

# Data Sources for U.S. Housing Research, Part 1: Public Sector Data Sources

Daniel H. Weinberg

Virginia Tech

*This article is the first of a two-part article about data sources for U.S. housing research. The second part, which will appear in the next issue of Cityscape (Volume 17, Issue 1), will address private sources, administrative records, and future directions.*

---

## Abstract

*For practitioners and policymakers to make a serious attempt to affect housing policy, they must cite evidence-based research. Part 1 of this article summarizes many of the government sources of housing data for researchers that can provide such evidence, such as the American Community Survey and the American Housing Survey.*

## Introduction

The basis for good housing policy is evidence-based research, and the only way to do good research on housing is to base that research on appropriate data. The principal research office of the U.S. Department of Housing and Urban Development (HUD)—the Office of Policy Development and Research (PD&R)—emphasizes such an approach. Its mission is—

To inform policy development and implementation to improve life in American communities through conducting, supporting, and sharing research, surveys, demonstrations, program evaluations, and best practices. [To carry out this mission,] PD&R compiles, analyzes, and disseminates data to support program operations, enable performance management, and inform program policy. PD&R sponsors major surveys to provide crucial intelligence about the operation of housing markets. (HUD PD&R, 2013a: 1)

One recent example of evidence-based research, conducted by the Joint Center for Housing Studies at Harvard University (JCHS, 2013), used more than 25 data sources for its report on rental housing. This article identifies those U.S. housing statistics data sources—and many more—and describes the suitability of those sources for research.<sup>1</sup>

## The Decennial Census of Population and Housing

The decennial census of population and housing, as its name implies, is conducted every 10 years and attempts, through extensive operations and thorough attention to detail, to gather information from every housing unit and group quarters in the United States.<sup>2</sup> Each census is based on the U.S. Census Bureau's Master Address File (MAF), a list of every residential address in the United States, including those sites for which building permits have been issued. The MAF is updated semiannually using the U.S. Postal Service Delivery Sequence File, a list of addresses to which mail is delivered. It is also updated before every census using two main additional techniques—nationwide address canvassing and the Local Update of Census Addresses Program. Although the multiplicity of operations ensures that the census reaches nearly every unit, some units are missed, and the Census Bureau conducts a Census Coverage Measurement program after each census to estimate the percentage of units that were missed. This operation estimated that the 2000 and the 2010 censuses both underestimated the number of housing units by 0.6 percent.<sup>3</sup>

The housing characteristics collected by the 2010 census were limited to only vacancy status and tenure. Vacancy was classified into seven categories and tenure into four.<sup>4</sup> Units that are vacant do not have residents to return the census form, thus an enumerator visited those units to determine their status. Units that appeared vacant were verified by consultation with neighbors, landlords, or other knowledgeable individuals (such as mail carriers), but vacancy status could not be confirmed

---

<sup>1</sup> I discuss neither international housing statistics nor data on homelessness in this article. International statistics are published by the United Nations Statistics Division in *Compendium of Human Settlements Statistics/Compendium of Housing Statistics*; the latest such report is for 2011. HUD publishes an *Annual Homeless Assessment Report*. Website addresses for all data sources are in the Data References Appendix. Thrall and Thrall (2011) presented a recent annotated bibliography of data relevant to real estate analysis.

<sup>2</sup> Group quarters are living quarters that are not housing units, such as prisons and nursing homes.

<sup>3</sup> The housing unit undercounts for the 1990, 2000, and 2010 censuses were 0.96, 0.61, and 0.60 percent, with standard errors of 0.24, 0.16, and 0.20 percent, respectively (Mule and Konicki, 2012). The estimated net undercount rate for occupied units in 2010 was 0.03 percent (0.14 percent standard error), whereas vacant units had a net undercount rate of 4.80 percent (1.06 percent standard error). Neither undercount rate was different from its corresponding 2000 rate, but the latter was significantly greater than zero.

<sup>4</sup> The vacancy categories were (1) for rent; (2) for sale; (3) rented but not occupied; (4) sold but not occupied; (5) for seasonal, recreational, or occasional use; (6) for migrant workers; and (7) other vacant. The tenure categories were whether the unit is (1) rented for cash rent, (2) occupied without the payment of cash rent, (3) owned with a mortgage, or (4) owned without a mortgage.

and was imputed for some units.<sup>5</sup> According to the 2010 census, 131.7 million housing units were in the United States on April 1, 2010. Of those housing units, 116.7 million (88.6 percent) had people living in them. The remaining 15.0 million units (11.4 percent) were vacant.<sup>6</sup>

The Minnesota Population Center's website, <https://www.ipums.org/>, provides access to a set of data files—the Integrated Public Use Microdata Sample (IPUMS) files—that include “harmonized [micro]data on people in the U.S. census and American Community Survey, from 1850 to the present.”<sup>7</sup> Housing characteristics are included on the microdata files only from the 1960 through 2000 censuses, and geography on the IPUMS files (and on the PUMS files provided by the Census Bureau) is limited to geographic areas constructed to have a population of 100,000 or more; internal files have the full geography. The Census Bureau has tabulated key housing characteristics for every census from 1940 through 2000 (see the Data References Appendix).

Basing their analysis on the census results, the Population Division of the Census Bureau issues annual population and housing unit estimates for states and counties. Housing unit estimates are used as “controls for several Census Bureau surveys, including the American Community Survey (ACS), the American Housing Survey (AHS), and the Current Population Survey (CPS). In addition to state and county housing unit estimates, [the Census Bureau] also produce[s] subcounty housing unit estimates [which] are central to the production of population estimates for cities and towns across the nation” (U.S. Census Bureau, 2014: 1).

## **The American Community Survey**

To simplify the 2010 census, the Census Bureau began the American Community Survey to replace the census long form, last fielded in 2000 to a one-in-six sample of households.<sup>8</sup> After undergoing a decade of testing, the ACS began in 2005 and was sent to a sample of about 2.9 million housing units each year, which increased to 3.3 million in 2011. From 2005 to 2010, interviews were completed in the United States and Puerto Rico at about 1.9 million housing units each year, which increased to 2.1 million housing units in 2011. The ACS accumulates 5 years of data (approximately an 11 percent sample) to provide detailed information for small geographic areas (for example, census tracts or small towns) and for small population groups (for example, those younger than 18 years old with a disability in a particular metropolitan area).<sup>9</sup>

By contrast with the two housing variables (vacancy status and tenure) collected on the 2010 decennial census short form, the ongoing ACS collects 35 housing variables (exhibit 1). Tabulations of these variables individually and cross-classified with other demographic, economic, and social

---

<sup>5</sup> In the 2010 census, 2.1 and 3.6 percent of occupied and vacant units, respectively, had their status imputed (for an overall status imputation rate of 2.3 percent of all housing units). See Summary File 1, Tables H3, H20, and H21; these tables are available at <http://factfinder2.census.gov>.

<sup>6</sup> See Mazur and Wilson (2011) for more information about these housing characteristics.

<sup>7</sup> Quoted from <https://www.ipums.org/> homepage.

<sup>8</sup> The first administration of questions to only a sample of people was in 1940.

<sup>9</sup> The Census Bureau oversamples small geographic areas.

**Exhibit 1****Housing Variables on the American Community Survey, 2005–2013**

Acreage	Meals included in rent	Second or junior mortgage payments or home equity loan
Agricultural sales	Mobile home costs	Selected monthly owner costs
Bedrooms	Monthly housing costs	Telephone service available
Business on property	Monthly housing costs as a percentage of household income	Tenure*
Condominium status and fee	Mortgage payment	Units in structure
Contract rent	Mortgage status	Utilities
Gross rent	Occupants per room	Vacancy status*
Gross rent as a percentage of household income	Plumbing facilities	Value
Homeowner vacancy rate	Real estate taxes	Vehicles available
House heating fuel	Rent asked	Year householder moved into unit
Insurance for fire, hazard, and flood	Rental vacancy rate	Year structure built
Kitchen facilities	Rooms	

\* Also on the 2010 decennial census.

characteristics are published annually.<sup>10</sup> The ACS provides 1-year estimates for geographic areas and population groups of 65,000 or more, 3-year estimates for areas and groups of 20,000 or more, and 5-year estimates for all areas and groups.<sup>11</sup>

HUD now uses ACS data to create a custom dataset with information on housing needs (particularly the housing needs of low- and moderate-income households). HUD provides that dataset to state and local governments to help them create their Comprehensive Housing Affordability Strategy and Community Development Plans (required for governments participating in several large HUD grant programs). These data help grantees to “assess their affordable housing and community development needs and market conditions, and to make data-driven, place-based investment decisions” (HUD PD&R, 2013b: 5).<sup>12</sup>

<sup>10</sup> See [http://www.census.gov/acs/www/Downloads/data\\_documentation/SubjectDefinitions/2011\\_ACSSubjectDefinitions.pdf](http://www.census.gov/acs/www/Downloads/data_documentation/SubjectDefinitions/2011_ACSSubjectDefinitions.pdf). The variables used for tabulations include several calculated (derived) from these basic variables (and others), such as “Selected Conditions”—defined for owner- and renter-occupied housing units as having at least one of the following conditions: (1) a lack of complete plumbing facilities, (2) a lack of complete kitchen facilities, (3) 1.01 or more occupants per room, (4) selected monthly owner costs as a percentage of household income greater than 30 percent, and (5) gross rent as a percentage of household income greater than 30 percent.

<sup>11</sup> For a limited number of variables, 5-year estimates are available for block groups and larger areas; others are available only for census tracts and small jurisdictions and for larger areas. Recent housing analyses using ACS by Census Bureau staff include Flanagan and Schwartz (2013) on rental housing market conditions, Mazur (2013) on the physical characteristics of housing, and Trevelyan, Acosta, and De La Cruz (2013) on homeownership by foreign-born residents.

<sup>12</sup> The 2005–2009 ACS 5-year estimates were the first ACS estimates that HUD used as a replacement for the long form estimates from the 2000 census to determine Community Development Block Grants.

## The American Housing Survey

The American Housing Survey, sponsored by HUD and carried out by the Census Bureau since 1973, is the centerpiece for detailed housing analysis in the United States.<sup>13</sup> According to its website, the intention of the AHS is to provide “current information on a wide range of housing subjects, including size and composition of the nation’s housing inventory, vacancies, fuel usage, physical condition of housing units, characteristics of occupants, equipment breakdowns, home improvements, mortgages and other housing costs, persons eligible for and beneficiaries of assisted housing, home values, and characteristics of recent movers” (U.S. Census Bureau, n.d.: 1).

The 2011 AHS had a sample size sufficient to provide detailed housing analyses for the nation as a whole (excluding Puerto Rico), for regions, and for 29 of the nation’s 388 metropolitan areas. A broad overview of housing conditions using data from all the AHS national surveys from 1973 to 2005 can be found in Eggers and Thackeray (2007).

The *Codebook for the American Housing Survey, Public Use File: 1997–2011* (Eggers and Wheelock, 2013) indicates that the 2011 AHS PUMS contains 186,448 cases compared with the 73,222 cases in the 2009 PUMS.<sup>14</sup> The 2013 sample started, as usual, with the 2011 national sample supplemented by a sample of new construction. This sample was enhanced by a different supplementary sample for 25 metropolitan areas (versus 29 in 2011). The total interviewed sample size was 167,911. Data from the 2013 AHS became available in 2014. Many researchers, including those at HUD, have used these data extensively for analyses of housing markets. To perform analyses that require that the exact location of the unit be known, researchers with approved projects can access the internal versions of the AHS through the Census Bureau Research Data Centers.

Two special features of the AHS have been relatively underused—its longitudinal nature and the clustering of neighborhood samples. First, the AHS has had the same national sample from 1985 to 2013, with periodic additions because of new construction (and subtractions because of demolition and conversion). Such a design is critical to the analysis of the Components of Inventory Change, or CINCH, and of housing filtering, but it has been used for relatively few research studies.<sup>15</sup> An entirely new sample will be drawn for the 2015 AHS.

Second, the AHS national surveys in 1985, 1989, and 1993 included a “Neighbor Sample”—the 10 housing units nearest a fraction of the national survey sample. This sample has been used to examine neighborhood effects on housing prices (Ioannides and Zabel, 2003). The 1985, 1989, and 1993 AHS national samples consisted of approximately 61,000, 56,500, and 59,000 addresses, respectively. According to Ioannides and Zabel (2003), 630, 769, and 1,018 units were selected as kernel units for clusters of interviewed housing units in 1985, 1989, and 1993, respectively.

---

<sup>13</sup> AHS was originally the Annual Housing Survey and was conducted once a year from 1973 to 1981. The name changed in 1985 after the survey became biennial in 1983.

<sup>14</sup> One important change was to add a sample of HUD-assisted rental housing (about 5,250 units).

<sup>15</sup> See Weicher, Eggers, and Moumen (2010), however, for an excellent study of the filtering of low-income affordable housing.

## Data for Analysis of Housing Prices

Moench and Ng (2011) provide an excellent summary of the various sources of data on housing prices, which follows in shortened form.

The Federal Housing Finance Administration [sic] (FHFA) provides price indices that include only homes with mortgages that conform to Freddie Mac and Fannie Mae guidelines [(excluding “jumbo” loans)]. Data are available at the national, regional and state levels, as well as for the major metropolitan areas. They are based on transactions and appraisals, and are then adjusted for appraisal bias. . . . The S&P/Case-Shiller home price indices, published by Fiserv Inc., are based on information from county assessor and recorder offices. The index started with data from 10 cities in 1987 but was extended to cover 20 cities in 2000. The Case-Shiller indices do not use data from 13 states and have incomplete coverage for 29 states. Compared to the FHFA, the Case-Shiller indices thus have a narrower geographical coverage but homes purchased with subprime and other unconventional loans are included in the indices. . . . The FHFA and the Case-Shiller indices are both based on repeat sales. In contrast, the NAR [National Association of REALTORS®] reports the median purchase prices of homes directly. The NAR surveys a fixed subset of its [2,000 local] associations. Based on reported transactions from the sample, the NAR calculates a median price for each of the four Census Bureau regions. . . .

The Census [Bureau] publishes several house price series. A monthly national series is available since 1963, but the regional data are available only quarterly. The Census also provides an average price of new homes of constant quality from 1977 onwards on a quarterly basis, both for the United States and for the four regions. The indices are based on a monthly survey of residential construction activity for single-family homes. These indices are also subject to compositional effects that might arise from the sales sample rather than any true changes in price. The Census Bureau also publishes an index of one-family homes sold based on the hedonic approach.

The Conventional Mortgage Home Price Index (CMHPI) is provided by Freddie Mac. It is calculated on a quarterly basis at both the national and regional level from 1975 onwards. The index is based on conventional conforming mortgages for single-unit residential houses that were purchased or securitized by Freddie Mac or Fannie Mae. (Moench and Ng, 2011: C4–C6)

The Conventional Mortgage Home Price Index (CMHPI) was replaced by the Freddie Mac House Price Index (FMHPI), recalculated using a repeat transactions methodology back to 1975. Freddie Mac publishes the monthly index values each quarter. Index values are available for the nation as a whole and for the 50 states, the District of Columbia, and 367 metropolitan statistical areas.<sup>16</sup> The Census Bureau house price indexes are discussed in the subsection “Residential Construction Surveys.”

---

<sup>16</sup> The two Freddie Mac national price indexes differ in two important ways. First, the national FMHPI is a weighted average of state indexes, whereas the national CMHPI was a weighted average of nine census region indexes. Second, the FMHPI uses Freddie Mac portfolio share weights to construct the national index, whereas the CMHPI used census region counts of single-family housing. The FMHPI also differs from the CMHPI in its treatment of refinance transactions. The original CMHPI included refinance transactions in the estimation but did not account for disparities between appraisals for refinance and for purchase. The FMHPI includes these transactions and uses statistical methods to account for the possibility that appraisal values might systematically differ from purchase prices. The purchase-only CMHPI excluded refinance transactions.

The Bureau of Labor Statistics (BLS) Consumer Price Index (CPI) Housing Survey provides the data needed to measure price change for the two housing indexes that are components of the CPI: (1) owners' equivalent rent of primary residence and (2) rent of primary residence. The BLS measure of rental equivalence for homeowner costs was introduced to the CPI in 1983. That measure attempts to estimate the flow of services for an owner-occupied dwelling based on market rents for rented dwellings based on actual market rents collected from a sample of renter-occupied housing units that are identified to be representative of owner-occupied housing. This method measures the rate of change in the amount an owner would need to pay to rent on the open market (Ptacek and Rippy, 2013).

Real Capital Analytics Inc. (RCA) and Moody's Investors Service, Inc., jointly publish several monthly Commercial Property Price Indexes (CPPI). Their indexes cover several property types—office (central business district and suburban), industrial, retail, apartment, commercial (combined office, retail, and industrial), and hotel—for the United States, for “major markets” (an aggregate of Boston, Chicago, Los Angeles, New York City, San Francisco, and Washington, D.C.), and for “non-major markets” (an aggregate of everything outside major markets). As their website indicates—

RCA records commercial properties valued at over \$2.5M (in constant dollars, inflation adjusted to December 2010) that have sold at least twice from 1988 to the current period. Filters are applied to exclude ‘flipped’ properties (sold twice in 12 months or less), transactions that are not arm's length, properties where size or use has changed, and transactions with extreme price movements (more than 50% annual gain or loss). The Moody's/RCA CPPI™ uses advanced Repeat-Sale Regression methodology of qualified repeat sale observations to measure price change in commercial real estate. (Real Capital Analytics, 2013: 1)

## **Other Government Housing Surveys**

### **Housing Vacancy Survey**

The Housing Vacancy Survey (HVS), a supplement to the monthly BLS-Census Bureau Current Population Survey designed to estimate the unemployment rate, provides information on rental and homeowner vacancy rates and on characteristics of units available for occupancy. Estimates of the total housing inventory and percentage distributions of vacant for-rent and for-sale-only units are available for the United States and the four census regions. In addition, the rental vacancy rate is a principal economic indicator, as designated by the Office of Management and Budget. The HVS also provides information on homeownership rates and the composition of the housing stock. Rental and homeowner vacancy rates and homeownership rates are available for the United States, regions, states, and 75 largest metropolitan areas. Data for all geographies are available quarterly and annually. Homeownership rates are also tabulated by age of householder and by family status for the United States and regions and are tabulated by race or ethnicity of householder and by Median Family Income for the United States.

### **Survey of Market Absorption**

The Survey of Market Absorption (SOMA), sponsored by HUD, is used to estimate the rate at which newly completed multifamily rental units (in buildings with five or more units), condominiums, and

cooperatives are absorbed (that is, rented or sold). It uses the Census Bureau's Survey of Construction (SOC) as its sampling base. The initial 3-month SOMA interview collects information on amenities, rent or sales price levels, number of units, type of building, and number of units taken off the market. Field representatives conduct subsequent interviews, if necessary, at 6, 9, and 12 months after completion. Beginning in 2002, the survey started collecting information on housing designed for the elderly.

### **The Rental Housing Finance Survey**

Immediately after every decennial census from 1950 to 2000, the Census Bureau conducted a Residential Finance Survey (RFS) to collect, process, and produce information about the financing of all nonfarm residential properties. (HUD sponsored the 2001 RFS.) The Rental Housing Finance Survey (RHFS) was designed to replace the RFS and create a nationally representative sample of data on the financing of multiunit rental housing. The first RHFS was conducted in 2012; the next is scheduled for 2015.

### **The New York City Housing and Vacancy Survey**

The New York City Housing and Vacancy Survey, sponsored by the New York City Department of Housing Preservation and Development, is conducted roughly every 3 years to comply with New York State and New York City (NYC) rent regulation laws. The Census Bureau has conducted the survey for NYC since 1965. Detailed data from the survey cover many characteristics of the NYC housing market, including characteristics of the population, households, housing stock, and neighborhoods.

The rental vacancy rate is the primary focus of the survey, because that value is crucial to the current rent regulation laws. Other important survey data on housing include rent regulatory and home-ownership status, structural conditions, unit maintenance, and neighborhood conditions; crowding, rents, utility costs, type of heating fuel, rent/income ratios; owner purchase price and estimated value, mortgage status, and interest rate; number of stories and units in building, cooperative/condominium status, wheelchair accessibility, and much more about housing and households in New York City. In addition to housing data, extensive information on characteristics of the population and households occupying housing units is collected, such as age, sex, race, ethnicity, education, household composition, birth region, when/why household moved, income, employment, and labor force status.<sup>17</sup> A New York City Housing and Vacancy Survey PUMS is available.

### **Residential Construction Surveys**

Once every 5 years, the Census Bureau conducts a Census of Construction to collect information about employer establishments (those with a payroll and employees) in the construction industry. The census collects only basic information—including sales (revenue), payroll, and number of employees—and uses administrative records from the Internal Revenue Service to collect information about small firms instead of using interviews. More frequent and detailed information on residential construction is collected by two surveys, the Building Permits Survey (BPS) and the SOC. These

---

<sup>17</sup> See <http://www.census.gov/housing/nychvs/data/2011/overview.pdf>.



surveys produce estimates that include national and regional data on the number of new housing units authorized by building permits, authorized but not started, started, under construction, and completed.

The purpose of the BPS is to provide national, state, and local statistics on the number and valuation of new privately owned housing units authorized by building permits in the United States. The statistics from the BPS are based on reports that local building permit officials submit in response to a voluntary mail survey. About one-half of the permit-issuing places—jurisdictions that issue building or zoning permits—in the United States are surveyed monthly. The remainder of the places are surveyed annually. Building permits are public records collected from individual permit offices, most of which are municipalities. From local area data, estimates are tabulated for counties, states, metropolitan areas, census divisions, census regions, and the nation. The BPS covers all permit-issuing places and collects data on zoning permits for areas that do not require building permits. Areas for which no authorization is required to construct a new privately owned housing unit are not included in the survey.

The purpose of SOC is to provide national and regional statistics on starts and completions of new single-family and multifamily housing units and statistics on sales of new single-family houses in the United States; HUD partially funds this survey. SOC also provides statistics on characteristics of new privately owned residential structures. The data included are various characteristics of new single-family houses completed, new multifamily housing completed, new single-family houses sold, and new contractor-built houses started. SOC includes two parts: (1) the Survey of Use of Permits, which estimates the amount of new construction in areas that require a building permit, and (2) the Non-Permit Survey, which estimates the amount of new construction in areas that do not require a building permit. (According to the Census Bureau, less than 2 percent of all new construction takes place in nonpermit areas.)<sup>18</sup>

The Manufactured Homes Survey (MHS) is conducted by the Census Bureau and sponsored by HUD. MHS produces monthly regional estimates of new manufactured home placements, average sales prices, and dealers' inventories and more detailed annual estimates, including selected characteristics of new manufactured homes. Statistics on the shipment of new manufactured homes are produced by the Institute for Building Technology and Safety and published by the Manufactured Housing Institute.

## **Residential Energy Consumption Survey**

The Residential Energy Consumption Survey (RECS) is conducted by the Energy Information Administration (EIA). Interviewers collect energy characteristics on the housing unit, usage patterns, and household demographics. Illustrative data items include main heating fuel, use of cooling equipment, electricity use and expenditures, energy consumption by end use (air-conditioning, heating, and appliances), housing type, and year of construction. This information is combined with data from energy suppliers to these homes to estimate energy costs and usage for heating, cooling, appliances, and other end uses. First conducted in 1978, the RECS was conducted for the 13th time in 2009. The 2009 survey collected data from 12,083 households in housing units statistically selected to represent the U.S. housing units that are occupied as primary residences. Data from the

---

<sup>18</sup> See [http://www.census.gov/construction/nrs/how\\_the\\_data\\_are\\_collected/](http://www.census.gov/construction/nrs/how_the_data_are_collected/).

2009 RECS were tabulated for the 4 census regions, the 9 census divisions, and 16 states.<sup>19</sup> The results of each RECS include data tables, a microdata (PUMS) file, and a series of reports. RECS and EIA supplier surveys are ingredients for some of EIA's more comprehensive data products and reports, such as the *Annual Energy Outlook* and *Annual Energy Review*.

## **The Panel Study of Income Dynamics**

The Panel Study of Income Dynamics (PSID) is conducted by the University of Michigan Institute for Social Research. It was sponsored originally by the U.S. Office of Economic Opportunity but is now largely funded by the National Science Foundation (with the occasional support of other federal agencies). The PSID began in 1968 with approximately 5,000 families containing more than 18,000 individuals. As the PSID website notes, "Information on these individuals and their descendants has been collected continuously [that is, every year or every other year], including data covering employment, income, wealth, expenditures, health, marriage, childbearing, child development, philanthropy, education, and numerous other topics."<sup>20</sup> Each PSID family in every year through 1995 (except 1969, in which the addresses are unavailable) has been identified if they were living in housing units subsidized by HUD, by the Farmers Home Administration, through tax credits administered by the U.S. Department of the Treasury (hereafter, Treasury), or through state housing programs. A project is currently under way (as of 2013) to update this match for every PSID family from 1995 through the most recent wave. For a fuller description of the restricted housing data, see Newman and Schnare (1997).

## **Discontinued Surveys**

In 1995 and 1996, the Census Bureau conducted the Property Owners and Managers Survey, which was designed to learn more about rental housing and the providers of rental housing. A nationwide sample of approximately 16,300 housing units that were rented or vacant-for-rent in the 1993 AHS was selected, and a questionnaire was mailed to the property owner, manager, or other agent of the owner of each property containing a selected unit. Detailed information was collected about maintenance, management practices, tenant policy, financial aspects of rental property ownership, owner characteristics, and related topics.

From 1966 to 2007, the Census Bureau also conducted the Survey of Residential Alterations and Repairs. The survey provided quarterly data on expenditures for maintenance and repairs, alterations and additions, and major replacements.

## **Other Government Sources for Housing Data**

### **HUD Low-Income Housing Tax Credit Database**

HUD has updated the Low-Income Housing Tax Credit (LIHTC) Database annually since the mid-1990s to include new and rehabilitated properties placed in service that used the credit. These data

---

<sup>19</sup> The 16 states are Arizona, California, Colorado, Florida, Georgia, Illinois, Massachusetts, Michigan, Missouri, New Jersey, New York, Pennsylvania, Tennessee, Texas, Virginia, and Wisconsin.

<sup>20</sup> Quoted from <http://psidonline.isr.umich.edu/> homepage.

were initially collected by a contractor, who worked closely with the state allocating agencies to compile the information. After the passage of the Housing and Economic Recovery Act<sup>21</sup> in 2008, and in keeping with its mandate that the state housing finance agencies provide HUD with demographic and economic data on tenants in LIHTC units, HUD began collecting LIHTC property and tenant information directly. The property data are available through an interactive query system and are updated annually.

## **Government-Sponsored Enterprises Data**

HUD publishes information on the single-family conventional mortgage purchases of Fannie Mae and Freddie Mac, the two government-sponsored enterprises (GSEs) that are involved with the housing market and that HUD oversees. Under the Federal Housing Enterprises Financial Safety and Soundness Act of 1992,<sup>22</sup> Fannie Mae and Freddie Mac are required to meet specified goals for purchases of mortgages that finance housing for very low-, low-, and moderate-income families and families living in areas traditionally underserved by the mortgage market. These datasets are intended to shed light on the effectiveness of GSEs and provide additional data for mortgage research. The single-family datasets include detailed data about the income, race, and gender of the borrower and about the census tract location, loan-to-value ratios, and affordability of the mortgage. The multifamily datasets include information about the number, type, and affordability of units and the size of the property, mortgage balance, and type of organization that sold the mortgage to the GSE.

The Federal Housing Finance Agency (FHFA) conducts the Monthly Interest Rate Survey to provide monthly information on interest rates, loan terms, and house prices by property type (all, new, and previously occupied), by loan type (15- or 30-year fixed rate or adjustable rate), and by lender type (savings associations, mortgage companies, commercial banks, and savings banks). In addition, the survey provides quarterly information about conventional loans by major metropolitan area and by Federal Home Loan Bank district. To conduct this survey, FHFA asks a sample of mortgage lenders to report the terms and conditions on all single-family, fully amortized, purchase-money, nonfarm loans that they close during the last 5 business days of each month. (The survey excludes Federal Housing Administration-insured and Department of Veterans Affairs-guaranteed loans, multifamily loans, mobile home loans, and loans created by refinancing another mortgage.)<sup>23</sup>

Freddie Mac's Primary Mortgage Market Survey® collects information from lenders each week on the rates and points for their most popular mortgage products: 30-year fixed-rate, 15-year fixed-rate, 5-1 hybrid amortizing adjustable-rate, and 1-year amortizing adjustable-rate mortgages. The survey is based on first-lien prime conventional conforming mortgages with a loan-to-value ratio of 80 percent. In addition, the adjustable-rate mortgage (ARM) products are indexed to Treasury yields and lenders are asked to provide the initial coupon rate, points, and margin on ARM products. The survey began with questions about conventional mortgages in 1971. About

---

<sup>21</sup> Public Law 110-289.

<sup>22</sup> Public Law 102-550.

<sup>23</sup> FHFA also provides a good website for downloadable housing data sources at <http://www.fhfa.gov/DataTools/Downloads>.

125 lenders currently are surveyed each week, and the mix of lender types—thrifts, credit unions, commercial banks, and mortgage-lending companies—is roughly proportional to the level of mortgage business that each type commands nationwide.

Fannie Mae conducts a monthly National Housing Survey (NHS) using random-digit dialing. The NHS interviews approximately 1,000 U.S. adults who are sole or joint financial decisionmakers to assess their attitudes about homeownership, renting a home, the economy, and household finances. Fannie Mae publishes 11 attitudinal indicators monthly and occasional analyses using quarterly data. The survey, which has been conducted since June 2010, is weighted to the ACS “to account for known biases resulting from non-response” (Fannie Mae, 2013: 2).

### **Board of Governors of the Federal Reserve System**

From its Flow of Funds accounts, the Federal Reserve issues tabulations of the net change in mortgages outstanding and of total outstanding mortgages once a quarter. The Federal Reserve also conducts the Survey of Consumer Finances (SCF), which is normally a triennial cross-sectional survey of U.S. families, but, during the 1983-to-1989 and the 2007-to-2009 periods, the survey collected panel data. The survey oversamples high-income households, and its data include information on families' balance sheets, pensions, income, and demographic characteristics. The SCF is a good source of data on home equity, first-lien and junior mortgages, refinance mortgages, and home equity lines of credit and loans. It distinguishes between primary residences and other residential and nonresidential property. In 2010, 82 percent of all debt was for residential properties and mortgage or real estate lenders held 27 percent of all debt, down from 42 percent in 2007 (Bricker et al., 2012).

### **U.S. Department of Agriculture, Office of Rural Development**

The U.S. Department of Agriculture administers programs that provide homeownership opportunities and home renovation and repair programs to Americans living in rural areas. The programs also provide financing to residents in multiunit buildings who are elderly or disabled and to those who have low incomes to ensure that they are able to make rent payments. These rural development programs include Business Loans and Grants, Cooperative Grants and Other Programs, Single Family Housing Loans and Grants, Multi-Family Housing Loans and Grants, Community Facilities Loans and Grants, Electric Loans and Grants, Telecommunications Loans and Grants, Water Loans and Grants, and Community and Economic Development Programs.<sup>24</sup> Aggregate funding obligations for these programs (and others) for 2009 through 2012 by state are available on line (USDA, 2013).

### **More Information**

Part 2 of this article, to be published in the next issue of *Cityscape*, summarizes the challenges of using administrative records (AR) and proposes to construct new and useful data sources by matching survey data with AR and by constructing synthetic databases. Part 2 concludes with a brief discussion of some data issues.

---

<sup>24</sup> See <http://www.rurdev.usda.gov/ProgramsAndOpportunities.html>.

## **Data References Appendix**

American Community Survey (ACS): <http://www.census.gov/acs/www/>; see also U.S. Census Bureau, American FactFinder.

American Housing Survey (AHS): <http://www.census.gov/programs-surveys/ahs/> and <http://www.huduser.org/portal/datasets/ahs.html> (1997–2011 AHS); see also U.S. Census Bureau, American FactFinder.

Census of Construction: <http://www.census.gov/econ/census/>; see also U.S. Census Bureau, American FactFinder.

Construction statistics (monthly, quarterly, and annual surveys): <http://www.census.gov/econ/construction.html>.

Construction statistics (New Residential Construction, Manufactured Homes Survey, New Residential Sales, Residential Improvements, Characteristics of New Housing, Construction Price Indexes, and Value of Construction Put in Place): <http://www.census.gov/mcd/>.

CoreLogic, Inc.: <http://www.corelogic.com/solutions/property-information-analytic-solutions.aspx>.

Decennial Census of Housing, 1940–2000 (tabulations): <http://www.census.gov/housing/census/data/>.

Decennial Census of Population and Housing, selected historical decennial census population and housing counts, 1790–1990: <http://www.census.gov/population/www/censusdata/hiscendata.html>.

Decennial Census of Population and Housing, 1990: <http://www.census.gov/main/www/cen1990.html>.

Decennial Census of Population and Housing, 2000 and 2010: <http://factfinder2.census.gov>.

Fannie Mae, National Housing Survey (NHS): <http://www.fanniemae.com/portal/research-and-analysis/housing-survey.html>.

Federal Financial Institutions Examination Council, Community Reinvestment Act (CRA), Distressed and Underserved Tracts, <http://www.ffiec.gov/cra/distressed.htm>.

Federal Financial Institutions Examination Council, Home Mortgage Disclosure Act (HMDA): <http://www.ffiec.gov/hmda/default.htm>.

Federal Housing Finance Agency: <http://www.fhfa.gov/DataTools>.

Federal Reserve Board of Governors, Flow of Funds Accounts, Net Change in Mortgages Outstanding and of Total Outstanding Mortgages: <http://www.federalreserve.gov/releases/z1/about.htm>.

Freddie Mac, House Price Index (FMHPI): <http://www.freddiemac.com/finance/fmhpi/archive.html>.

Freddie Mac, Primary Mortgage Market Survey (PMMS): <http://www.freddiemac.com/pmms/>.

Housing Vacancy Survey (HVS): <http://www.census.gov/housing/hvs/>.

Integrated Public Use Microdata Series (IPUMS): <http://www.ipums.org>.

Lender Processing Services (LPS): <https://www.lpsdefault.com/fnds/home.asp>.

Manufactured Homes Survey (MHS): <http://www.census.gov/construction/mhs/mhsindex.html>.

Mortgage Bankers Association, Weekly Applications Survey: <http://www.mortgagebankers.org/ResearchandForecasts/ProductsandSurveys/WeeklyApplicationSurvey>.

MPF Research: <http://www.realpage.com/apartment-market-research/rental-market-trends>.

National Association of Home Builders: <http://www.nahb.org/>.

National Association of REALTORS®: <http://www.realtor.org/>.

National Council of Real Estate Investment Fiduciaries (NCREIF), Data and Products Guide: [http://www.ncreif.org/public\\_files/NCREIF\\_Data\\_and\\_Products\\_Guide.pdf](http://www.ncreif.org/public_files/NCREIF_Data_and_Products_Guide.pdf).

New Residential Construction (Building Permits Survey, Quarterly Starts and Completions by Purpose and Design, Annual Characteristics of New Housing, Length of Time from Authorization to Start and from Start to Completion, Construction Price Indexes): <http://www.census.gov/construction/nrc/>.

New York City Housing and Vacancy Survey (NYCHVS): <http://www.nyc.gov/html/hpd/html/pr/vacancy.shtml> and <http://www.census.gov/housing/nychvs/about/>.

Panel Study of Income Dynamics (PSID): <http://psidonline.isr.umich.edu/>.

Property Owners and Managers Survey (POMS): <http://www.census.gov/housing/poms/>.

Real Capital Analytics Inc.: [https://www.rcanalytics.com/Public/rca\\_cpqi.aspx](https://www.rcanalytics.com/Public/rca_cpqi.aspx).

RealtyTrac® Inc.: <http://www.realtytrac.com/>.

Rental Housing Finance Survey (RHFS): <http://www.huduser.org/portal/datasets/rhfs/home.html>.

Residential Energy Consumption Survey (RECS): <http://www.eia.gov/consumption/residential/>.

Residential Finance Survey (RFS): <http://www.census.gov/housing/rfs/>.

Survey of Consumer Finances (SCF): <http://www.federalreserve.gov/econresdata/scf/scfindex.htm>.

Survey of Market Absorption (SOMA): <http://www.census.gov/hhes/www/housing/soma/overview.html>.

Survey of Residential Alterations and Repairs (SORAR): <http://www.census.gov/construction/c50/c50index.html>.

United Nations Statistics Division, Compendium of Human Settlements Statistics/Compendium of Housing Statistics: <http://unstats.un.org/unsd/demographic/sconcerns/housing/housing2.htm>.

U.S. Bureau of Labor Statistics, Consumer Price Index (CPI): <http://www.bls.gov/cpi/>.

U.S. Census Bureau, American FactFinder (2000 Census of Population and Housing, 2010 Census of Population and Housing, 2005–2012 ACS, 2011 AHS, 2002–2012 Census of Construction [sector 23]): <http://factfinder2.census.gov>.

U.S. Census Bureau, Housing Unit Estimates (Vintage 2012): <http://www.census.gov/popest/data/housing/totals/2012/index.html> and its methodology, at <http://www.census.gov/popest/methodology/2012-hu-meth.pdf>.

U.S. Census Bureau, Longitudinal Employer-Household Dynamics (LEHD): <http://lehd.ces.census.gov>.

U.S. Census Bureau, Research Data Centers (RDCs): <http://www.census.gov/ces/rcresearch/index.html>.

U.S. Department of Housing and Urban Development (HUD), Homelessness Data Exchange: <http://www.hudhdx.info/>.

U.S. Department of Housing and Urban Development (HUD), National Low-Income Housing Tax Credit (LIHTC) Database: <http://LIHTC.huduser.org>.

U.S. Department of Housing and Urban Development (HUD), Public and Indian Housing Information Center (PIC) system and Tenant Rental Assistance Certification System (TRACS): aggregate data (A Picture of Subsidized Households) at <http://www.huduser.org/portal/datasets/assthsg.html>; a 5-percent microdata sample from those systems for researchers at <http://www.huduser.org/portal/pumd/index.html>.

U.S. Postal Service, Vacancies: Aggregate data quarterly at the census tract level for government and not-for-profit researchers from HUD at <http://www.huduser.org/portal/datasets/usps.html>.

Zillow Home Value Index: <http://www.zillow.com/howto/api/APIOverview.htm>.

## **Acknowledgments**

An earlier version of this article was prepared for the January 2014 conference of the Weimer School of Advanced Studies in Real Estate and Land Economics. The author thanks Arthur Cresce, George Galster, Sandra Newman, Henry Pollakowski, Mark Shroder, and Dav Vandenbroucke for their comments and suggestions.

## **Author**

Daniel H. Weinberg is currently a visiting scholar in the Social and Decision Analytics Laboratory, Virginia Bioinformatics Institute at Virginia Polytechnic Institute and State University. He formerly held several positions at the U.S. Census Bureau, including Assistant Director for the Decennial Census and ACS (responsible for the 2010 census and the American Community Survey) and Chief of the Housing and Household Economic Statistics Division (responsible for the American Housing Survey, Housing Vacancy Survey, Residential Finance Survey, and many other housing surveys).

## References

- Bricker, Jesse, Arthur B. Kennickell, Kevin B. Moore, and John Sabelhaus. 2012. "Changes in U.S. Family Finances From 2007 to 2010: Evidence From the Survey of Consumer Finances," *Federal Reserve Bulletin* 98 (2): 1–80.
- Eggers, Fred, and Tobey Wheelock. 2013. *Codebook for the American Housing Survey, Public Use File: 1997–2011*. Version 2.1. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research. [http://www.huduser.org/portal/datasets/ahs/AHS\\_Codebook2.1\\_Final.pdf](http://www.huduser.org/portal/datasets/ahs/AHS_Codebook2.1_Final.pdf).
- Eggers, Frederick J., and Alexandre Thackeray. 2007. *32 Years of Housing Data*. Report prepared for the U.S. Department of Housing and Urban Development. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research. [http://www.huduser.org/datasets/ahs/AHS\\_taskC.pdf](http://www.huduser.org/datasets/ahs/AHS_taskC.pdf).
- Fannie Mae. 2013. *Fannie Mae National Housing Survey—Technical Notes*. Washington, DC: Fannie Mae. <http://www.fanniemae.com/resources/file/research/housingsurvey/pdf/nhstechnicalnotes.pdf>.
- Flanagan, Christine, and Mary Schwartz. 2013. "Rental Housing Market Condition Measures: A Comparison of U.S. Metropolitan Areas From 2009 to 2011." *2010 American Community Survey Brief ACSBR/11-07* (April). Washington, DC: U.S. Census Bureau. <http://www.census.gov/prod/2013pubs/acsbr11-07.pdf>.
- Ioannides, Yannis M., and Jeffrey E. Zabel. 2003. "Neighbourhood Effects and Housing Demand," *Journal of Applied Econometrics* 18 (5): 563–584.
- Joint Center for Housing Studies of Harvard University (JCHS). 2013. *America's Rental Housing: Evolving Markets and Needs*. Cambridge, MA: Joint Center for Housing Studies of Harvard University. [http://www.jchs.harvard.edu/sites/jchs.harvard.edu/files/jchs\\_americas\\_rental\\_housing\\_2013\\_1\\_0.pdf](http://www.jchs.harvard.edu/sites/jchs.harvard.edu/files/jchs_americas_rental_housing_2013_1_0.pdf).
- Mazur, Christopher. 2013. "Physical Characteristics of Housing: 2009–2011." *2010 American Community Survey Brief ACSBR/11-20* (April). Washington, DC: U.S. Census Bureau. <http://www.census.gov/prod/2013pubs/acsbr11-20.pdf>.
- Mazur, Christopher, and Ellen Wilson. 2011. "Housing Characteristics: 2010." *2010 Census Brief C2010BR-07* (October). Washington, DC: U.S. Census Bureau. <http://www.census.gov/prod/cen2010/briefs/c2010br-07.pdf>.
- Moench, Emanuel, and Serena Ng. 2011. "A Hierarchical Factor Analysis of US Housing Market Dynamics," *The Econometrics Journal* 14 (1): C1–C24.
- Mule, Thomas, and Scott Konicki. 2012. "Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Housing Units in the United States." *DSSD 2010 Census Coverage Measurement Memorandum Series, #2010-G-02*. [https://www.census.gov/coverage\\_measurement/pdfs/g02.pdf](https://www.census.gov/coverage_measurement/pdfs/g02.pdf).
- Newman, Sandra J., and Ann B. Schnare. 1997. "...'And a Suitable Living Environment': The Failure of Housing Programs To Deliver on Neighborhood Quality," *Housing Policy Debate* 8 (4): 703–741.



Ptacek, Frank, and Darren A. Rippy. 2013. "Owners' Equivalent Rent and the Consumer Price Index: 30 Years and Counting," *Beyond the Numbers* 2 (14): 1–9. [http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=2169&context=key\\_workplace](http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=2169&context=key_workplace).

Real Capital Analytics. 2013. "Factsheet: Moody's/RCA Commercial Property Price Indices." New York: Real Capital Analytics. [https://www.rcanalytics.com/public/pdf/MOODYSCPPI\\_FactSheet.pdf](https://www.rcanalytics.com/public/pdf/MOODYSCPPI_FactSheet.pdf).

Thrall, Grant Ian, and Susan E. Thrall. 2011. "Data Resources for Real Estate and Business Geography Market Analysis: A Comprehensive Structured Annotated Bibliography, Version 2.0," *Journal of Real Estate Literature* 19 (2): 415–468.

Trevelyan, Edward N., Yesenia D. Acosta, and Patricia De La Cruz. 2013. "Homeownership Among the Foreign-Born Population: 2011." *2010 American Community Survey Brief ACSBR/11-15* (January). Washington, DC: U.S. Census Bureau. <http://www.census.gov/prod/2013pubs/acsbr11-15.pdf>.

U.S. Census Bureau. 2014. *Methodology for State and County Total Housing Unit Estimates (Vintage 2013): April 1, 2010 to July 1, 2013*. Washington, DC: U.S. Census Bureau. <http://www.census.gov/popest/methodology/2013-hu-meth.pdf>.

———. n.d. "What Is the American Housing Survey?" <http://www.census.gov/programs-surveys/ahs/about/what-is-the-american-housing-survey.html>.

U.S. Department of Agriculture (USDA). 2013. *USDA Rural Development: 2012 Progress Report*. Washington, DC: U.S. Department of Agriculture. <http://www.rurdev.usda.gov/Reports/RDProgressReport2012Feb2013.pdf>.

U.S. Department of Housing and Urban Development, Office of Policy Development and Research (HUD PD&R). 2013a. *HUD Research Roadmap, FY 2014–FY 2018*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research. [http://www.huduser.org/portal/pdf/Research\\_Roadmap.pdf](http://www.huduser.org/portal/pdf/Research_Roadmap.pdf).

———. 2013b. *Summary of Key PD&R Research and Housing Survey Activities: FY 2013*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research. [http://archives.huduser.org/about/pdf/ongoing\\_research\\_2013.pdf](http://archives.huduser.org/about/pdf/ongoing_research_2013.pdf).

Weicher, John C., Frederick J. Eggers, and Fouad Moumen. 2010. *The Long-Term Dynamics of Affordable Rental Housing*. A report to the John D. and Catherine T. MacArthur Foundation. Washington, DC: Hudson Institute.