2021 New York City
Housing and Vacancy Survey

Sample Design, Weighting, and
Error Estimation

U.S. Census Bureau, Department of Commerce
New York City Department of Housing Preservation & Development

1. Overview

This document describes the sample design, weighting, and error estimation for the 2021 New York City Housing and Vacancy Survey (NYCHVS). The NYCHVS is sponsored by the New York City Department of Housing Preservation and Development (HPD) and conducted by the U.S. Census Bureau.

The City of New York is required by law to conduct a survey periodically to determine if rent regulations should be continued. A primary tool in this decision is the *"vacant available for rent"* rate, which is defined as the ratio of the vacant available for rent units to the total number of renter-occupied and vacant available for rent units for the entire city. The NYCHVS measures rental and homeowner vacancy rates, as well as various household and person characteristics. The design requires the standard error of the rental vacancy rate for the entire city to be no more than one-fourth of one percent if the actual rate was three percent.

2. Sample Design

The NYCHVS is a longitudinal survey that is conducted about every three years. The main sample of the survey is selected every decade, and additional new units are selected in each subsequent NYCHVS cycle. For the decade 2020-2030, the NYCHVS was conducted in 2021, with plans for additional survey years of 2023, 2026 and 2029. The main sample was selected using the 2020 July Master Address File (MAF); additional sample units plan to be selected in 2023, 2026, and 2029.

2.1. Eligible Universe

The universe of interest for the NYCHVS consists of the residential housing units (HUs) located within the five boroughs of New York City (Bronx, Brooklyn, Manhattan, Queens, and Staten Island). The principal exclusions are living quarters at locations that are classified as group quarters. These include:

- Correctional facilities,
- Mental health institutions,
- Hospitals,
- Military installations,
- Convents, monasteries, and rectories,
- Shelters, group homes, communes, and halfway houses,
- Home for the aged, disabled, homeless, or needy, and
- Dormitories for students or workers.

2.2. Sampling Frames

The 2021 NYCHVS frame is constructed using several files, including the July 2020 Master Address File (MAF) extract, the American Community Survey (ACS) 2014-2018 5-year file, and several administrative files from HPD. Valid, residential HUs on the July 2020 MAF extract comprise the frame for the 2021 NYCHVS sample selection process, and all other files are used to add variables needed for sorting and strata assignments. MAFIDs (the

primary HU identifying variable on the MAF) were appended to the administrative files. The administrative files were merged using MAFID with a block-level version of the ACS file and the July 2020 MAF to create the 2021 NYCVHS sample frame.

The frame was created based on the 2020 July MAF instead of the 2020 Census because the 2020 Census was not available at the time the sample was selected.

2.3. Sample Selection

Housing units on the frame were divided into several strata based on subsidized program participation and Condo or COOP status. A minimum sample size was determined using 2017 NYCHVS data with a target goal of 30,000 sampled units, which was later reduced to 12,000 (due to constraints caused by COVID). The sampling design used two variables, Strata 1 and Strata 2, each of which had several categories within them. The sample sizes for each category are presented in Tables 2.1, 2.2, and 2.3 below, first shown by the marginal counts in each variable in Tables 2.1 and 2.2 and then by the cross-sectional counts in Table 2.3.

Table 2.1 Strata 1: First Housing Type Allocation

Program Type	Frame Total	30,000 Sample	12,000 Sample
Affordable Owner	21,983	1,144	458
Affordable Renter	92,622	1,205	482
Condo	287,259	2,776	1,110
COOP	405,779	3,650	1,460
Other	2,616,491	17,692	7,077
Public Housing	161,926	2,658	1,063
Rent Control	12,294	880	352

Table 2.2 Strata 2: Second Housing Type Allocation

Program Type	Frame Total	30,000 Sample	12,000 Sample
Mitchell Lama COOP	62,669	1,309	524
Mitchell Lama Renter	28,256	1,085	434
Other	2,432,618	19,869	7,947
Rent Stabilized	1,074,811	7,742	3,097

Table 2.3 Allocation of Minimum Sample

Strata 1	Strata 2	Frame Total	30,000 Sample	12,000 Sample
Affordable Owner	Mitchell Lama COOP	12,551	562	225
Affordable Owner	Other	7,962	507	203
Affordable Owner	Rent Stabilized	1,470	75	30
Affordable Renter	Mitchell Lama COOP	509	10	4
Affordable Renter	Mitchell Lama Renter	7,486	288	115
Affordable Renter	Other	28,098	288	115
Affordable Renter	Rent Stabilized	56,529	619	248
Condo	Mitchell Lama Renter	219	8	3
Condo	Other	221,661	2,135	854
Condo	Rent Stabilized	65,379	633	253
COOP	Mitchell Lama COOP	49,609	737	295
COOP	Mitchell Lama Renter	1,614	62	25
COOP	Other	290,376	2,290	916
COOP	Rent Stabilized	64,180	561	224
Other	Mitchell Lama Renter	18,937	727	291
Other	Other	1,711,215	11,176	4,470
Other	Rent Stabilized	886,339	5,789	2,316
Public Housing	Other	161,926	2,658	1,063
Rent Control	Other	11,380	815	326
Rent Control	Rent Stabilized	914	65	26

A systematic random sample of housing units was selected within each cross-sectional program type in Table 2.3, sorting housing units by
- Borough
- Sub-borough
- Tract
- Median Income by Tract based on 2014-2018 ACS
- Block number
- Binary variable indicating building size as "big" or "small"
- Basic street address
- Unit designation

2.4. Interviews and Response Rates
The total number of sample housing units selected for the 2021 NYCHVS was 12,002. Table 2.4 provides the weighted and unweighted response rates by borough, as well as the distribution of completed interviews and noninterviews.

Table 2.4. Interview Activity for the 2021 New York City Housing and Vacancy Survey

Borough	Unweighted Response Rate	Weighted Response Rate	Selected	Completed Interviews	Type A Non-interviews	Type B & C Non-interviews
Bronx	73%	##%	1,978	1,414	513	51
Brooklyn	74%	##%	3,573	2,501	884	188
Manhattan	74%	##%	3,320	2,354	835	131
Queens	72%	##%	2,607	1,773	683	151
Staten Island	71%	##%	524	352	142	30
Total	73%	##%	12,002	8,394	3,057	551

Source: U.S. Census Bureau, 2021 New York City Housing and Vacancy Survey.
Note: The data are subject to error arising from a variety of sources.

In past cycles (prior to 2017), the NYCHVS conducted proxy or last resort interviews, where a proxy interview consisted of interviewing a real estate agent, building manager, or someone else knowledgeable about the HU and a last resort interview involved accepting an abbreviated questionnaire as complete for reluctant respondents. In 2021, NYCHVS did not conduct proxy or last resort interviews, which resulted in higher noninterviews.

In 2021, Type A noninterviews included occupied housing units where the occupants:
- Refused to be interviewed,
- Absent due to covid,
- Unable to locate,
- Were not at home after repeated visits, or
- Were unavailable for some other reason.

Type A noninterviews also include vacant units. In these cases, an interview was not obtained if no informed respondent could be found after repeated visits.

Type B noninterviews were not interviewed because they cannot be inhabited, such as under construction or set to be demolished.

Type C noninterviews were not interviewed because they did not meet the definition of a housing unit, such as the unit no longer existed or was uninhabitable.

The response rate is calculated as the total number of interviews divided by the total eligible sample, which can be written as:

$$Response\ Rate = \frac{Total\ Sample - Type\ A\ noninterviews - Type\ B \wedge C\ noninterviews}{Total\ Sample - Type\ B \wedge C\ noninterviews}$$

Note that the weighted response rate just applies the base weight of each HU to the counts.

For calculating response rates, enough of the interview had to be completed for it to be considered a valid interview. For vacant interviews, the entire interview must be completed. For non-vacant interviews, all of the following must be answered to be considered as a completed interview:

- Occupancy/vacancy status,
- Tenure,
- Type of vacant unit, and
- Reason Unit is not available for rent or sale

AND two of the following five items answered from the household roster for each person:

- Sex,
- Age,
- Relationship to householder,[1]
- Hispanic origin, and
- Race.

If these criteria were not met, the sampled unit was classified as a Type A noninterview, following the definitions above.

For evaluation of interviews, a second interview was conducted of all vacant units and five percent of all occupied units. The questions asked during the reinterview included information about the previous FRs that collected data, the time, date, and length of that interview, tenure, and vacancy status.

3. Weighting

To estimate HU and person characteristics for the 2021 NYCHVS, sample weights are calculated for each sample HU and each sample person. The final weight for each HU is the product of the following weights and adjustments:

3.1.   Base Weight

The base weight is the reciprocal of the probability of selecting the unit. This is simply the inverse of the rate at which HUs are selected within the cross-sectional breakdowns in Table 2.3. Note that this sample design resulted in an unequal probability of selection for HUs.

3.2.   Nonresponse Adjustment

---

[1]   In very rare situations, having only relationship to householder was considered sufficient if the persons could be verified as real people through other methods.

The base weight of each interviewed HU was adjusted to account for the eligible units that did not respond (Type A noninterviews). This nonresponse adjustment was applied using a noninterview adjustment factor (NAF), which was applied to all interviewed HUs to account for Type A noninterviews. The factor was calculated using the following ratio:

$$NAF = \frac{Interviews + Type\ A\ noninterviews}{Interviews}$$

A new method of calculating the NAF was introduced for the 2021 survey cycle. This involved estimating the probability of response from the responding and nonresponding HUs and grouping HUs with similar response propensities together for this adjustment. Note that some nonresponding HUs were excluded from this modeling because they could not be found within the 2020 Census. More information can be found in the XXXX INSERT LINK TO NAF MODELING DOC

3.3. Ratio Adjustment Factors for Housing Unit Weights

New methods for implementing ratio adjustment factors (RAFs) within NYCHVS were also introduced. The HU weights were adjusted using two main sources of known totals:
- The July 2021 MAF
- Totals by program type from HPD administrative files

At each step in the ratio estimation procedure, the factors were equal to the following ratio:

$$\frac{Known\ Totals}{NYCHVS\ Sample\ Estimate}$$

The denominators of the ratios are equal to the sum of the weights of HUs (or persons) with all previous factors applied. Appendix A includes more information on the ratio adjustment factors and examples on how the process works.

The updated process creates three partitions of the sample based on borough membership, subsidized program status, and Condo/COOP, similar but not identical to the sample design. The weighting program then iterates through these three partitions until the RAF factor at each iteration stabilizes and final estimated totals of the groups within these partitions equal their known totals. The partitions are listed below.
- Partition 1: Affordable Owners, Affordable Renters, Remainder – City-wide
- Partition 2: Mitchell Lama Renter, Mitchell Lama COOP, Remainder – by Borough
- Partition 3: Public Housing, Condo, COOP, Remainder – by Borough

Note that Partition 1 was done city-wide while Partitions 2 and 3 were done by borough. "Remainder" means any HU not fitting into the other categories in the partition. Estimates of total HUs made of these particular program types at the specified level of geography match their known totals.

3.4. Ratio Adjustment Factors for Person Weights

4. Nonsampling Errors

All numbers from the NYCHVS, except for sample size, are estimates. As in other surveys, two types of general errors occur: sampling errors and nonsampling errors. Sampling errors are discussed in Section 5. The definition of nonsampling errors is—

> Nonsampling errors arise mainly due to misleading definitions and concepts, inadequate sampling frames, unsatisfactory questionnaires, defective methods of data collection, tabulation, coding, incomplete coverage of sample units, and so on. These errors are unpredictable and not easily controlled. Unlike sampling error, this error may increase with increases in sample size. If not properly controlled, nonsampling error can be more damaging than sampling error for large-scale household surveys.[2]

The various types of nonsampling errors are discussed in the following sections.

4.1. Coverage Error

Coverage errors arise from the failure to give some units in the target population any chance of selection into the sample (undercoverage) or giving units more than one chance of selection (overcoverage). To calculate the coverage, we used the sample base weight, which is the weight prior to any sample adjustments. The sample adjustments, described in Section 3, help to mitigate the undercoverage identified in this section.

The coverage rate is the ratio of the survey population or HU estimate of a group or an area and the independent estimate (or the known totals). The undercoverage rate is calculated as:

$$Undercoverage = \left(1 - \frac{NYCHVS\ Sample\ Estimate}{Known\ Totals}\right) * 100$$

Table 4.1a indicates the estimated undercoverage for both HUs and persons. Based on the July 2021 MAF, we missed about eight percent of the housing units in the five boroughs. Overall, we missed about ## percent of the people in sample households.

Table 4.1a Overall Undercoverage for Housing Units and Persons

	Estimated from 2021 NYCHVS Base Weights	Known Total	Undercoverage
Housing Units	3,406,477*	3,644,065	7.53%
Persons	###	###	###%

Source: U.S. Census Bureau, 2021 New York City Housing and Vacancy Survey.
Note: The data are subject to error arising from a variety of sources.

---

[2] https://unstats.un.org/unsd/demographic/meetings/egm/Sampling_1203/docs/no_7.pdf

Table 4.1b provides the various sources of undercoverage for HUs.

Table 4.1b Undercoverage by Source for Housing Units

Source of Undercoverage	Undercoverage
Growth from July 2020 MAF to July 2021 MAF	1.26%
Type B&C Drops	5.26%
Nonrespondents not matched to 2020 Census	1.01%
Total	**7.53%**

Source: U.S. Census Bureau, internal data files.

The within-household undercoverage varied by age, race, sex, and borough. Table 4.1c gives the undercoverage of the various race-sex groups for the city as a whole.

Table 4.1c Undercoverage by Race/Ethnicity-Sex Group

Race/Ethnicity-Sex Group	Undercoverage
White & Other Females	#%
White & Other Males	#%
African American Females	#%
African American Males	#%
Asian Females	#%
Asian Males	#%
Hispanic Females	#%
Hispanic Males	#%

Source: U.S. Census Bureau, 2021 New York City Housing and Vacancy Survey.
Note: The data are subject to error arising from a variety of sources.

We adjusted for this undercoverage through the HU and person ratio adjustment factors. These factors adjust the sample weights to population totals provided by the Census Bureau, so the resulting final weight accounts for the undercoverage identified in Tables 4.1a and 4.1b. For more information on the sample adjustment process, see Sections 3.3 and 3.4. NYCHVS data users do not have to take any additional steps to account for coverage error.

4.2. Nonresponse Error

Some respondents refuse the interview or cannot be located. The Census Bureau mitigated the error due to nonresponse by applying the noninterview adjustment factors into the weighting process, as discussed in Section 3.2. NYCHVS data users do not have to take any additional steps to account for nonresponse error.

4.3. Measurement Error from Missing Responses to Questions

Some respondents participate in an interview but refuse to answer questions or do not know a particular answer. For certain questions, the Census Bureau imputes missing responses. When imputing, the Census Bureau tries to find households or persons with similar characteristics to fill in missing data. For each imputation, records are divided into 'donors' and 'recipients'.

For the demographic items, the Census Bureau first tries to impute based on other household information or household members. Every household must have some demographic information for each person in the household or it would be made a Type A. It is rare that a household is missing all demographic information for one item.

For imputing the housing items (including housing quality questions), units with similar characteristics are grouped. For example, when imputing contract rent, a unit with a similar year moved, year-built range, number of bedrooms, and input control status (stabilized, public housing, unregulated, etc.) is found and unit's contract rent is used to impute the recipient's contract rent. If no such unit can be found, contract rent is imputed based on the median value for units in 2017 (adjusted for inflation) with the same input control status in the particular borough. In 2021, contract rent was imputed in 11.6 percent of the renter-occupied units. For some characteristics (like mortgage information), where we didn't collect similar data in the 2017 NYCHVS, we used data from the 2019 American Housing Survey, for starting values when, in rare cases no donor could be found.

For economic items, such as income and employment status, the best possible match between donors and recipients is achieved through a statistical match with key items. The items used for matching donors and recipients are public assistance/non public assistance, borough, tenure, gender, race, ethnicity, age, relationship, education, worked last week, hours worked, type of work, government/nongovernment, and rent/value. All of these criteria are used to get the best statistical match possible. There are 33 income variables in 2021; in rare cases where a suitable donor was not found, income amount is imputed based on the median value of that income category. All donors and recipients have the same borough, tenure, and either receive public assistance or do not.

Appendix B provides the list of variables being imputed. Variables that can be used to determine imputation rates are in the public use files (PUF) and are defined on the record layout. Variables shown in Table B1 are for occupied units, Table B2 for persons, and Table B3 for vacant units. For example, using these variables from the PUF, users can see that

summer gas and electricity costs were imputed for 11.4 percent of occupied units, age was imputed for 6.3 percent of all persons, and stories was imputed for 0.1 percent of vacant units.

The Census Bureau does not know how close the imputed values are to the actual values. For other items, "not reported" is used as an answer category. NYCHVS data users do not have to take any additional steps to account for measurement error from missing responses to questions.

4.4.   Quality Validity Error

In order to design a survey question that accurately measures the constructs of interest, the Census Bureau carefully tests each new survey question to ensure it is measuring the construct of interest. While the questionnaire is provided in multiple languages, sometimes the respondent does not speak those languages. In these cases, the interview must be rescheduled so that a field representative (FR) that speaks the same language as the respondent can administer the interview. Although some respondents might misinterpret questions, the Census Bureau does not have any additional information to estimate validity error rates. NYCHVS data users do not have to take any additional steps to account for validity error.

4.5.   Processing Error

After the data are collected, errors that can be introduced include data capture errors and data editing and imputation errors. The Census Bureau carefully tests all aspects of the data capture and the editing and imputation procedures. Although mistakes are possible, the Census Bureau believes they are minimal. If a processing error is discovered, the Census Bureau will let NYCHVS data users know and, in some cases, will publish revised estimates. NYCHVS data users do not have to take any additional steps to account for processing error.

4.6.   Additional Considerations

The NYCHVS is a longitudinal survey conducted about every three years. Many NYCHVS users compare current year NYCHVS with prior year estimates. Users should be aware that HPD and the Census Bureau often make changes to the text of various questions between surveys and sometimes to the underlying weighting methodology or sample design. NYCHVS data users comparing estimates with prior year surveys should consult the 'Overview' document on the NYCHVS website (https://www.census.gov/programs-surveys/nychvs/about/overview.html).

5. Sampling Errors and Replicate Weights

Sampling error is a measure of how estimates from a sample vary from the actual value. By the term "actual value," we mean the value we would have gotten had all HUs been interviewed, under the same conditions, rather than only a sample.

Users of NYCHVS PUF should use replicate weights to estimate errors for any estimate. This is different from prior survey cycles, where the Census Bureau provided generalized variance function (GVF) parameters as an alternative method to estimate variance.

Variance estimation for surveys refers to the variation of an estimate due to selecting a sample from the set of all possible samples for a given sample design. To estimate the variance, multiple samples are needed but only one is observed. Replication allows small changes to a single probability sample to create a set of replicate samples, which can then be used to measure the variation of the estimates. Replication is done through subsets selected from the original sample in a process that mimics the original sample design. Each replicate sample, r, is then fully weighted, using the same process as the original sample, to ensure it represents the population of interest. This process forms the set of final replicate weights $\{w_r \mid r = 1, \dots, R\}$. Considering a particular estimate of interest, each replicate weight, w_r, can be used to create a replicate estimate $\hat{\theta}_r$. The set of replicated estimates $\{\hat{\theta}_r \mid r = 1, \dots, R\}$ represents the variability, or dispersion, of the estimate of interest under multiple samples of the population.

The Hadamard matrix was used to derive replicate factors to apply to individual full sample weights in creating replicate weights. Please note that for 2021 NYCHVS, the weights in Replicate 1 equal full sample weights, the weight used to derive sample estimates.

The 2021 NYCHVS uses a replicate variance estimator derived from a variance equation called the successive differences estimator. This estimator was first introduced by Fay and Train (1995) and then expanded for replication by Ash (2014). Using the replicated estimates, data users can calculate an estimated variance of an estimate of interest using the replication variance estimator:

$$\hat{v}\left(\hat{\theta}\right) = \frac{4}{80} \sum_{r=1}^{80} \left(\hat{\theta}_r - \hat{\theta}_0\right)^2$$

where $\hat{\theta}$ is the weighted point estimate of the statistic of interest, such as a total, median, mean, proportion, regression coefficient, or log-odds ratio, using the weight for the full sample and $\hat{\theta}_r$ is the replicate estimate for replicate r of the same statistic using the replicate weights. $\hat{\theta}_0$ is the full sample estimate. The value of 80 in $\hat{v}\left(\hat{\theta}\right)$ is the number of replicates used (NYCHVS uses 80 replicates).

There are two sets of replicate weights. One set of replicate weights is used for computing standard errors of housing unit characteristics and the second set is used for computing standard errors of person characteristics.

To calculate a standard error, the measure of dispersion when parameter estimates are calculated through repeated sampling from the population, obtain the square root of the variance estimate.

The following example illustrates how a statistic would be estimated, replicated, and combined to form a variance estimate. The goal of this example is to estimate the total number of renter-occupied HUs in Queens for 2021 and its corresponding estimate of variance.

For 2021, there are 893 completed interviews that are renter-occupied HUs in Queens (sample cases in Queens with responses to tenure status question as renters). Table 6.1 displays the first four and last one renter-occupied sample units in Queens. Note that the ordering in Table 6.1 is based on the variable *CONTROL*.

Table 6.1: Example of Estimating Variances with Replication of Renters in Queens

Sample HU	Full Sample Weight	Replicate Weights				
		Replicate 1	Replicate 2	Replicate 3		Replicate 80
1	432.920	432.920	430.180	724.691	...	128.270
2	1,223.797	1,223.797	2,071.712	351.266	...	365.787
3	36.056	36.056	35.119	11.083	....	61.436
4	476.777	476.777	470.671	793.062	...	485.903
...	...	...	...	...	...	...
893	432.899	432.899	125.991	724.655	...	747.576

Source: U.S. Census Bureau, 2021 New York City Housing and Vacancy Survey.
Note: The data are subject to error arising from a variety of sources.

In NYCHVS, the full sample weight and the full sample estimate are referred to as replicate weight 0 (w_0) and replicate estimate 0 ($\hat{\theta}_0$), respectively.

Step 1: Calculate the full sample weighted survey estimate.

The statistic of interest is the total number of renter-occupied housing units in Queens for 2021. Add the full sample weights of the sample cases that meet your criteria of interest. Therefore, the total number of renter-occupied housing units in Queens is calculated as follows:

Full Sample Renter-Occupied HUs in Queens Estimate:

$$\hat{\theta}_0 = \sum_{i=1}^{893} w_{0,i} = 432.920 + 1{,}223.797 + \ldots + 432.899 = 467{,}730$$

Step 2: Calculate the weighted survey estimate for each of the replicate samples.

The replicate survey estimates are as follows:

Rep 1 Rent-Occ Estimate $\quad \hat{\theta}_1 = \sum_{i=1}^{893} w_{1,i} = 432.920 + 1{,}223.797 + \ldots + 432.899 = 467{,}730$

Rep 2 Rent-Occ Estimate $\quad \hat{\theta}_2 = \sum_{i=1}^{893} w_{2,i} = 430.180 + 2{,}071.712 + \ldots + 125.991 = 468{,}933$

Rep 3 Rent-Occ Estimate $\quad \hat{\theta}_3 = \sum_{i=1}^{893} w_{3,i} = 724.691 + 351.266 + \ldots + 724.655 = 476{,}408$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

Rep 80 Rent-Occ Estimate $\quad \hat{\theta}_{80} = \sum_{i=1}^{893} w_{80,i} = 128.270 + 365.787 + \ldots + 747.576 = 469{,}330$

Step 3: Use the replicate estimates $\hat{\theta}_r$ in the formula below to calculate the variance estimate for the total renter-occupied HUs in Queens.

$$\hat{v}(\hat{\theta}) = \frac{4}{80} \sum_{r=1}^{80} (\hat{\theta}_r - \hat{\theta}_0)^2$$
$$¿ 0.05 \times \left[(467{,}730 - 467{,}730)^2 + (468{,}933 - 467{,}730)^2 + (476{,}408 - 467{,}730)^2 + \ldots + (469{,}330 - 467{,}7 \right.$$
$$¿ 0.05 \times \left[0 + 1{,}449{,}144 + 75{,}322{,}223 + \ldots + 2{,}560{,}800 \right]$$
$$¿ 117{,}126{,}870$$

The survey estimate for total renter-occupied population in Queens is 467,730 housing units, with an estimated variance of 117,126,870 or a standard error of 10,822.5 housing units.

6. References

Fay, R. E. and Train, G. F. (1995). Aspects and Survey and Model-based Postcensal Estimation of Income and Poverty Characteristics for States and Counties. *Proceeding of the Sections on Government Statistics,* American Statistical Association, 154-159.

Ash, S. E. (2014) Using Successive Difference Replication for Estimating Variances. *Survey Methodology,* Statistics Canada, Catalogue no.12-001-X Business Survey Method Division, Vol. 40, No.1, pp.47-59.

Lohr, S.L. (2007). Recent developments in multiple frame surveys. Proceedings of the Survey Research Methods Section, American Statistical Association, 3257-3264. Accessed online at http://www.amstat.org/sections/srms/Proceedings/ on September 1, 2015.

Lohr, S. (2010). "Dual Frame Surveys: Recent Developments and Challenges," paper presented at the Scientific Meeting of the 45th Italian Statistical Society, Padua, Italy, June 16-18.

U.S. Census Bureau (2014). Overview. https://www2.census.gov/programs-surveys/nychvs/about/overview/overview-2014.pdf. Date retrieved April 23, 2018.

U.S. Census Bureau (2017a). Methodology for United States Population Estimates: Vintage 2017. https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2010-2017/2017-natstcopr-meth.pdf Date retrieved April 23, 2018.

U.S. Census Bureau (2017b). Overview. https://www2.census.gov/programs-surveys/nychvs/about/overview/overview-2017.pdf. Date retrieved April 23, 2018.

Appendix A. Example of Ratio Adjustments

This appendix provides one hypothetical example that demonstrates how the sample weights are ratio-adjusted so that they are consistent with a set of control totals.

For this example, assume weights were calculated for a sample, including all weighting adjustments up to a nonresponse adjustment. With these weights, totals by two categories – simply identified as A or B for Category 1 and C or D for Category 2 – were created. Table A1 summarizes the estimated totals resulting from the hypothetical sample and weights, and Table A2 shows the hypothetical control totals.

Table A1: Example Estimated Totals

Cat1\Cat2	C	D	Total
A	110	91	201
B	97	107	204
Total	207	198	405

Table A2: Example Control Totals

Cat1\Cat2	C	D	Total
A	115	105	220
B	95	105	200
Total	210	210	420

The control totals of Table A2 are used to improve the weights by making the estimates from the weights consistent with the control totals. Table A3 shows the Ratio Adjustment Factor (RAF) that will make the estimated totals consistent with the control totals.

Table A3: Example Ratio Adjustment Factors

Cat1\Cat2	C	D
A	115/110 = 1.0455	105/91 = 1.1583
B	95/97 = 0.9794	105/107 = 0.9813

If the factors from Table A3 are multiplied to the weights of the sample units, then the estimates from the revised weights will be consistent with the totals of Table A2.

For example, the ratio-adjusted weights for the combination of Category 1 = A and Category 2 = C is the product of the original weight and the RAF for the A/C combination:

$$\text{Ratio-adjusted weight} = \text{original weight} \; ¿ \; 1.0455$$

The ratio-adjusted weights for the other combinations of Categories 1 and 2 are calculated in the same way, using the corresponding RAF for each combination.

Appendix B: List of Variables Imputed for 2017 New York City Housing and Vacancy Survey

Occupied Units	
Item Name	**Variable Name**
Number of stories in building	STORIES_FRONT
Need stairs from sidewalk to elevator	ELEV_NO_STEPS
Need stairs from sidewalk to unit	UNIT_NO_STEPS
Broken heat last winter	NOHEAT
Leaks in unit in the last year	LEAKS
Rodents in unit in the last 3 months	RODENTS_UNIT
Rodents in building in last 3 months	RODENTS_BUILD
Toilets not working for six or more hours in the last 3 months	TOILET_BROK
Cracks or holes in walls or ceiling of unit	WALLHOLES
Holes in the floors of unit	FLOORHOLES
Broken plaster or peeling paint on ceilings or walls of unit	PEELPAINT
Broken plaster or peeling paint is larger than an 8.5x11 piece of paper	PEELPAINT_LARGE
Number of times no heat in winter for six hours or more	NOHEAT_NUM
Used additional heating sources in winter	ADDHEAT
Number of bedrooms	BEDROOMS
Number of rooms	ROOMS
Number of full bathrooms in unit	FULLBATH_NUM
Number of half bathrooms in unit	HALFBATH_NUM
Amount paid for rent, including fees	RENT_AMOUNT
Was amount reported in RENT_AMOUNT paid to landlord	RENTPAID
Amount paid was different than the amount owed last month	RENTPAID_AMOUNT
rent paid by outside sources	RENTOUTSIDE
Amount of rent paid by outside sources	RENTOUTSIDE_AMOUNT
Rent paid by rental assistance programs - Section 8/Housing Choice Voucher	RENTASSIST_VOUCHER
Rent paid by rental assistance programs - Shelter Allowance/City FHEPS	RENTASSIST_SA
Rent paid by rental assistance programs - SCRIE/DRIE	RENTASSIST_RIE
Rent paid by rental assistance programs - Other assistance that pays part of your rent	RENTASSIST_OTHER
Rent paid by rental assistance programs - None	RENTASSIST_NONE
Amount paid by rental assist programs	RENTASSIST_AMOUNT
Year apartment/house was purchased or inherited	PURCHASEYEAR
Purchase price of apartment/house	PURCHASEPRICE
Housing Debt - First Mortgage	HDEBT_FIRSTMORT
Housing Debt - Second Mortgage	HDEBT_SECONDMORT

Occupied Units	
Item Name	*Variable Name*
Housing Debt - Heloc Mortgage	*HDEBT_HELOC*
Housing Debt - Home equity Mortgage	*HDEBT_HOMEEQUITY*
Housing Debt - Reverse Mortgage	*HDEBT_REVMORT*
Housing Debt - Other Mortgage	*HDEBT_OTHER*
Housing Debt - No Current Mortgage	*HDEBT_NONE*
Amount of most recent payment for first mortgage	*PAY_FIRSTMORT*
Frequency of loan payments on first mortgage	*FREQPAY_FIRSTMORT*
Frequency of loan payments - specify	*FREQPAYOTH_FIRSTMORT*
Outstanding principal balance	*TOTAL_FIRSTMORT*
Current interest rate on loan - whole number	*INT1_FIRSTMORT*
Current interest rate on loan - fraction	*INT2_FIRSTMORT*
Fixed interest rate on loan	*FIXED_FIRSTMORT*
Utilities paid – Electricity	*UTIL_ELECTRIC*
Utilities paid - Cooking Gas	*UTIL_GAS*
Utilities paid – Heat	*UTIL_HEAT*
Utilities paid - Water/Sewer	*UTIL_WATER*
Utilities paid - None of these	*UTIL_NONE*
Utilities paid - All utilities are included in the rent or condo/co-op fees	*UTIL_INCLUDED*
Gas and Electric costs – Summer	*UTILCOSTS_SUMMER*
Gas and Electric costs – Winter	*UTILCOSTS_WINTER*
Total cost of heat	*UTILCOSTS_HEAT*
Total cost of water and sewer	*UTILCOSTS_WATER*

Source: U.S. Census Bureau, 2021 New York City Housing and Vacancy Survey.

Table B2: List of Variables Imputed for Persons

Persons	
Item Name	**Variable Name**
Gender of person	GENDER_P
Race of person – White	RACE_W_P
Race of person - Black or African American	RACE_B_P
Race of person - American Indian or Alaska Native	RACE_AIAN_P
Race of person - Asian or Asian American	RACE_A_P
Race of person - Native Hawaiian or Other Pacific Islander	RACE_NHOP_P
Race of person – Other	RACE_OTH_P
Which Asian group – Chinese	ASIANORIG_CH_P
Which Asian group - Asian Indian	ASIANORIG_AI_P
Which Asian group – Filipino	ASIANORIG_FIL_P
Which Asian group – Korean	ASIANORIG_KOR_P

Persons	
Item Name	**Variable Name**
Which Asian group – Japanese	ASIANORIG_JAPA_P
Which Asian group – Vietnamese	ASIANORIG_VIET_P
Which Asian group - Something else	ASIANORIG_ELSE_P
Hispanic origin of person	HISP_P
Identify with an indigenous people or tribal group	HISPINDIG_P
Hispanic, Latino, or Spanish heritage - Puerto Rican	HISPORIG_PR_P
Hispanic, Latino, or Spanish heritage - Dominican	HISPORIG_DR_P
Hispanic, Latino, or Spanish heritage - Cuban	HISPORIG_CUB_P
Hispanic, Latino, or Spanish heritage - South/Central American	HISPORIG_SCA_P
Hispanic, Latino, or Spanish heritage - Mexican-American, Mexican, Chicano	HISPORIG_MEX_P
Hispanic, Latino, or Spanish heritage - Something else	HISPORIG_ELSE_P
Age of person	AGE
Educational level of person	EDUC_P
Specify grade currently attending	EDUCSP_P
Asks household members if they were enrolled in school last week.	SCHLNOW_P
Type of school/program person is attending	SCHLNOW_TYPE_P
Specify grade currently attending	SCHLNOW_TYPE_SP
Year when person moved into household	MOVEIN_P
Amount of rent paid by each person	RENTPAID_P
Did person work for pay	WORK_P
How many part-time and full-time jobs	WORKJOBS_P
When did you last work	WORKLAST_P
work every week of the last year	WORK52_P
How many weeks did person work	WORKWEEKS_P
How many hours did person unusually work a week	WORKHOURS_P
Type of Government Work	WORKGOV_P
Does person work daytime schedule	WORKSCHED_P
Is person an owner or partner in business	BUSINESS_P
Is business incorporated	BUSINESSINC_P
Is Person employee of that business	BUSINESSEMP_P
Type of employment - For-profit company or organization	WORKTYPE_PROFIT_P
Type of employment - Non-profit organization (including tax-exempt and charitable organizations)	WORKTYPE_NONPROFIT_P
Type of employment – Government	WORKTYPE_GOV_P
Type of employment - Self-employed or contract work	WORKTYPE_SELF_P
Type of primary employment - For-profit company or organization	WORKTYPEPRIM_PROFIT_P

Persons	
Item Name	**Variable Name**
Type of primary employment - Non-profit organization (including tax-exempt and charitable organizations)	WORKTYPEPRIM_NONPROFIT_P
Type of primary employment - Government	WORKTYPEPRIM_GOV_P
Type of primary employment - Self-employed or contract work	WORKTYPEPRIM_SELF_P
Most recent type of employment - For-profit company or organization	WORKTYPELAST_PROFIT_P
Most recent type of employment - Non-profit organization (including tax-exempt and charitable organizations)	WORKTYPELAST_NONPROFIT_P
Most recent type of employment - Government	WORKTYPELAST_GOV_P
Most recent type of employment - Self-employed or contract work	WORKTYPELAST_SELF_P
Did person receive income from: income from a job did person have	INC_JOB_P
Did person receive income from: salary did person have	INC_SALARY_P
Did person receive income from: wages did person have	INC_WAGES_P
Did person receive income from: tips did person have	INC_TIPS_P
Did person receive income from: income from self-employment did person have	INC_SELF_P
Did person receive income from: income from business did person have	INC_BUSINESS_P
Did person receive income from: additional income did person have	INC_ADD_P
Did person receive income from: income from bonuses or commissions did person have	INC_BONUS_P
Did person receive income from: income did person have from stipends	INC_STIPEND_P
Did person receive income from: income from renting some or all of your home did person have	INC_RENTINC_P
Did person receive income from: income from renting a property that isn't your home did person have	INC_RENTPERS_P
Did person receive income from: income from retirement did person have	INC_RETIRE_P
Did person receive income from: social security did person have	INC_SS_P
Did person receive income from: pension did person have	INC_PENSION_P
Did person receive income from: railroad retirement did person have	INC_RAIL_P
Did person receive income from: other retirement income did person have	INC_OTHRETIRE_P
Did person receive income from: disability and leave did person have	INC_DIS_P

Persons	
Item Name	**Variable Name**
Did person receive income from: workers compensation did person have	INC_WORKCOMP_P
Did person receive income from: paid family leave did person have	INC_PFL_P
Did person receive income from: paid family medical leave did person have	INC_FMLA_P
Did person receive income from: cash payment from supplemental insurance did person have	INC_SUPPINS_P
Did person receive income from: interest and payments did person have	INC_INVEST_P
Did person receive income from: interest of $500 or more did person have	INC_INTEREST_P
Did person receive income from: dividends did person have	INC_DIVIDEND_P
Did person receive income from: annuities did person have	INC_ANNUITY_P
Did person receive income from: estates and trusts did person have	INC_ESTATE_P
Did person receive income from: royalties did person have	INC_ROYALTY_P
Did person receive income from: other income did person have	INC_OTHERINC_P
Did person receive income from: unemployment did person have	INC_UNEMPL_P
Did person receive income from: child support and alimony did person have	INC_CHILDSUPP_P
Did person receive income from: survivor benefits did person have	INC_SURVIVOR_P
Did person receive income from: veterans payments did person have	INC_VET_P
Did person receive income from: other regular source of income did person have	INC_OTHERSOURCE_P
Did person receive income from: none	INC_NO_P
How much income from a job did person have	INCAMT_JOB_P
How much salary did person have	INCAMT_SALARY_P
How much wages did person have	INCAMT_WAGES_P
How much tips did person have	INCAMT_TIPS_P
How much income from self-employment did person have	INCAMT_SELF_P
How much income from business did person have	INCAMT_BUSINESS_P
How much additional income did person have	INCAMT_ADD_P
How much income from bonuses or commissions did person have	INCAMT_BONUS_P
How much income did person have from stipends	INCAMT_STIPEND_P
How much income from renting some or all of your home did person have	INCAMT_RENTINC_P

Persons	
Item Name	**Variable Name**
How much income from renting a property that isn't your home did person have	INCAMT_RENTPERS_P
How much income from retirement did person have	INCAMT_RETIRE_P
How much social security did person have	INCAMT_SS_P
How much pension did person have	INCAMT_PENSION_P
How much railroad retirement did person have	INCAMT_RAIL_P
How much other retirement income did person have	INCAMT_OTHRETIRE_P
How much disability and leave did person have	INCAMT_DIS_P
How much workers compensation did person have	INCAMT_WORKCOMP_P
How much paid family leave did person have	INCAMT_PFL_P
How much paid family medical leave did person have	INCAMT_FMLA_P
How much cash payment from supplemental insurance did person have	INCAMT_SUPPINS_P
How much interest and payments did person have	INCAMT_INVEST_P
How much interest of $500 or more did person have	INCAMT_INTEREST_P
How much dividends did person have	INCAMT_DIVIDEND_P
How much annuities did person have	INCAMT_ANNUITY_P
How much estates and trusts did person have	INCAMT_ESTATE_P
How much royalties did person have	INCAMT_ROYALTY_P
How much other income did person have	INCAMT_OTHERINC_P
How much unemployment did person have	INCAMT_UNEMPL_P
How much child support and alimony did person have	INCAMT_CHILDSUPP_P
How much survivor benefits did person have	INCAMT_SURVIVOR_P
How much veteran's payments did person have	INCAMT_VET_P
How much other regular source of income did person have	INCAMT_OTHERSOURCE_P

Source: U.S. Census Bureau, 2021 New York City Housing and Vacancy Survey.

Table B3: List of Variables Imputed for Vacant Units

Vacant Units	
Item Name	**Variable Name**
Number of stories in building	STORIES_FRONT
Need stairs from sidewalk to elevator	ELEV_NO_STEPS
Need stairs from sidewalk to unit	UNIT_NO_STEPS
Number of Bedrooms	V_BEDROOMS
Number of Rooms	V_ROOMS
Number of full bathrooms in unit	V_FULLBATH_NUM
Number of half bathrooms in unit	V_HALFBATH_NUM
Share bathroom with other apartments	V_SHAREDBATH
Does unit have complete bath facilities	V_COMPLETEBATH
Does unit have fridge	V_APP_FRIDGE
Does unit have stove	V_APP_STOVE
Does unit have dishwasher	V_APP_DISHWASH
Does unit have a dryer	V_APP_DRYER
Sink in unit	V_SINK
Is fridge replaced in unit	V_NEW_FRIDGE
Is stove replaced in unit	V_NEW_STOVE
Is dishwasher replaced in unit	V_NEW_DISHWASH
Is washing machine replaced in unit	V_NEW_WASHMACH
Is dryer replaced in unit	V_NEW_DRYER
Are cabinets replaced in unit	V_NEW_CABINETS
Is counter replaced in unit	V_NEW_COUNTER
Does unit share a kitchen	V_SHAREDKITCH
Asking price of unit	V_ASKINGPRICE
Monthly asking rent for unit	V_ASKINGRENT

Source: U.S. Census Bureau 2021 New York City Housing and Vacancy Survey.