The Quality of America's Assisted Housing Stock: Analysis of the 2011 and 2013 American Housing Surveys

Multidisciplinary
Research Team

U.S. Department of Housing and Urban Development | Office of Policy Development and Research

The Quality of America's Assisted Housing Stock: Analysis of the 2011 and 2013 American Housing Surveys

Multidisciplinary Research Team

U.S. Department of Housing and Urban Development
Office of Policy Development and Research

March 2017

Submitted by:

Economic Systems Inc.
Falls Church, Virginia

Prepared by:
Sandra Newman, Ph.D.
C. Scott Holupka, Ph.D.
Johns Hopkins University

About MDRT

PD&R developed the Multidisciplinary Research Team (MDRT) vehicle to manage a team of qualified researchers. Researchers are selected for their expertise to produce an array of high quality, short-turnaround research. MDRT researchers use a variety of HUD and external data sources to answer research questions relating to HUD's priority policies and strategic goals.

Disclaimer

Contents

List of Tables

List of Figures

Executive Summary

The substantial literature on housing quality demonstrates both the importance and the challenges of conceptualizing and measuring housing quality. The growing interest in "how housing matters," primarily in the low-rent unassisted housing stock, and the ongoing concerns about housing quality standards in the assisted stock, make this an opportune time to revisit housing quality.

In this paper, the research team reviews the relevant literature, develops alternative housing quality indices, tests their validity, and applies them to both the assisted and comparable unassisted housing stock. We focus on indicators of the physical integrity of housing systems and exclude measures that are more likely to reflect residents' housekeeping or behaviors.

Because no consensus exists about the features of the housing bundle that should be included in the definition of a dwelling's quality and how each should be weighted in determining overall quality, we rely on three external criteria first suggested by research on the Experimental Housing Allowance Program (Merrill, 1980): market value, consumer rating, and normative standards. The market value criterion is the standard of choice for housing economists. The consumer rating criterion identifies the dwelling features most associated with a resident's assessment of the dwelling as a good place to live regardless of the market price of these features. This criterion is consistent with the renewed interest by economists in happiness and subjective well-being. The normative standards criterion reflects community concerns and policy decisions such as building codes. Our main data are the 2011 and 2013 national American Housing Surveys (AHS). Both surveys provide augmented national samples and assisted housing samples, with assisted housing identified by address matches to the U.S. Department of Housing and Urban Development's (HUD's) administrative data, not respondent self-report.

Despite its intuitive appeal, the market value criterion for housing quality performs poorly and we drop it from further analysis. We construct one consumer rating index using the results from the multiple variable analysis as the weights. We also construct three indices using the normative standards criterion: a weighted index based on a previous analysis for HUD's Office of Policy Development and Research by Eggers and Moumen (2013b); an unweighted index; and an index with weights derived from an analysis that examines which components of the index are most closely associated with one another.

Consistent with much past AHS housing quality research, the prevalence rate of almost all housing quality problems is very low, with most dwelling units having no problems. We also find that the quality of assisted housing is comparable to the quality of unassisted housing. Further, the incidence of housing problems persisting over 2 years, 2011 and 2013, is very low, and repairs are made promptly.

Some heterogeneity in housing quality exists, however. Geographically, assisted housing quality in central cities and in the Northeast region is considerably lower than in their counterpart locations (that is, suburban or rural areas and the South, Midwest, and West regions). Demographically, three household subgroups emerge as particularly likely to live in assisted

housing with lower-than-average housing quality: nonelderly disabled persons, non-White persons, and large households. The analysis that simultaneously considers the strongest predictors of assisted housing quality for each of these three groups generally corroborates these risk factors and adds further insights. For example, disabled persons enjoy better housing quality using a voucher compared with those living in multifamily housing. For large households, living in the South and living in public housing are associated with considerably worse housing quality. The analysis also provides hard evidence supporting the current 40th percentile of rents (that is, the point on the rent distribution where 40 percent of rents fall below this amount and 60 percent of rents are higher than this amount) definition of the Fair Market Rent (FMR). We find that the FMR is set at a level that is roughly at the inflection point for maximizing assisted housing quality as measured by our housing quality indices.

Overall, these positive findings demonstrate that the current inspection and quality control systems appear to be achieving the goal of providing physically adequate housing to assisted housing residents. They also lend support to two policy shifts—biennial inspections in the voucher program and biennial and triennial inspections for standard and high performers, respectively, in the public housing program—and to serious consideration of proposals to streamline inspections to encourage participation in the voucher program by private owners of rental properties.

Suggestions for future research include expanding the measures to include additional aspects of the full housing bundle and comparing these results with housing inspection scores from HUD's administrative data—that is, the Public Housing Assessment System and Real Estate Assessment Center housing inspection ratings for public housing and multifamily housing, respectively. Additionally, a consideration of measures that are currently missing from the AHS, but should be included in the future, is beyond the scope of the current effort but worth serious attention.

Introduction

A key role of the U.S. Department of Housing and Urban Development (HUD) is designing, implementing and monitoring most of the nation's assisted housing programs. Central to this responsibility is ensuring that the units receiving HUD assistance are physically adequate. This, in turn, verifies that recipient households live in decent and safe dwellings and reassures the public that tax dollars are not supporting deficient housing, or worse. To accomplish this objective, HUD imposes a set of housing quality standards (HQS) that assisted units must meet, requires periodic inspections to confirm that standards are being met, and, when necessary, issues citations of violations that must be corrected within a specified time frame.[1] Although the broad outline of this assisted housing quality assurance system has remained essentially the same over several decades, HUD has modified it from time to time in an effort to improve the efficiency of inspections and the effectiveness of achieving the statutory goals of decent and safe assisted housing units.

Several recent circumstances prompt a reexamination of assisted HQS. First, HUD's *Strategic Plan 2014–2018* calls for the development of a "uniform asset risk assessment management model," which requires systematic evidence on the most meaningful approaches to measuring the quality of the assisted housing stock (HUD, 2014: 19). Second, the FY13 Senate Appropriation Committee Report raises concerns about violations of HQS in housing units participating in the Section 8 voucher program and "directs HUD to take meaningful and timely steps to strengthen oversight and quality control" of the public housing agency (PHA) inspection process (U.S. Senate, 2012: 92). An additional concern is that reports by HUD's Office of the Inspector General and the U.S. Government Accountability Office (GAO) note poor reliability of assisted housing inspections using the HQS (GAO, 2000; HUD OIG, 2008). This conclusion is based on a comparison of the PHA inspector scores with those collected by an independently trained rater.

This paper is designed to contribute to the reexamination of assisted housing *quality*. We do so by developing composite measures of housing quality, or housing quality indices, testing their validity, and then applying these indices to both the assisted and comparable unassisted housing stock. Assisted housing refers to housing subsidized by the federal government to reduce monthly rent and that is targeted on income-eligible households. Subsidies take two forms: demand side, through a rent voucher provided to income-eligible households to use in the private rental market; and supply side, through either public housing developed and managed by local PHAs, concessionary financing to private entities to develop or rehabilitate housing, or tax credits under the U.S. Department of the Treasury's Low-Income Housing Tax Credit program. For simplicity in the analysis, we will disaggregate assisted housing into three main programs: public housing; privately owned, publicly subsidized housing (referred to by HUD as multifamily); and vouchers.

---

[1] Inspection protocols and processes differ by program. Public housing and multifamily housing—for example, Section 8 new construction; Section 221(d)(3)—must meet property standards, while Section 8 vouchers must meet housing quality standards, or HQS. Inspectors employed by the local public housing agency conduct inspections on public housing and voucher housing, while inspectors contracted by the HUD regional offices inspect multifamily housing.

It is important to acknowledge at the outset that the concept of housing quality is not based on definitive criteria and has no precise quantifiable definition of where "bad" ends and "good" begins (Newman, 2008). As explained more fully below, because we lack a consensus definition of housing quality, we construct measures that characterize the dwelling's physical integrity (for example, holes in the floor) or housing systems (for example, heating system breakdown) and exclude those that are more likely to reflect the resident's housekeeping or behavior (for example, leaving unsafe chemicals within a child's reach).[2]

We use rich data from the 2011 and 2013 American Housing Surveys (AHS) to describe the quality of the assisted housing stock and to highlight geographic areas, types of households, housing types, and housing assistance programs most likely to experience quality problems. The AHS asks respondents about a wide range of conditions in their housing units. This allows us to develop a composite measure or index of quality, thereby capturing the multiple dimensions of housing quality. The 2011 and 2013 AHS samples have also been matched to administrative records on assisted housing receipt, alleviating serious concerns about the validity of self-reported housing assistance receipt.

We find that the quality of assisted housing is comparable to the quality of unassisted housing, the incidences of persistent problems are very low, and requested repairs are made promptly. Some heterogeneity in housing quality exists, however. Geographically, assisted housing quality in central cities and in the Northeast region is lower than in their counterpart locations (that is, suburban or rural areas and South, Midwest, and West regions). Demographically, three household subgroups emerge as particularly likely to live in assisted housing with lower than average housing quality: nonelderly disabled persons, non-White persons, and large households. Multivariate analysis of the strongest predictors of assisted housing quality for each of these three groups generally corroborates these risk factors and adds further insights. For example, disabled persons enjoy better housing quality using a voucher compared with those living in multifamily housing. For large households, living in the South and living in public housing are associated with considerably worse housing quality. The analysis also provides hard evidence supporting HUD's definition of the Fair Market Rent (FMR) as the 40th percentile of rents. We find that the FMR is set at a level that is roughly at the inflection point for maximizing assisted housing quality as measured by our housing quality indices.

In the next section, we review the literature on housing quality, emphasizing past research on assisted housing quality. The Literature Review section is followed by a discussion of our research approach, including a description of the 2011 and 2013 AHS data, analysis samples, methods, and alternative measures of housing quality, along with tests of their validity. We then examine assisted housing quality from several perspectives including a comparison with unassisted housing quality, an exploration of whether assisted housing quality varies depending on the characteristics of the housing or the characteristics of the residents. We also explore how

---

[2] Some features, such as mold, fall into a gray area.

well assisted housing quality aligns with the FMR, the persistence of assisted housing quality problems, and the related issue of timeliness of repairs. In the final section, we summarize the key results and discuss their implications for policy and for future research.

Literature Review

The substantial literature on housing quality spans nine decades and demonstrates both the importance and the challenges of conceptualizing and measuring housing quality. We review three relevant strands in past work: housing quality measurement and data collection methods; the AHS measurement of housing quality; and the quality of assisted housing.

Housing Quality Measurement and Data Collection Methods

The American Public Health Association provided some of the earliest contributions to the housing quality literature. Its *Basic Principles of Healthful Housing* (APHA, 1938) highlighted the connection between housing conditions and health, while *An Appraisal Method for Measuring the Quality of Housing* (APHA, 1945) recommended that data be collected through a field survey of many individual features of each dwelling unit, with penalty scores for each feature that falls below an established standard. The sum of all scores represents the quality of the dwelling. This methodology is roughly similar to that used for physical inspections of assisted housing under HUD's Uniform Physical Condition Standards (UPCS).[3]

Another important early contribution was the U.S. Census Bureau's methodological study of housing quality measurement (Census Bureau, 1967).[4] The historical context of this landmark report is worth reviewing. For decades, the decennial census included interviewer observations of housing features. In 1940, housing condition was measured by a dwelling's "state of repair," with trained enumerators rating the structure as either needing "major" repairs or not. In 1950, this approach was replaced by another dichotomous classification of structures as either "dilapidated" or "not dilapidated." This dichotomy was refined in 1960 by further classifying those structures designated as "not dilapidated" as either "sound" or "deteriorating." Following the 1960 census, the Census Bureau launched a detailed and thorough evaluation of its approach to measuring housing conditions that resulted in the 1967 publication. Its unambiguous conclusion was that housing conditions collected through interviewer observations are unreliable and, therefore, inaccurate. As a result, subsequent decennial censuses that relied on interviewers to administer the survey dropped the interviewer observations of housing unit condition. The AHS followed suit in 1997.[5]

---

[3] UPCS currently exists for public housing and for multifamily housing. A version for the voucher program, UPCS-V, is under development and will replace the current HQS system (Arcara, 2016).
[4] This discussion draws on Newman (2008).
[5] AHS began interviewing returning households by phone, when possible, in 1997. In 2011, a phone-first policy was instituted for both new and returning households (Vandenbroucke, 2016).

Alongside concerns in the literature about the best way to collect data on housing quality is the issue of the best way to measure housing quality. Curiously, much more attention has been paid to developing a summary measure than to identifying the individual housing features that should comprise the summary measure. The pioneering work of Kain and Quigley (1970) established the feasibility of using housing-unit-level measures of quality as predictors of house prices and rents, also known as hedonic models. The coefficients in these models can be viewed as weights in a hedonic price index. Kain and Quigley's work led to a burgeoning of hedonic modeling over the ensuing decades (Coulson and Li, 2013; Kriström, 2008; Merrill, 1980; Thibodeau, 1995). Three features of Kain and Quigley's approach are particularly relevant to this current paper. First, they apply factor analysis to reduce the 39 separate measures of housing quality in their St. Louis survey data to a manageable number. Second, they find that the 7 survey measures pertaining to the quality of the individual dwelling unit interior formed a single index or factor.[6] And third, in multivariate hedonic regressions, the dwelling unit quality factor has a statistically significant effect on rent. Consistent with most of the literature in this area, the authors do not take on the question of how best to conceptualize housing quality and, instead, assume that this concept is captured by their 39 variables pertaining to "the physical or visual quality of the bundle of residential services" (Kain and Quigley, 1970: 534).

AHS Measurement of Housing Quality

The AHS is the most comprehensive data source on the U.S. housing stock.[7] Policymakers, practitioners, and researchers seeking answers to questions about the conditions, costs, and various other attributes of the nation's housing rely on it. It is also relied on as a source of housing questions for those developing their own surveys. A prominent example is the Moving to Opportunity for Fair Housing Demonstration (Shroder, 2001). Of particular interest to many users is the AHS composite measure of housing inadequacy available on the public use database, a variable labeled ZADEQ. The measure combines multiple items on housing conditions into an index, setting numerical thresholds for the presence or absence of physical deficiencies in the dwelling to distinguish among "adequate," "moderately inadequate," and "severely inadequate" units.[8] Both the AHS and data users refer to this composite as the AHS "housing quality" measure.[9] Numerous published articles include the AHS measure in their analyses (Carter, 2011; Friedman and Rosenbaum, 2004; Khadduri, 2007; Ross et al., 2012). It plays a prominent role in HUD's *Worst Case Housing Needs* reports (for example, HUD, 2011b) and is also included in

[6] Their survey included many other items focusing on the condition of adjacent structures, parcels and block faces, along with the structure's exterior condition.
[7] The initial part of this discussion draws from Newman and Garboden (2013).
[8] The number of items varies from one AHS to another. Eggers and Moumen (2013a) indicate it is 14 items, Newman and Garboden (2013) indicate it is 15 items, and we count 20 items in the 2011 AHS.
[9] The AHS codebook notes: "This three-scale index, in which one is adequate and three is severely inadequate, is a summary measure of housing quality." The composite measure is also listed in the AHS documentation under the category "Unit Quality" (HUD, 2011a).

the frequently cited Joint Center for Housing Studies' *The State of the Nation's Housing* reports (for example, JCHS, 2010) and U.S. Millennial Housing Commission Report (2002). However, not until the last few years was the AHS quality measure subjected to careful examination.

Eggers and Moumen's (2013a) analysis of the 2005, 2007, and 2009 AHS data implicitly raises some concerns about whether the ZADEQ measure accurately reflects the quality of the housing stock. The measure produces a very low incidence of severe inadequacy (2 percent); only two items—sharing a bathroom and heating problems—account for most of the cases considered severely inadequate, and these problems generally do not persist over a 2-year period. The authors are particularly troubled by the lack of persistence in the shared bathroom item, investigate it in greater detail, and conclude that this item is likely to have been measured incorrectly. More generally, they conclude that the AHS quality measure may provide a reasonable cross-sectional estimate of the most severely inadequate units, but provides little information on the roughly 91 percent of units considered adequate.[10]

In a second paper (2013b), Eggers and Moumen propose an alternative to ZADEQ that is designed to provide more information about gradations within the adequate housing stock. A major motivation is their particular interest in being able to study filtering, which requires a measure to reveal increases in deficiencies or inadequacies over time as a unit deteriorates and is presumably filtered down from higher income to lower income residents. They develop an alternative measure, the poor quality index (PQI) that is a numeric scale of housing defects that draws on additional measures in the AHS (for example, assessments of the structure's exterior), as well as those included in ZADEQ. Lacking a reliable source on how to weight each item in the index, they assign weights based on a combination of ZADEQ's definitions and their own judgment. The PQI appears to achieve the goals of its creators. By contrast to ZADEQ, which estimates that a large majority of units had no problems, 47 percent of units had at least one PQI inadequacy.[11] The stability of the classification of the unit also differs for the two indices (Eggers and Moumen, 2013b: 9). According to ZADEQ, 95 percent of units designated as adequate in one survey remain adequate 2 years later, while roughly 30 to 35 percent of units categorized as moderately or severely inadequate in one survey remained inadequate in the subsequent survey. Using the PQI, a smaller share, 63 percent of units, remained adequate from one survey to the next, and a greater share of inadequate units, roughly 60 percent, retained that designation over 2 years.

Emrath and Taylor (2012) examine the AHS ZADEQ index using a hedonic model. Because of the multicollinearity among the individual measures that comprise ZADEQ, the authors test each ZADEQ item separately, along with other features of the dwelling (for example, number of rooms, geographic region, and square footage). They report that none of the ZADEQ items reach statistical significance and, in some cases, have an unexpected sign. A major policy concern of

---

[10] Authors' estimate based on the 2011 AHS.

[11] The PQI rate is based on the 1993 AHS. Unfortunately, the two Eggers and Moumen reports (2013a; 2013b) rely on different AHS years, so it is impossible to make direct comparisons between ZADEQ and PQI results.

the authors is that the very small rate of housing units meeting the definition of physically inadequate using ZADEQ leads to the conclusion that the nation's housing stock has no serious housing problems. They challenge this conclusion by identifying measures in the AHS, many of which are not included in ZADEQ, that have a strong effect on rents and prices. These items are similar to those included in Eggers and Moumen's (2013b) PQI. It is likely that Emrath and Taylor's ZADEQ results occur because of the very low variance of each individual item. This was part of Kain and Quigley's (1970) motivation for using factor analysis, which produced a single dwelling unit quality factor.[12]

A fourth recent paper assesses the reliability, consistency, and validity of the AHS ZADEQ index (Newman and Garboden, 2013). Like Eggers and Moumen (2013a; 2013b) and Emrath and Taylor (2012), the authors conclude that the index identifies only a very small share of units with multiple inadequacies and provides little information about variations among units classified as adequate. They also find that the items included in the index do not appear to be tapping the same underlying construct of housing quality. However, the two subindices within ZADEQ, moderate inadequacy and severe inadequacy, are strong and statistically significant predictors of residents' housing satisfaction.

Quality of Assisted Housing

The research literature on the physical quality of the assisted housing stock is sparse, at least in part, because the AHS, the main data source on housing, typically relies on respondent self-reports of the receipt of housing assistance, which are known to be unreliable (Shroder, 2002). The present paper can utilize the 2011 and 2013 AHS data, because the identification of assisted housing receipt, by program type, relies on administrative records, not self-reports. Validation of assisted housing receipt was previously done in the 1989 AHS. One paper used these validated data and a version of ZADEQ to study the assisted housing profiles of households with children (Newman and Schnare, 1993). The authors report that 15 percent of public housing units occupied by households with children had either a moderate or severe defect, compared with 5 percent of multifamily housing and 12 percent of voucher units. The average number of defects, however, was generally similar across the programs.

A more recent study examined the quality of housing in the voucher program (Buron et al., 2000). Data on voucher housing come from the 2000 Customer Satisfaction Survey (CSA). The authors developed two measures of housing quality, one relying on all quality-related items in the CSA and another using CSA items that align with those in the AHS. The CSA-based measure was used to explore voucher housing quality, and the CSA-AHS measure was used to compare housing quality in the voucher program with housing quality in a matched comparison sample of unassisted renters. The CSA-based summary measure combined items into four categories:

---

[12] Merrill (1980) used a somewhat similar approach in her hedonic modeling using data from the Experimental Housing Allowance Program demand data.

severely inadequate quality, moderately inadequate quality, adequate quality, and high quality. Based on voucher respondent reports to the CSA, 41 percent of voucher housing was considered high quality, 33 percent adequate quality, and 23 percent severely inadequate. The rate of severe inadequacy is higher than the 12- to 21-percent range in Gray et al. (2008), HUD's report on the first-year results of the CSA, which relied on similar though not identical quality measures. The analysis of voucher and comparable nonvoucher housing quality was based on a statistical match between the households in the CSA's voucher sample and households in the AHS. They use two different measures of housing quality, one a simple count of problems aggregated into four categories (0, 1–2, 3–4, 5+ problems) and another indicating whether at least one problem was reported for each of five housing dimensions (for example, kitchen and bathrooms; electrical). Both measures yield the similar finding of lower-quality housing of voucher users than housing occupied by unassisted renters. For example, 59 percent of voucher renters reported no housing problems, compared with 66 percent of unassisted renters. The authors caution that differences between the CSA and AHS may account for some or most of these disparities.

Research Approach

Data

Our main sources of data are the 2011 and 2013 AHS. The AHS—begun in 1973, sponsored by HUD, and conducted by the Census Bureau—is the primary source of data on housing in the U.S. The national AHS sample typically includes about 60,000 nationally representative residential dwellings. The 2011 and 2013 surveys differ in using larger samples. In 2011, to reduce costs and improve the precision of national estimates, more than 115,000 units from the 29 metropolitan area samples were merged with the national AHS into a single "combined national sample" of more than 186,000 housing units (Census Bureau, 2015). The 2013 national AHS includes 84,355 units, a smaller sample size than in 2011 because it does not combine national and metropolitan area samples except for the "big 5" metropolitan areas—Chicago, Detroit, New York City, Northern New Jersey, and Philadelphia—that have historically been included in the AHS national samples.

Another unique feature of the 2011 and 2013 data is that addresses of assisted housing are identified based on matching sample addresses to HUD administrative data on HUD assisted housing programs.[13] This greatly reduces the reliability and validity problems of self-reported receipt of housing assistance. In addition, an oversample of roughly 5,000 assisted housing units was added to the national AHS beginning in 2011. Oversampling ensures adequate representation

[13] The match to HUD data excludes housing units assisted by state and local programs and the federal Low-Income Housing Tax Credit (LIHTC) program, which is under the auspices of the Department of the Treasury. However, because a sizable share of LIHTC units also receives HUD subsidies, such as vouchers, these units are included in the HUD administrative files. O'Regan and Horn (2012) estimate that 46 percent of LIHTC households receive some form of rental assistance, Buron et al. (2000) puts the estimate at 37 percent, and GAO's (1997) estimate is 39 percent.

of individuals (in this case, assisted housing units) that do not occur frequently in the population. The 2011 AHS sample includes 9,721 assisted housing units and the 2013 sample includes 6,922 assisted housing units.[14]

Samples

Our first analytic step is to develop indices that are designed to tap the concept of housing quality. We base this analysis on the unassisted stock of rental housing, excluding all units designated by the administrative address match as part of the assisted housing stock. Because we will ultimately apply the quality indices to the assisted stock, we limit the unassisted cases to units in a single-family or multifamily property and exclude unassisted cases that are rare or nonexistent in the assisted stock (for example, manufactured housing; reduced rents because of relationship between renter and landlord). We also exclude vacant or vacation units and units where no interview was conducted. Table A1 highlights these selection criteria. After excluding these cases, the 2011 analysis sample for index development is 40,830 unassisted housing units. We develop the housing quality indices using the 2011 AHS to take advantage of its much larger sample size than the 2013 AHS.[15] We base the indices on 33 AHS items on housing conditions. Two-thirds of these items have no missing data, and only 1 percent of observations are missing on 25 percent of the items. Because such low rates are very unlikely to affect the estimates, we make no adjustments for missing data. Diagnostics on the full analysis sample revealed virtually no missing data or outlier values, and the AHS staff assigned reasonable top codes to continuous measures.[16] Top coding may be used to prevent a small number of very high values from distorting analysis.

Methods

Because the concept of housing quality is not based on explicit criteria, the large number of housing quality indices that have been developed with the AHS yield dramatically different prevalence rates (Newman and Schnare, 1988). The core challenge is well known: a housing unit is a bundle of attributes that extend beyond the dwelling itself, and it is unclear which of these attributes should be included in the definition of the dwelling's quality and how each should be weighted in determining overall quality (Merrill, 1980; Aaron, 1972).

In the absence of consensus responses to these questions, the next best option is to rely on an external criterion, as suggested by Merrill (1980). We examine three alternatives: market value, consumer rating, and normative standards. We test the market value criterion with a hedonic approach. As described earlier, it assumes that the unit's rent is correlated with the quantity and quality of housing such that higher rents reflect better quality. The consumer rating criterion

[14] Based on sample design appendices to 2011 and 2013 AHS documentation.
[15] We did not combine the 2011 and 2013 AHS samples because the 2013 AHS includes a subsample of units from the 2011 AHS. Because two waves of data are not available for every unit, combining the 2011 and 2013 AHS samples would overweight the importance of housing units included in both AHS samples.
[16] These measures include breakdowns in heating, toilets, or sewage; blown fuses; and rent.

identifies the dwelling features that are most closely associated with the resident's assessment of the dwelling as a good place to live regardless of what the market price of these features might be. This criterion broadens the concept of housing quality beyond specific housing features to the welfare of residents as they themselves report it (Goodman, 1978). It is consistent with the renewed interest by economists in happiness and subjective well-being as a measure of the utility an individual derives from goods and services (Dolan, Peasgood, and White, 2008).[17] The normative standards criterion is designed to reflect community concerns and policy decisions about housing quality, such as state building codes and assisted housing physical inspection standards. We test three versions of a normative standards index: a weighted index, an unweighted index, and a factor analysis index. This analytic strategy allows us to systematically examine housing quality from three different perspectives, driven by the empirical evidence. It contrasts with ad hoc approaches, such as assigning weights to individual housing quality items based on intuition or setting a priori thresholds on the number of breakdowns in key housing systems, such as heating and plumbing, which distinguish severe housing quality problems from moderate problems.

Although the AHS contains a rich array of measures plausibly associated with housing quality, some measures are not repeated from one wave to the next, and not all questions are asked of all housing types (that is, single-family or multifamily). Because our analysis includes a longitudinal component in which we assess whether housing inadequacies persist over the 2-year period from the 2011 survey to the 2013 survey, we only incorporate measures included in both surveys.[18] We also limit the analysis to measures collected from all rental units, including single-family properties and the housing units in multifamily buildings that are most relevant to our focus on the assisted housing stock. Using this sample of rental housing units, we construct housing quality indices using the consumer rating and normative standards perspectives, as described in more detail below. Because the market value index does not perform well, we exclude it from further analysis. We construct one consumer rating index and three indices using the normative standards criterion. In general, the indices are calculated by assigning a weight to each housing quality measure and then adding up the weighted scores of all quality measures identified for each housing unit. For the consumer rating index, we apply the coefficients from regression models used to predict house rating as weights. We treat each of the three normative standards indices differently. We assign weights to one index based on a previous analysis by Eggers and Moumen (2013b), assign the same weight of 1 to all individual items in a second index, which is equivalent to an unweighted index, and use the results of a factor analysis to assign weights to a third normative standards index.

---

[17] The market value model and the consumer rating model are also associated with other elements of the housing bundle that extend beyond the dwelling itself, such as neighborhood features. Therefore, we also control for neighborhood in these models.
[18] See Appendix B on the Healthy Homes module in the 2011 AHS.

We assess the validity of the four well-performing indices we construct by establishing whether the indices are strongly related to each other and to other measures that we can reasonably expect would be associated with housing quality, by studying their “convergent validity” and their “predictive validity” (Carmines and Zeller, 1979). Because each index is designed to tap the concept of housing quality, the indices should be strongly related to each other and to other measures that we can reasonably expect would be associated with housing quality. Further, if each index is effective at tapping housing quality, then it should predict quality-related outcomes. We estimate convergent validity by examining the correlations, both among the four indices and between each index and other attributes with which the index should be correlated, such as rent and building age. We estimate predictive validity by testing multivariate models of the predictive power and significance of the indices in predicting such outcomes as rent and the resident’s housing satisfaction as measured by rating the house as a place to live on a 10-point scale.

Results

Table 1 lists the AHS housing quality measures in this analysis and their mean values.[19] Consistent with much past AHS housing quality research, the prevalence rates of almost all problems are very low. Most (55 percent) dwellings have no problems, and fewer than 5 percent of units account for more than 75 percent of problems.[20]

[19] This table includes a few measures (for example, mold; broken steps) that we look at descriptively but do not include in any of the indices, because they do not meet all criteria.
[20] Because the AHS included no reports of “number of times completely without running water,” we dropped this item from all subsequent analyses.

Table 1. 2011 AHS Housing Quality Measure, Mean Percent of All Housing

		Mean Prevalence Rate	Mean Times for Counts
1.	Not all rooms have plugs	1	
2.	# times blown fuses last 3 months	9	0.23
3.	Exposed wiring	2	
4.	Unit does not have electricity	.03	
5.	Unvented room heaters	1	
6.	No heating equipment	1	
7.	Use stove/oven for heat	0.1	
8.	# heating breakdowns last winter	3	0.08
9.	Unit cold 24+ hours last winter	10	
10.	Cold due to utility interruption last winter	1	
11.	Cold due to inadequate heating capacity last winter	2	
12.	Cold due to inadequate insulation last winter	2	
13.	Cold due to other reason last winter	2	
14.	Roof leak last 12 months	5	
15.	Leak in wall/closet last 12 months	3	
16.	Leak in basement last 12 months	1	
17.	Leak other source last 12 months	1	
18.	Leaking pipes last 12 months	5	
19.	Leaking plumbing fixture last 12 months	2	
20.	Leak unknown source last 12 months	4	
21.	Crack in wall	7	
22.	Holes in floor	1	
23.	Peeling paint	3	
24.	Signs of rodents last 12 months	3	
25.	Signs of rats last 12 months	1	
26.	Signs of mice last 12 months	9	
27.	Signs of cockroaches last 12 months	5	
28.	Incomplete plumbing	0.3	
29.	# times toilet broke 6+ hours last 3 months	2	0.05
30.	Share plumbing facility	2	
31.	Incomplete kitchen	4	
32.	# times no water last 3 months	—	
32.	# sewage disposal breakdowns last 3 months	1	0.03
33.	No working elevator	5	
34.	Any mold	5	
35.	Broken stairs	1	
36.	Broken steps	1	

Notes:

1. N = 40,830 unassisted rental units from 2011 AHS, excludes manufactured housing and units where a relationship exists between renter and landlord. See text and table A-1 for more details.
2. Weighted data.
3. “Mean times for counts” = average for entire sample, not just those reporting the problem.
4. “# times no water last 3 months” never reported.
5. Items 34, 35, and 36 from Healthy Homes module.
6. AHS = American Housing Survey.

Housing Quality Indices

Market Value Index

We estimate the market value index with a hedonic approach. Following standards established in highly regarded housing economics research (Coulson and Li, 2013; Kriström, 2008; Thibodeau, 1995), we specify the dependent variable as the natural log of rent and add a granular level of controls for multiple features of the housing unit (for example, size; amenities) and geographic location (for example, region; city; suburb; rural). Because of the established importance of neighborhood quality in market values, we include the respondent's rating of the neighborhood, the only proxy measure available in the 2011 AHS, and include HUD's FMR for the metropolitan area as a gauge of the local housing market. Although our main interest is in the contribution of housing quality to market value, this effect could depend on the quality of the neighborhood, for example, or whether the unit is located in a central city, suburb, or rural area (often referred to as "urbanicity"). Therefore, excluding these additional measures from the model could provide misleading results. The full list of controls (or covariates) is shown in table 2.

Table 2. Control Measures Included in Market Value (Hedonic) Models

Systems	**Unit**
Room air conditioner	Single-family
Central air conditioner	Single-family attached
Dishwasher	Multifamily < 3 units
Garbage disposal	Multifamily 3–4 units
Clothes dryer	Multifamily 5–9 units
Washing machine	Multifamily 10–19 units
Electric heat	Multifamily 20+ units
Gas heat	
Oil heat	**Utility costs**
	Electric in rent
Physical	Gas in rent
# Baths	Oil in rent
# Bedrooms	Other fuel in rent
Basement	
# Floors to unit	**Neighborhood/location**
Den/library/TV room	Neighborhood rating
Dining room	Region/metropolitan area
Family room	HUD Fair Market Rent
Working fireplace	
Garage or carport	**Other**
# half bathrooms	Connected public sewer
Laundry/utility room	Use well water
Porch/deck/balcony/patio	# months live in unit
# Rooms	
Building age	
Building age-squared	
Building age-cubed	
# Floors in building	
Other fuel in rent	

Notes:
1. Control measures from 2011 American Housing Survey.
2. Models also include the 33 individual housing quality measures, see text for details.
3. # baths topcoded at 3.
4. # bedrooms and # floors topcoded at 4.

Because the 33 housing quality items in the analysis are often closely related to each other (statistically referred to as "multicollinearity"), the best method to assess the contribution of housing quality reflected in these individual quality items to rent is to combine them in a composite measure. Unfortunately, this defeats our purpose of developing a weight for each quality item to represent its individual contribution to rent. This leaves two alternative approaches, neither of which is ideal. The first option is to examine each of the 33 quality items separately. Although this approach provides a weight for each individual quality measure, it runs the risk of bias from omitting the other quality items. The second option is to include all 33 items in a single model. Although this approach also provides weights for individual quality items, it is less than ideal because of the substantial correlation among the 33 quality items. Recognizing that neither approach is ideal from a methodological standpoint, we test both.

To estimate the effect of each individual measure of housing quality on market value, we tested 33 hedonic models, one for each housing quality measure. The results are shown in table 3. Excluding any measure of housing quality, the other predictor variables account for 36 to 42 percent of the variance (that is, the R^2). This estimate remains unchanged after each individual housing quality measure is added, indicating virtually no contribution to the explanatory power of the model. Of the 33 dwelling quality items, 12 reach statistical significance at least at the .10 level, but 7 of these coefficients have an unexpected sign. That 85 percent of a large number of items purporting to represent some aspect of dwelling quality either do not reach statistical significance or, in 7 cases, operate in the opposite direction of expectations, raises serious questions about the usefulness of the market value approach in this context. Further, the large number of multiple comparisons required to operationalize the market value approach runs the risk of generating some statistically significant estimates purely by chance. Therefore, we adjust the threshold required for statistical significance using the Bonferroni correction.[21] After the correction, four measures remain statistically significant: seeing roaches daily; incomplete plumbing; no working elevator; and number of fuses blown. The first three are plausible and have reasonable interpretations. Incomplete plumbing has, by far, the most sizable effect, reducing the average monthly gross rent of about $1,000 by $263.[22] The absence of an elevator reduces the rent by $81, and the persistent presence of roaches by $45. Blown fuses have a positive sign, indicating that the rent is higher in dwellings with more blown fuses. It is difficult to interpret this odd relationship. However, despite the statistical significance of this item, it increases the rent by only about $9.[23]

---

[21] The Bonferroni correction is used to reduce the likelihood of reporting a false positive (for example, reporting a random difference as real). The correction divides the original significance level by the number of variables examined (here .10/33 = .003) so that the likelihood of reporting a false positive stays at .10 across the entire set of comparisons.

[22] Dollar interpretations based on: Exp(B)*mean - mean = difference in price. Thus, for incomplete plumbing, where B = -.295: exp (-.295) = .745; times $1,028.8 = $765.974, which is subtracted from the mean rent of $1,028.80, yielding $-262.826.

[23] One reviewer suggested that blown fuses may interact with historic buildings in high-cost urban areas.

Table 3. Market Value (Hedonic) Model Results, Housing Quality Measures

	Separate			Combined		
	Coef.	*p*-value		Coef.	*p*-value	
Not all rooms have plugs	– .043	.049	*	– .042	.060	+
# times blown fuses last 3 months	.009	.000	***	.010	.000	***
Exposed wiring	.008	.586		.001	.985	
Unit does not have electricity	.332	.171		—		
Unvented room heaters	– .063	.038	*	– .068	.029	*
No heating equipment	.011	.740		—		
Use stove/oven for heat	– .072	.283		– .044	.521	
# heating breakdowns last winter	.002	.624		.001	.815	
Unit cold 24+ hours last winter	.011	.153		.005	.741	
Cold due to utility interruption last winter	.028	.253		.015	.584	
Cold due to inadequate heating capacity last winter	.018	.285		.016	.424	
Cold due to inadequate insulation last winter	.028	.093	+	.016	.417	
Cold due to other reason last winter	– .008	.648		– .013	.550	
Roof leak last 12 months	.025	.018	*	.024	.024	*
Leak in wall/closet last 12 months	.025	.058	+	.019	.178	
Leak in basement last 12 months	.040	.034	*	.034	.079	+
Leak other source last 12 months	.018	.383		.013	.524	
Leaking pipes last 12 months	.014	.171		.013	.230	
Leaking plumbing fixture last 12 months	.006	.685		.009	.516	
Leak unknown source last 12 months	– .003	.806		– .006	.631	
Crack in wall	– .008	.400		– .010	.320	
Holes in floor	– .018	.367		– .011	.594	
Peeling paint	– .022	.103		– .032	.028	*
Signs of rodents last 12 months	– .021	.155		– .025	.158	
Signs of rats last 12 months	.023	.242		.038	.079	+
Signs of mice last 12 months	– .007	.422		– .001	.856	
Signs of cockroaches last 12 months	– .045	.000	***	– .049	.000	***
Incomplete plumbing	– .295	.000	***	– .242	.000	***
# times toilet broke 6+ hours last 3 months	– .004	.450		– .006	.247	
Share plumbing facility	– .040	.026	*	– .039	.033	*
Incomplete kitchen	– .011	.288		– .007	.537	
# sewage disposal breakdowns last 3 months	.016	.049	*	.014	.039	*
No working elevator	– .081	.000	***	– .073	.000	***

Notes:
1. Separate = Results from 33 ordinary least squares (OLS) regressions on log rent with individual housing quality measures.
2. Combined = Results from single OLS regression on log rent with all housing quality measures in model.
3. See table 2 for other measures included in models.
4. "—" = dropped from combined analysis because perfectly correlated with other measures.
5. N = 36,833.
6. *p*-values: + < .10, * < .05, ** < .01, *** < .001.

To test the effects of all 33 quality measures together, we estimated a hedonic model that included all 33 items simultaneously. These results are also shown in table 3. There is remarkable consistency between the two sets of analyses. Of the 12 statistically significant measures identified in the separate analyses, 10 are also statistically significant in the combined model. The same measures also emerge from both analyses as the strongest and most statistically significant: the number blown fuses, signs of cockroaches, incomplete plumbing, and no working elevator. Several of the unexpected results in the individual item models are also replicated in the simultaneous model, such as the positive sign on blown fuses, leaking roofs, and leaking basements, indicating that these problems increase the rent instead of decreasing the rent, as anticipated.

These hedonic results contrast with those of Kain and Quigley (1970) and Merrill (1980). The difference is that we are interested in the effect of each individual measure of housing quality on rent and, therefore, we test each measure separately. Kain and Quigley and Merrill combine measures of housing quality into a composite indicator or factor.[24] Our empirical results indicate that, despite a very large sample, 85 percent of the AHS measures are not statistically significant, and the rarity of each individual quality problem provides too little variance to make a meaningful contribution to market values. Therefore, despite the appeal of the external criterion of market value for assessing the relative importance of individual housing quality features, this approach is not feasible. Although estimating the relative contribution of a composite indicator to market value is feasible, it does not achieve the objective of identifying the most important individual measures of housing quality.

Consumer Rating Index

To assess the contribution of each housing quality measure to the household's preferences, we rely on the AHS question asking the respondent to provide a house rating. The specific item asks: "On a scale of 1 to 10, how would you rate your unit as a place to live?" Although the original coding designates 10 as best and one as worst, we reverse these codes for consistency with the normative standards indices we discuss later. Thus, a higher value indicates lower housing quality.

We test the consumer rating model using ordered logistic regression, because the house rating is an ordinal scale with only 10 points. As a result, the estimated contribution of each of the quality measures, or parameter estimates, is expressed as odds ratios: how much a unit changes in the housing quality item changes the house rating. As with the market value criterion, we test consumer rating two ways: individual quality items separately, and all quality items combined. Because the results are similar, we present the results from the separate quality measure tests for simplicity.[25]

The results are shown in table 4. In contrast to the market value analysis, each quality measure is shown to contribute to the consumer's rating. Virtually all of the odds ratios are greater than one, indicating that the presence of the condition worsens the rating, as hypothesized.[26] The five measures that make the largest contribution to house rating are: holes in the floor; peeling paint; cracks in the walls; presence of rodents;[27] and cold due to inadequate insulation. For example, the presence of holes in the floor makes it 3.5 times more likely that the consumer's house rating is poorer. Peeling paint makes a poorer score 3.3 times more likely, and rodents

---

[24] Creating a composite indicator also addresses the multicollinearity among the individual quality items, as does testing these items one at a time.

[25] In the model including all quality measures simultaneously, the odds ratios tend to be roughly 25 percent lower because of multicollinearity, and some nearly perfectly collinear measures do not reach statistical significance (for example, exposed wiring; reasons for cold unit; sewage disposal breakdown).

[26] One exception is the use of an oven for heat, which affects a very small proportion of rental units (see table 1).

[27] The item on rodents is asked because some respondents cannot be certain whether they are seeing mice or rats. These three items are very highly correlated, as shown in table 13. Arguably, all three items are substantively important, because they overlap.

make this 2.7 times more likely.[28] It is worth noting that these items do not overlap at all with the few statistically significant and qualitatively important measures that emerge from the market value model. The consumer rating model also attains good explanatory power, the large majority of the quality items are highly statistically significant, and they operate in the expected direction. We use the odds ratio for each quality measure as a weight to create a housing quality index based on the consumer rating criterion.

Table 4. Consumer Rating Model Results, Housing Quality Measures

	Odds Ratio	*p*-value	
Not all rooms have plugs	1.973	.000	***
# times blown fuses last 3 months	1.175	.000	***
Exposed wiring	1.363	.000	***
Unit does not have electricity	1.770	.573	
Unvented room heaters	1.075	.588	
No heating equipment	1.261	.106	
Use stove/oven for heat	0.725	.269	
# heating breakdowns last winter	1.225	.000	***
Unit cold 24+ hours last winter	2.017	.000	***
Cold due to utility interruption last winter	1.731	.000	***
Cold due to inadequate heating capacity last winter	2.212	.000	***
Cold due to inadequate insulation last winter	2.656	.000	***
Cold due to other reason last winter	1.531	.000	***
Roof leak last 12 months	1.987	.000	***
Leak in wall/closet last 12 months	1.801	.000	***
Leak in basement last 12 months	1.921	.000	***
Leak other source last 12 months	1.541	.000	***
Leaking pipes last 12 months	1.678	.000	***
Leaking plumbing fixture last 12 months	1.904	.000	***
Leak unknown source last 12 months	1.482	.000	***
Crack in wall	2.708	.000	***
Holes in floor	3.509	.000	***
Peeling paint	3.253	.000	***
Signs of rodents last 12 months	2.657	.000	***
Signs of rats last 12 months	2.127	.000	***
Signs of mice last 12 months	1.626	.000	***
Signs of cockroaches last 12 months	2.052	.000	***
Incomplete plumbing	1.430	.197	
# times toilet broke 6+ hours last 3 months	1.275	.000	***
Share plumbing facility	1.111	.172	
Incomplete kitchen	1.208	.000	***
# sewage disposal breakdowns last 3 months	1.243	.000	***
No working elevator	1.262	.002	**

Notes:
1. Results from 33 separate ordinal logistic regressions on subjective housing rating (reverse coded so high score = poorer quality). See table 2 for other measures included in model.
2. N = 36,833.
3. *p*-values: + < .10, * < .05, ** < .01, *** < .001.
4. Odds ratio = e^{β} where "β" is the coefficient from the logistic model.

[28] The odds ratios apply identically across the house rating distribution (that is, across the 1–10 rating). Thus, there are equal odds of a higher score if the original score is 2 versus an original score of 8, for example.

Normative Standards Index

We test three versions of a normative standards index: a weighted index, an unweighted index, and a factor analysis index. Each of these indices is based on the same array of housing quality measures included in the market value and consumer rating models.

Weighted Index. Because no consensus criteria exist for assigning weights to individual housing conditions presumably associated with housing quality, we adopt the weights assigned by Eggers and Moumen (2013b), which are based on a combination of the AHS ZADEQ housing quality measure and the authors' judgment.[29] For example, Eggers and Moumen assign their highest weight, 10, to each deficiency classified by ZADEQ as a severe inadequacy (that is, incomplete plumbing, incomplete kitchen, lacking electricity, lacking heating equipment, using a stove or oven for heat). Conversely, Eggers and Moumen assign their lowest weight, 2, to deficiencies that ZADEQ includes in its "moderately inadequate" category. The second column in table 5 lists the Eggers and Moumen weights for the items included in the present analysis.

To address items that measure how many times a particular problem arises (for example, number of heating system breakdowns), we compute the total weight by multiplying the number of times the problem occurred by the Eggers and Moumen weight. We also standardize the weights to account for minimal missing data and multiply the result by 100, which produces an index value that ranges from 0 to 100.[30]

Unweighted Index. The unweighted index directly acknowledges the absence of consensus criteria for assigning weights. Each item in this index has the identical weight of 1, as shown in the third column of table 5. As in the weighted index, we standardize for nominal missing data and multiply the result by 100.

Factor Analysis Index. Our third methodology for creating a normative standards index is factor analysis. This method provides a more systematic approach for weighting each of the 33 measures purported to reflect housing quality than the relatively ad hoc approach of the weighted and unweighted indices. Factor analysis examines the correlations among measures to determine the amount of common variance among them.[31] The analysis produces factor "loadings," which indicate how much variance is shared among the observed measures and the unobserved construct (in the present case, housing quality). The loadings or scores can be used as weights to create the factor analysis index.

---

[29] We exclude measures that do not meet our criteria noted earlier. Therefore, our measures do not entirely overlap with those of Eggers and Moumen's PQI.

[30] Imagine, for instance, that all of the housing measures are equally weighted with a score of 1 and that we are comparing two housing units, one with responses on all 33 housing quality measures and the other with responses to only 30 measures. Each household reports having two problems, so each has an initial index score of 2. For the first unit, we adjust this score by dividing it by the number of observed problems (that is, 2/33). For the second unit, this same procedure reflects that fewer questions were answered (that is, 2/30). The weighted index is computed using this procedure but adjusting the weights to account for the different values assigned by Eggers and Moumen.

[31] As Newman and Garboden (2013) expressed it, how well the measures "hang together."

Table 5. Normative Standards Weighted and Unweighted Indices

	Weighted	Unweighted
Not all rooms have plugs	3	1
# times blown fuses last 3 months	1	1
Exposed wiring	4	1
Unit does not have electricity	10	1
Unvented room heaters	4	1
No heating equipment	10	1
Use stove/oven for heat	10	1
# heating breakdowns last winter	2	1
Unit cold 24+ hours last winter	4	1
Cold due to utility interruption last winter	2	1
Cold due to inadequate heating capacity last winter	2	1
Cold due to inadequate insulation last winter	2	1
Cold due to other reason last winter	2	1
Roof leak last 12 months	2	1
Leak in wall/closet last 12 months	2	1
Leak in basement last 12 months	2	1
Leak other source last 12 months	2	1
Leaking pipes last 12 months	2	1
Leaking plumbing fixture last 12 months	2	1
Leak unknown source last 12 months	2	1
Crack in wall	2	1
Holes in floor	2	1
Peeling paint	2	1
Signs of rodents last 12 months	1	1
Signs of rats last 12 months	2	1
Signs of mice last 12 months	1	1
Signs of cockroaches last 12 months	3	1
Incomplete plumbing	10	1
# times toilet broke 6+ hours last 3 months	2	1
Share plumbing facility	2	1
Incomplete kitchen	10	1
# sewage disposal breakdowns last 3 months	2	1
No working elevator	4	1

Notes:
1. Weighted index values based on Eggers & Moumen (2013b), see text for details.
2. For continuous variables = # times * weight value.
3. Both weighted and unweighted scores are summed, divided by the total possible score and multiplied by 100 to create the final index.

Because many of the AHS quality measures are dichotomous, based on "yes or no" questions (for example, is there a leak in the roof? has the resident seen rodents?), we use a polychoric correlation matrix.[32,33] The results are shown in table A2. One challenge in estimating a polychoric factor analysis is that the standard approaches for selecting the number of factors (that is, examining the scree plot or comparing chi-square values) are either inappropriate or

[32] *Polychoric correlation* is a technique for estimating the correlation between two theorized normally distributed continuous latent variables from two observed ordinal variables.
[33] We used Mplus to estimate the polychoric matrix.

unavailable. However, Preacher et al. (2013) recommend that when the goal is to identify the number of factors that most accurately summarize the underlying structure of the data, the best approach is to select the smallest number of factors for which the root mean square error of approximation (RMSEA) is below .05.

Table 6 shows the RMSEA values for one-, two-, three-, and four-factor models. The table shows the lower and upper bound values at the 90 percent confidence interval, along with the average RMSEA scores. Preacher et al. advise that one can select the number of factors when either the average RMSEA score falls below .05, or the lower bound of the RMSEA score falls below this threshold. Fortunately, both the average and lower bound of the RMSEA scores are below .05. The results support the use of a one-dimensional index for measuring housing quality and provide weights for an index based on the factor scores.

Table 6. Normative Standards Factor Analysis: Model Fit Test Results to Select Number of Factors

	Number of Factors Tested			
	1	**2**	**3**	**4**
RMSEA	.026	.019	.013	.010
90% CI lower bound	.026	.018	.013	.010
90% CI upper bound	.027	.019	.014	.011

Notes:
1. Factor analyses estimated in Mplus using polychoric correlations.
2. RMSEA = root mean square error of approximation.
3. CI = confidence interval.
4. Lower and upper bounds based on 90 percent confidence interval for RMSEA, see text for details.
5. Preacher et al. (2013) recommend selecting the smallest number of factors with RMSEA < .05.

Table 7 shows the factor loadings from the single-factor model. Factor models assign one variable to be the reference variable, which in this case is whether all rooms in the dwelling have electric receptacles. This reference variable is automatically assigned a factor score of 1. The highest factor loadings occur for presence of rodents, number of times the dwelling was cold for 24 hours or longer, presence of mice, holes in the floor, peeling paint, and cracks in the walls. Examples of items with the lowest factor loadings are using the oven for heat, lacking heating equipment, number of toilet breakdowns lasting 6 hours or longer, and incomplete plumbing and sharing plumbing facilities.[34] There is notable overlap between the measures with the highest and lowest weights in the factor analysis normative standards model and the consumer rating model.

[34] All of these problems are reported by no more than 2 to 3 percent of units, and incomplete plumbing and using stove for heat are reported by less than 1 percent of units.

Table 7. Normative Standards Index: Factor Analysis Weights

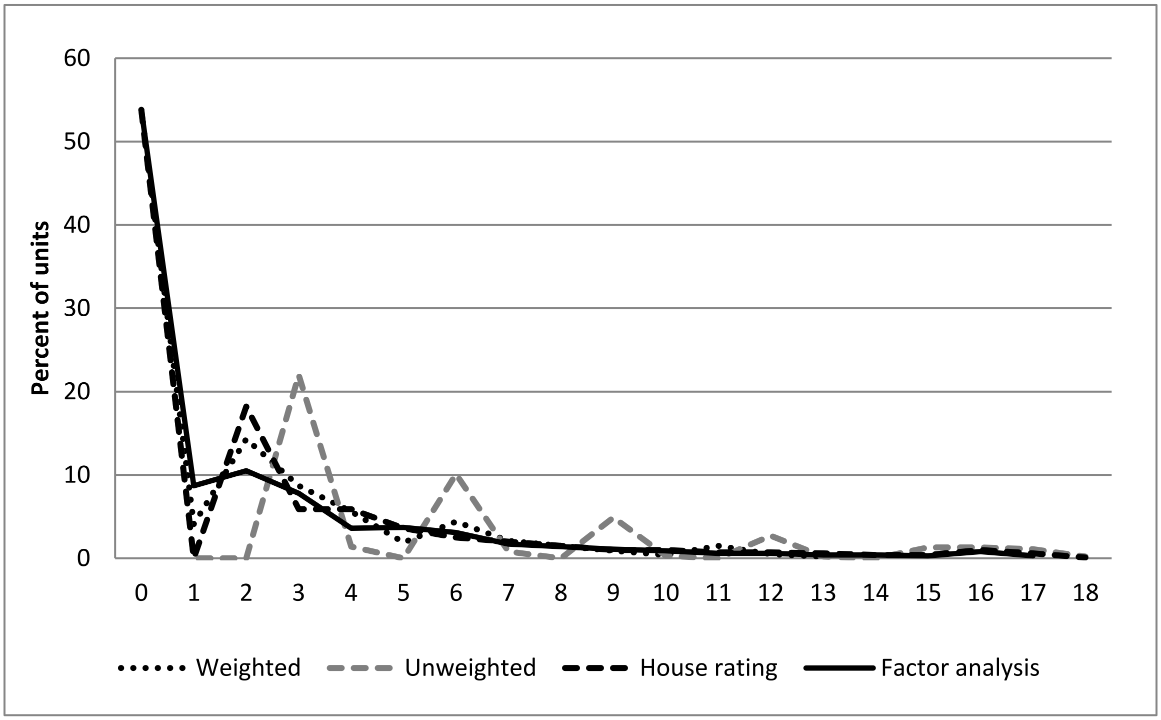
	Factor Analysis Weight
Not all rooms have plugs	1.000
# times blown fuses last 3 months	0.921
Exposed wiring	1.465
Unit does not have electricity	1.649
Unvented room heaters	0.968
No heating equipment	– 1.435
Use stove/oven for heat	– 1.497
# heating breakdowns last winter	0.747
Unit cold 24+ hours last winter	3.340
Cold due to utility interruption last winter	2.593
Cold due to inadequate heating capacity last winter	2.626
Cold due to inadequate insulation last winter	—
Cold due to other reason last winter	2.184
Roof leak last 12 months	2.019
Leak in wall/closet last 12 months	1.920
Leak in basement last 12 months	1.755
Leak other source last 12 months	1.454
Leaking pipes last 12 months	1.812
Leaking plumbing fixture last 12 months	1.787
Leak unknown source last 12 months	1.263
Crack in wall	2.667
Holes in floor	2.702
Peeling paint	2.685
Signs of rodents last 12 months	3.629
Signs of rats last 12 months	2.441
Signs of mice last 12 months	2.783
Signs of cockroaches last 12 months	2.071
Incomplete plumbing	1.629
# times toilet broke 6+ hours last 3 months	0.717
Share plumbing facility	0.270
Incomplete kitchen	0.707
# sewage disposal breakdowns last 3 months	0.207
No working elevator	0.482

Notes: 1. Factor analysis estimated in Mplus using polychoric correlations.
2. "—" = dropped from factor analysis because perfectly correlated with other measures.

A few key features of the foregoing discussion of index creation are worth summarizing at this point. First, despite its intuitive appeal, we are forced to drop the hedonic index reflecting the market value criterion because of its poor performance. In addition, the scores on the four housing quality indices that perform well are highly skewed. This is vividly demonstrated in figure 1 and can also be seen in table A3, which shows the distribution of each of the four housing quality indices. Most housing units have none of the 33 housing quality problems included in this analysis, with only a small fraction experiencing one or more problems.[35]

[35] It is important to note that some of the rarest housing conditions among the 33 measures, such as incomplete plumbing, are also a priori, serious problems and certainly related to housing quality. The only index that accounts for this a priori acknowledgment is the weighted index. However, the high correlation between this index and each of the other three indices in this analysis should allay concerns about this issue.

Figure 1. Distribution of Indices



Notes:
1. For weighted index 11 = 10 to 15, 12 = 15 to 20, 13 = 20 to 30, and 14 = 30+.
2. For consumer rating, normative standards unweighted, and factor analysis indices: 16 = 15 to 20; 17 = 20 to 30; and 18 = 30+.

Index Validity

Convergent Validity

Analytical techniques to check for convergent validity revealed that several indices point to the same findings, suggesting that they are likely to be measuring the same underlying phenomenon or construct.

Table 8 shows the correlations among the four indices. All of the correlations exceed 0.80. The strongest correlations are among the normative standards factor analysis index, consumer rating index, and normative standards unweighted index, which range between 0.94 and 0.97. Relatively speaking, the weighted index is less strongly correlated with the others, with a range of 0.81 to 0.88. Nonetheless, these are still very high correlations.

Table 8. Convergent Validity: Correlations Among Indices

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
Consumer rating	1			
Weighted	.877 (.000)	1		
Unweighted	.937 (.000)	.871 (.000)	1	
Factor analysis	.967 (.000)	.815 (.000)	.949 (.000)	1

Notes:
1. Weighted data.
2. Correlation (*p*-value) two-tailed significance.

Table 9 provides the correlations between each of the four housing quality indices and other AHS measures that are plausibly associated with housing quality. In addition to rent (both actual and logged), house rating, and building age, we also test the relationship of each index to ZADEQ, the housing quality measure included in the AHS data set. Although the house rating measure is the basis for the consumer rating index, it is distinct from the index. The index assigns weights to each of the 33 quality measures based on the respondent's rating of the dwelling, while the housing rating measure is the respondent's rating, from 1 to 10, of the dwelling as a good place to live, without any direct reference to the 33 quality measures.

Table 9. Convergent Validity: Correlations With AHS Measures Related to Housing Quality

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
Rent	– .017 (.000)	– .025 (.000)	– .021 (.000)	– .022 (.000)
Log rent	– .015 (.000)	– .028 (.000)	– .023 (.000)	– .019 (.000)
House rating	– .341 (.000)	– .302 (.000)	– .342 (.000)	– .338 (.000)
Building age	.214 (.000)	.205 (.000)	.235 (.000)	.223 (.000)
ZADEQ	.419 (.000)	.585 (.000)	.436 (.000)	.375 (.000)

Notes:
1. Weighted data.
2. Correlation (*p*-value) two-tailed significance.
3. ZADEQ is measure of housing unit quality computed in the American Housing Survey (AHS).

All of the correlations are statistically significant and all operate in the expected direction. Each of the indices is correlated with ZADEQ and, in fact, these correlations are the highest in the table. On the other hand, the range of these correlations also indicates that our four indices and ZADEQ differ from each other. It is no surprise that the consumer rating index correlates highly with the house rating outcome (measured ordinally, with values ranging from 1 to 10). What is worthy of note is that the correlations of the house rating outcome with the factor analysis index and with the unweighted normative standards index are nearly as high. This suggests that all three indices are likely to be measuring the same underlying phenomenon or construct.

Correlations assume a linear relationship between two measures. But it is also possible that one or more indices have a nonlinear relationship with one or more of the outcomes. We test for nonlinearities by focusing on three outcomes: rent, building age, and house rating. We divide the distribution of index scores for each of the four housing quality indices into five categories, ranging from zero (best quality) to 90+ percent (worst quality).[36] We then calculate the mean value for each of the three outcomes for each of the five categories of each housing index.

Table 10 summarizes the results. The patterns are generally linear, with adequate housing units (zero problems) consistently having the best scores on each of the quality indices and housing units with greater age or the most problems having the worst scores. For building age and house rating, the pattern is also monotonic across the middle of the distribution, consistent with a linear effect. But for rent, the distribution across the middle categories is relatively flat. This suggests that the effect of housing quality on rent is limited to housing units at the highest and lowest ends of the distributions for each of the housing quality indices.[37]

[36] Because there is no assumption of a linear relationship between rent and the categories of the index, we use rent in its unlogged form in this analysis.
[37] The nonlinearity in the rent models helps explain why a market value index using the hedonic specification performs poorly.

Table 10. Convergent Validity: Variation across the Distribution of Each Index

Consumer Rating Index

	0	**1–30%**	**30–70%**	**70–90%**	**90+**	***p*-value**
Rent ($)	1,035	1,048	1,007	1,037	976	.000
Building age (years)	42.5	50.5	50.6	55.8	64.9	.000
House rating	8.12	7.96	7.50	6.94	5.68	.000

Normative Standards Indices:
Weighted

	0	**1–30%**	**30–70%**	**70–90%**	**90+**	***p*-value**
Rent ($)	1,035	1,009	1,061	1,002	943	.000
Building age (years)	42.5	47.8	54.3	54.1	64.1	.000
House rating	8.12	7.76	7.45	7.03	5.95	.000

Unweighted

	0	**1–30%**	**30–70%**	**70–90%**	**90+**	***p*-value**
Rent ($)	1,035	1,026	1,029	1,029	989	.016
Building age (years)	42.3	48.9	53.6	56.3	63.5	.000
House rating	8.12	7.77	7.40	7.09	5.96	.000

Factor Analysis

	0	**1–30%**	**30–70%**	**70–90%**	**90+**	**p-value**
Rent ($)	1,037	1,044	1,006	1,033	968	.000
Building age (years)	42.5	49.0	51.9	56.1	65.6	.000
House rating	8.12	7.89	7.49	6.92	5.73	.000

Notes:
1. Weighted data.
2. Unequal intervals reflect skewed distributions of each index as shown in table A3.

Predictive Validity

Table 11 summarizes the results from testing the predictive power of each of the four indices on two outcomes: house rating and rent. The house rating equation uses an ordered logistic specification, and the rent equation is estimated using an ordinary least squares (OLS) hedonic specification. The strongest results are for the consumer rating index. Poorer housing quality, indicated by higher scores on each of the indices, is associated with a higher, and therefore, worse house rating even after controlling for household and geographic location characteristics. The coefficient on each index is also strongly statistically significant ($p < .001$) relative to house rating. These coefficients suggest that a one-standard-deviation increase in each of the four indices produces nearly a 20 percent reduction (improvement) in house rating. The predictive validity using rent as the outcome is somewhat weaker. Although the coefficient on each housing quality index is negative, supporting the expectation that more housing problems are associated with lower rent, none of the index coefficients is statistically significant at the .05 level, and the consumer rating index does not reach significance at the more liberal .10 level. The somewhat flat relationship between the housing quality measures and rent in the middle part of each distribution shown earlier in table 10 probably contributes to the relatively small size and more modest statistical significance of the effect of each index on rent. Further, the hedonic models explain only about one-third of the variance in housing cost (R^2 ranges from .357 to .358), less than half the total variance explained in, for example, the Kain and Quigley (1970) models (R^2 ~.72).[38] Because Kain and Quigley studied only a single housing market, St. Louis, they are better able to control for features of the market plausibly associated with rents. By contrast, our analysis is based on a national sample comprising widely varying housing markets. Although we apply controls for attributes associated with housing markets, specifically geographic location, urbanicity, and FMR, these controls are relatively coarse.

Table 11. Predictive Validity of Housing Quality Indices for Rents and Occupant Ratings Using Multivariate Regression Analysis

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
Log rent	– 1.518 (.157)	– 2.505 (.086)	– 1.550 (.097)	– 2.376 (.055)
House rating	.848 (.000)	.831 (.000)	.874 (.000)	.830 (.000)

Notes:
1. Top number shown is regression coefficient (for rent) or odds ratio (for house rating). Bottom number is *p*-value.
2. Log rent uses ordinary least squares, house rating uses ordinal logistic regression.
3. Log rent models also control for log household income, household head's age, race and ethnicity, air conditioning, washer and dryer, type of heat, fireplace, garage, laundry room, porch, number of rooms, pay for utilities, number of months in unit, age of building, area Fair Market Rent, number of bedrooms, number of bathrooms, geographic location (region and central city, suburban, or rural).
4. House rating models also control for log household income, household head's age, gender, race and ethnicity, geographic location (region and central city, suburban, or rural).

Assisted Housing Quality

We rely on the 2011 AHS to assess the physical quality of assisted housing in two ways. First, we examine differences between the assisted and unassisted housing stock for each of the 33 individual housing quality measures that comprise the four housing quality indices.[39] We also include 3 additional measures from the 2011 AHS Healthy Homes module in this analysis. We then compare differences in index ratings of assisted housing by program type, structure type, geography, and household characteristics. Because most of these analyses are based on the large sample sizes available in the AHS, measures of statistical significance are not very useful to gauge substantive importance. Therefore, we rely heavily on the size of the effect, measured by a statistical test known as Cohen's *d*.[40]

[38] Many other hedonic analyses are similarly based on one or only a few housing markets.
[39] We apply the same exclusion criteria to the unassisted stock used earlier to develop the housing quality indices.
[40] Cohen's *d* is the difference in means between two groups divided by the standard deviation for the pooled sample of the two groups.

Assisted and Unassisted Housing Quality

Based on the 36 individual housing quality measures in the 2011 AHS, the quality of assisted housing is comparable to the quality of unassisted housing. This conclusion is based on a comparison of prevalence rates and applies whether we limit the unassisted stock to units with rents at or below the FMR, to all unassisted rental units, or to housing units emerging from matching the assisted and unassisted units using propensity score matching (PM), a statistical procedure.[41] Table 12, which displays these prevalence rates, shows the comparisons to low-rent units in the first set of columns, followed by comparisons to the entire rental stock in the remaining columns. Even for the most highly significant measures, such as those pertaining to heating, roof leaks, rodents and mice, broken toilet, and nonworking elevator, Cohen's *d* never achieves 0.2, which defines the threshold for a small effect.

In light of the absence of large differences in individual items that constitute the housing indices, coupled with the predominance of "no problems" in both the assisted and unassisted stock, it is, therefore, not surprising that whether a housing unit is assisted or not is only weakly associated with its rating on the housing quality indices. As shown in tables 13a and 13b, the effect sizes fall below the level required for even a small effect whether the comparison is to units with comparable rents or all comparable rental units, respectively.

[41] In supplementary analysis, we used PM in two ways: (1) to match assisted housing units to unassisted housing units that are comparable based on a range of housing characteristics; and (2) to match assisted housing households to comparable low-income households not receiving housing assistance based on multiple household characteristics. Both of these PM samples produced similar results to those reported here, with not even small effect size.

Table 12. Prevalence Rates of 2011 Housing Problems by Assistance Status

		Rentals ≤ FMR				All Rentals		
		Assisted Housing	Unassisted Housing		Cohen's *d*	Assisted Housing	Unassisted Housing	
1.	Not all rooms have plugs	1.3	1.1	+		1.3	0.9	**
2.	Ever blown fuses	8.6	8.9			8.6	9.3	+
	# times blown fuses	22.1	22.3			22.1	23.3	
3.	Exposed wiring	2.8	2.3	*		2.8	2.2	**
4.	Unit does not have electricity	0.00	0.02			0.00	0.03	
5.	Unvented room heaters	0.3	1.3	***	.091	0.3	1.0	***
6.	No heating equipment	0.2	1.0	***	.094	0.2	0.7	***
7.	Use stove or oven for heat	0.1	0.2			0.1	0.1	
8.	Ever heating breakdowns	4.6	3.4	***	.060	4.6	3.1	***
	# times heating broke down	11.3	9.4	+		11.3	8.1	***
9.	Unit cold 24+ hours last winter	17.5	12.3	***	.100	17.5	11.3	***
10.	Cold: utility interruption	1.2	1.1			1.2	1.0	
11.	Cold: inadequate heating	3.0	2.2	**		3.0	1.9	***
12.	Cold: inadequate insulation	2.3	1.8	*		2.3	1.7	***
13.	Cold: other reason	2.3	1.8	*		2.3	1.5	***
14.	Roof leak last 12 months	3.2	5.3	***	.098	3.2	4.9	***
15.	Leak in wall/closet	3.4	3.0	+		3.4	2.9	*
16.	Leak in basement	1.3	1.2			1.3	1.4	
17.	Leak other source	1.4	1.3			1.4	1.3	
18.	Leaking pipes	5.4	5.2			5.4	4.9	
19.	Leaking plumbing fixture	2.4	2.4			2.4	2.4	
20.	Leak unknown source	5.0	3.7	***	.069	3.6	5.0	***
21.	Cracks in walls	7.5	7.7			7.5	6.9	
22.	Holes in floor	2.1	1.6	*		2.1	1.4	***
23.	Peeling paint	3.9	3.7			3.9	3.1	***
24.	Signs of rodents	5.3	4.0	***	.066	5.3	3.1	***
25.	Signs of rats last	1.1	1.3			1.1	1.2	
26.	Signs of mice last	12.8	10.2	***	.087	12.8	9.0	***
27.	Signs of cockroaches	5.8	6.3			5.8	4.5	***
28.	Incomplete plumbing	0.2	0.5	**		0.2	0.3	
29.	Toilet ever broke	3.3	2.5	***	.048	3.3	2.3	***
	# times toilet broke	6.7	4.9	**		6.7	5.0	**
30.	Share plumbing facility	2.3	1.7	**		2.3	1.6	***
31.	Incomplete kitchen	4.8	4.3			4.8	4.0	**
32.	Ever sewage breakdown	1.5	1.5			1.5	1.3	
	# sewage breakdowns	3.8	3.2			3.8	3.0	
33.	No working elevator	6.1	4.7	***	.066	6.1	5.0	***
34.	Any mold	8.0	8.5			8.0	7.9	
35.	Broken stairs	0.6	0.8			0.6	0.8	
36.	Broken steps	0.9	1.2			0.9	1.3	+

Notes:
1. Assisted housing N = 8,472; All rentals unassisted housing N = 40,830; Rentals ≤ FMR unassisted housing N = 24,190.
2. Weighted data.
3. Average for counts ("#") = average for entire sample, not just those reporting the problem.
4. *p*-values: + < .10, * < .05, ** < .01, *** < .001.
5. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.
6. FMR = Fair Market Rent.

Table 13a. Housing Quality Ratings of 2011 Assisted and Unassisted Housing Units, Comparable Rents

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
Assisted housing	2.49	1.94	3.51	2.24
Unassisted housing	2.22	1.78	3.14	1.95
p-value	.000	.000	.000	.000
Cohen's *d*	.069	.054	.075	.069

Notes:
1. Assisted housing N = 8,472; Unassisted housing N = 24,190.
2. Weighted data.
3. Comparable unassisted housing = rents at or below HUD's Fair Market Rent.
4. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.

Table 13b. Housing Quality Ratings of 2011 Assisted Housing and All Unassisted Housing Units

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
Assisted housing	2.49	1.94	3.51	2.24
Unassisted housing	2.03	1.62	2.88	1.77
p-value	.000	.000	.000	.000
Cohen's *d*	.119	.109	.136	.141

Notes:
1. Assisted housing N = 8,472; Unassisted housing N = 40,830.
2. Weighted data.
3. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.

Subgroup Variation in Assisted Housing Quality

Despite the lack of notable differences in the housing quality of the assisted and unassisted housing, there still may be sizable variations in quality within the assisted housing stock.

Table C1 displays selected characteristics of assisted housing by program type. Even a cursory glance demonstrates considerable variation in some basic features of assisted housing including its geographic location, racial and ethnic composition, and household size, among others. Unsurprisingly, we find differences in housing quality ratings for assisted housing that pertain to both the assisted stock's housing-related attributes and the characteristics of households residing in assisted housing.

Housing Characteristics. Differences of housing quality by program type are shown in table 14. Although housing quality index ratings are always worst for public housing and best for multifamily housing, the effect size approaches but never reaches the 0.2 threshold for a small effect. By contrast, sizable disparities in quality are evident among assisted housing for elderly, nonelderly disabled, and nonelderly families, as shown in table 15.[42] Across all four housing quality indices, elderly housing has the best housing quality and housing in which a nonelderly disabled person resides has the worst. Three of the four indices reach a medium effect size and all are at least twice the criterion for a small effect size. Despite the relatively large variations in structure type by assisted housing program (see table C1), table 16 indicates that no notable disparities in assisted housing quality are associated with the type of structure.

Table 14. Housing Quality Ratings of 2011 Assisted Housing Units by Program Type

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
All assisted housing	2.49	1.94	3.51	2.24
Public housing	2.96	2.20	4.07	2.67
Multifamily	2.19	1.82	3.14	1.96
Voucher	2.48	1.91	3.50	2.23
p-value	.006	.000	.000	.000
Cohen's *d*	.182	.121	.186	.194

Notes:
1. All assisted housing N = 8,472; Public housing N = 2,007; Multifamily N = 3,848; Voucher N = 2,617.
2. Weighted data.
3. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.

Table 15. Housing Quality Ratings of 2011 Assisted Housing Units by Household Type

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
Elderly	1.71	1.49	2.57	1.57
< 62 disabled	3.79	2.76	5.14	3.45
< 62 family	2.57	1.88	3.56	2.31
p-value	.000	.000	.000	.000
Cohen's *d*	.500	.409	.512	.514

Notes:
1. Weighted data.
2. Elderly N = 3,165; < 62 disabled N = 1,597; < 62 family N = 2,648.
3. Excludes 14 percent of assisted housing cases where head < 62, not disabled and no children.
4. Cohen's *d* compares highest/lowest categories; values < 0.2 indicate virtually no difference.

[42] The AHS does not identify housing for the elderly, families, or young disabled. To construct these categories, we assume a household head 62 years of age or older is living in elderly housing, that families with children 18 or younger and without a disabled member are living in family housing, and that nonelderly persons younger than 62, even if they are living with family members, are in housing for the disabled. This is admittedly a very blunt approach but is the best that can be done with the AHS data.

Table 16. Housing Quality Ratings of 2011 Assisted Housing Households by Structure Type

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
All assisted housing	2.49	1.94	3.51	2.24
Single-family	2.64	1.75	3.56	2.42
Multifamily ≤ 4 units	2.41	2.08	3.47	2.14
Multifamily 5+ units	2.45	1.97	3.50	2.20
p-value	.306	.021	.909	.118
Cohen's *d*	.055	.111	.017	.076

Notes:
1. Weighted data.
2. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.

Tables 17 and 18 indicate that geography matters for assisted housing quality. A central city location is associated with worse assisted housing quality compared with a suburban or rural location. Because assisted housing quality in suburbs and rural areas is extremely close, the small effect size in three of the four indices is driven entirely by the disparity between assisted housing quality in central cities versus other locations. Nearly 56 percent of assisted housing is located in central cities, and 65 percent of public housing units are located in central cities (see). Table 18 shows that assisted housing quality is worst in the Northeast, with the South, Midwest, and West following in descending order. While these three latter regions differ somewhat from each other in the housing quality index ratings, these differences are far smaller than those between the Northeast and the three other regions. The disparities are large enough to approach a medium effect size in three of the four indices. Here, too, a larger share of public housing is located in the Northeast compared with the fraction of multifamily and of voucher units in the Northeast (see table C1).

Table 17. Housing Quality Ratings of 2011 Assisted Housing Units by Metropolitan Location

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Weighted
All assisted housing	2.49	1.94	3.51	2.24
Central city	2.93	2.18	4.08	2.67
Suburb	1.93	1.65	2.82	1.69
Rural	1.95	1.64	2.76	1.73
p-value	.000	.000	.000	.000
Cohen's *d*	.237	.168	.253	.245

Notes:
1. Weighted data.
2. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.

Table 18. Housing Quality Ratings of 2011 Assisted Housing Units by Census Region

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
All assisted housing	2.49	1.94	3.51	2.24
Northeast	3.55	2.68	4.86	3.27
Midwest	2.04	1.63	2.92	1.83
South	2.29	1.73	3.27	2.00
West	1.73	1.49	2.53	1.55
p-value	.000	.000	.000	.000
Cohen's *d*	.414	.366	.451	.446

Notes:
1. Weighted data.
2. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.

Household Characteristics. The household characteristic with the greatest variation in assisted housing quality is household size. As shown in table 19, large households of six persons or more, which constitute roughly 15 percent of assisted housing households, have the worst quality ratings on all four indices, one-person households enjoy the best quality ratings, and two- to five-person households fall between these two poles. The weighted and factor analysis normative standards ratings also follow a roughly monotonic improvement from the largest households to single-person households. By contrast, the consumer rating and unweighted normative standards indices display a larger gap between two- to five-person households and the largest households. For all indices except the weighted normative standards index, the effect size exceeds the threshold for a medium effect; for the weighted index, *d* exceeds the threshold for a small effect. Public housing has a larger fraction of very large households compared with the other two programs (see table C1).

Table 19. Housing Quality Ratings of 2011 Assisted Housing Households by Household Size

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
All assisted housing	2.49	1.94	3.51	2.24
1 person	1.85	1.62	2.78	1.66
2–5 persons	2.85	2.13	3.90	2.55
6+ persons	4.12	2.67	5.49	3.87
p-value	.000	.000	.000	.000
Cohen's *d*	.651	.383	.604	.683

Notes:
1. Weighted data.
2. Cohen's d compares highest/lowest categories; values < 0.2 indicate virtually no difference.

Race and ethnicity also appear to have some association with assisted housing quality, with White households enjoying better housing quality, on average, than non-White households. These results are shown in table 20. This disparity generally reaches a small effect size and, thus, is considerably weaker than the relationship of housing quality and household size. Non-White households include Black people, Hispanic people, and people of other races and ethnicities. Initial analysis including each race and ethnicity separately demonstrated that the main disparity in housing quality is between White households and non-White households. Table 21 shows no association between household income, here measured in quartiles, and housing quality.[43] The presence of a household member who is disabled tends to be associated with worse housing quality, as shown in table 22. However, the effect approaches but does not quite achieve the standard for a small effect. This weak result most likely occurs because elderly households and families without a disabled member are combined into the "not disabled" category, as can be seen by comparing this table with table 15, which separates "elderly" and "family."

Table 20. Housing Quality Ratings of 2011 Assisted Housing Households by Race and Ethnicity of Head

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
All assisted housing	2.49	1.94	3.51	2.24
White	1.92	1.61	2.82	1.69
Non-White	2.78	2.11	3.86	2.52
p-value	.000	.000	.000	.000
Cohen's *d*	.210	.163	.209	.229

Notes:
1. Weighted data.
2. Cohen's d compares highest/lowest categories; values < 0.2 indicate virtually no difference.
3. Non-White includes all races and ethnicities other than White.

Table 21. Housing Quality Ratings of 2011 Assisted Housing Households by Household Income Quartiles

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
All assisted housing	2.49	1.94	3.51	2.24
Lowest	2.52	2.00	3.55	2.22
> 25–50%	2.51	1.98	3.60	2.37
> 50–75%	2.36	1.80	3.40	2.14
Highest	2.57	1.97	3.48	2.23
p-value	.569	.308	.742	.432
Cohen's *d*	.050	.066	.040	.062

Notes:
1. Weighted data.
2. Lowest quartile = ≤ $7,200; 25–50% = >$7,200 to $11,475; 50–75% = >$11,475 to $21,142; Highest = >$21,142 (2011$)
3. Cohen's d compares highest/lowest categories; values < 0.2 indicate virtually no difference.

[43] We use quartiles to divide the income distribution of the sample into four equally sized groups, each representing 25 percent of the sample. Thus, the lowest quartile includes the 25 percent of the sample with the lowest incomes.

Table 22. Housing Quality Ratings of 2011 Assisted Housing Households by Disabled Household Member

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
All assisted housing	2.49	1.94	3.51	2.24
No disability	2.23	1.77	3.16	1.98
With disability	2.94	2.23	4.11	2.69
p-value	.000	.000	.000	.000
Cohen's *d*	.171	.152	.192	.195

Notes:
1. Weighted data.
2. Designation of disabled household member based on response to American Housing Survey variable HDSB (household has a disabled person), which is a recoded variable that is computed based on responses to disability questions for individual household members.
3. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.

Multivariate Analysis

Assisted and Unassisted Housing Quality

In an effort to look comprehensively at the contribution of the array of attributes discussed thus far to the quality of the assisted housing stock, we estimated two sets of multiple regression models. In these models, the housing quality index ratings are the dependent variables, and housing, geographic, and household characteristics are the independent or explanatory variables (also referred to as covariates) that are plausibly related to housing quality. For simplicity of presentation, we use two of the four housing quality indices, the consumer rating index and the factor analysis normative standards index, as the dependent variables to be explained.

The first set of models is limited to the assisted housing sample, and we include assisted housing program type as one of the explanatory variables.[44] Our goal is to determine whether, once multiple attributes are accounted for, the type of assisted housing program or any of the other explanatory variables makes a meaningful contribution to housing quality. The second set of models broadens the sample to include both assisted and unassisted housing. We first test whether knowing a unit is assisted or unassisted is associated with a substantial difference in housing quality. In a second test, we distinguish assisted housing by program type to estimate whether a particular program has a substantively important relationship with housing quality even if assisted housing, as a whole, does not. We estimate the models twice—once OLS, the most common approach for estimating multivariate models, and a second time using a different estimation approach, Negative Binomial Modeling. This second technique is appropriate when the data are heavily skewed and when the distributions have considerable dispersion. Both

[44] Because program type is a categorical variable with three values—public housing, multifamily, and voucher—we use voucher as the excluded reference category.

conditions apply in the present case, casting doubt on the utility of the OLS estimates. More than half of the samples report no problems using either index, and the standard deviation is larger than the mean.[45]

Table 23 displays the results. The coefficients shown, betas for the OLS regression and odds ratios for the negative binomial models, can be interpreted as effect sizes. None of the odds ratios, the more appropriate coefficients for this analysis, achieves the standard for a small effect, despite several being statistically significant (Chen, Cohen, and Chen, 2010).[46]

Table 23. Multivariate Models Predicting Housing Quality

	Ordinary Least Squares		**Negative Binomial**	
	Beta	***p***	**Odds Ratio**	***p***
Assisted housing only				
(1) Consumer rating index				
Public housing	– .002		.952	
Multifamily (vouchers excluded)	– .030	+	.987	
(2) Factor analysis index				
Public housing	– .002		.924	
Multifamily (vouchers excluded)	– .032	*	.962	
Assisted and unassisted (≤ FMR) housing				
(1) Consumer rating index				
Assisted housing	– .032	**	.925	*
(2) Factor analysis index				
Assisted housing	– .023	***	.956	
(3) Consumer rating index				
Public housing	– .011	+	.869	*
Multifamily	– .028	***	.942	
Voucher (unassisted excluded)	– .020	**	.934	+
(4) Factor analysis index				
Public housing	– .006		.894	*
Multifamily	– .024	***	.966	
Voucher (unassisted excluded)	– .013	*	.970	

Notes:
1. Total unweighted N = 25,808.
2. Weighted data.
3. *p*-values: *** < .001, ** < .01, *< .05, + < .10.
4. Covariates are census region, metropolitan location, head's age, race, gender and marital status, number of persons in household, whether anyone in household disabled, income quartile and structure type.
5. Because negative binomial models cannot use decimals, dependent variables multiplied by 100 and rounded.
6. Beta coefficients can be interpreted like Cohen's *d*, with values < 0.2 indicating virtually no difference. Odds ratio > 1.4 if greater than 1 and < .710 if less than 1 considered the threshold for a small effect (equivalent to a Cohen's *d* of 0.2).
7. FMR = Fair Market Rent.

---

[45] Both indices have dispersion ratios of approximately 1.8 or a standard deviation that is nearly twice as large as the mean.
[46] Chen, Cohen, and Chen (2010) demonstrate that odds ratios above 1.4 or below 0.71 are the equivalent of a Cohen's *d* at 0.2, a small effect.

Assisted Housing Subgroups With the Most Housing Quality Problems

Three demographic subgroups that are particularly likely to have housing quality problems emerge from the foregoing analysis: nonelderly disabled households, non-White households, and large households. In an attempt to further elucidate the characteristics most closely associated with the housing quality of these three subgroups, we estimated a multivariate regression model for each subgroup in which the consumer rating housing quality index is the dependent variable and selected housing, geographic, and household characteristics are the independent variables. We again rely on negative binomial estimation because of the skew and dispersion in the dependent variables. Because the sample sizes of these subgroups are much smaller than those used in the previous analysis, we can rely on both the statistical significance and the effect size for interpretation.

Table 24 summarizes the results. Location in the Northeast region is statistically significant for all groups and produces a quality rating that is more than twice as bad as that for the West. For nonelderly disabled persons, being part of a large household (six or more persons) is also significant and associated with worse housing quality. Voucher receipt is associated with housing quality about 50 percent better than those living in multifamily housing, and living in a large household is associated with a housing quality rating that is about 50 percent worse than smaller households.

Table 24. Subgroup Regression Results on Consumer Rating Index

	< 62, Disabled		Non-White		Large households	
	Odds Ratio	***p***	**Odds Ratio**	***p***	**Odds Ratio**	***p***
Public housing	1.255		.929		1.449	+
Voucher	.679	**	.875	+	.808	
Central city	1.089		1.445	***	2.848	***
Northeast	2.049	***	2.323	***	2.173	**
Midwest	.882		1.350	**	.888	
South	1.184		1.397	***	1.421	
White	.900				2.053	**
Disabled			1.432	***	2.422	***
Large household	1.433	**	1.387	**		
N	1,082		3,663		261	

Notes:
1. Negative binomial regressions using weighted samples
2. Excluded categories: multifamily housing and western region. Central city, White, disabled, and large households are all dichotomous variables
3. *p*-values: *** < .001, ** < .01, * < .05, + < .10.
4. **Bold** = Odds ratio > 1.4 if greater than 1 and < .710 if less than 1 considered the threshold for at least a small effect (equivalent to a Cohen's *d* of 0.2).

For non-White households, beyond living in the Northeast, both a central city location and having a disabled household member is significantly associated with worse housing quality. In both cases, the consumer housing quality rating is about 50 percent worse than the respective excluded category (that is, non-central-city location and no disabled person in the household).

The large household model produces the largest number of covariates that are statistically significant and substantively meaningful. Beyond living in the Northeast, the largest effect is living in a central city, which is associated with a nearly three times worse housing quality rating than living in a suburban or rural area. The next largest effect is if a nonelderly disabled person is a member of the household, which is associated with more than two times worse housing quality. Living in the South (compared with living in the West) and living in public housing (compared with living in multifamily housing) each worsens the housing quality rating by about 50 percent.

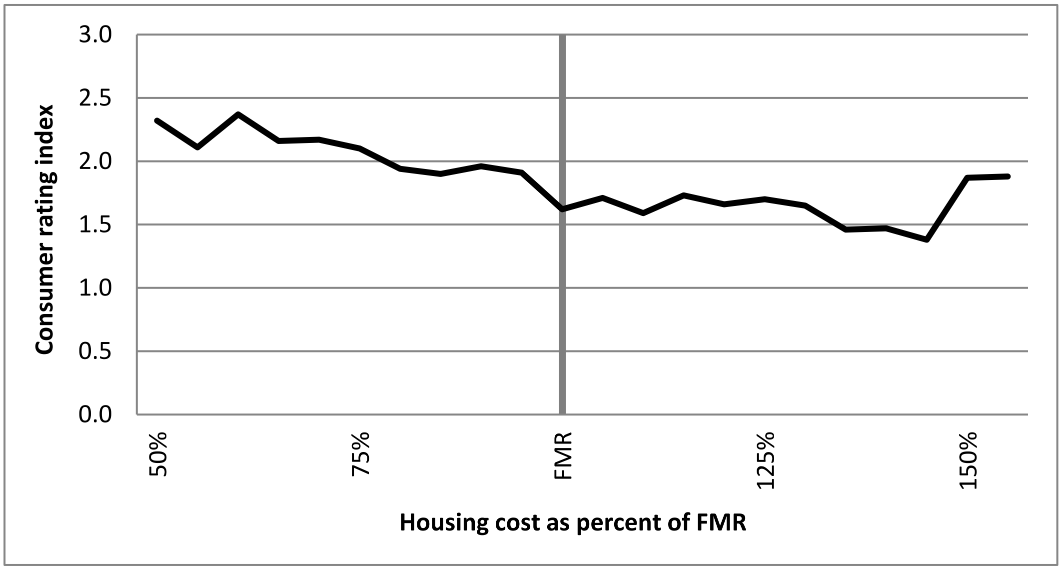
Assisted Housing Quality and Fair Market Rents

One policy question is how well the housing quality rating for assisted housing aligns with the FMR, now generally set at the 40th percentile of rents in each housing market (typically a metropolitan area). One objective in setting the FMR at a particular point in the distribution of rents is the household's ability to find physically decent rental housing at or below the FMR threshold. We explore this alignment by examining the relationship between the housing quality rating and the relative housing cost of the assisted housing sample of households. The relative housing cost is calculated by adding rent and utility costs and dividing by the FMR. We then divide the housing cost scale into equal units (for example: 40 to 45 percent of the FMR, 45 to 50 percent of the FMR, and so on) so that the FMR (defined here as 100 percent of the FMR) sits in the middle of the distribution and can be specifically designated.

Figure 2 graphically displays the results for the consumer rating housing quality index.[47] The most important result is that housing quality appears to be maximized roughly at the FMR, as indicated by the vertical line. The graph also suggests that the worst housing quality occurs at about the 24th percentile of rent (note the peak at approximately the 60th percentile; 60 x 40 = 24). In some cases, HUD approves payment standards up to 120 percent of the FMR, such as when a disabled household member requires reasonable accommodations in the voucher program. Figure 2 shows no appreciable difference in housing quality at 120 percent of the FMR. The same analysis using the factor analysis normative standards index produces a very similar pattern.

[47] We topcoded consumer rating index scores at the 99th percentile to avoid problems with outliers. We also limit the housing cost scale to range from the 5th to the 95th percentile of cases, again to avoid outlier problems at both ends of the scale.

Figure 2. Consumer Rating Index by Relative Cost



Notes:
1. Housing cost = rent + utilities (obtained from HUD's Housing Affordability Data System).
2. Housing cost relative to Fair Market Rent (FMR) = (Housing Cost / FMR) * 100.
3. The consumer rating index topcoded at the 99th percentile to avoid problems with outliers.
4. FMR scale excludes top and bottom 5 percent to avoid outliers at both ends of this distribution.

Persistence of Housing Quality Problems in Assisted Housing

Although housing quality problems are rare in assisted housing, one objective of the housing inspection and resolution regulations is to identify and address these problems in a timely manner. To shed some light on the timeliness of repairs and its opposite, the persistence of problems, we examine changes in reports of a problem for each of the 33 individual items that comprise the housing quality indices between the 2011 and 2013 AHS. Table 25 displays the full pattern of changes over the 2-year period (problem reported in both years; reported in 2011, not 2013; not reported in 2011 but reported in 2013; and not reported in either year). The table demonstrates, again, that the large majority of respondents report none of these problems in either year (0,0 column). The smallest group in terms of prevalence is persistent problems; that is 1,1—a problem that is reported in both 2011 and 2013. These data suggest that the prevalence of persistent problems, defined as housing quality problems that are reported to be present in both years, is very low.

The 1,0 category (present in 2011, not present in 2013) addresses the timeliness of repairs. These rates, while relatively modest, are nonetheless considerably larger than those for persistent problems (the 1,1 column). This suggests that problems are more likely to be addressed than to persist over a 2-year period. The 0,1 category (not present in 2011, present in 2013) provides further evidence of the low prevalence of these individual housing quality problems in assisted housing.

Table 25. Persistence of Housing Quality Problems in Assisted Housing Units, 2011 and 2013

		1,1	1,0	0,1	0,0
1.	Not all rooms have plugs	0.0	1.3	1.4	97.3
2.	Ever blown fuse	0.8	5.6	3.7	89.8
3.	Exposed wiring	0.0	2.6	2.0	95.4
5.	Unvented room heaters	0.1	0.2	0.1	99.6
6.	No heating equipment	0.2	0.0	0.0	99.8
7.	Use stove/oven for heat	0.1	0.0	0.1	99.8
8.	Ever heating breakdown	0.5	3.6	3.0	92.9
9.	Unit cold 24+ hours last winter	3.7	10.4	7.2	78.6
10.	Cold: utility interruption	0.1	1.2	1.0	97.7
11.	Cold: inadequate heating	0.6	2.4	1.2	95.7
12.	Cold: inadequate insulation	0.2	2.0	1.4	96.4
13.	Cold: other reason	0.1	2.2	1.6	96.0
14.	Roof leak last 12 months	0.5	2.8	2.3	94.4
15.	Leak in wall/closet	0.4	3.1	2.3	94.2
16.	Leak in basement	0.5	0.8	0.7	98.0
17.	Leak other source	0.1	1.3	0.7	97.9
18.	Leaking pipes	0.6	4.9	4.0	90.6
19.	Leaking plumbing fixture	0.1	2.2	3.2	94.5
20.	Leak unknown source	0.3	4.7	2.2	92.8
21.	Crack in wall	2.1	5.6	5.8	86.5
22.	Holes in floor	0.1	2.2	1.5	96.2
23.	Peeling paint	0.7	3.3	2.7	93.3
24.	Signs of rodents	1.0	4.7	2.8	91.4
25.	Signs of rats last	0.1	0.8	0.9	98.1
26.	Signs of mice last	4.8	8.7	5.3	81.2
27.	Signs of cockroaches	2.0	4.1	3.8	90.1
29.	Toilet ever broke	0.3	3.2	2.7	93.8
30.	Share plumbing facility	0.1	2.2	1.1	96.6
31.	Incomplete kitchen	0.2	4.6	3.8	91.3
32.	Ever sewage breakdown	0.1	1.4	1.4	97.2
33.	No working elevator	4.0	2.2	0.8	92.9

Notes:
1. Unweighted sample = 5,362 assisted housing cases in both 2011 and 2013 American Housing Survey (AHS)
2. Weighted data.
3. “1,1” = problem reported in 2011 and 2013 AHS.
4. “1,0” = problem reported in 2011 but not in 2013 AHS.
5. “0,1” = problem reported in 2013 but not in 2011 AHS.
6. “0,0” = problem not reported in 2011 or 2013 AHS.
7. “Unit does not have electricity” (4) and “Incomplete plumbing” (28) not shown because no problems reported in 2011 or 2013 for this sample.

Because a change in household or in the household respondent could affect housing quality reporting, we conducted additional analysis (not shown) on the effect of household reporter status on the change in the consumer rating index between 2011 and 2013. The change in the index rating is small and is unaffected by the household reporter (< .1 of a standard deviation).

Promptness of Repairs in Assisted Housing

We can also gain insights into the timeliness of repairs from a set of questions asked in the 2011 AHS about the promptness and effectiveness of repairs to rental housing units and whether owners were polite and considerate during the repair process.[48] We compared responses from assisted and comparable unassisted housing and find only weak effects. These results are summarized in table 26. The large majority of residents in assisted and unassisted housing report that both major and minor repairs are started soon enough, solved quickly, and handled politely.

Table 26. Comparison of Repair Reports on Assisted Housing and Comparable Unassisted Housing

	Assisted Housing	Unassisted Housing	*p*-Value	Cohen's *d*
Major repairs				
Start repairs soon enough	77.8	79.6	.008	.043
Solve repairs quickly	82.0	83.8	.002	.049
Handled politely	90.5	92.2	.000	.061
Minor repairs				
Start repairs soon enough	79.9	81.9	.001	.052
Solve repairs quickly	84.8	86.6	.001	.053
Handled politely	91.1	92.5	.002	.050

Notes:
1. Comparable unassisted housing based on rents at or below HUD's Fair Market Rent.
2. Cohen's *d* computed for all effects significant at .001 or less. Values < 0.2 indicate virtually no difference.

This section on promptness, and the preceding one on persistence, lend support to the biennial inspections in the voucher program and the biennial and triennial inspections for standard and high performers, respectively, in the public housing program.[49]

Discussion

The substantial literature on housing quality demonstrates both the importance and the challenges of conceptualizing and measuring housing quality. The growing interest in how housing matters—primarily in the low-rent unassisted housing stock—and the ongoing concerns about housing quality standards in the assisted stock, make this an opportune time to revisit housing quality.

In this paper, we review the relevant literature, develop alternative housing quality indices, test their validity, and apply them to both the assisted and comparable unassisted housing stock. We focus on indicators of physical integrity of housing systems and exclude measures that are more likely to reflect residents' housekeeping or behaviors.

[48] The questions, which are repeated for "major maintenance or repairs" and "minor maintenance or repairs" are: (1) Do they start quickly enough? (2) Do they solve the problem quickly once they start? (3) Are they polite and considerate of your home?

[49] The one exception is the inspection of the unit when it changes hands from one household to another.

Because there is no consensus about the features of the housing bundle that should be included in the definition of a dwelling's quality and how each should be weighted in determining overall quality, we rely on three external criteria first suggested by research on the Experimental Housing Allowance Program (EHAP; Merrill, 1980): market value, consumer rating, and normative standards. We test the market value criterion with a hedonic approach. The consumer rating criterion identifies the dwelling features most associated with a resident's assessment of the dwelling as a good place to live regardless of what the market price of these features might be. This criterion is consistent with the renewed interest by economists in happiness and subjective well-being. The normative standards criterion reflects community concerns and policy decisions such as building codes. Our main data are the 2011 and 2013 national AHS. Both surveys provide augmented national samples and assisted housing samples, and assisted housing is identified by address matches to HUD administrative data, not respondent self-report.

Despite its intuitive appeal, the market value criterion performs poorly. This likely occurs because the national sample comprises widely varying housing markets, and our controls for market attributes are relatively coarse. Another explanation is that the individual dwelling quality measures may not be the main drivers of rents, as suggested by Merrill (1980). Therefore, we drop this market value criterion from further analysis. We construct one consumer rating index and three indices using the normative standards criterion: a weighted index based on a previous analysis for HUD's Office of Policy Development and Research by Eggers and Moumen (2013b); an unweighted index; and an index with weights derived from a factor analysis. For the consumer rating index, we use the odds ratios from ordered logistic regressions as the weights. Consistent with much past AHS housing quality research, the prevalence rate of almost all problems is very low, with most dwelling units having no problems. There is notable overlap between the measures that are the strongest predictors of the consumer ratings index and the factor analysis normative standards index. These are presence of rodents, cold dwelling unit, holes in the floor, peeling paint, and cracks in the walls. Both the convergent validity and predictive validity of the four indices are strong.

Although the four indices we construct are similar to each other, the consumer ratings index and the factor analysis normative standards index emerge as consistently the strongest. Thus, the weighted index, which applied the weights developed by Eggers and Moumen (2013b), while reasonable on its face, is not quite as good as these two other approaches. Our results also call into question the applicability of hedonic models using individual measures of physical inadequacies as predictors for rental properties, as is done in Emrath and Taylor (2012). Merrill (1980) raised this same issue in her EHAP analysis.

We find that the quality of assisted housing is comparable to the quality of unassisted housing. This conclusion applies whether we limit the unassisted stock to units with rents at or below the FMR, to all unassisted rental units, or to housing units emerging from statistically matching the assisted and unassisted units. Consistent with the sound quality of most of the assisted housing stock is the very low incidence of housing problems persisting over 2 years, 2011 and 2013, and the promptness of repairs.

The type of assisted housing program does not appear to have an appreciable effect on housing quality. Although housing quality index ratings are always worst for public housing and best for multifamily housing, the effect size falls below the threshold for a small effect. The quality of assisted housing located in central cities and in the Northeast region is considerably lower than counterpart locations (that is, suburban or rural areas; South, Midwest, and West regions). Large households of six persons or more have the worst assisted housing quality while one-person households enjoy the best quality ratings. This is consistent with our finding that elderly headed households enjoy better housing quality than families. But it is households that include a disabled member that have much lower housing quality than either elderly households or families without a disabled member. Non-White households also are more likely than White households to live in lower-quality assisted housing. Thus, three household subgroups emerge as particularly likely to live in assisted housing with lower than average housing quality: nonelderly disabled households, non-White households, and large households. More detailed analysis revealed that disabled persons enjoy better housing quality using a voucher compared with living in multifamily housing. For large households, living in the South and living in public housing are associated with considerably worse housing quality.

This research provides hard evidence supporting the current 40th percentile of rents definition of the FMR. We find that the FMR is set at a level that is roughly at the inflection point for maximizing assisted housing quality as measured by our housing quality indices.

Overall, these positive findings demonstrating that the quality of the assisted housing stock suggest that the current inspection and quality control systems appear to be achieving the goal of providing physically adequate housing to assisted housing residents. They also lend support to the shift to biennial inspections in the voucher program and the biennial and triennial inspections for standard and high performers, respectively, in the public housing program. The lack of persistent problems over a 2-year period and the promptness of repairs further suggest that proposals to streamline inspections to encourage participation in the voucher program by private owners of rental properties should be seriously considered.[50] The findings reported may also be useful to HUD as it finalizes plans for a demonstration program to test a new approach to physical inspections including a single inspection protocol for public housing and voucher units.[51]

We view this research as one step along the path toward improving our understanding of housing quality. One important enhancement of this work is expanding the measures to include additional aspects of the full housing bundle, particularly neighborhood features. This could be done by linking the AHS data via confidential geocodes to census tract data and an array of administrative data at the neighborhood level. Another extension would be to compare these results with housing inspection scores from HUD's administrative data (that is, the Public

---

[50] See H.R. 3700, Housing Opportunity Through Modernization Act.
[51] See the joint Explanatory Statement accompanying the Consolidated Appropriations Act, 2016 (Pub. L. 114-113).

Housing Assessment System and Real Estate Assessment Center housing inspection ratings for public housing and multifamily housing, respectively). At the more conceptual end of the continuum, this report does not focus on what measures *should be* included in a measure of housing quality, only on how well the measures included in the AHS appear to be reliable and valid and form a coherent index. A consideration of what measures currently missing from the AHS should be included in the future is beyond the scope of the current effort but worth serious attention.

References

Aaron, Henry. 1972. Shelter and Subsidies: Who Benefits from Federal Housing Policies. Washington, DC: The Brookings Institution.

Arcara, James S. 2016 (February 16). Personal communication (e-mail). U.S. Department of Housing and Urban Development, Real Estate Assessment Center.

American Public Health Association (APHA). 1938. *Basic Principles of Healthful Housing*. Washington, DC: Committee on Hygiene of Housing.

———. 1945. "An Appraisal Method for Measuring the Quality of Housing: A Yardstick for Health Officers, Housing Officials and Planners," *American Journal of Public Health* 35 (8): 866–867.

Buron, Larry, Sandra Nolden, Kathleen Heintz, and Julie Stewart. 2000. *Assessment of the Economic and Social Characteristics of LIHTC Residents and Neighborhoods*. Cambridge, MA: Abt Associates Inc.

Carmines, Edward, and Richard Zeller. 1979. *Reliability and Validity Assessment: Quantitative Applications in the Social Sciences*. Thousand Oaks, CA: Sage Publications, Inc.

Carter, George R. III. 2011. "From Exclusion to Destitution: Race, Affordable Housing, and Homelessness," *Cityscape* 13 (1): 33–70.

Chen, Henian, Patricia Cohen, and Sophie Chen. 2010. "How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies," *Communications in Statistics-Simulation and Computation* 39 (4): 860–864.

Coulson, N. Edward, and Herman Li. 2013. "Measuring the External Benefits of Homeownership," *Journal of Urban Economics* 77: 57–67.

Dolan, Paul, Tessa Peasgood, and Mathew White. 2008. "Do We Really Know What Makes Us Happy? A Review of the Economic Literature on the Factors Associated with Subjective Well-Being," *Journal of Economic Psychology* 29 (1): 94–122.

Eggers, Frederick, and Fouad Moumen. 2013a. *American Housing Survey: Housing Adequacy and Quality as Measured by the AHS*. Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

———. 2013b. *American Housing Survey: A Measure of (Poor) Housing Quality*. Report prepared for the U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Emrath, Paul, and Heather Taylor. 2012. "Housing Value, Costs, and Measures of Physical Adequacy," *Cityscape* 14 (1): 99–125.

Friedman, Samantha, and Emily Rosenbaum. 2004. "Nativity-Status and Racial/Ethnic Differentials in Access to Quality Housing: Does Home Ownership Bring Greater Parity?" *Housing Policy Debate* 15 (4): 865–902.

Goodman, J.L. (1978). "Causes and Indicators of Housing Quality," *Social Indicators Research* 5 (1): 195–210.

Gray, Robert W., Barbara A. Haley, and Brent D. Mast. 2008. Tell Us About Your Home: Three Years of Surveying Housing Quality and Satisfaction in the Section 8 Housing Choice Voucher Program. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research, Program Monitoring and Research Division.

Joint Center for Housing Studies of Harvard University (JCHS). 2010. *The State of the Nation's Housing 2010*. Cambridge, MA: Harvard University, Joint Center for Housing Studies.

Kain, John F., and John Quigley. 1970. "Measuring the Value of Housing Quality," *Journal of the American Statistical Association* 65 (330): 532–548.

Khadduri, Jill, and Kathryn P. Nelson. 2007. "Targeting Housing Assistance," *Journal of Policy Analysis and Management* 11 (1): 24–41.

Kriström, Bengt. 2008. "Applying Hedonics in the Housing Market: An Illustration." In *Hedonic Methods in Housing Markets: Pricing Environmental Amenities and Segregation*, edited by Baranzini, Andrea, José Ramierz, Caroline Schaerer, and Philippe Thalmann. New York: Springer: 247–258.

Merrill, Sally. 1980. Hedonic Indices as a Measure of Housing Quality. Cambridge, MA: Abt Associates Inc.

Millennial Housing Commission. 2002. *Report of the Bipartisan Millennial Housing Commission*. Washington, DC: The Millennial Housing Commission.

Newman, Sandra. 2008. "Does Housing Matter? A Critical Summary of Research and Issues Still to be Resolved," *Journal of Policy Analysis and Management* 27 (4): 895–925.

Newman, Sandra, and Philip Garboden. 2013. "Psychometrics of Housing Quality Measurement in the American Housing Survey," *Cityscape* 15 (1): 293–306.

Newman, Sandra, and Ann Schnare. 1993. "Last in Line: Housing Assistance for Households With Children," *Housing Policy Debate* 4 (3): 417–455.

———. 1988. Subsidizing Shelter: The Relationship Between Welfare and Housing Assistance. Washington, DC: Urban Institute.

O'Regan, Katherine, and Keren Horn. 2012. "What Can We Learn About the Low-Income Housing Tax Credit Program by Looking at the Tenants?" New York: New York University, Furman Center for Real Estate and Housing Policy.

Preacher, Kristopher, Guangjian Zhang, Cheongtag Kim, and Gerhard Mels. 2013. "Choosing an Optimal Number of Factors in Exploratory Factor Analysis: A Model Selection Perspective," *Multivariate Behavioral Research* 48 (1): 28–56.

Ross, Lauren M., Anne B. Shlay, and Mario G. Picon. 2012. "You Can't Always Get What You Want: The Role of Public Housing and Vouchers in Achieving Residential Satisfaction," *Cityscape* 14 (1): 35–54.

Shroder, Mark. 2001. "Moving to Opportunity: An Experiment in Social and Geographic Mobility," *Cityscape* 5 (2): 57–67.

———. 2002. "Does Housing Assistance Perversely Affect Self-Sufficiency? Review Essay," *Journal of Housing Economics* 11 (4): 381–417.

Thibodeau, Thomas. 1995. "House Price Indices from the 1984–1992 MSA American Housing Surveys," *Journal of Housing Research* 6 (3): 439–481.

U.S. Bureau of the Census (Census Bureau). 1967. *Measuring the Quality of Housing: An Appraisal of Census Statistics and Methods*. Washington, DC: Census Bureau.

———. 2015. *Identifying and Weighting the National and Metropolitan Samples for the 2011 and 2013 American Housing Survey*. Washington, DC: Census Bureau. https://www.census.gov/programs-surveys/ahs/tech-documentation/help-guides/identifying_samples.html.

U.S. Department of Housing and Urban Development (HUD). 2011a. *Codebook for the American Housing Survey Public Use File: 1997 and Later*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

———. 2011b. *Worst Case Housing Needs 2009: Report to Congress*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

———. 2014. *Strategic Plan 2014–2018*. Washington, DC: U.S. Department of Housing and Urban Development.

U.S. Department of Housing and Urban Development, Office of Inspector General (HUD OIG). 2008. *HUD Lacked Adequate Controls Over the Physical Condition of Section 8 Voucher Program Housing Stock*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Inspector General.

U.S. Government Accountability Office (GAO). 1997. *Tax Credits: Opportunities to Improve Oversight of the Low-Income Housing Program*. Washington, DC: U.S. General Accounting Office.

———. 2000. HUD Housing Portfolios: HUD Has Strengthened Physical Inspections but Needs to Resolve Concerns About Their Reliability. Washington, DC: U.S. General Accounting Office.

U.S. Senate. 2012. *Senate Committee Report to Accompany the T-HUD Appropriations Act for FY 2013*. S. Rept. 112–157 (April 19). 92 p. https://www.congress.gov/112/crpt/srpt157/CRPT-112srpt157.pdf.

Vandenbroucke, David. 2016 (February 17). Personal communication (e-mail). Senior Economist, U.S. Department of Housing and Urban Development.

Appendix A: Tables

Table A1. Selection of Analysis Sample: Units Excluded by Type or Category

	2011 AHS	2013 AHS
Total units in sample	186,448	84,355
URE[1], vacant or noninterviews	51,530	24,258
Occupied unit interviews	134,918	60,097
Occupied unit interviews:		
Mobile home	4,736	2,282
Other type of housing unit	850	160
House/apartment/flat	129,332	57,655
House/apartment/flat:		
Owner-occupied	78,291	34,004
Occupied without rent	1,739	809
Rental unit	49,302	22,842
Rental units:		
Unassisted	40,830	16,602
Assisted	8,472	6,240

Notes:
1. URE = usual residence elsewhere (for example, vacation home).
2. AHS = American Housing Survey.

Table A2. Polychoric Correlation Matrix, Housing Quality Measures

	Plugs	Fuses	Wiring	No Elec	Unvented	No Heat	Oven	# Heat	Cold 24	Util	Cap	Insul	Other
Plugs	1.000												
Fuses	0.061	1.000											
Wiring	0.417	0.068	1.000										
No elec	0.516	– 0.176	0.603	1.000									
Unvented	0.118	0.030	0.115	– 0.766	1.000								
No heat	0.054	0.044	0.096	– 0.743		1.000							
Oven	0.095	– 0.093	0.266		– 0.797	– 0.781	1.000						
# Heat	0.184	0.102	0.196	– 0.828	0.047			1.000					
Cold 24+	0.180	0.105	0.167	0.262	0.117		0.012	0.836	1.000				
Cold util	– 0.069	0.050	– 0.037	– 0.760	– 0.381		– 0.823	0.086	0.775	1.000			
Cold cap	0.248	0.103	0.184	– 0.793	0.223		0.239	0.131	0.798	0.214	1.000		
Cold insul	0.147	0.092	0.161	0.133	0.205		0.032	0.266	0.785	0.214	0.643	1.000	
Cold other	0.128	0.062	0.181	– 0.783	0.096		– 0.213	0.169	0.792	0.172	0.228	0.306	1.000
Leak roof	0.142	0.083	0.118	– 0.849	0.040	– 0.012	0.172	0.273	0.262	0.112	0.279	0.227	0.254
Leak wall	0.178	0.081	0.122		– 0.262	0.057	– 0.850	0.282	0.261	0.075	0.255	0.267	0.183
Leak base.	– 0.039	0.067	0.051	0.268	– 0.150	– 0.383	.	0.133	0.175	0.132	0.100	0.189	0.014
Leak other	0.054	0.055	0.028	– 0.762	– 0.001	– 0.085	0.135	0.223	0.166	0.077	0.172	0.135	0.138
Leak pipe	0.107	0.078	0.142	– 0.837	0.085	– 0.067	0.173	0.293	0.261	0.100	0.294	0.295	0.146
Leak fix.	0.180	0.078	0.087	– 0.781	– 0.024	0.161	– 0.255	0.244	0.226	0.010	0.167	0.295	0.165
Leak unk	0.093	0.065	0.080	– 0.833	– 0.089	– 0.051	– 0.004	0.258	0.216	0.048	0.232	0.187	0.212
Crack	0.285	0.120	0.182	0.107	0.209	0.040	– 0.156	0.381	0.383	0.159	0.350	0.364	0.271
Hole	0.266	0.093	0.274	0.269	0.263	0.200	0.046	0.304	0.345	– 0.106	0.374	0.340	0.227
Paint	0.227	0.104	0.209	– 0.791	0.140	0.040	0.011	0.352	0.377	0.109	0.373	0.368	0.325
Rodents	0.183	0.090	0.206	0.122	0.191	0.170	0.232	0.357	0.367	0.128	0.385	0.265	0.368
Rats	0.071	0.054	0.229	– 0.760	0.267	0.202	0.263	0.229	0.306	0.073	0.300	0.231	0.238
Mice	0.114	0.090	0.110	0.298	0.182	0.027	0.135	0.329	0.305	0.123	0.326	0.267	0.216
Roaches	0.121	0.070	0.180		0.173	0.335	0.192	0.309	0.288	0.001	0.302	0.270	0.252
Inc. plumb	0.065	– 0.012	0.242		0.232	0.314	0.460	0.044	0.082	– 0.037	0.075	– 0.855	0.218
# Toilet	0.166	0.094	0.196	– 0.792	0.035	0.175	– 0.846	0.246	0.254	0.001	0.278	0.265	0.192
Share bath	0.193	– 0.011	0.175	– 0.800	0.006	0.032	– 0.045	– 0.042	– 0.020	– 0.140	0.034	0.114	– 0.074
Inc. kitchen	0.116	0.016	0.094	– 0.114	– 0.054	0.074	0.150	0.005	0.023	– 0.042	0.140	0.060	0.052
# Sewer	0.080	0.093	0.136	– 0.790	0.241	0.135		0.317	0.252	0.037	0.193	0.264	0.137
Mold	0.149	0.051	0.195	0.244	– 0.030	0.000	– 0.132	0.305	0.414	0.165	0.208	0.235	0.203
Stair	0.285	0.066	0.201	0.088			– 0.109	0.129	0.231	– 0.299	0.090	0.096	0.197

Table A2. Polychoric Correlation Matrix, Housing Quality Measures (continued)

	Roof	Wall	Base	Lk: Oth	Pipe	Fixture	Lk: Unk	Crack	Hole	Paint	Rodents	Rats	Mice
Leak roof	1.000												
Leak wall	0.351	1.000											
Leak base.	0.289	0.335	1.000										
Leak other	0.204	0.310	0.273	1.000									
Leak pipe	0.211	0.235	0.108	0.251	1.000								
Leak fix.	0.152	0.263	0.143	0.147	0.220	1.000							
Leak unk	0.099	0.192	0.095	0.246	– 0.021	– 0.061	1.000						
Crack	0.379	0.317	0.238	0.256	0.346	0.306	0.191	1.000					
Hole	0.316	0.236	0.239	0.185	0.336	0.287	0.155	0.677	1.000				
Paint	0.427	0.356	0.247	0.239	0.343	0.231	0.270	0.590	0.527	1.000			
Rodents	0.313	0.263	0.228	0.207	0.263	0.187	0.179	0.452	0.437	0.425	1.000		
Rats	0.218	0.210	0.012	0.233	0.218	0.253	0.080	0.369	0.382	0.303	0.681	1.000	
Mice	0.271	0.209	0.318	0.200	0.226	0.176	0.117	0.356	0.359	0.344	0.851	0.296	1.000
Roaches	0.255	0.226	– 0.097	0.199	0.243	0.240	0.254	0.355	0.391	0.379	0.528	0.406	0.340
Inc. plumb	0.021	0.058	– 0.309	– 0.189	– 0.113	– 0.194	0.034	0.121	0.220	0.091	0.282	0.197	0.170
# Toilet	0.254	0.229	0.116	0.181	0.233	0.418	0.100	0.345	0.283	0.326	0.329	0.247	0.204
Share bath	0.010	– 0.013	– 0.050	0.139	0.019	0.015	– 0.175	0.054	0.002	0.098	0.057	0.065	0.025
Inc. kitchen	0.039	0.102	– 0.128	– 0.027	0.043	– 0.024	– 0.018	0.076	0.106	0.044	0.089	0.089	– 0.022
# Sewer	0.235	0.215	0.218	0.153	0.249	0.435	0.216	0.276	0.250	0.280	0.247	0.192	0.190
Mold	0.263	0.345	0.299	0.229	0.235	0.218	0.214	0.365	0.357	0.412	0.328	0.239	0.256
Stair	0.266	0.127	0.307	– 0.164	0.096	– 0.064	0.092	0.217	0.109	0.094	0.173	0.107	0.172

	Roaches	Plumb	# Toilet	Bath	Kitchen	# Toilet	Mold	Stairs
Roaches	1.000							
Inc. plumb	0.204	1.000						
# Toilet	0.315	– 0.131	1.000					
Share bath	– 0.005	0.221	0.034	1.000				
Inc. kitchen	0.114	0.437	0.112	0.002	1.000			
# Sewer	0.327	0.086	0.662	0.002	0.063	1.000		
Mold	0.322	0.001	0.225	– 0.154	0.058	0.200	1.000	
Stair	0.041	– 0.064	0.100	– 0.172	0.114	0.109	0.266	1.000

Table A3. Distribution of Housing Quality Indices

	Consumer Rating Index	Normative Standards Indices		
		Weighted	Unweighted	Factor Analysis
N	40,654	40,654	40,654	40,654
Mean	2.03	1.62	2.88	1.77
Median	0	0	0	0
Standard Deviation	3.80	2.88	4.58	3.30
Skewness	3.630	3.239	2.645	3.136
Kurtosis	20.669	17.317	10.077	13.876
Percentiles				
10	0	0	0	0
20	0	0	0	0
30	0	0	0	0
40	0	0	0	0
50	0	0	0	0
60	1.330	1.124	2.941	.604
70	1.887	1.685	2.941	1.683
80	3.360	2.809	5.882	2.896
90	6.025	5.618	8.823	5.636
95	9.424	7.303	11.765	8.491
99	18.201	13.483	20.588	15.595

Notes:
1. Weighted data.
2. Skewness > 1 indicates non-normally distributed distribution.
3. Kurtosis > 3 indicates more cases than expected in the tail of the distribution.

Appendix B: Healthy Homes Module

The 2011 American Housing Survey includes a topical module, Healthy Homes, designed to collect information on hazards and safety features in a home. Most of the questions in this module are not applicable to this study because they pertain either to a person's health (for example, does anyone in the household have asthma) or to a household's behavior (for example, are hazardous materials stored out of children's reach), not housing quality *per se*. We limit this discussion to the three questions that most likely pertain to housing unit quality: presence of mold; broken steps; and missing railings on stairs.

Our analysis of these items strongly suggests that had we been able to include these three measures in the analysis, results would not have been qualitatively different, as summarized in tables A3 and B1. In particular:

1. The polychoric correlation between presence of mold and several other items included in the 33 housing quality measures used in this paper's analysis is ≥.30, reasonably large. These items are: unit was cold for 24 hours or more; cracks in the walls; holes in the floor; presence of rodents; and seeing roaches daily.
2. Including the three Healthy Homes items along with the 33 individual quality measures in the factor analysis still yields a single "housing quality" factor.
3. The correlations between the presence of mold and each of the four housing quality indices are 0.23 with the weighted index, 0.28 with the unweighted index, 0.28 to the factor analysis index, and 0.28 with house rating.

The correlations are much smaller between each of the four indices and the steps and stairs items, primarily because very few units report either of these problems. However, rental units that report any of the three Healthy Homes problems have much higher scores on all four indices than units that do not report these problems. The index scores are roughly twice as high for units with steps or stairs problems and more than three times higher for units with mold.

Table B1. Relationship between Indices and Healthy Homes Measures

Mold

	Consumer Rating Index	Normative Standards Indices		
		Weighted	**Unweighted**	**Factor Analysis**
No mold	1.800	1.476	2.593	1.565
With mold	6.502	4.490	8.406	5.717
p-value	(.000)	(.000)	(.000)	(.000)

Broken Steps

	Consumer Rating Index	Normative Standards Indices		
		Weighted	**Unweighted**	**Factor Analysis**
No mold	2.023	1.617	2.869	1.761
With mold	5.036	3.513	6.409	4.450
p-value	(.000)	(.000)	(.000)	(.000)

Broken Stair Rail

	Consumer Rating Index	Normative Standards Indices		
		Weighted	**Unweighted**	**Factor Analysis**
No mold	1.995	1.602	2.836	1.736
With mold	4.153	3.003	5.359	3.675
p-value	(.000)	(.000)	(.000)	(.000)

Notes:
1. Weighted data.
2. Mean value for each group.

Appendix C: Characteristics of Assisted Housing by Program Type

Table C1. Percent Distribution of Characteristics of Assisted Housing for Each Program Type

	Public Housing	Multifamily	Voucher	Total
Region				
Northeast	34.0	28.0	27.7	29.0
Midwest	15.8	27.7	19.0	20.8
South	35.6	30.1	27.3	29.7
West	14.5	14.2	26.0	20.5
Metropolitan location				
Central city	65.0	56.7	51.1	55.3
Suburban	16.6	27.0	34.0	28.7
Rural	18.2	16.4	14.9	15.9
Structure type				
Single-family	23.7	8.8	30.2	23.0
Multifamily ≤ 4 units	27.9	13.6	22.8	21.2
Multifamily 5+ units	48.4	77.5	47.0	55.8
Household type				
Elderly	21.8	45.7	23.5	29.4
< 62, disabled	20.7	15.2	21.5	19.6
< 62, family	41.9	27.8	39.0	36.4
< 62, no children	15.6	11.4	16.0	14.6
Race and ethnicity				
White	28.9	37.0	33.8	33.7
Non-White	71.1	63.0	66.2	66.3
Household size				
1 person	38.0	54.3	36.6	41.8
2–5 persons	55.8	42.9	58.2	53.5
6+ persons	6.2	2.8	5.3	4.7
Disabled household member				
Not disabled	67.0	58.9	64.8	63.6
Disabled	33.0	41.1	35.2	36.4