Consistency and Accuracy of USPS-Provided Undeliverable Codes: Implications for Frame Construction, Data Collection Operational Decisions, and Response Rate Calculations

Kristine Wiant[1], Joe McMichael[1], Joe Murphy[1]
Katie Morton[1], Megan Waggy[1]
[1]RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27707

Abstract

When survey mailings are returned as "undeliverable" in the US, we rely on information codes provided by the U.S. Postal Service (USPS) to determine the eligibility of a sampled household. Although little is published about the consistency and accuracy of USPS undeliverable codes, they are often accepted at face value and used to make operational decisions about further contact and the ultimate disposition of a sample case. For the Residential Energy Consumption Survey (RECS), we examine the quality of the USPS's application of 8 different undeliverable codes. First, we review USPS codes from 3 mailings for the self-administered RECS National Pilot sent within a 12-day period to all 9,650 sampled households. We are particularly interested in codes that influence whether a case is disposed as an eligible occupied household or an ineligible vacant or seasonal household. Next, we compare USPS codes to field-verified status collected from the 2015 RECS main, field study. Then, we assess a model to predict the occupancy status of a case using auxiliary data that enhances known characteristics from the sampling frame and we assess the impact of that model on occupancy, coverage, and bias. Finally, we discuss the implications for operational tracking and decision-making on sample cases, the use of specific undeliverable codes in determining case eligibility, and the construction of sampling frames.

Key Words: address based sampling, undeliverables, frame construction, mail surveys, field studies

1. Introduction

When conducting a survey using an address-based sample (ABS), researchers have available some information on the units in the sampling frame prior to data collection. ABS frames typically contain indicators from the U.S. Postal Service (USPS) for whether an address is residential, whether it is the physical location of post-office box (PO Box) or a housing unit, and whether units have been flagged as vacant or seasonal. This information is potentially of value to the researcher to avoid the cost and effort of contacting units likely to be ineligible for the study at hand. Depending on the criteria for eligibility, these flags may be well suited for pre-screening cases; however, there is an open question regarding the quality of some indicators included on the frame.

In this paper, we are particularly interested in the frame indicator for vacant and seasonal homes. For the 2015 Residential Energy Consumption Survey (RECS) and RECS Household Pilots, eligibility was based on homes being non-vacant and non-seasonal. The extent to which the frame indicator could reliably and accurately predict the status of units represented a potential savings of cost and effort. However, if the frame information was not reliable or accurate, we ran the risk of excluding units that were actually eligible. This would result in undercoverage and the possibility of biased estimates obtained in the survey.

Some recent studies have suggested that a fair proportion (around 40%) of units flagged as vacant are actually occupied (Amaya et al. 2014; Kalton et al. 2014). The error in the frame may be due to the fact that housing units can change status rapidly and information on the frame may be outdated. Harter et al. (2016) caution against pre-screening to remove addresses classified as vacant or seasonal.

A second source of information about the status of a household comes when advance letters or surveys are mailed to a unit on the frame and the USPS returns the mailing as undeliverable. The USPS employs a variety of codes (e.g. Vacant, temporarily away, no such number, insufficient address) to indicate the reason a mailing may be undeliverable. For the purposes of RECS, we generally classify these into codes suggesting the unit is eligible (e.g. occupied) or ineligible for the survey. However, we did not have specific information regarding the reliability and accuracy of the USPS coding of undeliverable status and type. This is important to determine if the undeliverable information is to be used for estimating case eligibility.

This paper takes a critical look at the information on vacancy and/or seasonality status included in the ABS frame and returned by the USPS when cases are coded undeliverable. We examine the consistency of the frame indicator of vacancy, USPS returned mail, and verified outcomes from the field survey component. We consider whether the USPS is consistent in returning mail as undeliverable and whether their codes indicating the reason for non-delivery are consistent. Next, we fit a model to predict the occupancy status of a case using auxiliary data that enhances known characteristics from the sampling frame and we assess the impact of that model on occupancy, coverage, and bias. Finally, we discuss the implications for operational tracking and decision-making on sample cases, the use of specific undeliverable codes in determining case eligibility, and the construction of sampling frames.

2. Descriptions of Data

To address the research questions, we analyze data from two different but related sources: the Residential Energy Consumption Survey (RECS) National Pilot (RECS-NP) and the 2015 RECS. The RECS-NP is a pilot study designed to test the feasibility, cost-effectiveness, time efficiency, and response validity of conducting the RECS using a mixture of web and paper questionnaires delivered by mail. The second data source, the 2015 RECS, is the most recent of a periodic field study designed to collect data on household energy use for the target population. The target population for both surveys is all housing units (HUs) occupied as primary HUs in the United States. Both surveys were sponsored by the U.S. Energy Information Administration (EIA) and were conducted by IMG-Crown Energy Services and RTI International.

For the RECS-NP, a sample of 9,650 HUs was selected from an ABS frame derived from the USPS's Computerized Delivery Sequence (CDS) file. The CDS is a file of mail delivery points made available through a vendor that qualifies for updates to its proprietary address list. The CDS has very high coverage of occupied HUs in the United States (Shook-Sa et al., 2013). Prior to selecting the sample, drop points[1] and only way to get mail (OWGM) PO Boxes were removed from the frame.

Based on the recommendations from the tailored-design method (Dillman et al., 2014), a 7-step contact strategy was used for the RECS National Pilot. The first 3 mailings were sent via USPS to the full sample on the following schedule: the prenote was sent on day 1, the first invitation on day 6, and a reminder postcard on day 12. The data for the first part of our analysis includes the USPS undeliverable codes from the first 3 mailings, data on the ABS sampling frame, and the final case disposition codes.

To maximize coverage while controlling costs, the 2015 RECS sampling frame was constructed using a combination of ABS, RTI's frame-linking procedure, Check for Housing Units Missed (CHUM) (McMichael et al., 2008, 2013), and traditional field enumeration. From this hybrid frame, a sample of 6,522 addresses was selected. Like the RECS-NP, the frame for the 2015 RECS excludes OWGM PO Boxes; however, it includes drop points. In field studies, one or more units associated with a drop point can be randomly selected and contacted for interview.

The contact strategy for the 2015 RECS included an advance letter sent by USPS first class mail followed by in-person contacts. Therefore, the data available for the second part of the analysis include the undeliverable code from the prenotice letter, field codes indicating whether or not a dwelling unit is occupied, the number of contact attempts made, auxiliary data including data that was on the frame, census data, and additional data purchased from Acxiom. Specifically, the Acxiom data includes an indicator as to whether a name is appended to the address and whether the appended name is of Hispanic origin. Data analyses are of ABS addresses for which occupancy has been determined (i.e., excludes field enumerated addresses, added addresses from CHUM, and addresses with unknown occupancy).

3. Research Questions

The following research questions were posed to guide the analysis of these data. Findings and implications are discussed in subsequent sections.

1) In advance of a mail survey, can we predict whether or not a mailing will be returned as vacant?
2) Is the USPS consistent in their use of undeliverable codes?
3) Is an undeliverable letter an accurate predictor of vacancy as confirmed by field staff?
4) Can we build a model to better predict occupancy at the sampling stage that would be more effective than simply using the Vacant/Seasonal flag that is on the frame?

[1] A drop point is a delivery point on the CDS corresponding to multiple units. Because they are not uniquely identified, individual drop point units can be difficult to contact via mail, especially in repeated attempts.

5) When subsetting the national CDS file to create an ABS frame, how does both the vacancy/seasonal indicator and the proposed occupancy model impact: 1) coverage of occupied housing units; 2) Expected Vacancy rate; 3) coverage bias?

4. Findings

Question 1: In advance of a mail survey, can we predict whether or not a mailing will be returned as vacant?

We compared the frame indicator of vacancy with the USPS undeliverable information that we received in response to our first 3 mailings on the RECS-NP. Approximately 3% of the sample was flagged as vacant, and our actual overall undeliverable rate was around 5% for the first 3 mailings. However, most of our undeliverables came from those cases indicated as “not vacant/seasonal” on the frame.

Figure 1 shows that frame indicator of vacancy has limitations in predicting receiving a vacant code from the USPS. Less than 40% of the cases coded as “vacant/seasonal” on the frame were returned as undeliverable, and 20% of the households with the “vacant” indicator on the frame provided some sort of reply to our mailings.

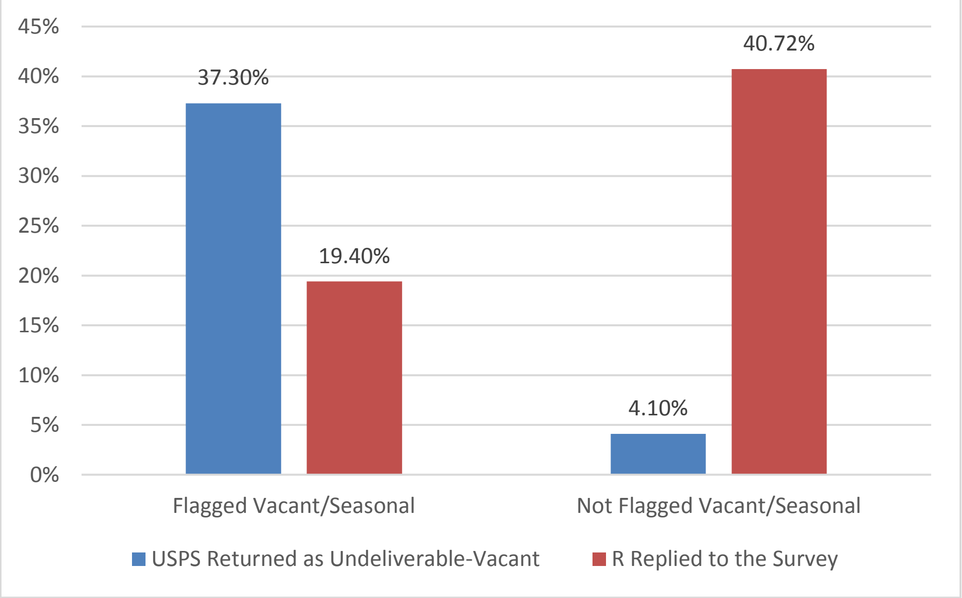

Figure 1: *Frame Indicator of Vacancy by USPS Returned Mail and Respondent’s Reply*

The pattern of data shown in Figure 1 is likely due to limitations of the “vacant/seasonal” flag on the sample frame; the vacancy status of a sampled household may change between the time the sample is procured and the package is sent. However the pattern shown in Figure 1 could also reflect issues with the quality of feedback received from the USPS. This possibility led to our second research question.

Question 2: Is the USPS consistent in their use of undeliverable codes?

We compared mailing results for 3 mailings sent to the same sampled addresses over the course of a 12-day period. Figure 2 shows that at the individual level the USPS is not consistent. Over the course of 12 days we would expect there to be little to no change in the actual deliverable status of a particular mailing address. However, 31% of the undeliverable cases came back as undeliverable only 1 out of 3 times. Another 29% of the undeliverables were returned as undeliverable in 2 out of 3 of the mailings. Perfect consistency, where all 3 came back undeliverable, was seen with only 40% of the undeliverable cases.

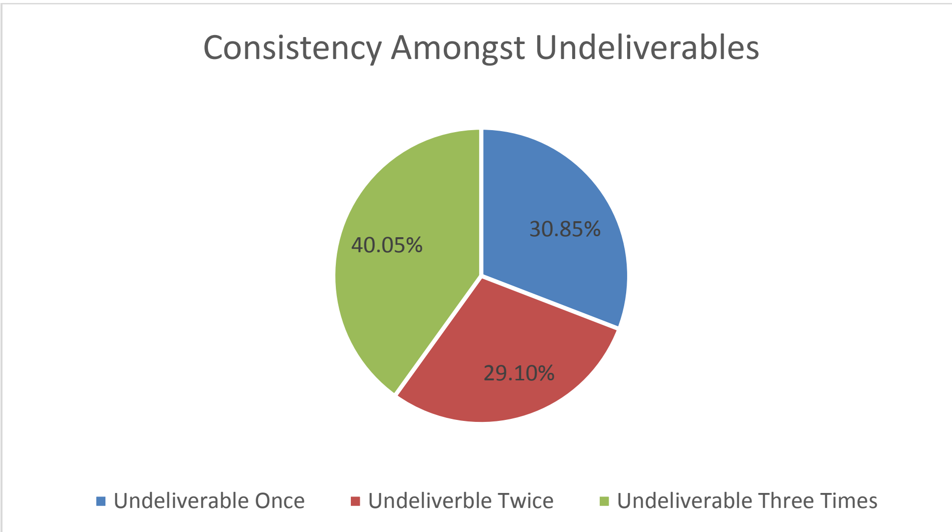

Figure 2: Within Address Inconsistency in USPS's Handling of Returned Mail for 3 Consecutive Mailings Sent over a 12-Day Period

In Figure 3, we see that cases that were undeliverable more than once over a 12-day period did not necessarily receive the same undeliverable code by the USPS each time. Amongst cases that were undeliverable twice, over 30% were returned with a different undeliverable code on the second package than on the first. Amongst cases that were undeliverable three times, less than 50% were consistently returned with the same explanation for the undeliverable status. This is a concern, because in AAPOR response rate calculations vacant households are treated as not eligible. The "vacant" undeliverable code was the most frequently used code at each of our 3 mailings; however, Figure 3 reveals considerable inconsistency in whether the USPS coded mail addressed to a particular address as vacant from one mailing to the next.

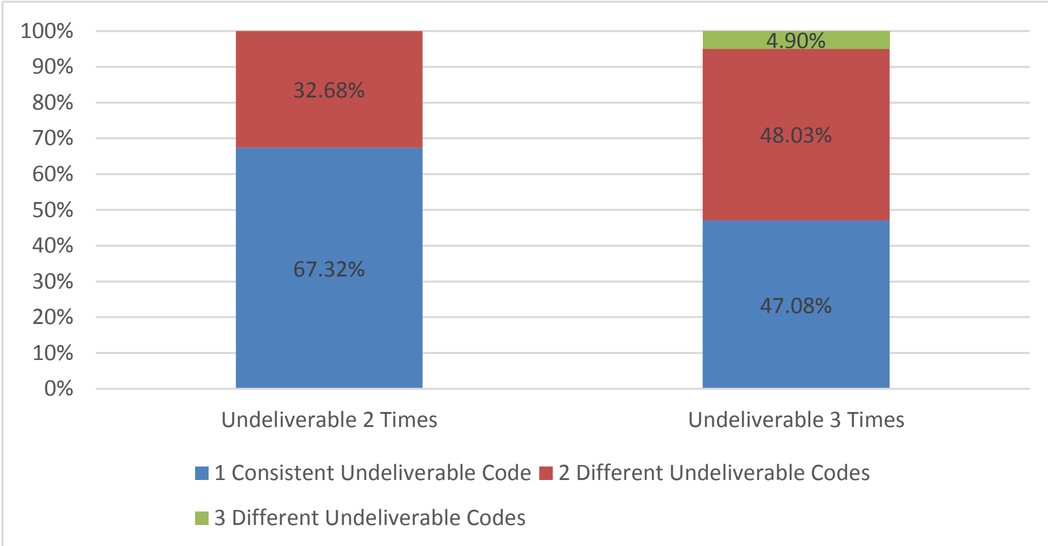

Figure 3: Within Address Inconsistency in USPS's Use of Particular Undeliverable Codes for 3 Consecutive Mailings Sent over a 12-Day Period

These inconsistencies raised questions over the household's true occupancy status, leading to question three.

Question 3: Is an undeliverable letter an accurate predictor of vacancy as confirmed by field staff?

The weighted percent of vacant housing units as confirmed by field staff in the 2015 RECS was computed by undeliverable type among households in which occupancy was determined. As shown in Figure 4, 62.4% of addresses returned by the USPS as undeliverable and 63.2% of addresses returned as vacant were actually confirmed as vacant or seasonal by field staff. This implies that approximately one-third of addresses that the USPS flags as vacant were actually occupied. Thus, the USPS undeliverable status has limitations in predicting actual vacancy.

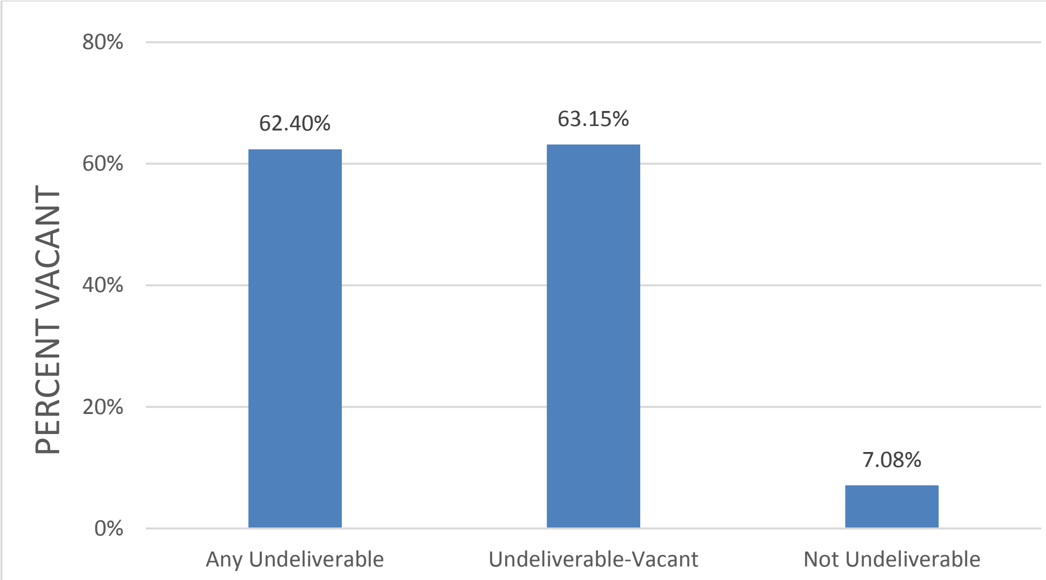

Figure 4: Percent Vacant by Undeliverable Type

We have shown that both frame vacant/seasonal (Figure 1) and undeliverable status (Figure 4) are individually poor predictors of actual vacancy. Our next question was then whether there are other variables we can use at the sampling stage in combination with the vacant/seasonal flag to predict occupancy.

Question 5: Can we build a model to better predict occupancy at the sampling stage that would be more effective than simply using the vacant/seasonalflag that is on the frame?

Using field-verified occupancy from the 2015 RECS, we modeled frame variables and other indicators in a stepwise logistic regression model to see whether we could improve the ability to predict occupancy. Variables assumed to be associated with occupancy were added to the model one at a time and dropped if not significant. Variables were added in the order shown in Table 1. Multi-family DUs and addresses with a Hispanic surname appended were originally predictive of occupancy, but became insignificant after subsequent variables were added to the model. We attempted to add HU density and percent owner-occupied HUs to the model, but these variables were not significant and were dropped. The cube root transformations of Census occupancy rate and Census low response score were used to get better linear relationships with field-verified occupancy.

As shown in Table 1, several variables are predictive of occupancy. Thus, using the occupancy prediction model, we are able to predict occupancy more effectively than when using the frame vacancy flag alone. According to the model, addresses flagged as vacant and/or seasonal on the sample frame, in rural areas, and with a higher Census low response score (Erdman and Bates, 2014) are less likely to be occupied. Addresses in the Midwest and West, addresses with a name appended, and addresses in block groups with a higher percentage of occupied dwelling units (DUs) are more likely to be occupied.

Table 1: Predicting Occupancy in a Stepwise Logistic Regression Model

Effect (in order added to the model)	*P-Value*	*Odds Ratio (95% CI)*
Vacant/Seasonal (CDS)	<.0001	0.12 (0.08,0.18)

Rural (Census)	0.0007	0.63 (0.48,0.82)
Census Region	0.0177	-
Midwest vs. Northeast	-	1.46 (1.01,2.11)
South vs. Northeast	-	1.15 (0.81,1.64)
West vs. Northeast	-	1.66 (1.14,2.43)
Multi-family Dwelling Unit (CDS)	0.1597	1.25 (0.92,1.69)
Name Append (Acxiom)	<.0001	3.00 (2.32,3.88)
Hispanic Name Append (Acxiom)	0.0687	1.47 (0.97, 2.23)
Census Occupancy Rate (cube root)	0.0256	2.56 (1.12, 5.83)
Census Low Response Score (cube root)	0.0001	0.36 (0.22, 0.58)

Question 5: When subsetting the national CDS file to create an ABS frame, how does both the vacant/seasonal indicator and the proposed occupancy model impact: 1) coverage of occupied housing units; 2) Expected Vacancy rate; 3) coverage bias?

When constructing an ABS frame, survey researchers make decisions on which addresses to include. Factors such as target population, survey mode and how to spend available resources will influence this decision. Ideally a sample frame represents the target population (occupied residential housing units) perfectly. Like most frames, ABS frames are not perfect and have ineligible units represented which causes inefficient use of resources since it takes time and money to determine eligibility. Survey researchers being a practical lot, may be willing to make trade-offs between frame coverage, coverage bias, and field costs.

Next we explore how subsetting an ABS frame using the CDS Vacancy/Seasonal indicator and the proposed Occupancy Prediction model impacts frame coverage of occupied housing units, vacancy rate, and coverage bias.

Coverage of Occupied Housing Units. Accurately estimating an ABS frame's coverage of occupied housing units directly is difficult or impossible with most survey data. However, estimating loss of coverage (percent occupied excluded) under differing scenarios is straightforward. We are able to do this by assuming the RECS sampling weight (calibrated and adjusted for non-response) provides an unbiased estimate of occupied housing units. This assumption seems reasonable since the RECS is designed to provide national energy consumption estimates for occupied residential households. As we exclude addresses from the sample to simulate their exclusion from the frame, we can calculate a weighted estimate of how many occupied addresses would have been excluded under that scenario.

To illustrate, we examine the Vacant/Seasonal indicator that is available on the ABS frame. We know 3.6% of addresses on the ABS frame are flagged as Vacant or Seasonal. At the same time, we know the weighted percent of known occupied 2015 RECS addresses that are flagged as Vacant or Seasonal on the ABS frame is 1.8%. This means if the RECS, or a survey using a similar frame, chose to exclude addresses that were flagged as Vacant or Seasonal they would be excluding an estimated 1.8% of occupied housing units from the frame. The blue circle in Figure 5 shows both the percent of occupied housing units flagged as Vacant or Seasonal (1.8%) and the percent of overall addresses flagged as Vacant or Seasonal on the frame (3.6%).

The Occupancy Predication Model applied to the ABS frame can be used in a similar way as the Vacant/Seasonal indicator but rather than being dichotomous, its continuous predicted probability allows for finer control in subsetting the frame. For example, if we opted to create an ABS frame excluding the top 10% of addresses least likely to be occupied, according to the model, we would also be excluding approximately 7% of occupied addresses. This result is seen represented by the blue line (labeled "Occupied Units Excluded") in Figure 5. While this example may not be practical for all surveys, we use it to explain how to interpret Figure 5.

Vacancy Rate. Vacancy Rate (=1- Occupancy Rate) is evaluated to gauge the potential for reductions in data collection costs. Intuitively, as the vacancy rate for a given sample decreases (or the occupancy rate increases) less effort needs to be expended determining a unit's eligibility. This could mean fewer doors to knock on in the case of in-person surveys or fewer mailings for mail surveys. Note we chose to display vacancy rate rather than its compliment, occupancy rate (Vacancy Rate = 1- Occupancy Rate), so it could easily be viewed on the same graph as the other measures discussed in this paper.

Vacancy Rate is calculated using the same assumptions used when calculating the percent of occupied housing units excluded. We estimate the vacancy rate to be 11% (or 89% occupancy rate). When examining the effects of excluding addresses from the ABS frame that are flagged as Vacant or Seasonal, the vacancy rate falls from 11% to 9.2%, a 1.8% decrease. This can be seen in Figure 5 represented by the red square.

When applying the Occupancy Prediction Model to subset the frame we can evaluate the impact on vacancy rate. For example, if we opted to create an ABS frame excluding the top 10% of addresses least likely to be occupied, according to the model, we would expect the vacancy rate to fall to 8.1%. This result is seen represented by the red line (labeled "Vacancy Rate") in Figure 5.

Coverage Bias. Coverage bias occurs when the population covered by a sampling frame is different than the population not covered. Naturally this differs by the characteristic we are interested in measuring. Most surveys attempt to correct for this type of bias through weight calibration but this is not always effective. Furthermore, measuring coverage bias directly is not always possible because it requires the characteristic of interest to be known for units both covered and not covered by the frame. We identified two characteristics known for all units on the frame, rural housing unit and multi-family housing unit, so as we subset the frame we can observe how the weighted estimate changes as we remove addresses from the frame.

We estimate 23.5% of occupied housing units are flagged as multi-family on the frame. In addition, we estimate 17.5% of occupied housing units are identified as rural. Again, we first examine the effects of excluding addresses from the ABS frame that are flagged as Vacant or Seasonal. The percent of multi-family falls slightly from 23.5% to 22.8%, a 0.7% decrease. Percent rural holds steady, only changing from 17.5% to 17.6%. Both of these estimates can be seen can be seen in Figure 5 where the green diamond represents multi-family and the black triangle represents rural.

When applying the Occupancy Prediction Model to subset the frame we can evaluate the impact on bias for these two characteristics. Again using the same example, if we opted to create an ABS frame excluding the top 10% of addresses least likely to be occupied, according to the model, we would also be excluding approximately 2.5% of occupied addresses identified as multi-family, and see no change in the estimated number of rural housing units. These results are seen represented by the green and black dotted lines in Figure 5.

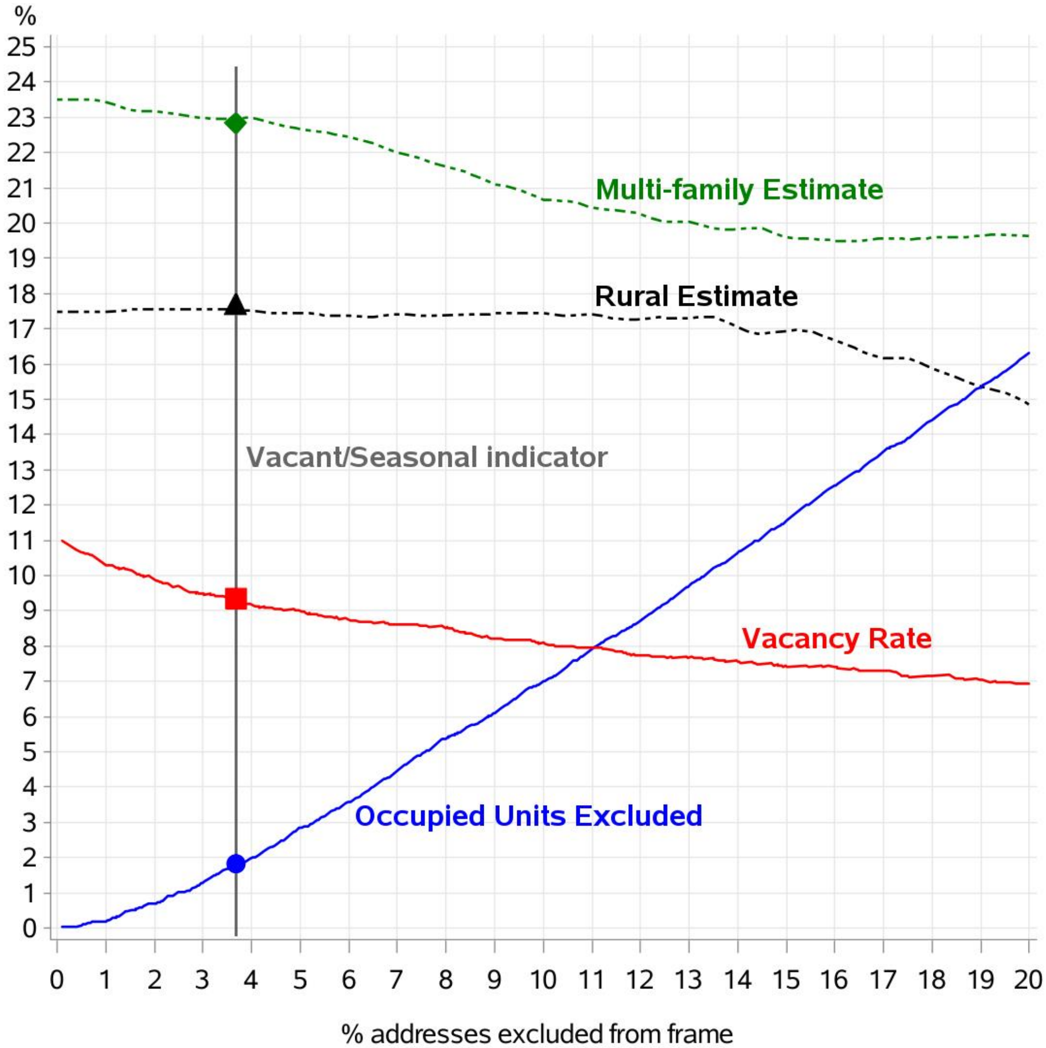

Figure 5: Subsetting an ABS frame by Vacant /Seasonal flag and Occupancy Prediction Model: Estimated impact on Frame Coverage of occupied housing units, multi-family and rural units as well as Vacancy Rate.

5. Summary and Implications

In this analysis, we have found that the USPS is inconsistent in handling and coding of undeliverable mail, and the presence of undeliverable mail and the "vacant/seasonal" flag that is available in address-based samples both have limitations in reliably predicting vacancy. It is possible to increase the accuracy of predicting occupancy with a model that includes frame variables (urban/rural, Census region, multi-family dwelling unit indicator, Census occupancy rate, and Census low response score) and auxiliary data (name and Hispanic name append acquired from Acxiom), in addition to the vacant/seasonalflag. The use of the proposed occupancy prediction model provides survey researchers more control over their ABS frame, allowing for tradeoffs between data collection costs (vacancy and occupancy rates) and the potential for coverage bias (% occupied housing units excluded).

These findings have a number of implications for operational tracking and decision-making on sample cases, the use of specific undeliverable codes in determining case eligibility, and the construction of sampling frames.

For mail surveys, we argue it is not worth the resources to continue to ship to a particular sampled address after receiving a package returned as undeliverable for that address. While most undeliverable addresses are not consistently returned as undeliverable, we found that the odds that an undeliverable address will be undeliverable a second or third time is 70%. Furthermore, we argue it is not meaningful to distinguish between cases returned as undeliverable-vacant and other “eligible” undeliverable codes, such as insufficient address or no such number. Both have similar outcomes in terms of occupancy status, as observed by field interviewers.

These findings also have implications for response rate calculations. For mail surveys, the USPS is so inconsistent in the particular undeliverable code they use for a given address that we caution against relying on a single USPS vacancy code as a single indicator of eligibility. For the RECS National Pilot, Biemer et. al (2016) addressed this challenge by employing a latent class analysis to estimate respondent eligibility using the USPS undeliverables codes from the first 3 mailings in combination with the frame variable that indicates vacancy.

References

Amaya A., LeClere F., Fiorio L., English, N. (2014). "Improving the utility of the DSF address-based frame through ancillary information." *Field Methods* 26:70-86.

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

Harter R., et al. (2016). AAPOR Report: Address Based Sampling. Available at http://www.aapor.org/Education-Resources/Reports/Address-based-Sampling.aspx

Dillman, Don A., Smyth, Jolene D., & Christian, Leah Melani. (2014). *Internet, phone, mail, and mixed-mode surveys : the tailored design method* (4th edition. ed.). Hoboken: Wiley.

Erdman, Chandra, and Nancy Bates. "The US Census Bureau Mail Return Rate Challenge: Crowdsourcing to Develop a Hard-to-Count Score." Statistics (2014): 08. http://www.census.gov/srd/papers/pdf/rrs2014-08.pdf

McMichael, J., Ridenhour, J., & Shook-Sa, B. (2008). A robust procedure to supplement the coverage of address-based sampling frames for household surveys. *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 4329–4335.

McMichael, J. P., Shook-Sa, B. E., Ridenhour, J. L., & Harter, R. (2013). The CHUM: A frame supplementation procedure for address-based sampling. In *2013 Research Conference Papers* (paper: http://www.fcsm.gov/events/papers2013.html, presentation:

http://www.copafs.org/UserFiles/file/fcsm/C1_McMichael_2013FCSM.pdf). Washington, DC.

Shook-Sa, B. E., Currivan, D. B., McMichael, J. P., & Iannacchione, V. G. (2013). Extending the coverage of address-based sampling frames: Beyond the USPS computerized delivery sequence file. *Public Opinion Quarterly, 77*(4), 994–1005.