the least likely to be resolved through additional contact attempts (Smith 2009). This, however, may not be the case with ABS surveys conducted by mail because it is possible that for certain address types (e.g., vacant addresses, unassigned P.O. Boxes), additional contact attempts would yield increased postmaster returns, allowing cases to be classified as known ineligible addresses. Furthermore, it is plausible that postmaster returns are quite effective at identifying ineligible addresses among certain types of addresses, in which case the proportional allocation method may, in fact, underestimate the number of eligible addresses for some ABS designs. Research does not exist, as of yet, that examines the proportion of ineligible addresses for which postmaster returns are received for different address types and geographic indicators. Therefore, it is not yet known how accurate a proportional allocation estimate of e is for ABS samples.

One recommendation that is becoming more widely used for ABS sample surveys, particularly those where the only or initial mode is mail, is to stratify the sample based on auxiliary variables that may be indicators of address eligibility and estimate e using a proportional allocation within strata. For example, addresses classified as "vacant" on the ABS frame are believed more likely to be ineligible and are likely to have a higher proportion of postmaster returns than those classified as "not vacant." Therefore, e can be estimated separately by stratum, and then the total number of eligible addresses can be calculated as the sum of the estimates within strata.

Whichever calculation method is used, it is crucial that the basis for the estimate be explicitly stated and detailed. In addition, it may be prudent to use multiple methods to estimate e and present a range of response rates when estimates of e vary. For example, one may wish to report response rates that compute e using the minimum and maximum allocation as well as a proportional allocation or stratified proportional allocation to present best- and worst-case estimates of nonresponse. Finally, for subpopulation surveys, multiple estimates of e should be computed for address- and household- or person-level eligibility.

5.3 Weights

Weighting an address-based sample follows many of the weighting procedures associated with list samples. A perfect address-based (list) sample would have a one-to-one correspondence between each residential address and a household, all residential addresses would correspond to occupied housing units, and each household in the population would be included in the sampling frame. None of these conditions will hold for an address-based sample. In addition, an address-based sample is almost certainly going to be subject to differential nonresponse. The weighting procedures that have been developed for list sampling take these factors into account, and the literature on this topic is directly relevant to weighting address-based samples (Valliant et al. 2013).

An important part of any weighting methodology is making use of the final disposition assigned to each address in the sample. This topic has been addressed previously, and the weighting procedures should make use of the classification of each sample address as, at a minimum, an eligible household with a completed interview, an eligible household without a completed interview, an ineligible sample address, or an address of unknown eligibility.

The steps in the weighting process typically involve the calculation of design weights to account for probabilities of selection, eligibility adjustments, unit nonresponse adjustments, and poststratification to control totals.

5.3.1 Design Weights

If an ABS design employs probability sampling, then each sampling unit (residential address) has a known probability of being included in the sample (i.e., an inclusion probability). The design (sampling) weight equals the reciprocal of the inclusion probability (Kish 1965). Consider a stratified address-based sample that uses equal probability sampling within each stratum. The population count of residential addresses in each stratum can be obtained from the sampling frame. The number of sample addresses drawn from each stratum is available as part of the sample design. The design weight (also called the base weight) for this ABS design equals the population count of addresses in the stratum divided by the sample count of residential addresses for the stratum. If the overall sample design is self-weighting, then the design weight is a constant for all sample addresses. This design weight can be referred to as an address-level weight.

The address-level weight may need further probability-based adjustments depending on how the address-based sample was designed. For example, if a drop-point address (Section 2.2.2) sampling unit contains 10 housing units and the ABS design uses a sampling procedure to randomly select one housing unit, then the household-level weight needs to be multiplied by the ratio of the number of housing units in the drop-point address to the number sampled (10:1 in our example). In addition to accounting for housing unit selection probabilities within drop-point addresses, other address types

may require further adjustment to the address-level weights. One needs to be particularly aware of the possibility of housing units with multiple file addresses (Section 2.2.1). An example would be a household that has a P.O. Box but also has home mail delivery. In this situation there is a “more than one” linkage of frame file addresses to the household. For this specific example, a question can be included in the survey to determine if the household has a P.O. Box (if the sample address is not a P.O. Box) or if the household has home mail delivery (if the sample address is a P.O. Box). After adjusting for multiple probabilities of selection, the design weight can be referred to as a household-level weight.

Once the household-level weight has been calculated, the sample design will dictate if further probability adjustments are required. The most common example of further adjustment is the calculation of a person-level weight based on the random selection of one adult from the household. This is referred to as accounting for within-household selection and is not unique to address-based samples. The most common situation involves multiplying the household-level weight by the ratio of the number of adults in the household to the number sampled from the household (generally one adult). In some situations a cap is put on the maximum number of adults in the household to avoid large within-household selection weight adjustment factors.

5.3.2 Multimodality Weighting Issues

When multiple modes are used for data collection in the same phase (e.g., mail for the initial attempt and telephone for nonresponse follow-up, provided a telephone number is available), the key weighting issues introduced by using multiple modes are specific issues for determining eligibility—confirming the address and identifying and reconciling multiple responses. See Section 5.1.1 for further discussion of these issues.

5.3.3 Adjusting for Unknown Eligibility

As discussed in Sections 5.1 and 5.2, careful attention needs to be paid to assigning final status codes to the multiple disposition codes possible with ABS surveys. As with the computation of response rates, adjustments must be made to sampling weights to properly account for the fact that some portion of unknown eligibility cases should be considered eligible nonrespondents. This adjustment is computed by multiplying the base weights of known eligible cases by the inverse of the assumed eligibility rate (typically e). This computation can be made overall or within strata or within weighting classes created using variables known for all sampled addresses such as address type, vacancy status, or home tenure (if those data are appended to the address sample). Details about forming weighting classes for ABS surveys are discussed in detail in Section 5.3.4.

5.3.4 Unit Nonresponse Adjustments

Most surveys use weights to adjust for unit nonresponse. List samples such as address-based samples are in a particularly good situation to implement weighting class

adjustments for unit nonresponse. Basically, sampling frame variables (i.e., auxiliary variables known for all sampled units) are used to form nonresponse weighting adjustment classes (Lohr 2010). Poststratification (or other calibration approaches) is an alternative that can be used in place of or in combination with weighting class adjustments, and these methods require only the auxiliary variables to be known for respondents, with population totals of these auxiliary variables available from a reliable external source; this approach is discussed further in Section 5.3.5.

With weighting class adjustments, the nonresponse adjustment factor is calculated for each class as the ratio of the sum of the unknown eligibility-adjusted weights for the eligible sampling units in the class to the sum of the eligibility-adjusted weights for the responding sampling units in the class. The eligibility-adjusted weight of each responding sampling unit in the class is then multiplied by the nonresponse adjustment factor for that class. In most situations the unit nonresponse adjustment is carried out at the household level because most of the sampling frame variables are at the address or household level.

Address-based samples have three levels of auxiliary variables that can potentially be used to form weighting classes. The first level consists of auxiliary variables available at the census tract, block group, census block, or ZIP Code level for each sample address. These variables are obtained from the most recent decennial census, the ACS, or commercial data sources. Examples include median household income, proportion of the adult population that graduated from college, and the race/ethnicity distribution of the population. The second level consists of auxiliary variables that are available for all addresses in the sampling frame. The primary examples are variables associated with the type of address (city-style, drop point, P.O. Box, vacant address, etc.). The third level consists of auxiliary variables that are usually appended after the sample of addresses has been selected. Examples include tenure status, presence of children in the household, and education of the head of the household (Buskirk et al. 2014). These appended variables are almost always not available for every sampling unit (i.e., not every address in the sample will have a match), so if these appended variables are used, the impact of missingness needs to be considered.

In general, weighting classes should be constructed so that the sampling units within each weighting class are similar on the key survey subject matter variables, and so that the response rates vary across the weighting classes. For ABS one should therefore review the available auxiliary variables to identify variables that exhibit response rate differences (i.e., differential nonresponse). Many auxiliary variables are categorical, and response rates can be calculated for each category of the variable. For other auxiliary variables it is necessary first to form categories to do this type of response rate analysis. Many surveys specify minimum category sample sizes in terms of the number of respondents or maximum nonresponse adjustment factors. Categories are collapsed together as needed to meet the minimum sample size or to stay within the maximum adjustment factor. The second step in the process is to take some of the key survey subject-matter variables and determine if the sample mean (or sample proportion) varies across the categories of each auxiliary variable under consideration. The ideal

weighting class variable exhibits response rate differences and differences in the sample means across the categories (Little and Vartivarian 2005).

Many implementations of ABS have found that for appended auxiliary variables, the category corresponding to “no match” (i.e., the auxiliary variable is not available for the sample address) often has a different response rate than the categories corresponding to the auxiliary variable being matched to the sample addresses. In other words, whether an auxiliary variable can be appended to a sample address is often a significant predictor of response (Brick et al. 2013; Montaquila 2014). Therefore, if one is using an appended auxiliary variable to form weighting classes, it is important to have a category for the “no match” sample addresses. It is also possible to form weighting classes based on the cross-classification of two or more auxiliary variables, although one needs to review the cross-classification weighting cells for small sample sizes and collapse cells as needed. For example, if one identifies two appended variables each with three categories, this would form nine weighting classes before any cell collapsing. However, each of these variables also has a category for “no match” sample addresses, and therefore, one actually has 16 weighting classes before any cell collapsing.

Other approaches to forming weighting classes for unit nonresponse adjustment of address-based samples should also be considered. These methods include propensity modeling, also known as propensity score adjustment, and classification tree methods (Buskirk and Kolenikov 2015; Lohr 2010; Valliant et al. 2013). These are multivariable approaches that model nonresponse and can be used to form weighting classes that reflect several variables. Compared to RDD, address-based samples have numerous sampling frame variables available, and may be better suited for the application of these methods to reduce potential unit nonresponse bias. However, further research is needed on how these auxiliary variables are associated with unit nonresponse and with key survey variables and, as a result, how useful they are in adjusting for nonresponse bias.

5.3.5 Poststratification to Control Totals

A full discussion of the types of variables to consider for poststratification to control totals is both complex and multifaceted. There are generally two approaches to poststratification. The first approach involves the direct use of explicit variable distributions (e.g., age categories, gender, race/ethnicity) and controlling to these population distributions using calibration techniques such as cell poststratification or marginal raking (Lohr 2010). The second approach makes use of propensity models to assign values via logistic regression or other more complex models such as classification trees (Valliant et al. 2013). Furthermore, there are at least two basic underlying models that form the basis of a poststratification system itself. One set of models attempts to adjust for differential nonresponse by modeling either the response rates or the propensity to either respond or not respond. The other set of models directly examines the relationship between key survey outcome measures (e.g., health risk factors) or sociodemographic variables correlated with the survey outcome measures

that are available for survey respondents and the overall population (Smith et al. 2004).

Level of Variables and Data Requirements

The choice of poststratification variables must explicitly recognize the analytical unit associated with the sample frame, the sampling process, the units of data collection, and the units of analysis (e.g., households versus adults). It must be understood that the particular variable(s) (including defined units) must be available for each completed sample case and the total population to which projection is desired.⁴

In some instances, the variable used for poststratification may be quite simple and straightforward (e.g., the use of age categories for weighting of sample individuals to available published Census Bureau age data). In other instances, it must be recognized that a hierarchy of units may be involved in a variable definition and in obtaining corresponding poststratification control totals. In some situations, the variables and population distributions used for weighting may be quite simple and easily obtained. In other situations, the structure and process may be quite complex and subject to error. For example, a sample design might involve the selection of census blocks, followed by the selection of households (addresses) within the blocks, and, finally, one or more individuals within the households. Assuming that weighting and analyses are to be carried out using both households and individual persons as units, possible poststratification weighting variables might include median household income at the block level. For the variable median household income, population distribution data may be required at the household level and at the person level (i.e., as an individual person characteristic).

⁴ *If the variable is also available for each sample selection, this may also be used to facilitate nonresponse analyses.*

Sources of Control Totals

There are a number of publicly available sources for control totals in general population surveys: these include the decennial census, the Census Bureau Population and Housing Unit Estimates, the ACS, and the Current Population Survey. These sources are primary sources for sociodemographic population control totals. There are also a number of large-scale, government-sponsored surveys such as the National Health Interview Survey that, under certain circumstances, may be useful for other types of control total variables such as type of telephone service.

Sometimes variable information at the appropriate unit level is available directly from published tables. However, for other control totals it may be necessary to do direct tabulations from full census population data (e.g., 2010 Census Summary Files) or to estimate the totals from a public-use micro data set (PUMS) such as the ACS PUMS.

5.4 Section Summary

In ABS studies, special consideration should be given to eligibility issues, response rate computations, and weighting methodologies. These aspects are not unique to ABS studies, but rules and procedures appropriate for other types of studies (e.g., RDD) do not necessarily apply or translate to ABS. A key point to bear in mind is that the sampling unit is the address, so regardless of the mode(s) used for contact and for data collection, the information obtained about the eligibility of that address (as opposed to the eligibility of a phone number associated with the address, for example) must be taken into consideration in computing response rates and survey weights. Additionally, while standard weighting methodologies may be applied to ABS samples, the sources of error and the variables available in ABS studies differ from other studies and thus, special consideration should be given to the approaches and variables used.

SECTION 6

REPORTING GUIDELINES

6.1 Current Best Practices for Reporting

Replicability is a key component of a scientific endeavor such as survey research. Toward this end, the [AAPOR Code of Professional Ethics and Practices](#) requires disclosure of key items about a study. This section first outlines several of these reporting guidelines and then discusses them in the context of ABS surveys.

To make valid inferences from a sample survey, it is crucial that survey researchers be well informed about the population under study and any limitation to how that population is defined. Given the difficulty of assembling an ABS frame, the task force anticipates that few researchers will create their own frame. Thus, most will purchase a sample from an existing frame maintained by a vendor. Section 2 discusses considerations in the construction of an ABS frame that researchers should be aware of when selecting an ABS frame vendor, and Section 4 discusses the decisions a researcher needs to make in specifying the sample from a vendor's frame. In this section, we focus on what should be reported with a release of data or analytic results from a study that uses an ABS frame. Guidance for best practices in this area comes from two sources, the AAPOR code and the Office of Management and Budget (OMB) standards for statistical agencies.

Section III.A of the AAPOR code discusses minimum standards for disclosure at the release of results from a study. Within this section there are four standards that directly address the reporting of information about the sample frame and would be relevant to studies conducted with an ABS frame.

1. A definition of the population under study and its geographic location (AAPOR Code Section III.A.3).
2. A description of the sampling frame(s) and its coverage of the target population, including mention of any subgroup of the target population that is not covered by the design. This may include, for example, excluding Alaska and Hawaii in U.S. surveys; excluding specific provinces or rural areas in international surveys; and excluding nonpanel members in panel surveys. If possible, the estimated size of

noncovered subgroups should be provided. If a size estimate cannot be provided, this should be explained. If no frame or list was used, this should be indicated (AAPOR Code Section III.A.5).

3. The name of the sample supplier, if the sampling frame or the sample itself was provided by a third party (AAPOR Code Section III.A.6).
4. A description of the sample design, giving a clear indication of the method by which the respondents were selected, recruited, intercepted, or otherwise contacted or encountered, along with any eligibility requirements or oversampling. If quotas were used, the variables defining the quotas should be reported. If a within-household selection procedure was used, it should be described. The description of the sampling frame and sample design should include sufficient detail to determine whether the respondents were selected using probability or nonprobability methods (AAPOR Code Section III.A.8).

The OMB, which sets policy for federal statistical agencies, goes further in its Standards and Guidelines for Statistical Surveys (U.S. Office of Management and Budget 2006):

Guideline 2.1.1: Describe target populations and associated survey or sampling frames. Include the following items in this description:

- the manner in which the frame was constructed and the maintenance procedures;
- any exclusions that have been applied to target and frame populations;
- coverage issues such as alternative frames that were considered, coverage rates (an estimation of the missing units on the frame (undercoverage), and duplicates on the frame (overcoverage); multiple coverage rates if some addresses target multiple populations (such as schools and children or households and individuals); what was done to improve the coverage of the frame; and how data quality and item nonresponse on the frame may have affected the coverage of the frame;
- any estimation techniques used to improve the coverage of estimates such as poststratification procedures; and
- other limitations of the frame including the timeliness and accuracy of the frame (e.g., misclassification, eligibility).

The AAPOR and OMB guidelines set minimum standards for reporting from a survey, but do not specify what or how to report specifically from an ABS study. Based on review of current ABS frames and studies, the task force has prepared the following recommendations for meeting the reporting guidelines in the context of an ABS study. As discussed in Sections 2 and 4, there are a number of decisions made by ABS frame vendors and by researchers specifying a sample from a vendor's frame that may affect the coverage and thus potential for bias in the resulting sample. Survey documentation should, therefore, include a clear description of the sampling frame and its coverage of the target population, including mention of any subgroup of the target population not covered by the research design. This may include, for example, exclusion of particular geographic areas or types of addresses (e.g., P.O. boxes, vacant addresses, drop

points, seasonal or educational addresses). Specific information about the ABS frame should include the following:

- Vendor name (if purchased from a vendor).
- Name of frame used (if vendor offers different frames).
- Date the frame was constructed and the time period it represents.
- Source(s) of addresses for the frame. If CDS, indicate what percent of ZIP Codes vendor qualifies for from USPS; if DSF, indicate what sources are used to ensure complete coverage. Describe any additional files used to augment the frame, such as the No-Stat file or other lists.
- Source and description of modifications to addresses that are used in sample selection (e.g., expanded drop points, auxiliary data such as census block information or household composition information).
- Methods the vendor uses to deduplicate addresses if multiple sources are used.

Similarly, survey researchers should provide a detailed description of the sample design, giving a clear indication of the method by which the addresses were selected from the frame (see Section 4 for a description of sample selection methods for ABS surveys). In addition, if a within-household selection procedure was used, this should be described. Other specific details about the selected sample should include the following:

- The sample size
- The date the sample was selected
- A comprehensive list of address types included (e.g., city-style addresses, rural route boxes, highway contract boxes, P.O. Boxes, OWGM P.O. Boxes, vacant addresses, drop points, seasonal or educational addresses, residential/business or mixed-use addresses, and single- and multifamily dwellings)
- Any procedures used to verify the representativeness of the selected sample to the population (or frame)
- Any procedures used to identify and resolve duplicate and invalid addresses during data collection

Data collection procedures from ABS studies can lead to unique case-specific dispositions. Section 5 discusses the issues associated with assigning final case status to surveys conducted using address-based samples, particularly those collected via mail. Summaries of the methodologies used to assign disposition codes to sample records should be included in study documentation as should the methods used to calculate response rates (also discussed in Section 5). Particular attention should be paid to the computation of e (the proportion of addresses with unknown eligibility assumed to be eligible). As noted in Section 5, some studies are computing e at the address type level, rather than overall for the study. If sample dispositions cannot be provided, this should be disclosed as a study limitation.

Finally, because most ABS surveys will be probability surveys, and estimates reported using the data will be weighted, study documentation should include a description of how the weights were calculated, including how e was computed, the variables used, the sources of weighting parameters, and the methods by which the weights are applied. Please refer to Section 5 for a discussion of the computation of weights for ABS surveys.

6.2 Section Summary

This section reviewed general guidelines from AAPOR and OMB for reporting on data collections. The task force then made specific recommendations for meeting these reporting guidelines in the context of a study using an ABS frame.

SECTION 7 QUALITY AND COST ISSUES FOR ADDRESS-BASED SAMPLES

As the previous sections have discussed, ABS offers many opportunities for survey research. With any survey design, there are common issues associated with quality and cost that have to be decided upon when using ABS. For example, the timeliness of the updates to frame variables or auxiliary variables may affect accuracy, quality, and cost as discussed in Sections 2 and 3. The frame source and creation procedures can affect coverage and cost as discussed in Section 2.

In this section, we give further consideration to the quality of the variables on ABS frames, methods to determine and improve coverage, and factors to consider when evaluating the budget for an ABS sample.

7.1 Frame Variable Quality

In discussing the quality of ABS frame variables, there are two aspects to consider: completeness and accuracy. To facilitate the discussion, we consider four broad classes of variables: (1) the address variables (street number and name, secondary street address, city, state, ZIP Code, and the various pre- and post-directions); (2) other variables originating from the USPS (e.g., the seasonal and vacant flags); (3) address-level appended variables; and (4) aggregate (i.e., area-level) appended variables.

We first consider the variables that originate from the USPS (i.e., CDS or DSF) files. The address variables are assumed to be complete and accurate, because these are essentially the unique identifiers used for the purpose of mail delivery.

Other variables originating from the USPS are coded and updated by the mail carrier, however. Thus, their completeness and accuracy depends in part on the diligence and timeliness of the mail carrier in recording these variables (or updates to them) and on these updates being processed in the USPS AMS, the database that is the source of

the CDS. The completeness and accuracy of these other nonaddressing variables also depends on the frequency that the ABS vendor receives updates to its files, and the amount of time it takes the vendor to process these updates. Although the USPS releases weekly updates, the vendor may have a license to obtain these updates less frequently. Roth, Han, and Montaquila (2013) estimated that nearly one-fifth of addresses flagged as vacant on a frame based on the CDS file were actually eligible addresses. Using addresses flagged as vacant on a DSF-based frame, Amaya et al. (2014, p. 78) found 8.8% to be households and another 26.0% to have unresolved residency status (and they note that “the vast majority of unresolved addresses are likely to be occupied based on results from previous work”). Kalton et al. (2014) reported similar results, finding 41% of sampled addresses flagged as vacant to be eligible, occupied housing units.

For the address-level variables appended to the ABS frame, the sources of variables vary considerably. Likewise, the rates of completeness and accuracy of the variables also vary considerably. Amaya and colleagues (2010) discuss issues related to matching telephone numbers to addresses. DiSogra et al. (2010); Roth et al. (2013); Valliant et al. (2014); and Ridenhour et al. (2014) discuss findings involving the completeness and accuracy of appended address-level variables. As third-party sources and ABS vendor methods change over time, continued evaluations of these variables (similar to those cited here) are recommended.

Aggregate-level variables are generally complete, provided the ABS vendor has geocoded the entire frame so that census geographic variables (e.g., tract, block group, and block) are available for all addresses. Aside from sampling and nonsampling error in the estimates from the source of these variable (e.g., the ACS), any error in these variables is generally because of geocoding error (i.e., associating the wrong area with the particular address). Vendors can provide information about their geocoding methods, which generally involve different levels of geocoding (depending on the ability to match the address to addresses in the geocoding database); see Eckman and English (2012a; 2012b), Fiorio and Fu (2012), and Dohrmann et al. (2012) for further discussion of this. In addition, because these are area-level variables, their association with characteristics of the particular household at the address may be limited.

7.2 Determining Coverage

“Coverage” is one of the major aspects when evaluating a survey design, and so being able to measure it is of key interest to researchers and practitioners. We can define net coverage as the ratio of the number of sampling units (in this case, households or persons) on our frame to actual households or people in the population. We decompose net coverage into undercoverage and overcoverage. Undercoverage represents units in the population that are either missing from the sampling frame (omissions) or erroneously filtered out of the frame (erroneous exclusions). Overcoverage represents units on the sampling frame that are out of scope for the survey (erroneous inclusions) or duplicated on the frame (duplicates). A net coverage rate close to 100% is not necessarily indicative of a high-quality frame. Consider that a frame with a large number

of omissions could be offset by an equal number of erroneous inclusions or duplicates to yield a net coverage rate close to 100%.

Until recently, RDD computer-assisted telephone interviewing (CATI) surveys had been an often-used approach for population studies. Call screening devices, mobile telephones, and the rise of mobile-only households have complicated CATI studies by both putting pressure on response rates and contributing to undercoverage of certain subgroups without landline phones (Blumberg and Luke 2007; Blumberg, Luke, and Cynamon 2006; Blumberg et al. 2008; Curtin, Presser, and Singer 2005; Link et al. 2006). Because of the exclusion of wireless-only households, wireline RDD frames have been estimated to exclude 41% of households (Blumberg and Luke 2014), indicating the need to use mobile phones in telephone survey designs. Telephone studies that use both wireline and mobile telephone exchanges are a special case of dual-frame designs (Brick et al. 2007; Keeter et al. 2007). Dual-frame designs do carry complexities related to weighting and cost. Pertinent to ABS studies, Link et al. (2008) indicate that the coverage of using USPS lists can be as high as 98%, but there are other studies that have had estimates in the low 90% range.

Coverage error in ABS frames can happen for a variety of reasons. For example, an ABS address is geocoded to the census block level, and geocoding errors will result in some households being assigned to the wrong census block. Second, apartment numbers may not be available for some drop-point addresses in the frame. Third, a household may have a P.O. Box as the only way of obtaining mail, and these households are not assigned to the census block where they are located. Fourth, new construction may result in occupied households at the time of the survey not being included in the frame. Finally, the CDS file used by vendors does not include residential units with simplified addresses (Section 2.2.1; approximately 200,000 in 2014) (Dohrmann et al. 2014). All of these issues contribute to frame coverage error. Some of these issues are discussed more fully in the following sections.

7.2.1 Effect of Sample Design on Coverage Error

We would expect the coverage of address frames to vary depending on the type of study and modes used, in addition to how one creates the sampling frame. For example, local studies could be complicated by the presence of non-city-style addresses, which are usually not geocodable, but such addresses might have less of an impact at the state or national scale. It would be up to an individual study to include ungeocodeable addresses or omit them as erroneous exclusions, with mode being an important factor in such a decision. The inclusion or exclusion of P.O. Boxes and rural route addresses may also depend on mode. For example, a survey may mail to P.O. Box and rural route addresses, but not send interviewers to such addresses. The decision to employ in-person, mail, telephone matching, web, or a combination may also partially depend on the target area for a particular study. When area probability designs are used, the frames corresponding to selected geographies can have coverage errors when addresses are geocoded to the wrong geographies. Lastly, the methods used to conduct within-household selection of a particular respondent will

be influenced by mode and sampling frame, because self-administered methods are more difficult to reliably execute member selection than those that are interviewer administered.

Decisions on how to create the sampling frame will also affect apparent coverage (e.g., whether to include and how to treat P.O. Box, Highway Contracts, and other ungeocodable addresses). Moreover, different results may be realized depending on the inclusion or exclusion of group quarters, military addresses, and remote areas. Estimates of coverage will also vary considerably depending on the source of a given benchmark, whether it is the Census/ACS, traditional listing (described below), or another source; the same is true if we are examining households, housing units, or another count.

7.2.2 Coverage Evaluations

Until the early 2000s, the only national source of address frames was listing. Traditional listing is a method of address frame generation created by field staff, known as listers, who record all residential addresses in defined geographies in a systematic manner (Kish 1965). This method of frame creation has been considered the “gold standard” in the survey research community since the early days of in-person studies (O’Muircheartaigh, English, and Eckman 2007; O’Muircheartaigh et al. 2006). Motivated by the high costs associated with traditional listing, survey research organizations have in recent years been undertaking research into using the USPS CDS file and DSF2 as a replacement, because it is a list of essentially all housing units receiving mail in the United States (Kennel and Li 2009; Montaquila et al. 2009; O’Muircheartaigh et al. 2007). In addition, research has shown that traditionally listed frames are themselves subject to undercoverage because of lister error (Kwiat 2009), with Eckman and Kreuter (2013) estimating an undercoverage rate for traditional listing efforts of 13.6%.

Because of the high costs involved, survey research organizations have been researching the use of an address frame as a complete or partial replacement for traditional listing (Battaglia et al. 2008; Iannacchione et al. 2003; Link et al. 2008; Montaquila et al. 2009; O’Muircheartaigh et al. 2007). One of the first examinations of an address frame was in Dallas County, TX, which showed promising coverage in an urban sample (Iannacchione et al. 2003; Staab and Iannacchione 2003). Others have since conducted assessments of the coverage properties of CDS-derived frames in different environments across a number of projects (English et al. 2009; Montaquila et al. 2009). Subnational coverage estimates have been explored by McMichael, Harter, Shook-Sa, Iannacchione, Ridenhour and Hutchison-Everett (2012). Additional research has been undertaken into where it is appropriate to use the address frame, where it is necessary to list, and where it may be possible to employ a hybrid approach to fill in gaps in a given frame (Montaquila, Hsu, and Brick 2011).

A consistent finding across studies is that the databases suffer from undercoverage of rural areas, in large part because of the type of addresses available in those areas (Dohrmann et al. 2006; Dohrmann et al. 2007; O’Muircheartaigh et al. 2007). City-style

addresses consist of a house number, street name, state, and ZIP Code. Non–city-style addresses, common in rural areas, consist of box numbers rather than house numbers and delivery route numbers rather than street names; they include post office, rural route, and highway contract boxes. The correspondence between a mailbox and a housing unit are not necessarily clear in rural areas, especially if several mailboxes are clustered at the end of a lane leading to multiple housing units. For in-person surveys, interviewers must be able to locate each selected address; thus, only city-style addresses are useful (Eckman and English 2012a). We expect coverage to improve in rural areas as addresses convert to city style for 911 services.

The current thinking is that a CDS-based address frame performs at least comparably to traditional listings in urban and suburban areas, especially those with regular block patterns, single-family homes, and relative housing stability (English et al. 2009; Iannacchione 2011). Rural areas, however, are known to contain a larger share of ungeocodable addresses, including P.O. Box and rural route addresses. Because survey research organizations are usually interested in targeting small areas, ungeocodable addresses are usually considered undercoverage for in-person studies. Consequently, the coverage of the CDS in rural areas is not yet adequate for in-person surveys that require a locatable housing unit address for sampling purposes, especially for those federal surveys following OMB guidelines (Eckman and English 2012a; WhiteHouse.gov 2006). Non–city-style addresses may be sufficient for mixed-mode surveys, especially in areas where addresses have not been converted to city-style addresses. Frame providers that maintain a historic record of addresses or that have problems matching non–city-style addresses may have both the non–city-style address and city-style address for the same housing unit.

7.2.3 Decisions to Make When Evaluating Coverage

When determining the power and size of a sample, statisticians often use target response rates and target margins of error or coefficients of variation for key estimates. In addition to these targets, it is also helpful to set a target coverage rate. Target coverage rates can be used as a project-specific standard for determining whether coverage improvement is necessary overall or in specific areas.

According to standards issued by OMB, the sampling frames for U.S. Federal Government surveys should have a coverage rate over 95% overall and for major strata. Furthermore, if the coverage rate falls below 85%, coverage improvement should be considered and a coverage bias study should be conducted (WhiteHouse.gov 2006). That said, the coverage rate can be defined differently based on the target population (e.g., total housing units vs. occupied housing units at this stage), and the methods of estimating coverage can vary.

Coverage thresholds are often used in determining whether coverage improvement is necessary and where the coverage improvement should be conducted. Thresholds are usually set for the smallest geographic unit of analysis, but can be set for key domains as well. The threshold is then used to determine whether coverage improvement is

necessary. These thresholds are study-specific decisions, not an absolute criterion. For example, consider an ABS frame with 93% national coverage. If this particular study has a requirement for 90% national coverage, then no coverage improvement would be necessary with this threshold. On the other hand, a survey specifying a 95% coverage threshold would require some coverage improvement with this frame. A survey producing state-level estimates and a coverage threshold of 85% per state would need coverage improvement only in states where the coverage was less than 85%.

Like target response rates, coverage thresholds do not directly address bias. Setting a target coverage rate threshold is most useful when the undercoverage rate is highly correlated with undercoverage bias. On the other hand, if the characteristics of the households omitted from the ABS frame are similar to the households that are covered by the ABS frame on key survey measures, then little coverage improvement is necessary. For an estimated mean, coverage bias can be expressed as the product of the undercoverage rate and the difference between the mean of covered units and the mean of the omitted units.

When information about coverage bias is available, then coverage thresholds can be set on the amount of bias that can be tolerated. If no information is available about the frame omissions, one can conduct a sensitivity analysis to see what coverage rate is needed to achieve a target amount of bias. For example, if a key proportion being measured is 50% for the covered population and 25% for the omissions, then the target coverage rate would need to be 80% to get a bias of 5 percentage points or less. However, if the key proportion is 50% for the covered population and 40% for the omissions, then the target coverage rate would need to be 50 percentage points to attain that same bias of 5 percentage points or less.

The coverage threshold can also be used to identify the size of the coverage improvement operation. Consider a state with an estimated 80% coverage rate and a study coverage threshold of 90%. If area segment coverage estimates exist, one may want to conduct coverage improvements in the area segments with the most ABS omissions until the state coverage rate increased from 80% to 90%. Note that this does not mean all area segments with a coverage rate less than 90% are improved; rather, the area segments with the most ABS omissions are improved until the 90% coverage rate at the state level is attained.

Frequently, area segment-level counts of omission are not available. Thus, area segment-level coverage is estimated using models or net coverage ratios. Shook-Sa (2014) describes a number of methods for predicting ABS coverage, including multiple regression models used by Montaquila, et al. (2011). A simpler approach is one described by Iannacchione et al. (2012), where they evaluate a coverage model that estimates coverage by taking the ratio of ABS addresses in the area segment to the estimated number of dwelling units. One could estimate the number of omissions in an area by comparing frame counts to a control such as the ACS, and decide what area segments need to be listed to achieve the overall target coverage rate. As noted previously, frame counts might include both erroneous exclusions and duplicates or

erroneous inclusions, so the ratio is not a definitive measure of coverage.

The choice of coverage thresholds (and the coverage enhancement procedures they trigger) can have an impact on the survey budget. In general, it is more expensive to reach a high coverage threshold because lots of coverage improvement may be needed. Thus, coverage thresholds should be set with the survey budget and analytical goals of the survey in mind.

7.2.4 Impact of Geocoding Error

Geocoding is a key step in using an address file as a sampling frame (Eckman and English 2012a; Eckman and English 2012b). The process of geocoding is discussed in Section 3.1.2. We do expect a degree of error in geocoding, which may or may not influence studies that employ geocoded addresses. For example, if addresses tended to be geocoded in particular Census blocks incorrectly, they may be subject to erroneous inclusion or exclusion. A 2012 evaluation found that 83% of all residential addresses and 94% of those that were city-style on a CDS-based address frame geocoded precisely enough that they were placed in the correct block (Eckman and English 2012b). It is also possible to geocode addresses to ZIP Code or ZIP+4 centroids, both of which are coarser units of geography than the block or individual address. Doing so requires less-sophisticated software than address-level geocoding, but can create relatively imprecise location information. Thus, smaller geographical areas have more geocoding error than larger areas. Eckman and English (2012b) found that the national net coverage for in-person surveys was 86.7% to 92.3% if one compared the count of addresses that geocoded at the block level to the housing unit counts from Census 2010. Their method excluded non-city-style addresses and other ungeocodable addresses, but they found that 6% of city-style addresses could not be geocoded to the block level with confidence. These results imply that removing ungeocodable addresses and those that do not geocode to the block level could introduce undercoverage.

7.3 Methods for Improving Coverage

Depending on coverage goals, budget, mode of data collection, quality of the ABS frame, the sample design, and analytic goals, coverage improvements may be necessary to mitigate undercoverage on the ABS frame. There are two common approaches to improve the coverage of frames: linking methods and supplementary frames (Kalton et al. 2014; Kish 1965, sections 2.8 and 9.4C). The first approach involves matching the frame to other sources, commonly additional files or field listings, to determine frame omissions. The omissions are then subsampled, sometimes with certainty. The second approach involves creating a stratum for units or areas that might have inadequate coverage on the ABS frame.

Before conducting coverage improvement operations, the frame creation processes should be considered and possibly enhanced. Section 2 discussed several techniques some data providers take to build ABS frames. Other commercially available files such as credit data, phone directories, and real estate databases can be matched to ABS

frames, but the benefit to matching to such frames is rather small (Kalton et al. 2014). By merging multiple address files, data providers potentially can reduce undercoverage of ABS frames. On the other hand, augmenting frames with additional addresses can also increase erroneous inclusions and duplicates on the frame, resulting in increased cost per interview. Depending on the quality of the augmented addresses and the augmenting procedures, frame augments with other sources can either reduce undercoverage or have very little impact on undercoverage. Thus, an augmented frame is not necessarily better than a frame that does not include additional addresses from other sources. One should balance the undercoverage reductions with the increased cost of the survey because of increased overcoverage. If combining multiple files and unduplicating addresses does not meet coverage goals, one may consider linking procedures or supplemental frames.

7.3.1 Linking Procedures

Once the vendor delivers the frame or sample, there are additional procedures and methods to improve the coverage of samples and sampling frames. Linking procedures require determining if a unit on the ground is on the ABS frame. The Half-Open Interval (HOI), CHUM, Coverage Improvement (CI) Frame, and Address Coverage Enhancement (ACE) are linking procedures.

The HOI procedure is implemented for in-person surveys after the sample is selected. Field staff are given their assigned sample units and the unit immediately after the sample case on the frame, whether the neighboring unit is in sample. If the interviewer finds a housing unit between the sample unit and the neighboring unit, the missed unit or units are also interviewed and assigned weights equal to the sample unit or they are subsampled and a subsampling factor is applied to the weights. The HOI requires that the frame be given a geographic sort so that geographic neighbors appear next to each other on the frame, and that can be challenging with a mail delivery sort order (McMichael et al. 2008a). The HOI also requires that neighbors be uniquely identified. Because the HOI requires some addresses to be present on the frame, the HOI cannot correct undercoverage where the data provider does not have any addresses in the area. For example, if the data provider does not own a ZIP Code and does not have any addresses in the ZIP Code or if the data provider does not provide simplified addresses in simplified ZIP Codes, the HOI cannot fix the undercoverage problem. In areas where blocks and parcels are not rectangular, the HOI can be difficult to implement. Also, because the correspondence between the address and the housing unit for rural route and highway route addresses is not always obvious, the HOI can be challenging to implement in some rural areas. Eckman and O'Muircheartaigh (2011) note other limitations of the HOI.

The CHUM procedure enhances the HOI procedure (McMichael et al. 2008b; McMichael et al. 2014; Pedrazzani et al. 2012). Unlike the HOI, the CHUM does not require the address list to be geographically sorted so that neighboring units appear next to each other on the list, but it does require the interviewer to search the selected housing unit and to follow a prescribed order around the block until another HU on the

ABS frame is encountered. The interviewer also searches a subset of selected blocks to ensure that housing units in blocks with no city-style addresses on the frame have a chance of selection.

The CI Frame and ACE procedure are two additional linking procedures for in-person surveys that require listing a sample of blocks to determine which addresses were omitted from the ABS frame. Linking is necessary to classify a listed address as an omission. In both methods, all or a sample of the omissions found are interviewed.

In the CI Frame, a model is fit to predict the total number of ABS omissions in each block. Blocks with more than some threshold of expected omissions are put into the CI block frame. This CI block frame exists only in states that are at high risk of having biased estimates because of undercoverage based on previous coverage bias studies. The threshold determining if a block is to be included in the coverage improvement block frame is based on the number of omissions needed to limit undercoverage bias from changing key estimates. For example, if the target omission rate is 5% for a state, but the state has an omission rate of 10%, then the blocks with the most omissions will be placed into the CI block frame until they contain 50% of the expected omissions in the state. A sample of blocks from the CI block frame is selected and listed using a dependent listing. All ABS frame omissions found during the listing process are interviewed. Several Census Bureau surveys, including the Current Population Survey, currently use this method.

The ACE procedure addresses both geocoding errors and noncoverage on ABS frames (Dohrmann and Sigman 2013; Kalton et al. 2014). With this procedure, a subsample of area segments is selected for coverage enhancement. In the subsampled area segments, field representatives are given the list of addresses that geocode into an assigned area and asked to determine if each housing unit on the ground is on the ABS frame or needs to be added. The ACE procedure may be done either prior to the sampling of addresses (by including the previously omitted addresses in the sampling frame) or after the original sampling of addresses (by sampling from a frame that contains only the previously omitted addresses, and adding the sampled addresses to the original sample). In practice, ACE has been done after the initial sample selection to make efficient use of resources.

7.3.2 Supplementary Frames

Supplemental addresses can be constructed at the unit or area segment level. If the supplementation is done at the unit level, then a separate file with addresses that are mutually exclusive to the ABS frame is created and sampled as a separate stratum. Supplementation at the area segment level relies on area probability samples. First, area segments are defined and sampled. Then the sample segments are partitioned into those with adequate coverage and those with inadequate coverage. In the area segments with inadequate coverage, alternative frame creation procedures are developed, usually listing. Listing is a method where field representatives canvass an area, writing down all of the housing units in the assigned area. Typically the listings are

sampled.

Data providers with a CDS license can receive updates for addresses on the No-Stat file (Shook-Sa 2014; 2013). The No-Stat files contain addresses that are not included in the regular CDS updates and not included in USPS performance measures and other mandated reports on mail delivery. More research into the No-Stat files is needed, but recent research suggests that the No-Stat file can improve rural coverage by up to 4% (Shook-Sa et al. 2013), but with some loss of efficiency from adding erroneous inclusions (inactive addresses) (Dekker et al. 2012; Dekker and Murphy 2014). By including only rural throwback addresses, coverage can be increased 2.2% without the inefficiency (Shook-Sa et al. 2013). With the exception of one address type (drop points), at any moment, the addresses on the regular CDS updates should be mutually exclusive of the addresses on the No-Stat file. Thus, the addresses on the No-Stat file can be placed into a separate stratum and sampled independently of the regular CDS updates. On the other hand, because units can move between the No-Stat and CDS files over time, there may be duplicates between these two files for vendors that maintain a cumulative list of historic updates.

Supplemental unit frames can also be used to improve the coverage of specific deficiencies. For example, a supplemental frame of group quarters, assisted living facilities, campgrounds, or marinas can be sampled. However, because addresses of these units may also exist on the main ABS frame, effort needs to be made to deduplicate between the main ABS frame and the supplemental frame. This can be done in weighting if one can estimate the amount of overlap between the frames. Or one can decide to not interview units on the main ABS frame if the unit is in the supplemental frame. The National Health Interview Survey uses this method for college housing. A supplemental frame of college housing is developed. Interviewers who visit a college dormitory that was sampled from the ABS frame are instructed not to interview at the dorm. They are instructed only to interview at dorms from the college housing frame.

In areas where the coverage of the frame from the data provider is insufficient, one may opt to canvass or list blocks. Typically, an area probability sample is selected. In areas where the coverage of the vendor file is adequate, addresses from the vendor file are sampled; otherwise, selected areas are listed and the listings are sampled. The unit frame contains all addresses in geographic areas where the ABS frame has adequate coverage. Adequate coverage is usually determined by comparing census counts to ABS file counts within area segment, although more complicated models exist (Montaquila et al. 2011). Addresses that are not filtered into the unit frame are put into an area frame and must be captured through other means, such as listing. Many federal national surveys such as the Current Population Survey, the Consumer Expenditure Survey, the National Crime Victimization Survey, the American Housing Survey, and the Survey of Income and Program Participation used this technique of frame construction until a 2014 frame design change revamped the frame methods for each of these studies. Although the unit frame blocks are mutually exclusive of the area frame blocks, geocoding errors can result in some overlap and duplication between the

frames. All listed units in sample area frame blocks are eligible for sample. The National Survey on Drug Use and Health experimented with an area frame where the unit frame was supplemented with the CHUM procedure (Iannacchione et al. 2010).

In dependent listing, the data provider's address list is used as a starting point for the listing. Field representatives delete, add, and modify addresses on the vendor file so that it reflects what they see "on the ground." In independent listing, the field representatives write down all addresses in the block from scratch. Listers can miss units in systematic or variable ways, which may result in coverage bias or coverage variance (Eckman 2010; Eckman and Kreuter 2011; Kwiat 2009).

Regardless of the method used to enhance an ABS frame, the goal of enhancements is to minimize the potential of undercoverage bias. Thus, enhancements are usually targeted in areas to minimize undercoverage bias of key estimates for key domains. For example, a survey making key estimates at the state level might only conduct enhancements in states where the undercoverage rate is high. And within those states, only the sample segments with the most ABS omissions might need to be enhanced. Because the goal is to minimize undercoverage bias, the area segments with the most omissions should be enhanced with higher priority than the area segments with the highest omission rates.

7.4 Cost and Quality Tradeoffs

As this report demonstrates, when using an address-based sample for research, there are many considerations that should be made. The decisions a researcher makes will have both cost and quality implications. To that end, there are questions that the researcher should be prepared to ask the sample vendor, for example:

- What is the source of the ABS frame: CDS, DSF, or other source?
- Is the ABS frame supplemented with variables from other databases? If so, what variables are available? Is there supporting research around the accuracy or coverage of the auxiliary variables?
- How often is the frame updated?
- What is the coverage of the frame (relevant to the researcher's sample design)?
- Can the vendor provide counts of the address frame at the level of geography needed to determine sample size (e.g., county or ZIP Code level)?

So before approaching an individual vendor, it is recommended that the study design, survey design, and budget are defined because these elements will drive the conversation between the researcher and the sample vendor. Moreover, because vendors of address frames may not be aware of the needs of survey research professionals, some education may be needed to differentiate the needs of direct mailers and other users of ABS lists.

7.4.1 Factors to Consider When Looking at Costs

When determining the specific approaches to be used for a particular study and the capabilities of the vendor, the study's budget is usually a key in the quality of the sample. There may be both variable and fixed costs that should be considered when looking to optimize a budget. Variable costs usually are related to the sample design while fixed costs are linked to the vendor processing costs and survey design.

The sample size can have a large effect on the variable cost of procuring a sample file from a vendor. So the survey organization should consider whether it is necessary to obtain the entire frame for an area (e.g., for use in coverage enhancement operations) or only to obtain sampled addresses. Section 4 discusses the potential for using variables appended by the vendor to stratify (or substratify) the sample. Some vendors charge additional costs for the presence of these variables on the sample or frame. Although variables can be appended to the frame, the missingness and accuracy of these variables should be considered as well as the cost. Also, coverage of the frame can impact the vendor's ability to deliver a requested sample size. For example, if a sample size of 1,000 households is targeted in a particular county, but the frame only contains 700 addresses the quality/effectiveness of your sample may be impacted. Knowing the frame counts in advance can help inform the design.

As with any vendor agreement, the ABS vendor may impose fixed processing fees (e.g., to cover the cost to process the frame to specification, like stratification, selection, and delivery). In addition, if the vendor uses supplemental data from other vendors, there may be licensing fees associated with that data that are passed on to the client. Consequently, it is important to understand the fee structure (including which costs are fixed and which are variable or depend on the data items requested) the particular vendor uses to limit the total expense.

As mentioned previously, the ABS frame may need augmentation if it has coverage issues. The utilization of field procedures, linking procedures, or supplemental frames usually requires additional funds. Field methods, in particular can add nontrivially to the cost. Also, if a researcher is changing the mode of recruitment (e.g., moving from a single mode to mixed mode design), there may be additional costs related to the change in the process or to response rates.

Expenditures for data collection, incentives, and materials generally include both fixed and variable expenses, and these expenditures depend on the survey design. These costs should be factored in when determining the sample design and size.

7.5 Section Recommendations

ABS offers many possibilities for survey research. The choices that are made can impact quality and costs. Consideration should be given to the source of the variables used because many rely on geocoding, vendor sources, or mail carriers. Additionally, coverage error can result from geocoding errors, lack of apartment numbers for drop boxes, P.O. Boxes, only way to get mail and simplified addresses. This section offers several techniques available to correct for coverage issues. When considering the

budget for a survey, the sample design, recruitment approach, and quality should be taken into account.

SECTION 8

SUMMARY OF THE CURRENT STATE OF ABS

Increasingly, survey researchers are using ABS successfully for household or person surveys in the following contexts:

- Address frames can largely replace traditional field enumeration of housing units for area probability, in-person surveys.
- Address frames are high quality for mail surveys, which can often achieve better response rates than RDD surveys.
- Mail contacts to samples from ABS frames can push contacted households to respond by web, potentially saving considerable costs.
- Address-based samples and mail mode can be part of other mixed-mode designs.
- With auxiliary data appended, often in a two-phase approach, designs can be targeted to oversample certain subpopulations.

Researchers have turned to ABS for multiple reasons:

- Decreasing coverage of landline telephone frames, and increasing costs and complexities with dual-frame RDD methods.
- Declining response rates to telephone and in-person surveys, along with increasing costs to counter nonresponse.
- Rising costs of traditionally enumerated frames for in-person surveys.
- The commercial availability of samples and frames based on mailing addresses.

Although ABS is a viable alternative in many situations, it should not be considered a panacea for all survey research situations. The survey researcher should understand the details, considerations, unknowns, and limitations of this methodology. Section 8.1 reviews the strengths and limitations of the ABS methodology with respect to both sampling and survey administration considerations. Section 8.2 summarizes the major points of the report, and Section 8.3 indicates areas for further research.

8.1 Strengths and Limitations of Address-based Sampling

ABS has proven to be a viable approach for many surveys of households and individuals, but it is not appropriate in all situations. All approaches have strengths and weaknesses that affect their suitability for a particular study. This section points out some of the limitations (and strengths) particular to ABS to be considered when determining an appropriate study approach.

8.1.1 Sampling Frame versus Delivery Database

Address lists that are updated with the USPS information are likely the best available frames for surveys of the U.S. residential population. Fundamentally, though, the USPS files that form the basis of most ABS sampling frames are derived from the AMS database for mail delivery (Section 2); this database itself is not developed and maintained as an effective sampling frame (Section 2). Vendors who license USPS products, however, may enhance these files to make them more viable as sampling frames.

First, the USPS files do not include geodemographic indicators for effective sample stratification—an issue of critical importance for complex designs. Of course, most addresses do include the city, state, and ZIP Code that may indirectly correspond to a county or other census geography, enabling geodemographic indicators for the geography to be appended. Sections 3 and 7 discuss issues with matching variables that may be used for stratification.

Second, surveys where data are to be collected in person require the exact location of all sample dwellings, and delivery information is not always adequate for this purpose. The ability to locate an address is of particular concern, for example, where a P.O. Box is the only means of delivery for a household. In addition, surveys of specific geographical areas also require addresses to be linked to the target areas. Geographical precision has always been an issue with unlisted RDD numbers in which telephone exchanges are located with limited precision. Thus, the inability to locate some addresses, although new for in-person surveys, is nevertheless familiar to the survey industry and often an improvement for the many addresses that can be located accurately.

Third, there are dwellings that remain uncovered by the address files despite the continuous updates from mail carriers and the conversion of simplified addresses (Section 2.2.1) to a standard city-style format. In addition, vendor files may have coverage issues in ZIP Codes that the vendor does not own (Section 2.1.2). Sections 2 and 7.2 discuss frame coverage, and Section 7.3 discusses some available methods for improving coverage. In general, the coverage of housing units provided by well-maintained address lists and prepared by knowledgeable vendors is quite good, but results vary by location and vendor.

Finally, the USPS “geographic” denominations (i.e., ZIP Codes) do not conform to those from the Census Bureau (e.g., block groups and tracts). Census geographies are often used for sample stratification. Unfortunately, the process of merging the two geographical systems via geocoding is subject to mapping and compilation errors (Sections 2 and 7). Although the art and science of geocoding is improving depending on the vendor and the process employed, users need to be aware of the possibility of miscoded locations when developing sample designs for their surveys. Naturally, this issue can be more pronounced for smaller geographic locations (Dohrmann et al. 2012). Even so, as noted above, the ability to match most addresses to the correct geography is a strength of ABS.

8.1.2 Reliability of Auxiliary Variables

Depending on the vendor that is used to secure an ABS list, various geodemographic variables can be appended to the addresses. The quality of such data can differ significantly based on the data type and level of aggregation used for appending, and some types of variables (e.g., name and telephone number) are available for only a portion of the addresses in the frame. (Section 3). Variables at the aggregate level (e.g., county, census tract, or census block group) can be matched to locatable addresses and are highly accurate in aggregate, but not necessarily for individual addresses.

Other data that can be appended to addresses include household race/ethnicity and income, and person-level data, such as age and education. In addition, the commercial databases may have various behavioral data about households and consumers. Although some of these data are obtained directly from households and householders, in many cases the appended data are inferred (or modeled) and likely less accurate. Researchers who rely on ancillary data for design and postsurvey adjustment must be cognizant of the limitations of the appended information (DiSogra et al. 2010; Roth et al. 2013). Such data are not always accurate and are present for only a nonrandom subset of households (which may also be true of auxiliary variables appended to other types of frames).

Even so, sometimes imperfect stratification can make a design much more efficient in terms of cost, if not in statistical efficiency. As noted in Section 4, some ABS studies make use of two-phase samples in which the stratification variables for the second phase sample are collected directly from the households that responded in the first phase.

8.1.3 Frame Multiplicity, Vacant Addresses, and Other Frame Issues

According to Iannacchione (2011), in 2010 there were nearly 14 million residential P.O. Boxes that were not the only means of delivery, virtually all of which correspond to households with a residential address already accounted for in the CDS or DSF. Most researchers exclude these boxes to reduce problems associated with frame multiplicity. P.O. Boxes that are a household's only way to get mail are often included for mail contact (if geographical location is not a concern), but excluded for surveys where precise location is required.

A small but not insignificant portion of mail sent to undeliverable addresses will be returned. In theory, the ABS frames should be relatively up to date because of constant maintenance by mail carriers who are observing the addresses on the ground. However, the return rate may vary by geography.

Returned mail can be accounted for, but mail that is not returned may or may not reach the intended households. Some addresses correspond to vacant housing units, for example, and the mail may be forwarded to a different address (unless the envelope is

marked “do not forward”), returned, or left to accumulate until the vacancy status is known. Alternatively, a household may simply discard the mail. Unless the survey researcher receives a postmaster return (Sections 5.1.1 and 5.2.1), the eligibility status of a nonresponding household is rarely known.

Multiplicity and ineligibility problems are not unique to ABS frames. Field enumerated lists are expected to have a number of vacant housing units that are ineligible for in-person surveys. RDD samples are expected to have telephone numbers that are not associated with working residential lines. RDD samples also have the issue of multiple telephone numbers for many households. The point is that all frames will have some cases of multiplicity or ineligibility, and the informed survey designer will take these shortcomings into account in planning.

8.1.4 Mode Effects

Multimode surveys are often an attempt to compensate for the shortcomings of any single mode. A strength of ABS is its adaptability to multiple modes of data collection. In multimode studies, systematic differences because of mode may be observed and should be examined (Dillman et al. 2009; Dillman et al. 1996). For example, studies have shown a greater likelihood for respondents to give socially desirable responses to sensitive questions in interviewer- administered surveys than in self-administered surveys (Aquilino 1994). Even when social desirability does not seem to be present, telephone respondents tend to give more extreme and more positive responses than web respondents (Christian, Dillman, and Smyth 2008). Furthermore, self-administered surveys often have more missing item responses than interviewer-administered surveys (Biemer and Lyberg 2003). Branching questions are more easily handled in web or other computer-assisted modes than in mail. Mode effects can be difficult to measure and might be confounded with other effects because of the interviewer, the respondent, or the survey content (Voogt and Saris 2005). Although mode differences are not specifically an ABS issue, researchers who move to multimode data collection for the first time with ABS may encounter this issue.

ABS improvements in coverage and multimode improvements in response rates over RDD tend to overshadow concerns about mode differences. Good visual design can usually overcome a significant portion of the missing data problem and mode differences (Messer, Edwards, and Dillman 2012; Millar and Dillman 2011). Some mode differences may remain, however.

8.1.5 Required Infrastructure, Time, and Costs

ABS requires a learning curve to understand and use the frames effectively. Researchers may need to review the literature on the best methodology, designs, and mode(s) for a particular application. In this regard, ABS is no different from other survey approaches. Knowledge is key to successful implementation.

Infrastructure is required to process and manage any type of sample(s), if not the frame,

and some investment is required at first use. In the context of multimode studies, infrastructure is necessary to manage data collection across modes. Depending on the mode, such infrastructure might include mailing, receiving, and scanning hardcopy survey materials, developing programmed instruments for telephone or in-person interviewing, developing and hosting online instruments, handling outbound and inbound telephone calls, and developing case management systems and data processing systems that cover multiple modes of data collection.

In-person, mail, and multimode ABS studies require a relatively long field period, especially when different modes are used sequentially to address nonresponse. ABS might not be appropriate for, say, a quick-turnaround political poll.

Survey costs can vary considerably by mode or by survey house, and costs are rarely published. Section 7.4 discussed some of the cost issues associated with ABS. Although it is not possible to set expectations accurately, some generalizations can be surmised. Switching to ABS for the first time will involve some startup costs. In-person surveys typically have higher costs but also higher response rates than other modes, and using ABS frames for in-person surveys can save considerable costs over field enumeration. Beyond in-person surveys, it is difficult to comment on ABS costs relative to other modes because of the many ways that ABS can be implemented and the lack of published cost information. The fact that ABS has already endured in some form for more than a decade indicates that it has a place in the spectrum of options.

Rushing into ABS without adequate preparation is ill advised, and ABS is not appropriate for all types of surveys. Nonetheless, when designed and administered properly, ABS may be superior in many applications.

8.2 Report Summary

ABS is rapidly becoming an important methodology in survey work. It is incumbent on survey researchers to understand the nuances required for ABS studies. For this reason, this report reviews in Section 2 many technical concepts about the address files and their implications for survey quality. In particular, survey researchers should be aware of the types of addresses available and the potential for duplications and coverage error. Knowing these aspects of the address files can help the researcher work with the ABS vendor to customize the frame optimally for a particular study. Vendors primarily maintain address lists for marketers. Knowing the nature of the vendor's licensing agreement with the USPS, the geographical areas for which the vendor is qualified, the update processes, and the vendor's familiarity with factors that affect surveys can all affect the coverage and quality of the frame and sample. Section 2 includes some specific topics to be discussed with an ABS vendor.

Address frames lend themselves well to mergers with other types of data (Section 3), including geocodes, telephone numbers, government area-level variables, and commercial household- and person-level demographic and market segmentation variables. The possibilities are seemingly endless. The auxiliary data open up many

possibilities for sample design, data collection methodology, and estimation. In particular, auxiliary variables may be useful for stratification, tailored data collection approaches, weighting, or estimation with auxiliary variables. For licensing reasons some variables may be obtained from a vendor only for a sample, not a frame, leading to two-phase sampling. The quality of the auxiliary variables may determine whether they improve the survey. Generally, government data are high quality and can be appended to locatable addresses, but area-level variables may or may not be associated with the specific characteristics of those living at a particular address. Household- and person-level auxiliary variables may not be complete in that they may not be available and adequately matched to all addresses. Furthermore, it is not possible to know the accuracy of all possible variables appended to the addresses, some of which are modeled and some of which are derived from other commercial sources.

Understanding the technical aspects of the frame and the auxiliary variables informs both the sample design and the data collection methodology (Section 4). For example, in most cases auxiliary variables are better suited for stratification than for restricting the available addresses in the frame. Sometimes auxiliary data are collected in a first phase of data collection for more efficient sampling in a second phase. Special consideration is given to surveys of local areas or surveys of nonresidential addresses.

ABS can be used for a single-mode survey by mail or in person, but ABS also lends itself well to mixed-mode surveys. Mail surveys require a reevaluation of instrument design, but a well-planned mail survey usually obtains improved response rates over RDD. Web-push designs have the potential to lower survey costs. Different modes can be applied sequentially, which works better than offering a sample household or person a choice initially. Section 4 summarizes recent research in the effectiveness of various options of single- and mixed-mode ABS surveys.

Section 5 covers the subtleties of eligibility and case dispositions, and how they are used in calculating weights and response rates. Mixed-mode studies in particular complicate the dispositions of sample cases because the sample unit is the address, and nonreturned mail or a disconnected telephone number does not necessarily mean the address is an ineligible sampling unit. Consequently, ABS studies often have higher proportions of cases of unknown eligibility than in-person surveys, which puts greater importance on the assumptions one uses in weighting. The eligibility assumptions for cases of unknown eligibility can differ by mode, which can make the assumptions more complex for a mixed-mode study. The calculation of response rates can also be more complex for a mixed-mode study.

Section 6 builds on other AAPOR and OMB standards by reviewing what should be included in study reports and reframes the existing standards for ABS studies.

Section 7 reviews quality aspects of variables in ABS frames. Of particular interest is a frame's coverage. Several methods have been developed to evaluate and improve the coverage of ABS frames, particularly for in-person surveys. There are cost and quality

tradeoffs in making design decisions with vendors. Although specific costs are proprietary, some generalizations are noted.

Section 8.1 gives some caveats to help set expectations for would-be ABS users, to balance the strengths and promise of ABS. Section 8.3 below summarizes some topics for further research.

8.3 Future Research

Overall, this report summarizes the current state of knowledge about ABS studies. Each section in the report covers a different aspect of ABS studies, and these aspects are highly interrelated. It is important for survey researchers to understand and explore these relationships because there are aspects of ABS that are not yet well understood. ABS studies are in need of additional research to understand the best ways to take advantage of the address lists, even as the vendors of frames expand their offerings and auxiliary data sources become more prevalent. In particular, the report identifies the following areas for future research:

- Coverage of ABS frames has been researched extensively, but coverage is not a static property. Some address types such as P.O. Boxes are useful for mail but not locatable for in-person surveys. The transition to emergency 911 city-style addresses in rural areas continues to improve the coverage of the locatable addresses. But even for mail surveys, some address types such as drop points can affect coverage. For a particular study, the inclusion or exclusion based on address type or on address flags such as those for vacant or seasonal housing can impact the coverage and the efficiency of the sample. Furthermore, vendors can vary in the areas they cover well, and coverage differences between primary and secondary vendors have not been explored. For any particular study, the coverage of the frame will be of interest (Sections 2 and 7.2).
- The expected coverage of an ABS frame helps determine the extent to which the frame can replace traditional enumeration of housing units for in-person surveys. Methods for estimating coverage in advance of sample selection have been developed (Section 7.2), but could be improved. Methods for supplementing coverage in the field have been developed and continue to evolve (Section 7.3).
- Drop points (Section 2.2.2) pose particular problems, and although the nature of the problems is known, the impact is not fully quantified. Drop points will affect some geographical areas more than others. The number of units that really exist at a drop point, relative to the information available in the frame, needs more exploration. Additional methods for working with drop points need to be developed and tested.
- The update patterns of the USPS files and the vendor frames are not well understood (Sections 2.1 and 2.3). The impact of the updates or lack thereof on coverage, efficiency, data quality, and costs have not been researched to any great extent.

- The No-Stat file is known to add some coverage to an address frame, and some guidance exists for which types of No-Stat addresses to include (Sections 2.1.2 and 2.3.7). Even so, more research into the No-Stat file is desirable to take full advantage of it.
- Many types of auxiliary variables can be appended to ABS frames or samples (Section 3). Auxiliary variables hold much promise for sample stratification, customized data collection strategies to reduce nonresponse and nonresponse bias, and possibly use in weighting and estimation, and techniques could be explored. Appropriate use of auxiliary variables depends on their match rate to the frame and their accuracy, but variables with inadequate match rates or accuracy may cause more harm to design effects without the desired cost savings. Additional research is needed to investigate the wide range of auxiliary variables available today. Variables from commercial sources, in particular, need more scrutiny.
- Two-phase data collection efforts have the potential to increase or decrease response rates. More investigation of two-phase approaches would help determine the best way to execute such designs (Sections 4.2 and 4.3).
- To date, relatively little research has been conducted on the use of ABS for business surveys (Section 4.7). Mailing addresses are likely less useful for reaching businesses than for households.
- The use of ABS has contributed to a rise in mixed-mode survey designs (Section 4.8.3 and 4.8.4). Mode effects are likely to be encountered more often as a result, and the impact of mode effects should be explored more fully. Costs of mixed-mode data collection options should be explored relative to changes in response rates and nonresponse bias. Mail and “web-push” surveys need to be refined to optimize response rates and reduce nonresponse bias. Mixed-mode surveys with sequential modes of collection lead to exploration of the impact of multiple incentives offered (Section 4.9). Other techniques for improving response rates in mixed-mode surveys should be explored as well.
- For mail surveys, little is known about rates of ineligible addresses and postmaster returns (Sections 5.1.1 and 5.2.1). Like RDD, mail surveys have higher rates of units with unknown eligibility status compared to in-person surveys, so the assumptions about the proportion of eligible units (e) among the unknown units can have a larger impact on response rates and weights.
- In general, cost and quality tradeoffs of many types impact the effectiveness of ABS, and the tradeoffs should be explored and documented (Section 7.4).

The authors sincerely hope that this report will lead to new research aimed at understanding the tradeoffs of ABS design decisions, keeping survey researchers at the forefront of survey design and methodology.

SECTION 9 REFERENCES

AAPOR Cell Phone Task Force. 2010. *New considerations for survey researchers when planning and conducting RDD telephone surveys in the U.S. with respondents reached via cell phone numbers*. Prepared for the AAPOR Council under the auspices of the AAPOR Standards Committee. Deerfield, IL: Author.

Alexander, Charles H. and Signe Wetrogan. 2000. "Integrating the American Community Survey and the Intercensal Demographic Estimates Program." Pp. 295-300 in *Proceedings of the Survey Research Methods Section*: American Statistical Association.

Alkire, Elise. 2010. "Handheld data collection and its effects on mapping." in *A Special Joint Symposium of ISPRS Technical Commission IV & AutoCarto in conjunction with ASPRS/CaGIS 2010 Fall Specialty Conference*. Orlando, FL.

Amaya, Ashley, Felicia LeClere, Lee Fiorio, and Ned English. 2014. "Improving the utility of the DSF address-based frame through ancillary information." *Field Methods* 26:70-86.

Amaya, Ashley, Ben Skalland, and Karen Wooten. 2010. "What's in a match?" *Survey Practice* 3.

American Association for Public Opinion Research (AAPOR). 2015. *Standard definitions: Final dispositions of case codes and outcome rates for surveys*. Deerfield, IL: Author.

Aquilino, William S. 1994. "Interview mode effects in surveys of drug and alcohol use: a field experiment." *Public Opinion Quarterly* 58:210-40.

Bailey, Justin T., Gretchen Grabowski, and Michael W. Link. 2010. "Your home was specially selected: Using address-based sampling as a recruitment technique." Pp. 5938-5948 in *JSM Proceedings, Survey Research Methods Section, American Statistical Association*. Vancouver, British Columbia.

Battaglia, Michael P., Michael W. Link, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2008. "An evaluation of respondent selection methods for household mail surveys." *Public Opinion Quarterly* 72:459-469.

Biemer, Paul P. and Lars E. Lyberg. 2003. *Introduction to survey quality*. New York: John Wiley & Sons, Inc.

Biemer, Paul P. and Andy Peytchev. 2012. "Census geocoding for nonresponse bias evaluation in telephone surveys: An assessment of the error properties." *Public Opinion Quarterly* 76:432-452.

—. 2013. "Using geocoded census data for nonresponse bias correction: An assessment." *Journal of Survey Statistical Methodology* 1:24-44.

Blumberg, Stephen J. and Julian V. Luke. 2007. "Coverage bias in traditional telephone surveys of low-income and young adults." *Public Opinion Quarterly* 71:734-49.

—. 2014, "Wireless substitution: Early release of estimates from the National Health Interview Survey, July – December 2013", Retrieved October 13, 2015, (<http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201407.pdf>).

Blumberg, Stephen J., Julian V. Luke, and Marcie L. Cynamon. 2006. "Telephone coverage and health survey estimates: Evaluating the need for concern about wireless substitution." *American Journal of Public Health* 96:926-31.

Blumberg, Stephen J., Julian V. Luke, Marcie L. Cynamon, and Martin R. Frankel. 2008. "Recent trends in household telephone coverage in the United States." Pp. 56-86 in *Advances in Telephone Survey Methodology*, edited by J. M. Lepkowski, C. Tucker, J. M. Brick, et al. New York: John Wiley and Sons, Inc.

Bonner, Matthew R., Daikwon Han, Jing Nie, Peter Rogerson, John E. Vena, and Jo Freudenheim. 2003. "Positional accuracy of geocoded addresses in epidemiologic research." *Epidemiology* 14:408-12.

Brick, J. Michael, William R. Andrews, and Nancy A. Mathiowetz. 2012, "A comparison of recreational fishing effort survey designs", Retrieved October 14, 2015, https://www.st.nmfs.noaa.gov/mdms/doc/08A_Comparison_of_Fishing_Effort_Surveys_Report_FINAL.pdf

Brick, J. Michael, Pat Dean Brick, Sarah Dipko, Stanley Presser, Clyde Tucker, and Yangyang Yuan. 2007. "Cell phone survey feasibility in the U.S.: sampling and calling cell numbers versus landline numbers." *Public Opinion Quarterly* 71:23-39.

Brick, J. Michael, Sharon Lohr, W. Sherman Edwards, Pamela Giambo, Pam Broene, Douglas Williams, and Sarah Dipko. 2013. *National Survey of Crime Victimization Companion Study-Summary of pilot results*. Prepared for U.S. Bureau of Justice Statistics. Rockville, MD: Westat.

Brick, J. Michael, Jill Montaquila, Mary Collins Hagedorn, Shelley Brock Roth, and Christopher Chapman. 2005. "Implications for RDD design from an incentive experiment." *Journal of Official Statistics* 21:571-589.

Brick, J. Michael, Douglas Williams, and Jill M. Montaquila. 2011. "Address-based sampling for subpopulation surveys." *Public Opinion Quarterly* 75:409-428.

Burks, Anu Thu and Trent D. Buskirk. 2012. "Can response propensities grow on trees? Exploring response propensity models based on random forests using ancillary

data appended to an ABS sampling frame." in *2012 Midwest Association of Public Opinion Research*. Chicago, IL.

Buskirk, Trent D. and Stanislav Kolenikov. 2015. "Finding respondents in the forest: A comparison of logistic regression and random forest models for response propensity weighting and stratification." in *Survey Methods: Insights from the Field*.

Buskirk, Trent D., David Malarek, and Jeffrey S. Bareham. 2014. "From flagging a sample to framing it: Exploring vendor data that can be appended to ABS samples." Pp. 111-124 in *Proceedings of the Survey Research Methods Section*: American Statistical Association.

Buskirk, Trent D., Brady West, and Anh Thu Burks. 2013. "Respondents: who art thou? Comparing internal, temporal, and external validity of survey response propensity models based on random forests and logistic regression models." in *2013 Joint Statistical Meetings of the American Statistical Association*. Montreal, Canada.

Cantor, David, Barbara C. O'Hare, and Kathleen S. O'Connor. 2007. "The use of monetary incentives to reduce nonresponse in random digit dial telephone surveys." in *Advances in Telephone Survey Methodology*, edited by J. M. Lepkowski, C. Tucker, J. M. Brick, et al. Hoboken, NJ: John Wiley & Sons, Inc.

carrierroutes.com. "Homepage", Retrieved October 14, 2015, (<http://www.carrierroutes.com/>).

Cayo, Michael R. and Thomas O. Talbot. 2003. "Positional error in automated geocoding of residential addresses." *International Journal of Health Geographics* 2:10.

Center for Behavioral Health Statistics and Quality. 2014. *National Survey on Drug Use And Health: Summary of Methodological Studies, 1971–2014*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Christian, Leah Melani, Don A. Dillman, and Jolene D. Smyth. 2008. "The Effects of Mode and Format On Answers to Scalar Questions in Telephone and Web Surveys." Pp. 250-275 in *Advances in Telephone Survey Methodology*, edited by J. M. Lepkowski, C. Tucker, J. M. Brick, et al. New York: Wiley-Interscience.

Church, Allan H. 1993. "Estimating the effect of incentives on mail survey response rates: A meta-analysis." *Public Opinion Quarterly* 57:62-79.

Cochran, William G. 1977. *Sampling Techniques*. New York: John Wiley & Sons.

Council of American Survey Research Organizations. n.d., "The Voice and Values of Research. Preface", Retrieved October 13, 2015, (http://c.ymcdn.com/sites/www.casro.org/resource/resmgr/docs/casro_on_definitions_of_resp.pdf).

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2005. "Changes in telephone survey nonresponse over the past quarter century." *Public Opinion Quarterly* 69:87-98.

de Leeuw, Edith D. 2005. "To mix or not to mix data collection modes in surveys." *Journal of Official Statistics* 21:233-255.

Dekker, Katie, Ashley Amaya, Felicia LeClere, and Ned English. 2012. "Unpacking the DSF in an attempt to better reach the drop point population." Pp. 4596-4604 in *Proceedings of the Section on Survey Research Methods*: American Statistical Association.

Dekker, Katie and Whitney Murphy. 2014. "Mailing to drop points in a multi-mode survey: Using the nostat file to supplement unit information." in *Paper presented at the American Association for Public Opinion Research Annual Conference*. Anaheim, CA.

Dillman, Don A., Glenn Phelps, Robert Tortora, Karen Swift, Julie Kohrell, and Jodi Berck. 2009. "Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response, and the Internet." *Social Science Research* 38:1-18.

Dillman, Don A., Roberta L. Sangster, John Tarnai, and Todd H. Rockwood. 1996. "Understanding differences in people's answers to telephone and mail surveys." Pp. 45-62 in *Current issues in survey research, new directions for program evaluation series*, edited by M. T. Braverman and J. K. Slater. San Francisco: Jossey-Bass.

Dillman, Don A., Jolene D. Smyth, and Leah M. Christian. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. New York: Wiley.

DiSogra, Charles, J. Michael Dennis, and Mansour Fahimi. 2010. "On the quality of ancillary data available for address-based sampling." Pp. 4174-4183 in *Proceedings of the Survey Research Methods Section*: American Statistical Association.

Dohrmann, Sylvia, Trent D. Buskirk, Ashley Hyon, and Jill Montaquila. 2014. "Address-based sampling frames for beginners." Pp. 1009-1018 in *JSM Proceedings, Survey Research Methods Section*. Alexandria, VA: American Statistical Association.

Dohrmann, Sylvia, Daifeng Han, and Leyla Mohadjer. 2006. "Residential address lists vs. traditional listing: Enumerating households and group quarters." Pp. 2959-2964 in *Proceedings of the Joint Statistical Meetings, Survey Research Methods Section*.

—. 2007. "Improving coverage of residential address lists in multistage area samples." Pp. 3219-3126 in *Proceedings of the American Statistical Association, on Survey Research Methods*: American Statistical Association.

Dohrmann, Sylvia, Graham Kalton, Jill Montaquila, Cindy Good, and Martha Berlin.

2012. "Using address-based sampling frames in lieu of traditional listing: A new approach." Pp. 3729-3741 in *Joint Statistical Meetings, Survey Research Methods Section*.

Dohrmann, Sylvia, Lin Li, and Leyla Mohadjer. 2011. "Updating the measures of size of local areas late in the decade using USPS address lists." Pp. 2891-2901 in *Proceedings of the Survey Research Methods of the American Statistical Association*.

Dohrmann, Sylvia and Richard Sigman. 2013. "Using an area linkage method to improve the coverage of abs frames for in-person household surveys." in *Proceedings of Federal Committee on Statistical Methodology Research Conference*.

Eckman, Stephanie. 2010. "Errors in housing unit listing and their effects on survey estimates. Joint Program in Survey Methodology (Doctoral dissertation)." Dissertation Thesis, Graduate School of the University of Maryland.

Eckman, Stephanie and Ned English. 2012a. "Creating housing unit frames from address databases: Geocoding precision and net coverage rates." *Field Methods* 24:399-408.

—. 2012b. "Geocoding to create survey frames." *Survey Practice* 5.

Eckman, Stephanie and Frauke Kreuter. 2011. "Confirmation bias in housing unit listing." *Public Opinion Quarterly* 75:139-150.

—. 2013. "Undercoverage rates and undercoverage bias in traditional housing unit listing." *Sociological Methods and Research* 42:264-293.

Eckman, Stephanie and Colm O'Muircheartaigh. 2011. "Performance of the half-open interval missed housing unit procedure." *Survey Research Methods* 5:125-131.

Edwards, Michelle L., Don A. Dillman, and Jolene D. Smyth. 2014. "An experimental test of the effects of survey sponsorship on internet and mail survey response." *Public Opinion Quarterly* 78:734-750.

English, Ned, Yung Li, Andrea Mayfield, and Alicia Frasier. 2014. "The use of targeted lists to enhance sampling efficiency in address-based sample designs: Age, race, and other qualities." in *2014 Proceedings of the American Statistical Association, Survey Research Methods [CD ROM]*. Alexandria, VA: American Statistical Association.

English, Ned, Colm O'Muircheartaigh, Stephanie Eckman, Katie Dekker, and Michael Latterner. 2009. "Coverage Rates and Coverage Bias in Housing Unit Frames." in *2009 Proceedings of the American Statistical Association, Survey Research Methods Section [CD ROM]*. Alexandria, VA: American Statistical Association.

Fahimi, Mansour and Dale Kulp. 2009. *Address-based sampling-Alternatives for*

surveys that require contacts with representative samples of households: Quirk's Marketing Research Review.

Finamore, John and Don A. Dillman. 2013. "An experimental evaluation of how mode sequence for offering internet, mail and telephone options affects responses to a national survey of college graduates." in *European Survey Research Association Conference 2013*. Ljubljana: ESRA.

Fiorio, Lee and Jizhou Fu. 2012. "Modeling coverage error in address lists due to geocoding error: The Impact on survey operations and sampling." Pp. 5588-5596 in *Joint Statistical Meetings, Survey Research Methods Section*.

Gelman, Andrew, Matt Stevens, and Valerie Chan. 2003. "Regression modeling and meta- analysis for decision making: A cost-benefit analysis of incentives in telephone surveys." *Journal of Business & Economic Statistics* 21:213-225.

Goldberg, Daniel W., John P. Wilson, and Craig A. Knoblock. 2007. "From text to geographic coordinates: The current state of geocoding." *URISA Journal* 19:33-46.

Griffin, Joan M., Alisha Baines Simon, Erin Hulbert, John Stevenson, Joseph P. Grill, Siamak Noorbaloochi, and Melissa R. Partin. 2011. "A comparison of small monetary incentives to convert survey non-respondents: a randomized control trial." *BMC Medical Research Methodology* 11:81.

Groves, Robert M., Mick P. Couper, Stanley Presser, Eleanor Singer, Roger Tourangeau, Giorgina Piani Acosta, and Lindsay Nelson. 2006. "Experiments in producing nonresponse bias." *Public Opinion Quarterly* 70:720-736.

Groves, Robert M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology*. Hoboken, NJ: John Wiley & Sons.

Han, Daifeng, David Cantor, Pat Dean Brick, Richard Sigman, and Maribel Aponte. 2010. "Findings from a Two Phase Mail Survey for a Study of Veterans." in *Paper presented at the 65th Annual Meeting of the American Association for Public Opinion Research*. Chicago.

Han, Daifeng, Jill M. Montaquila, and J. Michael Brick. 2013. "An evaluation of incentive experiments in a two-phase address-based sample mail survey." *Survey Research Methods* 7:207-218.

Harter, Rachel and Joseph P. McMichael. 2013. "Scope and coverage of landline and cell phone numbers appended to address frames." Pp. 3651-3665 in *JSM Proceedings, Survey Research Methods Section*. Alexandria: American Statistical Association.

Iannacchione, Vincent G. 2011. "The changing role of address-based sampling in

survey research." *Public Opinion Quarterly* 75:556-575.

Iannacchione, Vincent G., Joseph P. McMichael, Bonnie E. Shook-Sa, and Katherine B. Morton. 2012. "A proposed hybrid sampling frame for the National Survey on Drug Use and Health." Prepared for the Substance Abuse and Mental Health Services Administration, Office of Applied Studies, under Contract No. 283-2004-00022, RTI/0209009, Rockville, MD.

Iannacchione, Vincent G., Jennifer M. Staab, and David T. Redden. 2003. "Evaluating the use of residential mailing lists in a metropolitan household survey." *Public Opinion Quarterly* 67:202-210.

Iannacchione, Vincent, Katherine Morton, Joseph McMichael, Bonnie Shook-Sa, Jamie Ridenhour, Stephanie Stolzenberg, David Bergeron, James Chromy, and Arthur Hughes. 2010. "The best of both worlds: a sampling frame based on address-based sampling and field enumeration." in *Proceedings of the Survey Research Methods Section*: American Statistical Association.

JHSPHOPEN Courseware. 2011, "Issues in Survey Research Design. Lecture Materials", Retrieved October 20, 2015, (<http://ocw.jhsph.edu/index.cfm/go/viewCourse/course/SurveyResearchDesign/coursePage/lectureNotes/>).

Johnson, Clifford L., Sylvia M. Dohrmann, Vicki L. Burt, and Leyla K. Mohadjer. 2014. "National Health and Nutrition Examination Survey: Sample Design, 2011–2014." *National Center for Health Statistics* 2.

Juster, F. Thomas and Richard Suzman. 1995. "An overview of the Health and Retirement Study." *Journal of Human Resources* 30:S7-S56.

Kalton, Graham. 2010. "Developing Linkage Rules to Reduce Noncoverage." in *Presented at the AAPOR 65th Annual Conference*. Chicago, IL.

Kalton, Graham, Jennifer Kali, and Richard Sigman. 2014. "Handling frame problems when address-based sampling is used for in-person household surveys." *Journal of Survey Statistics and Methodology* 2:1-22.

Keeter, Scott, Courtney Kennedy, April Clark, Trevor Thompson, and Mike Mokrzycki. 2007. "What's missing from national landline RDD Surveys? The impact of the growing cell- only population." *Public Opinion Quarterly* 71:772-792.

Kennel, Timothy. 2012. "Evaluation of the Delivery Sequence File as a Survey Frame." edited by Ruth Ann Killion to Nancy Potok: Internal Census Bureau Memo.

Kennel, Timothy L. and Mei Li. 2009. "Content and coverage quality of a commercial

address list as a national sampling frame for household surveys." in *Proceeding of the Joint Statistical Meetings*.

Kish, Leslie. 1965. *Survey Sampling*. New York: John Wiley & Sons.

Kolenikov, Stanislav, Heather Hammer, Charles DiSogra, Rachel Martonik, David Finkelhor, and Heather Turner. 2013. "Adaptive design for ABS in a national CATI survey of households with children." in *Paper presented at MAPOR annual conference*. Chicago, IL.

Kwiat, Aliza. 2009. "Examining blocks with lister error in area listing." in *Proceedings of Survey Research Methods*: American Statistical Association.

Lepkowski, James M., William D. Mosher, Karen E. Davis, Robert M. Groves, and John Van Hoewyk. 2010. "The 2006-2010 National Survey of Family Growth: Sample design and analysis of a continuous survey." U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, and National Center for Health Statistics, Hyattsville, MD.

Lepkowski, James M., Clyde Tucker, J. Michael Brick, Edith D. De Leeuw, Lilli Japiec, Paul J. Lavrakas, Michael W. Link, and Roberta L. Sangster. 2007. *Advances in Telephone Survey Methodology*. Hoboken, NJ: John Wiley & Sons.

Lesser, Virginia M., Don A. Dillman, John Carlson, Frederick Lorenz, Robert Mason, and Fern Willits. 2001. "Quantifying the influence of incentives on mail survey response rates and their effects on nonresponse error." in *Proceedings of the Annual Meeting of the American Statistical Association*.

Link, Michael, Ali Mokdad, Dale Kulp, and Ashley Hyon. 2006. "Has the national do not call registry helped or hurt survey research efforts?" *Public Opinion Quarterly* 70:794-805.

Link, Michael W. 2010, "Address-based sampling: What do we know so far?", Retrieved October 14, 2015, (<http://www.amstat.org/sections/SRMS/AddressBasedSampling11-29-2010.pdf>).

Link, Michael W., Michael P. Battaglia, Martin R. Frankel, Larry Osborn, and Ali H. Mokdad. 2008. "A comparison of address-based sampling (ABS) versus random-digit dialing (RDD) for general population surveys." *Public Opinion Quarterly* 72:6-27.

Link, Michael W. and Anh Thu Burks. 2013. "Leveraging auxiliary data, differential incentives, and survey mode to target hard-to-reach groups in an address-based sample design." *Public Opinion Quarterly* 77:696-713.

Link, Michael W. and Jennie W. Lai. 2011. "Cell-phone-only households and problems of differential non-response using an address-based sampling design." *Public Opinion*

Quarterly 75:613-635.

Little, Roderick J. and Sonja Vartivarian. 2005. "Does weighting for nonresponse increase the variance of survey means?" *Survey Methodology* 31:161-168.

Lohr, Sharon L. 2010. *Sampling: design and analysis*. Boston: Brooks/Cole.

Loudermilk, Clifford and Timothy Kennel. 2005. "Deciphering the DSF: Which addresses from the Delivery Sequence File should be included in the sampling frames for demographic surveys?" in *Proceedings of the Joint Statistical Meeting*.

Lupia, Arthur, Jon A. Krosnick, Patricia Luevano, Matthew DeBell, and Darrell Donakowski. 2009. *User's guide to the ANES 2008 Time Series Study*. Ann Arbor, MI and Palo Alto, CA: University of Michigan and Stanford University.

Martin, Joel M. and Clifford L. Loudermilk. 2008. "Assessing the filter rules for extracting addresses from the master address file to construct a housing unit frame for current demographic surveys." Pp. 1359-1366 in *Proceedings of the Joint Statistical Meeting*.

Mathiowetz, Nancy A., Jesse M. Brick, Lynne Stokes, Rob Andrews, and Seth Muzzy. 2010. "A Pilot Test of a Dual-Frame Mail Survey as an Alternative to an RDD Survey." in *Paper presented at the Joint Statistical Meetings*. Vancouver, BC.

McElroy, Jane A., Patrick L. Remington, Amy Trentham-Dietz, Stephanie A. Robert, and Polly A. Newcomb. 2003. "Geocoding addresses from a large population-based study: Lessons learned." *Epidemiology* 14:399-407.

McMichael, Joseph P., Jamie L. Ridenhour, Susan Mitchell, Kristine Fahrney, and Wanda Stephenson. 2008a. "Evaluating the use and effectiveness of the half-open interval procedure for sampling frames based on mailing address lists in urban areas." in *Proceedings of the American Statistical Association, Section on Survey Research Methods*: American Statistical Association.

McMichael, Joseph P., Jamie L. Ridenhour, and Bonnie E. Shook-Sa. 2008b. "A robust procedure to supplement the coverage of address-based sampling frames for household surveys." Pp. 4329-35 in *Proceedings of the American Statistical Association, Survey Research Methods*.

McMichael, Joseph P. and David Roe. 2012. "ABS and cell phones: Appending both cell phone and landline phone numbers to an address-based sampling frame." in *American Association for Public Opinion Research (AAPOR) Annual Conference*. Orlando, FL.

McMichael, Joseph, Bonnie Shook-Sa, Jamie Ridenhour, and Rachel Harter. 2014. "The CHUM: A Frame Supplementation Procedure for Address-Based Sampling." in *2013 Federal Committee on Statistical Methodology (FCSM) Research Conference*.

McMichael, Joseph. P., Rachel M. Harter, Bonnie E. Shook-Sa, Vincent G. Iannacchione, Jamie L. Ridenhour, and Kibri Hutchison-Everett. 2012. "Sub-national coverage profile of U.S. housing units using an address-based sampling frame." in *2012 Joint Statistical Meetings of the American Statistical Association*. San Diego, CA: American Statistical Association.

Medway, Rebecca L. and Jenna Fulton. 2012. "When more gets you less: A meta-analysis of the effect of concurrent web options on mail survey response rates." *Public Opinion Quarterly* 76:733-746.

Messer, Benjamin L. 2012. "Pushing households to the web: Experiments of a 'Web+Mail' methodology for conducting general public surveys. (Doctoral dissertation)." Department of Sociology, Washington State University.

Messer, Benjamin L. and Don A. Dillman. 2011. "Surveying the general public over the internet using address-based sampling and mail contact procedures." *Public Opinion Quarterly* 75:429-457.

Messer, Benjamin L., Michelle L. Edwards, and Don A. Dillman. 2012. "Determinants of item nonresponse to web and mail respondents in three address-based mixed-mode surveys of the general public." *Survey Practice* 5.

Millar, Morgan M. and Don A. Dillman. 2011. "Improving response to web and mixed-mode surveys." *Public Opinion Quarterly* 75:249-269.

Montaquila, Jill M. 2014. "Use of vendor data in optimization of address-based sampling procedures: Discussion." in *Presented at the 2014 Joint Statistical Meetings*. Boston, MA.

Montaquila, Jill M., J. Michael Brick, Douglas Williams, Kwang Kim, and Daifeng Han. 2013. "A study of two-phase mail survey data collection methods." *Journal of Survey Statistics and Methodology* 1:66-87.

Montaquila, Jill M., Valerie Hsu, and J. Michael Brick. 2011. "Using a "match rate" model to predict areas where USPS-based address lists may be used in place of traditional listing." *Public Opinion Quarterly* 75:317-335.

Montaquila, Jill M., Valerie Hsu, J. Michael Brick, Ned English, and Colm O'Muircheartaigh. 2009. "A comparative evaluation of traditional listing vs. address-based sampling frames: Matching with field investigation of discrepancies." in *Proceedings of the Survey Research Methods Section of the American Statistical Association*. Alexandria, VA: American Statistical Association.

O'Muircheartaigh, Colm A., Stephanie Eckman, and Charlene Weiss. 2003. "Traditional and Enhanced Field Listing for Probability Sampling." in *Proceedings of the Survey*

Research Methods Section: American Statistical Association.

O'Brien, Eileen M. 2013. "A tale of two surveys: Learning from the application of address-based sampling in federal surveys." in *Federal Committee on Statistical Methodology Conference*. Washington, DC.

O'Muircheartaigh, Colm, Stephanie Eckman, and Charlene Weiss. 2002. "Traditional and enhanced field listing for probability sampling." Pp. 2563-2567 in *Proceedings of the Social Statistics of the American Statistical Association*.

O'Muircheartaigh, Colm, Ned English, and Stephanie Eckman. 2007. "Predicting the Relative Quality of Alternative Sampling Frames." in *2007 Proceedings of the American Statistical Association, Survey Research Methods Section [CD ROM]*. Alexandria, VA: American Statistical Association.

O'Muircheartaigh, Colm, Ned English, Stephanie Eckman, Heidi Upchurch, Erika Garcia Lopez, and James Lepkowski. 2006. "Validating a sampling revolution: Benchmarking address lists against traditional field listing." in *2006 Proceedings of the American Statistical Association, AAPOR Survey Research Methods Section [CD ROM]*. Alexandria, VA: American Statistical Association.

O'Muircheartaigh, Colm, Ned English, Michael Latterner, Stephanie Eckman, and Katie Dekker. 2009. "Modeling the need for traditional vs. commercially-available address listings for in-person surveys: Results from a national validation of addresses." in *2009 Proceedings of the American Statistical Association, AAPOR Survey Research Methods Section [CD ROM]*. Alexandria, VA: American Statistical Association.

Olson, Richard and Trent D. Buskirk. 2015. "Can I get your phone number? Examining the relationship between household, geographic and census-related variables and phone append propensity for ABS samples." in *70th Annual AAPOR Conference*. Hollywood, FL.

Pasek, Josh, S. Mo Jang, Curtiss L. Cobb, J. Michael Dennis, and Charles DiSogra. 2014. "Can marketing data aid survey research? Examining accuracy and completeness in consumer- file data." *Public Opinion Quarterly* 78:889-916.

Pedrazzani, Sue, Joe McMichael, Matthew Strobl, Gina Kilpatrick, Julie Feldman, and Neal Halfon. 2012. "Improving coverage of an address-based sampling frame for the National Children's study, Los Angeles County." in *American Public Health Association Conference*. San Francisco, CA.

Petrolia, Daniel R. and Sanjoy Bhattacharjee. 2009. "Revisiting incentive effects: Evidence from a random-sample mail survey on consumer preferences for fuel ethanol." *Public Opinion Quarterly* 73:537-550.

Pew Research Center. 2012, "Assessing the Representativeness of Public Opinion

Surveys", Retrieved October 14, 2015, (<http://www.people-press.org/files/legacy-pdf/Assessing%20the%20Representativeness%20of%20Public%20Opinion%20Survey%20s.pdf>).

Redline, Cleo D., Don A. Dillman, Araf Dajani, and Mary Ann Scaggs. 2003. "Improving navigational performance in U.S. Census 2000 by altering the visual languages of branching instructions." *Journal of Official Statistics* 19:403-420.

Ridenhour, Jamie L., Joseph P. McMichael, Rachel Harter, and Jill A. Dever. 2014. "ABS and demographic flags: Examining the implications for using auxiliary frame information." in *Joint Statistical Meetings*. Boston, MA.

Roth, Shelley B., Daifeng Han, and Jill M. Montaquila. 2013. "The ABS frame: Quality and considerations." *Survey Practice* 6:3779-3793.

Schwartz, Barry. 2004. *The paradox of choice: Why more is less*. New York: Harper Perennial. Shook-Sa, Bonnie E. 2014. "Improving the efficiency of address-based sampling frames with the USPS No-Stat File." *Survey Practice* 7:1-10.

Shook-Sa, Bonnie E. and Douglas Currivan. 2011. "Supplementing address-based sampling frames with physical addresses of housing units with unlocatable mailing addresses." Pp. 5798-5805 in *Proceedings of the Survey Research Methods Section*: American Statistical Association.

Shook-Sa, Bonnie E., Douglas B. Currivan, Joseph P. McMichael, and Vincent G. Iannacchione. 2013. "Extending the coverage of address-based sampling frames beyond the USPS computerized delivery sequence file." *Public Opinion Quarterly* 77:994-1005.

Singer, Eleanor, Mick P. Couper, Angela Fagerlin, Floyd J. Fowler, Carrie A. Levin, Peter A. Ubel, John Van Hoewyk, and Brian J. Zikmund-Fisher. 2014. "The role of perceived benefits and costs in patients' medical decisions." *Health Expectations* 17:4-14.

Singer, Eleanor and Cong Ye. 2013. "The use and effects of incentives in surveys." *Annals of the American Academy of Political and Social Science* 645:112-141.

Smith, Philip. J., David C. Hoaglin, J. N. K. Rao, Michael P. Battaglia, and Danni Daniels. 2004. "Evaluation of adjustments for partial nonresponse bias in the National Immunization Survey." *Journal of the Royal Statistical Society, Series A* 167:1-16.

Smith, Tom W. 2009. *A revised review of methods to estimate the status of cases with unknown eligibility*. Chicago: NORC at the University of Chicago.

Smyth, Jolene D., Don A. Dillman, Leah M. Christian, and Allison C. O'Neill. 2010. "Using the internet to survey small towns and communities: Limitations and possibilities in the Early 21st century." *American Behavioral Scientist* 53:1423-1448.

Staab, Jennifer M. and Vincent G. Iannacchione. 2003. "Evaluating the use of residential mailing addresses in a national household survey." Pp. 4028-4033 in *2003 Joint Statistical Meetings - Section on Survey Research Methods: Survey Research Methods of the American Statistical Association*.

Statistics Canada. 2002, "Maps and Mapping/Geographic Information Systems (GIS): Block- face", Retrieved October 14, 2015, (<http://www12.statcan.ca/English/census01/products/reference/dict/geo003.htm>).

Tomaszewski, Christine G. and Kevin Shaw. 2013, "2010 Census Evaluation of Address List Maintenance Using Supplemental Data Sources", Retrieved September 29, 2015, (https://www.census.gov/2010census/pdf/2010_Census_Evaluation_of_Address_Listing_Maintenance_Using_Supplemental_Data_Sources.pdf).

Tversky, Amos and Eldar Shafir. 1992. "Choice under conflict: the dynamics of deferred decision." *Psychological Science* 3:359-361.

U.S. Census Bureau. 2014, "2013 American Community Survey and Puerto Rico Community Survey 2014 Subject Definitions", Retrieved October 14, 2015, (http://www2.census.gov/programs-surveys/acs/tech_docs/subject_definitions/2014_ACSSubjectDefinitions.pdf).

—. n.d.-a, "Geographic Terms and Concepts - Census Tract", Retrieved October 20, 2015, (https://www.census.gov/geo/reference/gtc/gtc_ct.html).

—. n.d.-b, "Geography. ZIP Code™ Tabulation Areas (ZCTAs™)", Retrieved October 13, 2015, (<https://www.census.gov/geo/reference/pdfs/GARM/Ch11GARM.pdf>).

U.S. Congress. 1994. "Nondisclosure of lists of names and addresses." vol. 39.

U.S. Office of Management and Budget. 2006, "Standards and Guidelines for Statistical Surveys", Retrieved September 29, 2015, (www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf).

U.S. Postal Service. 2004, "National Change of Address. Frequently Asked Questions", Retrieved October 20, 2015, (<http://www.nationalchangeofaddress.com/FAQs.html>).

—. 2008, "Intelligent Mail® Barcode Questions & Answers", Retrieved October 20, 2015, (https://ribbs.usps.gov/onecode_solution/documents/tech_guides/USPSIMB_QandA.pdf).

—. 2013a, "CDS User Guide", Retrieved September 29, 2015,

(http://ribbs.usps.gov/cds/documents/tech_guides/CDS_USER_GUIDE.PDF).

—. 2013b, "Highway Contract Routes-Contract Delivery Service. Handbook SP-1", Retrieved September 29, 2015, (<http://about.usps.com/handbooks/sp1.pdf>).

—. 2013c, "Postal Terms", Retrieved September 28, 2015, (https://about.usps.com/publications/pub32/pub32_terms.htm).

—. 2014a, "2014 Annual Report to Congress", Retrieved September 29, 2015, (<https://about.usps.com/publications/annual-report-comprehensive-statement-2014/annual-report-comprehensive-statement-2014.pdf>).

—. 2014b, "LACS Link(R)", Retrieved October 20, 2015, (<https://ribbs.usps.gov/index.cfm?page=lacslink>).

—. 2015a, "508 Recipient Services", Retrieved October 20, 2015, (<http://pe.usps.gov/text/dmm300/508.htm>).

—. 2015b, "Address Information System Products Technical Guide", Retrieved September 29, 2015, (https://ribbs.usps.gov/addressing/documents/tech_guides/pubs/AIS.PDF).

—. 2015c, "DSF2® License Agreement. Version 19", Retrieved October 13, 2015, (https://ribbs.usps.gov/dsf2/documents/tech_guides/DSF2LICA.PDF).

—. n.d.-a, "CDS Brochure", Retrieved September 28, 2015, (https://ribbs.usps.gov/cds/documents/tech_guides/CDSBrochure.pdf).

—. n.d.-b, "U.S. Postal Service Standard Mailboxes, Curbside", Retrieved October 20, 2015, (<https://about.usps.com/publications/engineering-standards-specifications/spusps-std-7b01/welcome.html>).

Valassis Lists, Inc. 2011, "Understanding Saturation Mail Terminology", Retrieved October 14, 2015, (<http://www.valassislists.com/glossary.php>).

Valliant, Richard, Jill A. Dever, and Frauke Kreuter. 2013. *Practical Tools for Designing and Weighting Survey Samples*. New York: Springer.

Valliant, Richard, Frost Hubbard, Sunghee Lee, and Chiungwen Chang. 2014. "Efficient use of commercial lists in U.S. household sampling." *Journal of Survey Statistics and Methodology* 2:182-209.

Voogt, Robert J. J. and Willem E. Saris. 2005. "Mixed mode designs: Finding the balance between nonresponse bias and mode effects." *Journal of Official Statistics* 21:367-387.

Ward, Mary H., John. R. Nuckols, Jim Giglierano, Matthew R. Bonner, Calvin F. Wolter, Matthew Airola, Wende Mix, Joanne S. Colt, and Patricia Hartge. 2005. "Positional accuracy of two methods of geocoding." *Epidemiology* 16:542-7.

WhiteHouse.gov. 2006, "Office of Management and Budget Standards and Guidelines for Statistical Surveys", Retrieved October 14, 2015, (http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf).

Williams, Douglas, J. Michael Brick, Jill M. Montaquila, and Daifeng Han. 2014. "Effects of screening questionnaires on response in a two-phase postal survey." *International Journal of Social Research Methodology*:1-17.

Ying, Star. 2012, "Identifying excluded from delivery statistics records that elude the American Community Survey housing unit frame filters. Final Report", Retrieved September 28, 2015, (http://www.census.gov/content/dam/Census/library/working-papers/2012/acs/2012_Ying_01.pdf).

Yu, Julie and Harris Cooper. 1983. "A Quantitative Review of Research Design Effects on Response Rates to Questionnaires." *Journal of Marketing Research* 20:36-44.

Zandbergen, Paul A. 2008. "A comparison of address point, parcel, and street geocoding techniques." *Computers, Environment and Urban Systems* 32:214-232.

APPENDIX A

DEFINITIONS

Address-based sampling requires some terminology associated with the address frames and associated methods. Although many terms and definitions in this appendix are also appropriate to survey sampling or survey methodology generally, they are contextualized here in the design and implementation of ABS studies. The terms are listed alphabetically to simplify searching. Each term has an indicator describing the section(s) where the term is used. Each definition contains a reference, where appropriate. Useful sources for ABS terms include USPS (2013c), AAPOR (2015), and U.S. Census Bureau (2014).