Assessing the Effects of COVID-19 on Housing Vacancy Survey Estimates Using a Revised Nonresponse Adjustment Factor

Jonathan Spader
Daniel Truver
Peter Mateyka
Patricia Holley
Robert Callis
U.S. Census Bureau

Abstract

Rising COVID-19 case counts in early 2020 led to changes in the data collection procedures used for the Current Population Survey/Housing Vacancy Survey (CPS/HVS), an important source of information about vacancy rates and the homeownership rate in the United States. This report examines the implications of these data collection changes for CPS/HVS estimates. The analyses draw on multiple auxiliary data sources to understand the extent to which changes in nonresponse outcomes accompanied the changes in data collection procedures. The report then develops an alternative nonresponse adjustment factor that corrects for the observed changes in nonresponse. The results suggest that changes in nonresponse likely contributed to the sharp increase in the homeownership rate estimate for the second quarter of 2020. Conversely, the vacancy rate estimates are not similarly sensitive to the alternative nonresponse weighting adjustment; however, the results illustrate the potential for the vacancy rate estimates to underestimate the actual vacancy levels due to the weighting methodology's assumption that all nonresponding housing units are occupied. These results suggest that the CPS/HVS estimates of vacancy rates and the homeownership rate should be interpreted with caution for the period affected by the changes in data collection procedures.

Introduction

In response to rising COVID-19 case counts in the United States, the U.S. Census Bureau, on March 20, 2020, suspended in-person interview attempts for the Current Population Survey, including its Housing Vacancy Survey supplement (CPS/HVS).[1] This report examines the implications of this change in data collection procedures for the CPS/HVS estimates. The quarterly CPS/HVS estimate of the rental vacancy rate is a Principal Federal Economic Indicator used by the federal government and macroeconomists to evaluate current economic conditions. Additionally, the CPS/HVS estimates include closely watched measures of the homeownership rate, homeowner vacancy rate, and gross vacancy rate that provide timely information about housing market conditions and the housing inventory.

The report first provides a brief summary of the changes made to CPS/HVS data collection procedures in response to the suspension of in-person interviews. The analyses then develop an alternative nonresponse weighting adjustment factor and examine the sensitivity of CPS/HVS estimates. The results suggest that changes in sample composition likely contributed to the historically large increase in the CPS/HVS homeownership rate estimate for the second quarter of 2020. Conversely, the CPS/HVS vacancy rate estimates are not significantly affected by the use of the alternative nonresponse adjustment factor;[2] however, the analyses illustrate the potential for the CPS/HVS vacancy rate estimates to underestimate the actual levels of vacancy in 2020 due to the weighting methodology's assumption that all nonresponding housing units are occupied.

Changes to Housing Vacancy Survey Data Collection Procedures in 2020

In response to the rising numbers of COVID-19 cases in the United States, the Census Bureau suspended personal visits for the CPS/HVS on March 20, 2020. The suspension of personal visits continued in all areas of the United States for the CPS/HVS data collection periods in April, May, and June 2020. Beginning in July, personal visits began to be reintroduced in a subset of localities, with additional areas added in August. All areas of the country were eligible for personal visits in September. During this period, the Census Bureau continued to collect the CPS/HVS by telephone, making efforts to collect telephone interviews for all sample units, including vacant units and ineligible units.

---

[1] The rental vacancy rate and homeowner vacancy rate are produced from the data collected by the Housing Vacancy Survey supplement, whereas the homeownership rate is produced from the occupied units in the Current Population Survey. For ease of notation, this report uses the CPS/HVS label to refer to the combined set of vacancy rate and homeownership rate estimates. Additional information about the quarterly CPS/HVS estimates is available at: https://www.census.gov/housing/hvs/index.html

[2] Use of the term statistically significant in the text of this report indicates that a finding is significant at the 90 percent level or higher. All tables report significance at the 90, 95 and 99 percent levels.

The standard CPS/HVS data collection procedures use personal visits as the primary mode of data collection but allow telephone interviews when certain conditions are met.[3] The suspension of in-person interviews meant that telephone contact attempts replaced in-person interview attempts for all housing units in the sample. These telephone contacts relied on phone numbers identified through multiple sources. For housing units with a completed interview in a previous month, interviewers attempted to contact the occupant or knowledgeable proxy interviewed during the previous month. For other housing units, interviewers were encouraged to use the available resources to identify contact information for sample housing units and/or knowledgeable proxy respondents. These resources included internal resources such as purchased third-party telephone lookup databases, as well as public records databases such as tax assessor records. Interviewers could also use online searches to identify leasing offices or telephone contacts with knowledgeable local sources such as real estate agents, neighbors, and postal workers who might identify vacant units, provide contact information for the property owner, or complete a proxy interview.

In each month, CPS/HVS data collection generally begins at the start of the week containing the 19th and closes out early the following week. On March 20, 2020, the suspension of personal visits occurred on the Friday during the week of data collection. While interviewers were able to make at least one personal visit attempt to most sample units prior to the suspension of personal visits, the suspension occurred prior to the completion of data collection activities for March 2020—and therefore prior to the completion of data collection for the first quarter 2020.

While interviewers made extensive efforts to complete data collection using telephone-based contact attempts, response rates declined following the suspension of personal visits. The share of all sampled housing units for which data collection could not be completed (i.e., Type A nonresponses using CPS/HVS terminology) increased from 14–15 percent in each quarter of 2019 to 18 percent in the first quarter of 2020, 28 percent in the second quarter of 2020, and 24 percent in the third quarter of 2020.

Data and Methodology

The base dataset for the analyses is the monthly sample of housing units selected for CPS/HVS data collection. We append the monthly datasets from January 2019 through September 2020, categorizing the data into seven quarters to match the CPS/HVS quarterly releases. In each quarter, nonresponse is defined to include any housing unit where data collection could not be completed (e.g., no response, refusal, unreachable, etc.)—termed "Type A" nonresponses in CPS/HVS terminology. The set of "completed" responses conversely includes all housing units where data collection could be completed, which includes completed interviews for occupied and vacant

---

[3] The CPS/HVS sample design is a rotating panel. Once selected, a housing unit is in the sample for four consecutive months, out for 8 months, and then in the sample for 4 months. Under the standard CPS/HVS data collection procedures, the first and fifth interviews are required to be collected through personal visits. In other months, a telephone interview can be completed with HVS-eligible sample housing units if the unit was HVS-eligible in the previous month, the unit is located geographically distant from the interviewer's home and other remaining interviews, and the name and telephone number of a reliable respondent is available and a telephone interview is acceptable to that person. For additional information, see the Current Population Survey Interviewing Manual (2015): https://www.census.gov/housing/hvs/methodology/CPS_Manual_April2015.pdf.

units as well as units determined to be ineligible for interview.[4] To the extent that the suspension of personal visits limited the ability of interviewers to complete data collection for sample units, the likely result would be an increase in the count of Type A nonresponses and decreases in the counts of completed cases for occupied, vacant, and ineligible units. This classification, therefore, starts with the full sample of all units selected for data collection and defines nonresponse to include units where the data collection process could not be completed.

The base dataset is supplemented with several auxiliary data sources that provide information about the characteristics of both respondents and nonrespondents.[5] First, information from the 2010 Decennial Census and property records from Black Knight, Inc. are each merged to the base sample at the housing unit level, using the Census Bureau's master address file identifier (Brummet, 2014). The 2010 Decennial Census contains information about the vacancy status, tenure, and other unit attributes at the time of the 2010 Census. The vendor data from Black Knight, Inc. contain information about the housing unit compiled from county tax assessor records and other sources. Additionally, neighborhood attributes measured at the census tract level are added from the 2018 American Community Survey 5-year estimates. Exhibit 1 contains a summary of the data sources and variables added from each source.

Exhibit 1

Supplemental Data Sources and Variable Definitions (1 of 2)

Variable	Definition
Current Population Survey & Housing Vacancy Survey Supplement	
mis1–8	Month-in-sample (MIS) group for data collection. 1 = MIS 1 … 8 = MIS 8.
metro1–3	Metropolitan status: 1 = Principal city; 2 = Metropolitan area outside principal city; 3 = Nonmetropolitan area.
2010 Decennial Census	
decmis	Sample unit cannot be matched to 2010 Decennial housing units using MAFID.
vacant1–7	Vacant unit: 1 = For rent; 2 = Rented, not occupied; 3 = For sale only; 4 = Sold, not occupied; 5 = For seasonal/recreational use; 6 = For migrant workers; 7 = Other vacant.
tenure1–4	Tenure status: 1 = owned free and clear; 2 = owned with a mortgage; 3 = rented; 4 = occupied without payment of cash rent
bld	Building type: s = single-family home; m = multifamily structure; to = mobile home or other building type
hht1–7	Household type: 1 = family, married; 2 = family, male reference person, no spouse; 3 = family, female reference person, no spouse; 4 = nonfamily, male reference person, living alone; 5 = nonfamily, male reference person, not living alone; 6 = nonfamily, female reference person, living alone; 7 = nonfamily, female reference person, not living alone.

[4] In CPS terminology, the set of completed responses includes completed interviews for occupied housing units, Type B units, and Type C units. Type B units include HVS-eligible vacant units, as well as units that are occupied solely by persons not eligible for interview. Type C units include units that are not eligible for interview such as demolished units and units converted to a nonresidential use. For additional information, see the Current Population Survey Interviewing Manual (2015): https://www.census.gov/housing/hvs/methodology/CPS_Manual_April2015.pdf.

[5] This approach is inspired by other recent nonresponse analyses that use linked data to expand the set of attributes that can be observed for both respondents and nonrespondents (Bee, Gathright, and Meyer, 2015; Brummet, 2014; Brummet et al., 2018; Eggleston and Westra 2020; Rothbaum and Bee 2020; Sabelhaus et al., 2015; Wagner and Layne, 2014).

Exhibit 1

Supplemental Data Sources and Variable Definitions (2 of 2)

Variable	Definition
2010 Decennial Census	
hhldrage	Age of the householder, continuous
Hispanic	Hispanic origin of the householder: 1 = Hispanic; 0 = Non-Hispanic
White	Race of the householder: 1 = white
Black	Race of the householder: 1 = black
AIAN	Race of the householder: 1 = American Indian or Alaska Native
Asian	Race of the householder: 1 = Asian
NHOPI	Race of the householder: 1 = Native Hawaiian or Pacific Islander
other	Race of the householder: 1 = Other race
Black Knight Inc. Records Pulled in 2018.	
bkmis	Sample unit cannot be matched to units in Black Knight data pulled in 2018.
bkowner	Black Knight's measure of owner-occupancy: 1 = owner-occupied
bkrenter	Black Knight's measure of owner-occupancy: 1 = renter-occupied
2014–18 American Community Survey 5-Year Estimates	
acsmis	Census tract of the sample unit cannot be matched to tracts in 2014–2018 American Community Survey 5-year estimates
medval	Median home value in the tract.
medinc	Median household income in the tract.
phhpov	Percent of tract households with income below the poverty-level.
pvacs	Percent of tract housing units that are vacant.
pmover	Percent of tract population age 1 and over who moved during the previous year.
pown	Percent of tract housing units that are owner-occupied.
pa17	Percent of tract population age 17 or younger.
pa18	Percent of tract population age 18–34.
pa35	Percent of tract population age 35–54.
pa55	Percent of tract population age 55–74
pa75	Percent of tract population age 75 or older
phis	Percent of tract population: Hispanic
pnhw	Percent of tract population: Non-Hispanic White
pnhb	Percent of tract population: Non-Hispanic Black
pnha	Percent of tract population: Non-Hispanic Asian
poth	Percent of tract population: Non-Hispanic other race

This set of auxiliary data sources is unlikely to exhaustively capture all of the possible changes in nonresponse patterns following the changes in data collection procedures. Instead, it should be interpreted as the set of observable factors available for this analysis. In choosing supplemental

data sources for this analysis, we limited our search to data sources that could be quickly accessed and linked to the CPS/HVS sample, reflecting the desire to produce information quickly. We then sought to identify variables likely to be correlated with vacancy and homeownership while also seeking to include a broad set of demographic and housing characteristics.[6]

Spader et al. (2021) present the results of nonresponse analyses that use these data to compare the characteristics of responding versus nonresponding housing units in each quarter—and to test whether these nonresponse outcomes changed significantly following the suspension of personal visits. These analyses conclude that while the differences between nonrespondents and respondents did not change significantly across quarters in 2019, multiple significant changes appeared in the second quarter of 2020. Specifically, the nonresponse analyses suggest that the set of responding housing units in the second quarter of 2020 included fewer units that were rotating into the sample for their first or second month in the sample and more units that were in month-in-sample (MIS) 7 and 8; more units that were owned free and clear in the 2010 Decennial Census; more units identified as owner-occupied by Black Knight's measure and fewer units that could not be matched to Black Knight data;[7] and fewer units in neighborhoods with higher poverty rates.[8]

These changes over time in the relative characteristics of respondents versus nonrespondents may affect CPS/HVS estimates to the extent that they are not accounted for by the existing weighting methodology—which is described briefly in the following paragraphs and in greater detail in CPS Technical Paper 77 (U.S. Census Bureau, 2019).

Base Weights. Under the standard methodology, the base weights are the first component of the CPS/HVS weights and account for differences in sampling probabilities. As described in Technical Paper 77, the base weights are sufficient to produce unbiased estimates of vacancy rates and the homeownership rate under strong assumptions about ideal survey conditions such as zero frame error, zero nonsampling error, and nonresponse patterns that are independent of the variables used to produce the estimates (U.S. Census Bureau, 2019).

Standard NR Weights. The subsequent weighting adjustment factors applied by the standard methodology adjust for nonresponse in two ways. First, the CPS household weight applied to occupied units includes a nonresponse weighting adjustment factor that adjusts for differences in response across primary sampling units (PSUs) and central city location status.[9] This nonresponse adjustment is the first weighting adjustment factor applied to the base weights for occupied units. The HVS supplement weights applied to vacant units do not have any similar adjustment

---

[6] Initial analyses also merged postal data about change of address requests associated with the housing unit from the National Change of Address (NCOA) database. Surprisingly, none of the NCOA variables tested were correlated with nonresponse or selected for inclusion in the regressions so this data source is omitted from exhibit 1.

[7] The measure of nonmatches to Black Knight data should be interpreted to include a variety of factors related to both data coverage and the matching process. In particular, tax assessor data has higher coverage of single-family homes than of rental units in large multifamily buildings, so the measure of nonmatches to Black Knight data is likely correlated with renter-occupancy.

[8] More detailed results and discussion of the nonresponse analyses are available in Spader et al. (2021).

[9] The nonresponse adjustment factor groups PSUs within the same state that are similar in metropolitan status and size and then splits these clusters based on central city and non-central-city location to form the adjustment cells.

for nonresponse. In this report, we use the term 'Standard NR Weights' to refer to the weights constructed by applying this nonresponse adjustment factor to the base weights.

Second, the process of controlling the weights to independent population totals may also reduce the effects of differential nonresponse to the extent that nonresponse is correlated with the demographic subgroups used in the population controls.[10] Any adjustments for nonresponse from this second process occur in later weighting steps, so they are not reflected in the Standard NR Weights.

Alternative NR Weights. This report develops and tests an alternative nonresponse (NR) weighting adjustment factor using a propensity-score-based approach. This alternative nonresponse weighting adjustment factor is designed to replace the existing nonresponse weighting adjustment factor described previously, and we use the term "Alternative NR Weights" to refer to the weights constructed by applying the alternative nonresponse weighting adjustment factor to the base weights.

The propensity scores are constructed by estimating logistic regressions with the following form:

$$(1) \qquad \mathrm{Log}((\mathrm{Pr}(Y_i \,/\, (1 - \mathrm{Pr}(Y_i)))) = \alpha + X_i\beta_1 + G_i\beta_2 + M_i\beta_3$$

Where Y_i is an indicator for whether the housing unit is a Type A nonresponse, X_i is a vector of covariates from the supplemental data sources, G_i is a set of fixed effects that interact the 51 states with three metropolitan status categories (principal city, other MSA/CBSA, and nonmetropolitan), and M_i is a set of fixed effects that interact the three metropolitan status categories with 8 month-in-sample categories. Equation 1 is estimated separately for each quarter using logistic regressions. The estimated coefficients are shown in exhibit 2.[11]

---

[10] A full discussion of the weighting components included in the standard methodology is available in CPS Technical Paper 77 (U.S. Census Bureau, 2019).

[11] To determine the model specification, the authors started with the set of attributes that show either a significant difference between nonrespondents versus respondents or a significant change over time in these differences. Two of the authors then independently developed specifications using logistic regressions with Akaike Information Criterion (AIC) as a measure of model performance and using ordinary least squares (OLS) with r-squared as a measure of performance. The two resulting specifications were then compared and consolidated into the final specification. While few of the estimated coefficients are significant, the AIC values and other performance metrics suggest that the covariates shown in exhibit 2 add explanatory power. Additional information about the model specification is available in the appendix of Spader et al. (2021).

Exhibit 2

Logistic Regressions Modeling Response versus Type A Nonresponse on Selected Covariates (1 of 2)

	Q1 2019		Q2 2019		Q3 2019		Q4 2019		Q1 2020		Q2 2020		Q3 2020	
	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)
Intercept	1.894	1.592	1.965	1.552	1.588	1.570	1.833	1.661	1.823	1.418	1.490	1.286	1.228	1.287
decmis	0.129	0.370	0.179	0.335	0.229	0.325	0.154	0.351	0.169	0.299	0.140	0.274	0.116	0.251
vactype1	0.089	0.448	0.183	0.410	0.075	0.376	0.068	0.419	0.013	0.392	− 0.056	0.310	0.019	0.351
vactype2	0.339	1.677	0.601	2.042	0.313	1.979	0.344	1.983	0.804	1.801	0.057	1.098	− 0.016	1.341
vactype3	0.052	0.539	0.169	0.550	0.109	0.549	0.072	0.625	0.225	0.566	0.266	0.433	0.128	0.522
vactype4	0.286	1.384	0.569	1.332	0.311	1.450	0.557	1.514	0.533	1.170	0.279	0.935	0.320	0.918
vactype5	1.125	0.641*	1.131	0.676*	0.955	0.663	0.935	0.674	0.985	0.594*	1.003	0.500**	0.914	0.489*
tenure2	0.069	0.214	0.142	0.204	0.102	0.225	0.044	0.222	0.032	0.192	0.039	0.157	0.068	0.185
tenure3	− 0.144	0.205	-0.090	0.206	− 0.067	0.201	− 0.148	0.197	− 0.213	0.183	− 0.262	0.156*	− 0.167	0.168
tenure4	0.202	0.560	0.239	0.538	0.150	0.700	− 0.040	0.514	0.026	0.440	− 0.007	0.449	0.077	0.428
bldm	− 0.139	0.239	− 0.135	0.213	− 0.164	0.233	− 0.175	0.228	− 0.124	0.191	− 0.085	0.159	− 0.191	0.181
bldto	0.220	0.344	0.157	0.335	0.160	0.320	0.160	0.312	0.136	0.300	− 0.004	0.280	0.008	0.255
hht2	− 0.089	0.305	− 0.088	0.311	− 0.133	0.333	− 0.103	0.315	− 0.187	0.270	− 0.217	0.254	− 0.220	0.277
hht3	− 0.186	0.195	− 0.134	0.181	− 0.170	0.230	− 0.191	0.213	− 0.145	0.172	− 0.179	0.160	− 0.176	0.175
hht4	− 0.126	0.217	− 0.161	0.215	− 0.097	0.250	− 0.220	0.226	− 0.192	0.196	− 0.152	0.180	− 0.119	0.189
hht5	− 0.070	0.328	− 0.183	0.320	− 0.137	0.373	− 0.118	0.362	− 0.102	0.335	− 0.145	0.264	− 0.046	0.295
hht6	− 0.161	0.224	− 0.186	0.228	− 0.149	0.246	− 0.177	0.211	− 0.152	0.199	− 0.085	0.191	− 0.059	0.195
hht7	− 0.142	0.390	− 0.166	0.368	− 0.061	0.391	0.131	0.427	− 0.094	0.360	− 0.130	0.329	− 0.058	0.331
hhldrage	0.007	0.004*	0.008	0.004**	0.008	0.004*	0.007	0.004*	0.007	0.004*	0.008	0.003**	0.006	0.003*
Hispanic	0.015	0.278	0.008	0.308	− 0.077	0.275	− 0.089	0.265	− 0.050	0.264	− 0.068	0.226	− 0.176	0.241
Black	− 0.188	0.237	− 0.216	0.251	− 0.175	0.239	− 0.197	0.273	− 0.128	0.231	− 0.081	0.197	− 0.125	0.180
AIAN	− 0.160	0.686	− 0.186	0.687	− 0.133	0.642	− 0.121	0.649	− 0.104	0.599	− 0.166	0.523	− 0.143	0.492
Asian	0.024	0.332	0.035	0.341	− 0.134	0.348	− 0.044	0.310	− 0.023	0.337	− 0.031	0.265	− 0.066	0.299
NHOPI	− 0.220	1.250	− 0.110	1.333	− 0.230	1.220	− 0.224	1.206	− 0.230	1.245	− 0.150	1.237	− 0.118	1.115

Exhibit 2

Logistic Regressions Modeling Response versus Type A Nonresponse on Selected Covariates (2 of 2)

	Q1 2019		Q2 2019		Q3 2019		Q4 2019		Q1 2020		Q2 2020		Q3 2020	
	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)
other	− 0.037	0.402	− 0.157	0.394	− 0.052	0.398	− 0.042	0.383	− 0.083	0.349	− 0.125	0.337	0.019	0.309
bkmis	0.060	0.222	− 0.027	0.242	− 0.010	0.220	0.054	0.244	0.041	0.197	− 0.022	0.182	0.001	0.181
bkowner	− 0.031	0.199	− 0.073	0.214	− 0.006	0.234	0.017	0.230	0.006	0.193	0.172	0.188	0.185	0.169
acsmis	0.222	1.556	0.178	1.540	0.264	1.591	0.155	1.595	0.034	1.536	0.069	1.371	0.176	1.312
medval	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
medinc	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
phhpov	0.755	1.017	1.092	1.167	0.506	1.062	1.047	1.148	0.751	1.081	− 0.131	1.033	0.038	1.015
pvacs	0.792	0.887	0.964	0.971	0.937	0.847	0.775	0.887	0.620	0.888	0.734	0.720	0.913	0.761
pmover	0.156	1.378	0.232	1.252	− 0.086	1.415	− 0.223	1.429	− 0.085	1.280	− 0.356	1.256	− 0.137	1.123
pown	0.128	0.532	0.258	0.583	0.134	0.603	0.092	0.581	0.177	0.535	0.162	0.558	0.121	0.447
pa17	0.227	1.516	− 0.079	1.685	0.609	1.963	0.267	1.797	− 0.096	1.827	− 0.708	1.732	− 0.500	1.521
pa18	0.160	1.717	− 0.120	1.621	0.343	1.658	0.076	1.732	0.022	1.688	0.007	1.473	0.032	1.396
pa35	− 0.232	2.168	− 0.482	2.206	− 0.025	2.204	− 0.102	2.338	− 0.340	1.978	− 0.298	1.874	0.086	1.882
pa75	1.277	2.703	0.938	2.972	1.734	2.903	0.968	2.913	1.047	3.033	0.533	2.700	1.035	2.518
phis	− 0.196	0.536	− 0.210	0.555	− 0.295	0.636	− 0.268	0.531	− 0.253	0.494	− 0.132	0.502	− 0.176	0.497
pnhb	− 0.525	0.528	− 0.476	0.468	− 0.494	0.539	− 0.618	0.486	− 0.495	0.474	− 0.363	0.445	− 0.321	0.426
pnha	− 0.297	1.009	− 0.322	1.020	− 0.287	0.953	− 0.535	1.008	− 0.235	0.968	− 0.198	0.919	− 0.125	0.923
poth	− 1.012	1.379	− 0.582	1.414	− 0.823	1.726	− 0.678	1.946	− 0.597	1.217	− 0.907	1.141	− 0.725	1.498
State x Metro FE	Yes		Yes		Yes		Yes		Yes		Yes		Yes	
MIS x Metro FE	Yes		Yes		Yes		Yes		Yes		Yes		Yes	

Q =quarter

***p<.01; **p<.05; *p<.10. Asterisks are reported for all estimates.

Notes: This exhibit reports the results of logistic regressions that model an indicator for Type A nonresponse on the set of covariates identified in the exhibit. The dependent variable is defined so that positive coefficients correspond to increased likelihood of response (0 = Type A Nonresponse; 1 = Completed interview, Type B, or Type C). Separate logistic regressions are estimated for each quarter.

Sources: U.S. Census Bureau; Current Population Survey and Housing Vacancy Survey Supplement data for 2019–2020 linked to the supplemental data sources identified in exhibit 1

These regressions are used to calculate the predicted probability of response for each sample housing unit in each quarter, including both occupied and vacant units. The alternative nonresponse weighting adjustment factor is then calculated as the inverse of the predicted probability of response, and the alternative weights are constructed by multiplying this alternative nonresponse adjustment factor by the base weights to account for differences in response propensities.

Results

Exhibit 3 reports estimates of the homeownership rate, rental vacancy rate, homeowner vacancy rate, and gross vacancy rate using the weights described in the previous section. For ease of review, exhibits 4A–4D visualize these estimates in line charts, along with the published figures that rely on the CPS/HVS final weight (labeled "Published").

Comparison of the Alternative NR Weights estimates to the Base Weights estimates reveals the effects of the alternative nonresponse weighting adjustment, because the only difference between these estimates is the application of the nonresponse weighting adjustment to the base weights. The Alternative NR Weights estimates can similarly be compared with the Standard NR Weights estimates to examine the differences between the alternative nonresponse adjustment and the standard nonresponse adjustment. Lastly, the Published estimates provide an additional benchmark that shows the cumulative effects of the remaining weighting adjustment factors that are applied after the standard nonresponse weighting adjustment factor, such as the second-stage weighting adjustment factor that controls the weights to independent population totals.

Exhibit 3

Estimates of the Homeownership Rate, Rental Vacancy Rate, Homeowner Vacancy Rate, and Gross Vacancy Rate using Alternative Weighting Approaches

|	Estimates							Year-Over-Year Change		
	Q1 2019	Q2 2019	Q3 2019	Q4 2019	Q1 2020	Q2 2020	Q3 2020	Q1 2020 – Q1 2019	Q2 2020 – Q2 2019	Q3 2020 – Q3 2019
Homeownership Rate										
Alternative NR Weights	0.655	0.652	0.657	0.661	0.661	0.673	0.672	0.006	0.021	0.015
(S.E.)	0.019	0.018	0.019	0.018	0.020	0.018	0.019	0.018	0.019	0.020
Standard NR Weights	0.661*	0.660**	0.665**	0.670**	0.673***	0.702***	0.698***	0.012	0.042***	0.033***
(S.E.)	0.020	0.018	0.019	0.019	0.020	0.017	0.019	0.018	0.019	0.020
Base Weights	0.664**	0.661**	0.667***	0.673***	0.675***	0.705***	0.699***	0.011	0.044***	0.032***
(S.E.)	0.020	0.018	0.019	0.018	0.020	0.017	0.019	0.018	0.019	0.020
Rental Vacancy Rate										
Alternative NR Weights	0.086	0.084	0.084	0.080	0.085	0.089	0.093	– 0.001	0.005	0.009
(S.E.)	0.014	0.012	0.013	0.013	0.015	0.018	0.016	0.019	0.020	0.017
Standard NR Weights	0.072***	0.070***	0.070***	0.067***	0.069***	0.060***	0.068***	– 0.004	– 0.010**	– 0.002**
(S.E.)	0.012	0.011	0.011	0.012	0.013	0.012	0.012	0.016	0.015	0.014
Base Weights	0.086	0.084	0.083	0.080	0.085	0.087	0.093	– 0.001	0.004	0.009
(S.E.)	0.015	0.013	0.013	0.014	0.016	0.017	0.015	0.019	0.019	0.017
Homeowner Vacancy Rate										
Alternative NR Weights	0.016	0.015	0.017	0.016	0.014	0.013	0.013	– 0.002	– 0.002	– 0.004
(S.E.)	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.006	0.005	0.005
Standard NR Weights	0.013***	0.013***	0.014***	0.014***	0.011***	0.008***	0.009***	– 0.002	– 0.004	– 0.005
(S.E.)	0.003	0.003	0.003	0.004	0.003	0.003	0.003	0.004	0.004	0.004
Base Weights	0.016	0.015	0.017	0.016	0.014	0.013	0.013	– 0.002	– 0.003	– 0.004
(S.E.)	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.005	0.005	0.005
Gross Vacancy Rate										
Alternative NR Weights	0.139	0.140	0.140	0.132	0.136	0.139	0.134	– 0.003	– 0.001	– 0.007
(S.E.)	0.013	0.014	0.013	0.014	0.014	0.016	0.014	0.012	0.012	0.011
Standard NR Weights	0.120***	0.121***	0.121***	0.115***	0.113***	0.098***	0.100***	– 0.007**	– 0.023***	– 0.021***
(S.E.)	0.012	0.013	0.012	0.013	0.013	0.013	0.012	0.010	0.010	0.009
Base Weights	0.142	0.143*	0.142	0.134	0.139	0.139	0.135	– 0.003	– 0.004	– 0.007
(S.E.)	0.014	0.015	0.014	0.014	0.016	0.017	0.015	0.012	0.013	0.011

NR = nonresponse. Q= quarter. SE = standard error.*

**p<.01; **p<.05; *p<.10. The asterisks reflect significance tests that compare the estimates using the Base Weights and the Standard NR Weights, respectively, to the estimates using the Alternative NR Weights for the same outcome and quarter.

Notes: This exhibit reports estimates of each outcome using the propensity-score-based nonresponse adjustment developed in this report (Alternative NR Weights), the CPS base weights (Base Weights), and the CPS base weights multiplied by the existing nonresponse adjustment factor (Standard NR Weights).

Exhibit 4A

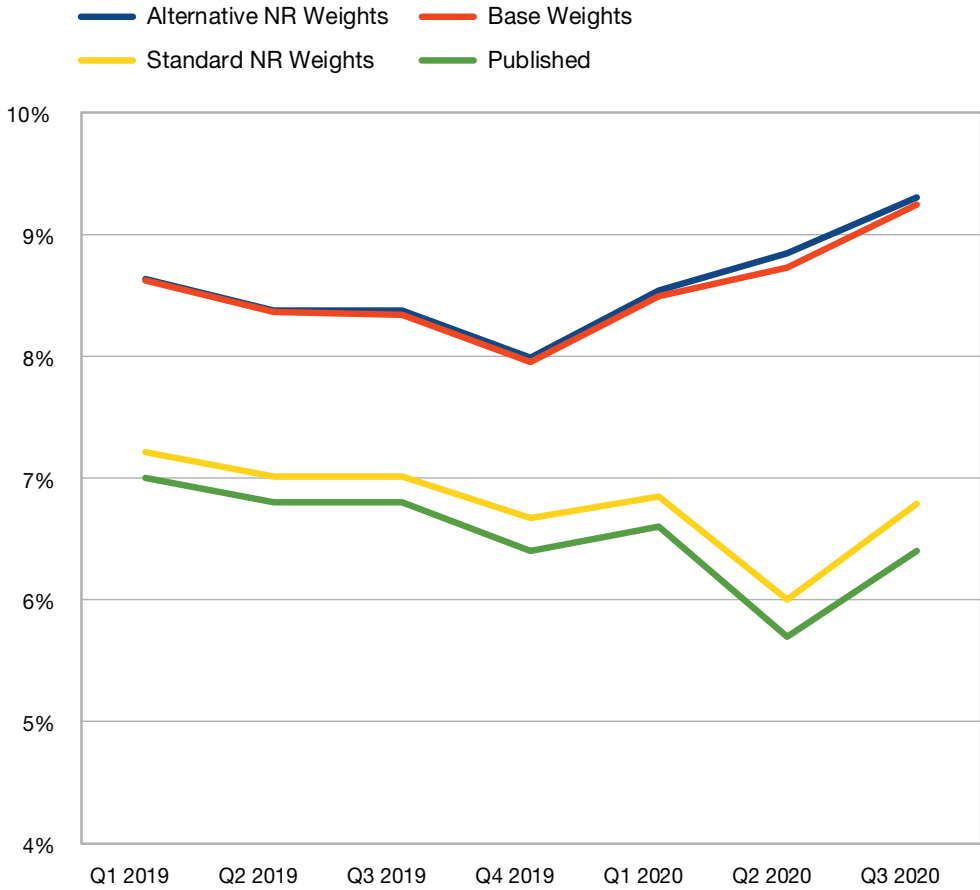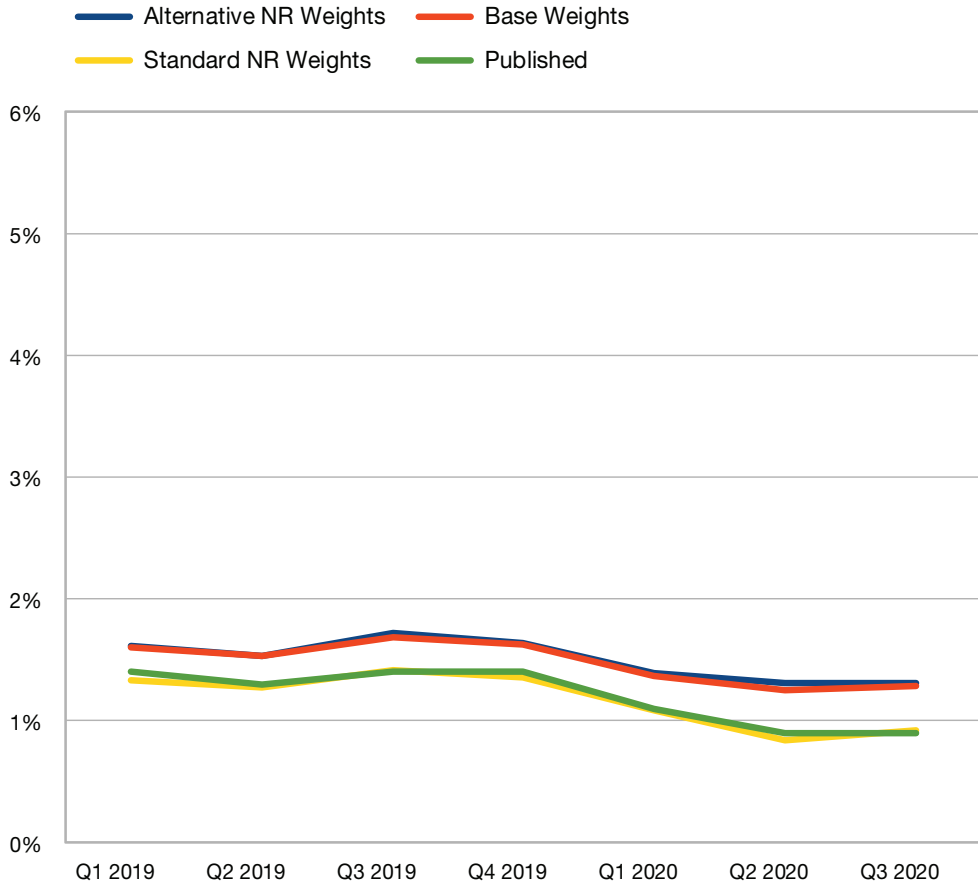Homeownership Rate Estimates by Quarter and Weighting Approach
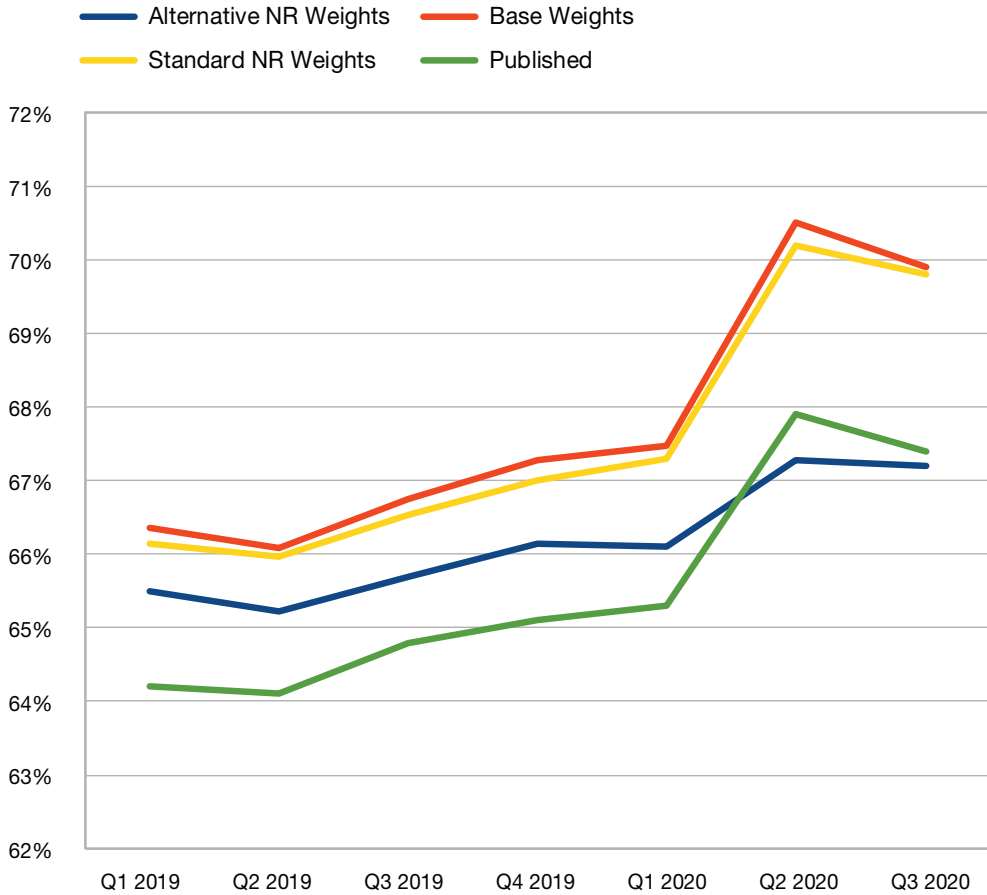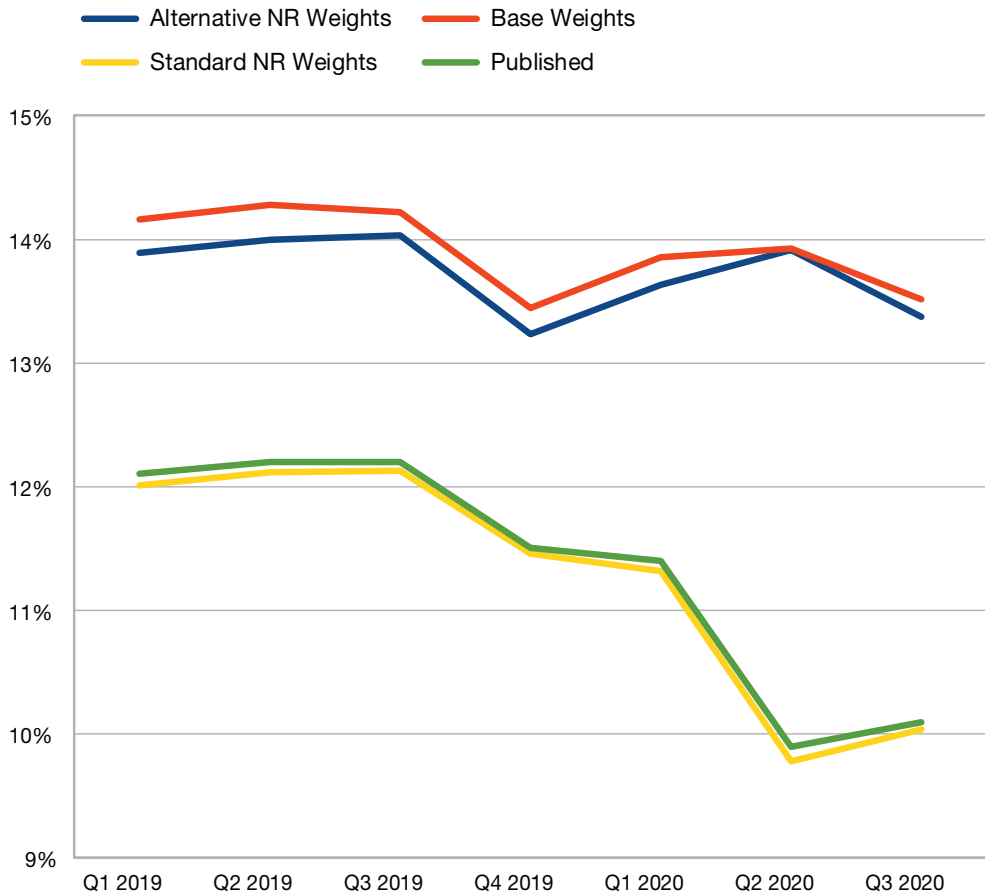

NR = nonresponse. Q = quarter.
Note: This exhibit visualizes the estimates reported in exhibit 3, along with the published Current Population Survey/Housing Vacancy Survey estimates using the final weights.

Exhibit 4B

Rental Vacancy Rate Estimates by Quarter and Weighting Approach

NR = nonresponse. Q = quarter.

Note: This exhibit visualizes the estimates reported in exhibit 3, along with the published Current Population Survey/Housing Vacancy Survey (CPS/HVS) estimates using the final weights.

Sources: U.S. Census Bureau; Current Population Survey and Housing Vacancy Survey Supplement data for 2019–2020 linked to the supplemental data sources identified in exhibit 1

Exhibit 4C

Homeowner Vacancy Rate Estimates by Quarter and Weighting Approach

NR = nonresponse. Q = quarter.

Note: This exhibit visualizes the estimates reported in exhibit 3, along with the published Current Population Survey/Housing Vacancy Survey (CPS/HVS) estimates using the final weights.

Sources: U.S. Census Bureau; Current Population Survey and Housing Vacancy Survey Supplement data for 2019–2020 linked to the supplemental data sources identified in exhibit 1

Exhibit 4D

Gross Vacancy Rate Estimates by Quarter and Weighting Approach

NR = nonresponse. Q = quarter.
Note: This exhibit visualizes the estimates reported in exhibit 3, along with the published Current Population Survey/Housing Vacancy Survey (CPS/HVS) estimates using the final weights.
Sources: U.S. Census Bureau; Current Population Survey and Housing Vacancy Survey Supplement data for 2019–2020 linked to the supplemental data sources identified in exhibit 1

Comparison of the homeownership rate estimates in exhibit 3 suggests that correcting for the observed sample composition changes using the alternative nonresponse adjustment factor significantly reduces the size of the homeownership rate increases estimated for the second and third quarters of 2020. The year-over-year increase in the Alternative NR Weights estimate of the homeownership rate is 2.0 percentage points in the second quarter of 2020, which is significantly smaller than the 4.4 percentage-point increase in the Base Weights estimates. Similarly, the year-over-year increase in the Alternative NR Weights estimate for the third quarter of 2020 is 1.5 percentage points, which is significantly smaller than the 3.2 percentage-point increase in the Base Weights estimates. Comparing the Alternative NR Weights estimates to the Standard NR Weights estimates, the year-over-year increases in the Alternative NR Weights estimates for the second and

third quarters of 2020 are also significantly smaller than the increases in the Standard NR Weights estimates, suggesting that the Alternative NR Weights estimates correct for sample composition changes that are not addressed by the existing nonresponse adjustment.

These results suggest that the homeownership rate increases in the second and third quarters of 2020 are influenced by the observed changes in sample composition in addition to changes in the true homeownership rate; however, there are important caveats. First, the alternative nonresponse weighting adjustment factor applied to the Alternative NR Weights estimates adjusts only for differences in observed attributes. There may be important additional changes in sample composition that are unobserved. Second, drawing inferences about the implications of these results for the published estimates based on the CPS/HVS final weight must also consider the additional adjustment factors and population controls applied to the published estimates.

Exhibit 3 presents similar estimates for the rental vacancy rate, homeowner vacancy rate, and gross vacancy rate. In contrast to the results for the homeownership rates, the Alternative NR Weights estimates for the vacancy rates closely track the Base Weights estimates in all quarters. The differences between the Alternative NR Weights estimates and the Base Weights estimates are less than 0.1 percentage points in all quarters for the rental vacancy rate and the homeowner vacancy rate and less than 0.3 percentage points in all quarters for the gross vacancy rate. Moreover, none of the Alternative NR Weights estimates of any vacancy rate are statistically different from the Base Weights estimates. One possible explanation for these results is that the observed changes in nonresponse captured by the regressions in exhibit 3 are not strongly correlated with vacancy— and therefore that adjusting for these changes does not substantially alter the vacancy rate estimates. However, an important caveat is that the covariates included in the logistic regressions may omit important attributes relevant to correcting for the impact of changing nonresponse patterns on vacancy.

An additional finding from exhibit 3 is that the Base Weights and Standard NR Weights estimates of the vacancy rates diverged in 2020 after moving roughly in tandem throughout 2019. For example, the difference between the Base Weights and Standard NR Weights estimates of the rental vacancy rate was either 1.3 or 1.4 percentage points in each quarter of 2019 before increasing to 2.7 percentage points by the second quarter of 2020. Because the only difference between the Base Weights and Standard NR Weights estimates is the choice of weights, these outcomes are due to changes over time in the effects of the nonresponse adjustment factor applied to occupied units in the current methodology. Specifically, the differences reflect the current methodology's assumption that all vacant units will be identified during in-person data collection attempts, so all Type A nonresponses are occupied units.

The current CPS/HVS weighting methodology calculates the nonresponse weighting adjustment using the pooled set of completed interviews and Type A nonresponses, excluding vacant units. Therefore, the resulting nonresponse weighting adjustment weights the occupied interviews up to the total of completed interviews plus Type A nonresponses, but does not include a similar adjustment for vacant units. This adjusted total of occupied units is then combined with the unadjusted total of vacant responses when the population control totals are applied to weight the units up to the total number of housing units in the United States. This sequence of steps relies

on the assumption that all vacant units will be identified during the in-person data collection attempts. The result is that any vacant unit that cannot be identified and is instead coded as a Type A nonresponse will increase the CPS/HVS estimate of occupied units and decrease the estimate of vacant units.

Under normal conditions, interviewers make multiple in-person data collection attempts with the goal of identifying as many vacant units as possible and minimizing the effect of this assumption on CPS/HVS estimates. Additionally, comparisons of the CPS/HVS estimates across quarters are made under the assumption that the effects of any remaining misclassifications are approximately constant across quarters, allowing for valid comparisons over time. The suspension of in-person data collection attempts in early 2020 has the potential to violate these assumptions, increasing the risks that vacant units might be missed and altering the data collection procedures used in different quarters. The consequence is that the Standard NR Weights and Published estimates of each vacancy rate will underestimate the true vacancy rate to the extent that vacant units were missed. The divergence between the Standard NR Weights and Base Weights estimates in exhibit 3 illustrates the extent to which these issues affected the CPS/HVS vacancy rate estimates for the second and third quarters of 2020 and must be considered when interpreting the estimates for that period.

Conclusion

The onset of the COVID-19 pandemic in early 2020 led to changes in the data collection procedures used for the HVS,, an important source of information about vacancy rates and the homeownership rate in the United States. This report describes the changes in data collection procedures and examines their implications for CPS/HVS estimates in 2020. The results illustrate the potential for the changes in data collection procedures to affect the CPS/HVS estimates of vacancy rates and the homeownership rate during this period. The results for the homeownership rate suggest that observed changes in sample composition likely contributed to the homeownership rate's increase in the second quarter of 2020. While the vacancy rate estimates are not found to be similarly sensitive to the observed changes in sample composition, the results suggest that the CPS/HVS vacancy rates likely underestimate the actual levels of vacancy in 2020 due to the weighting methodology's assumption that all nonresponding housing units are occupied. The conclusion is that data users should apply caution when interpreting CPS/HVS estimates for the period when in-person interviews were suspended.

Acknowledgments

Authors

Jonathan Spader is an assistant division chief in the Social, Economic, and Housing Statistics Division at the U.S. Census Bureau. Daniel Truver is a survey statistician in the Social, Economic,

and Housing Statistics Division at the U.S. Census Bureau. Peter Mateyka is a survey statistician in the Social, Economic, and Housing Statistics Division at the U.S. Census Bureau. Patricia Holley is a survey statistician in the Social, Economic, and Housing Statistics Division at the U.S. Census Bureau. Robert Callis is a branch chief in the Social, Economic, and Housing Statistics Division at the U.S. Census Bureau.

References

Bee, C.A., G.M.R. Gathright, and B.D. Meyer. 2015. Bias from Unit Non-Response in the Measurement of Income in Household Surveys. Unpublished paper, August 2015. Chicago, IL: University of Chicago.

Brummet, Q. 2014. Comparison of Survey, Federal, and Commercial Address Data Quality. Working paper 2014-06. CARRA. Suitland, MD: U.S. Census Bureau.

Brummet, Q., D. Flanagan-Doyle, J. Mitchell, J. Voorheis, L. Erhard, and B. McBride. 2018. "What Can Administrative Tax Information Tell Us About Income Measurement in Household Surveys? Evidence from the Consumer Expenditure Surveys," *Statistical Journal of the IAOS* 34 (4): 513–520.

Eggleston, J., and A. Westra. 2020. Incorporating Administrative Data in Survey Weights for the Survey of Income and Program Participation. Working paper 2020-07. SEHSD. Suitland, MD: U.S. Census Bureau.

Rothbaum, J., and C.A. Bee. 2020. Coronavirus Infects Surveys, Too: Non-response Bias During the Pandemic in the CPS ASEC. Working paper 2020-10. SEHSD. Suitland, MD: U.S. Census Bureau.

Sabelhaus, J., D. Johnson, S. Ash, D. Swanson, T. Garner, J. Greenlees, and S. Henderson. 2015. "Is the Consumer Expenditure Survey Representative by Income?" In *Improving the Measurement of Consumer Expenditures*, edited by Christopher D. Carroll, Thomas F. Crossley, and John Sabelhaus. Chicago, IL: University of Chicago Press.

Spader, Jonathan, Daniel Truver, Peter Mateyka, Patricia Holley, and Robert Callis. 2021. Is Homeownership Actually Rising and Vacancy Falling? Understanding Housing Vacancy Survey Estimates for 2020. Working paper 2021-04. SEHSD. Suitland, MD: U.S. Census Bureau.

U.S. Census Bureau. 2019. Current Population Survey Design and Methodology Technical Paper 77.

Wagner, D., and M. Layne. 2014. The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software. Working paper 2014-01. CARRA. Suitland, MD: U.S. Census Bureau.