

# Overview

## Introduction

The New York City Housing and Vacancy Survey (NYCHVS), sponsored by the New York City Department of Housing Preservation and Development, is conducted approximately every 3 years to comply with New York state and New York City's rent regulation laws. The Census Bureau has conducted the survey for the city since 1965. The sample size for the survey is approximately 18,000 housing units representing the five boroughs of the city.

The rental vacancy rate is the primary focus of the survey, because it is used to evaluate the current rent control and rent stabilization laws. Other important survey data include rent regulation status, number of stories, number of units in building, number of rooms in unit, type of heating fuel, monthly rent, estimated value and building condition. The survey also includes information concerning housing and neighborhood quality.

Although the main purpose of the survey is to collect housing data, information on the demographic status of the population and households of the city is also collected. Information collected includes age, sex, race, ethnicity, household composition, labor force status, income, employment, and education level.

In addition to rent regulation, analysis of the data will provide a basis for planning and policy decisions on housing preservation and development.

## 2002 NYCHVS Sample

Sample units for the 2002 NYCHVS came from two primary sources: 1) the 2000 decennial census address files, and 2) a file of addresses listing all residential units, citywide, issued Certificates of Occupancy for new construction between January 1, 2000 through October 31, 2001. Another sample was drawn from a list of housing units located in structures owned by the city because the owner failed to pay taxes on the property (in rem units). Refer to the Source and Accuracy Statement for more detail.

Approximately 18,000 units throughout the city were selected as a representative sample of the housing in the five boroughs of the City of New York. Each sample unit represents approximately 180 similar housing units. The Census Bureau attempts to obtain an interview at each sample address. In the 2002 NYCHVS, the interview rate was 98 percent.

## File Structure

There are three files: household, house, and person. There is a household record for each occupied housing unit, a house record for each vacant housing unit, and a person record for each person in occupied housing units (households).

Person records can be linked to their appropriate household record by means of a sequence number. The sequence number is unique to each housing unit and is the same as on all person records within a household.

There are three "weights" that appear on the various files. To generate estimates pertaining to occupied housing units or households, use the "household weight". Similarly, to generate estimates pertaining to vacant units, apply the "housing unit weight". Use the "person weight" to generate estimates of population in the housing units.

## Relationship to 2000 Census

Most of the concepts and definitions are the same for similar items that appear in the 2000 Census, and the 2002 NYCHVS.

The similar questionnaire items from the 2002 NYCHVS and the 2000 Census 100 percent items are household member relationships, race, age, and sex of household members; and tenure. The similar questionnaire items from the 2002 NYCHVS and the 2000 census sample items are year moved-in, bedrooms, complete plumbing facilities, heating fuel, units in structure, number of rooms, value, monthly rent, condominium fee, utility costs, mortgage status, real estate taxes, fire and liability insurance, education level, labor force characteristics, industry and occupation for household members aged 15 years or over, and income.

However, there are differences between the 2002 NYCHVS and the 2000 Census. The differences are in interviewing procedures, staff experience and training, processing procedures, sample design, the sampling variability associated with the NYCHVS and the sample data from the census, and the nonsampling errors associated with the NYCHVS and the census.

## Relationship to Previous NYCHVS Surveys

The 2002 NYCHVS questionnaire, concepts, definitions, and data products are basically the same as those in the 1991, 1993, 1996, and 1999 NYCHVSs. In addition, most of the concepts and definitions in 2002 are similar to those used in NYCHVSs prior to 1991. For more information on previous surveys, see the 1991, 1993, 1996, and 1999 NYCHVS Tabulation Reference Guides and the 1975, 1978, 1981, 1984, and 1987 NYCHVS Notes and Definitions.

Several differences between the 1999 NYCHVS and previous surveys are worth noting:

1. The 2002 NYCHVS reflects the selection of a new sample based on the 2000 Decennial Census. NYCHVSs taken between 1975 and 1987 were based on a sample from the 1970 Decennial Census, while those from 1991 - 1999 were drawn from the 1990 Census. With each new sample design, new weighting and variance estimation procedures are required.
2. A new questionnaire was designed in 1991 replacing the design used from 1975 to 1987. The new design resulted in the primary demographic items being asked of all household members rather than for just the householder as was done in the past. The 2002 questionnaire is basically the same as the one used in 1991, 1993, 1996, and 1999.
3. Detailed questions on labor force status, industry and occupation, and income level were added in 1991. These items were either not covered in surveys prior to 1991 or covered in a much less detailed manner.
4. For 2002, 1999, 1996, and 1993 (done retroactively) key items that were not answered were assigned to an answer category in order to reduce non-response. The Census Bureau uses similar imputation methodology for most of its other demographic surveys. The procedure used is generally referred to as the "hot deck" approach. This approach assigns values for nonresponses from sample persons and housing units who did respond and who have characteristics similar to the nonrespondents. Some of the characteristics used for imputation include for persons: age, race, worker status, relationship, and education level, and for households: year moved, year acquired, control status, tenure and borough. The items that include imputed responses cover housing, demographic, and economic characteristics from the survey.

- Housing imputation items: year moved-in, year acquired, contract rent, asking rent (vacant units), stories in structure, rooms and bedrooms, type of heating fuel, utility costs (specific), value (owner-occupied unit), plumbing facilities, kitchen facilities, and rental assistance/out-of-pocket rent.
  - Demographic imputation items: sex, Hispanic Origin, race, age, and education level (persons aged 15 and older).
  - Economic imputation items (persons aged 15 and older): labor force status, year last worked, type of business, class of worker, industry/occupation, hours worked, weeks worked, and income.
5. In the 1991 NYCHVS, persons of Hispanic origin who did not classify themselves into one of the major race categories were classified as "other race". Beginning with the 1996 survey (and retroactively applied to the 1993 data) all persons reporting "other race" were allocated into one of the major race categories, as were persons not reporting race. These changes were necessitated by new methodologies used to derive independent weighting control estimates for all Census Bureau surveys, and to make procedures used in the NYCHVS consistent with those used in other surveys, such as the Current Population Survey. As a result, the count of persons and households classified as "other race" will be nonexistent in 1993, 1996, 1999, and 2002. There will be a corresponding increase in the number of persons and households classified in the specific race categories (particularly "White" and "Black or African American"). Thus, caution should be used when comparing data on race between the 1991 (and previous) NYCHVSs and later surveys.
  6. For 2002, 1999 and 1996, the wording of the question on persons from a temporary residence was changed to focus more specifically on traditional homeless situations.
  7. Beginning in 1996, the question on receipt of a Senior Citizen Rent Increase Exemption (SCRIE) was incorporated into a more detailed question on whether government programs paid part of the household's rent. The question on receipt of public assistance and welfare payments was also similarly modified. The purpose of these changes was to learn more about the types of assistance persons and households received. Beginning in 1999, the format of both of these questions was further modified to make each component of the item a separate question rather than to have the respondent answer for "all that apply." The purpose of these changes is an attempt to increase the response rate to each component of the items. In addition, the questions on rental assistance now identify the date on which the assistance began. As a result, caution should be used when comparing data on rental assistance, and public assistance and welfare payments, between surveys.
  8. Beginning in 1996, a series of questions on wheelchair-accessibility were added. These items aim to determine whether a wheelchair user would be able to reside in a sample building and the specific sample units. Like the building condition items, some of these were interviewer-observation items and some were asked of the respondents.
  9. An additional question on ownership of cooperative shares was added in 1996 in order to ensure a more accurate count of owner-occupied cooperative units. The effect of this new question is to increase the number of owner-occupied and reduce the number of renter-occupied cooperatives, compared to the previous wording.
  10. Starting in 1999, the two countries India and Mexico became separate response categories in the question on the last place lived and place of birth. Previously, each was included as part of a grouping of neighboring countries, respectively.
  11. Beginning with 1999, a series of questions on immigration to the United States and year moved to New York City replaced a set of questions on the previous residence of the householder.
  12. For 2002, the race question followed the format of the 2000 Census in that each person could report more than one racial category. Additionally, the selection of the "other" category as a response was discouraged as respondents were asked to pick a specific category or categories.
  13. New questions on health were added for 2002. These questions dealt with asthma, smoking, and performing normal daily activities. There were also two new additional questions on the respondents' perception of the neighborhood.
  14. The housing unit definition was changed for the 2000 Census relative to prior versions, thus the 2002 NYCHVS had to change its definition, accordingly. The primary result of the change is the potential elimination of what were once termed "boarding/rooming houses." Previously, such living quarters may have been classified as a group quarters, and thus, eliminated from sample. This type of housing, in effect, no longer exists. Therefore, use caution when comparing data on the number of housing units (and persons) from the 2002 NYCHVS and prior years.

As a result of the differences noted above, comparisons between surveys should be made with caution particularly when comparing data with surveys prior to 1991.

### **Statement on Undercoverage**

All demographic surveys suffer from undercoverage. This undercoverage results from missed housing units and missed persons within sample households. For units that have existed since 1990, our coverage is about one percent lower than that of the 1990 Decennial Census. This undercoverage also varies by age, ethnicity, and race of householder; however, estimates of this undercoverage are not available. For some household composition (e.g. persons per household) and income characteristics, estimates are affected by missed persons within households. The weighting procedures used by the Census Bureau partially correct for the bias due to undercoverage. However, its final impact on estimates is unknown.

### **Capped Data**

See [topcoding](#) [PDF]