Data Shop

Data Shop, a department of Cityscape, *presents short articles or notes on the uses of data in housing and urban research. Through this department, the Office of Policy Development and Research introduces readers to new and overlooked data sources and to improved techniques in using well-known data. The emphasis is on sources and methods that analysts can use in their own work. Researchers often run into knotty data problems involving data interpretation or manipulation that must be solved before a project can proceed, but they seldom get to focus in detail on the solutions to such problems. If you have an idea for an applied, data-centric note of no more than 3,000 words, please send a one-paragraph abstract to david.a.vandenbroucke@hud.gov for consideration.*

Psychometrics of Housing Quality Measurement in the American Housing Survey

Sandra J. Newman
Philip M.E. Garboden
Johns Hopkins University

Abstract

This article assesses the test-retest reliability, internal consistency, and convergent and predictive validity of the American Housing Survey inadequacy index. We find that the index does not appear to tap a single underlying construct of housing quality and does not differentiate among the worst quality units. We conclude that it may be time to reconceptualize the elusive construct of housing quality.

Introduction

As the most comprehensive source of data on the U.S. housing stock, the American Housing Survey (AHS) is relied on by policymakers, practitioners, and researchers seeking answers to questions about the conditions, costs, and myriad other attributes of the nation's housing. For those developing their own surveys, the AHS is also a source of housing questions. Some of the most prominent social science studies of the past two decades, including the Moving to Opportunity

for Fair Housing Demonstration (Shroder, 2001), Welfare, Children, & Families, A Three-City Study (Winston et al., 1999), and the Fragile Families and Child Well-Being Study (Reichman et al., 2001), include questions that are strikingly similar if not identical to the AHS items on housing characteristics and conditions.

Of particular interest to many users is the AHS composite measure of housing inadequacy available on the public use database. This measure combines 15 individual questionnaire 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. Both the AHS and data users refer to this composite as AHS's "housing quality" measure.[1] Numerous published articles include the AHS measure in their analyses (for example, Carter, 2011; Friedman and Rosenbaum, 2004; Khadduri, 2007; Ross, Shlay, and Picon, 2012), the measure plays a prominent role in the U.S. Department of Housing and Urban Development's *Worst Case Needs* reports (for example, HUD, 2011b), and it is also included in the frequently cited Joint Center for Housing Studies' *State of the Nation's Housing* reports (for example, JCHS, 2010) and by the Millenial Housing Commission (2002).

Despite widespread reliance on the AHS inadequacy index by a broad audience of users, little is known about its reliability, internal consistency, and validity. These attributes are typically referred to as *psychometric* features, because these tests were originally developed to assess indicators within the purview of psychologists, such as cognitive achievement, attitudes, and personality (Nunnaly, 1978).

Our goal in this article is to shed light on each of these psychometric properties of the AHS inadequacy index. This information will enable users to assess, for example, (1) if the inadequacy index differentiates among dwellings of different housing quality, (2) if respondents can reliably answer the questions used to create the 15-item index, and (3) if we can be reasonably confident that the index is a valid representation of housing quality. The next section presents the composite index and its distribution. The subsequent sections present results on reliability and validity. We summarize and discuss the implications of this review in the final section.

The AHS Inadequacy Index

The AHS inadequacy index is shown in exhibit 1. A dwelling unit is deemed severely inadequate in one of four ways: (1) the existence of a single inadequacy (for example, electricity is not used); (2) the combination of two inadequacies (for example, the unit has fewer than two full bathrooms and does not have a bathtub or shower); (3) the combination of three inadequacies (exposed wiring, lack of working electrical plugs in all rooms, and fuses blown more than twice in the past 3 months); or (4) the combination of five inadequacies (including leaks, floor holes, cracks, peeling paint, and rats). Units are deemed moderately inadequate if they do not meet the criteria for severe inadequacy but have three or four (instead of five) of the problems listed under (4) or have one of three

[1] 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" (HUD, 2011a: 212). The composite measure is also listed in the AHS documentation under the category "Unit Quality."

Exhibit 1

AHS Housing Inadequacy Index

A unit is considered severely inadequate if it meets one of the following conditions:

1. Unit has less than 2 full bathrooms and the unit has at least one of the following (incomplete plumbing).
   a. Unit does not have hot and cold running water.
   b. Unit does not have a bathtub or shower.
   c. Unit does not have a flush toilet.
   d. Unit shares plumbing facilities.
2. Unit was cold for 24 hours or more and there have been more than 2 breakdowns of the heating equipment that lasted longer than 6 hours.
3. Electricity is not used.
4. Unit has exposed wiring and not every room has working electrical plugs, and the fuses have blown more than twice.
5. If the unit meets five or six of the following:
   a. Unit has had outside water leaks in the last 12 months.
   b. Unit has had inside water leaks in the last 12 months.
   c. Unit has holes in the floor.
   d. Unit has open cracks wider than a dime.
   e. Unit has an area of peeling paint larger than 8 x 11.
   f. Rats have been seen recently in the unit.

A unit is considered moderately inadequate if it is not severely inadequate and meets one of the following conditions:

1. Three or four of the conditions listed in item (5) above.
2. There have been more than 2 breakdowns of the toilet that lasted longer than 6 hours.
3. The main heating equipment is unvented room heaters burning kerosene, gas, or oil.
4. The unit is lacking complete kitchen facilities.

AHS = American Housing Survey.

Source: HUD (2011a)

additional problems (for example, unvented heating equipment). The complexity of the index suggests it is based on considerable statistical analysis, but no documentation of which we are aware reveals the nature of this analysis or, alternatively, the basis for selecting the 15 measures in the composite and the decision rules for identifying severe and moderate inadequacy.

Exhibit 2 shows the distribution of the components of the index and its three aggregate categories for the full sample, for rental units, and for owned units, using data from the 2007 national AHS.[2] Less than 4 percent of the housing stock is rated moderately inadequate and less than 2 percent is rated severely inadequate. The inadequacy rate is less for owned units (roughly 3 percent) and more for rentals (nearly 10 percent). Of the 15 measures, only 3 characterize more than 3 percent of all dwelling units: exterior water leak, interior water leak, and cracks in walls. By identifying the small share of dwelling units that have multiple physical inadequacies, the index characterizes what it considers the lowest quality units. On the other hand, it does not characterize units without these multiple inadequacies other than to deem them adequate. Therefore, it does not produce a distribution of units along a continuum ranging from best to worst quality.

[2] We use 2007 data because our predictive validity analyses rely on 2009 data for outcomes. These data are the latest available at the time of this writing.

Exhibit 2

Prevalence of 2007 AHS Inadequacy Index Components and Individual Items

	Full Sample (%)	Renters (%)	Homeowners (%)
Adequate	94.96	90.10	96.97
Moderately inadequate	3.48	6.84	2.08
Severely inadequate	1.56	3.06	0.94
Incomplete plumbing	1.08	1.88	0.73
Incomplete kitchen	1.39	4.03	0.35
Exterior water leak	10.74	9.62	11.12
Interior water leak	8.16	11.50	6.83
Cracks in walls	4.72	7.22	3.66
Holes in floor	0.91	1.63	0.60
Peeling paint or plaster	1.95	3.23	1.40
Rats in unit	0.72	1.08	0.55
Unvented room heaters	1.07	1.14	1.00
Frequent toilet breakdowns	0.20	0.41	0.11
Frequent heating breakdowns	0.45	1.04	0.20
Not all rooms have outlets	1.20	1.70	1.00
Frequent blown fuses	2.05	2.60	1.83
Exposed wiring	0.95	1.38	0.79
No electricity	0.00	0.00	0.00
N	35,128	9,721	24,781
Weighted N	99,090,591	27,151,173	70,129,019

AHS = American Housing Survey.

Notes: Values shown denote the existence of inadequacy (some items have been reverse coded). Weighted data. Listwise deletion of missing data.

Source: 2007 AHS

Reliability

A reliable index should score identical units in identical ways across different time points and survey modalities (Carmines and Zeller, 1979). For this condition to be met, the individual items need to be sufficiently clear to produce consistent responses, and the overall index should measure a single characteristic of the unit—in our case, housing quality.

Test-Retest Reliability

The most straightforward way to test the reliability of survey questions is to ask the same respondents the same questions again, either later in the survey or shortly thereafter, and to calculate the correlation between the two responses. In the case of housing conditions, asked about at a second point in time, the reasonable expectation is that they are unlikely to change during the brief hiatus between the initial survey and the followup (for example, the house has a basement in the initial survey but not in the second).

Although the Census Bureau routinely follows up with a subsample of AHS households, this followup is part of the quality assessment of field operations, not the survey instrument. The purpose is to determine whether particular questions were asked (for example, a question about income),

not to re-ask these questions (Cole, 2011). From the inception of the AHS in 1973 through 1981, however, the Census Bureau conducted reinterviews with roughly 20,000 households (Chakrabarty, 1996). The reinterview data files are not publicly available, but an analysis of the test-retest data is presented in a 1996 Census Bureau report (Chakrabarty, 1996). Exhibit 3 reproduces the report's results for 11 of the 15 items in the AHS inadequacy index. Starting from the first results column, the exhibit shows the percentage of responses that changed between the original survey and the reinterview survey; the fraction of "yes" responses to the original survey that changed to "no" in the retest; the fraction of "no" responses to the original survey that changed to "yes" in the retest; and, when applicable, the fraction of responses that were "don't know" in the initial survey that changed to a "yes" or "no" in the reinterview.[3]

Although a relatively small proportion of responses changed between the test and retest, the low prevalence of inadequate housing conditions in the initial AHS interview, as described in the previous section, means that even modest test-retest differences are meaningful for some analyses. The overall pattern suggests a greater tendency to report that a condition exists in the original AHS interview but that it does not exist in the reinterview than vice versa. In the 1974 AHS, for example,

Exhibit 3

Test-Retest Reliability: AHS Inadequacy Index Items, 1973–1981

Binary Response Variables

Item	All (%)	Yes (%)	No (%)	Don't Know (%)	Survey Year
Heating breakdown	6	54	4	NA	1977
Heating breakdown	5	40	2	NA	1976
Interior open cracks/holes	5	49	2	NA	1977
Interior open cracks/holes	5	51	3	NA	1976
Holes in floors	2	35	1	NA	1977
Holes in floors	2	58	1	NA	1976
Seen mice or rats	9	40	4	NA	1976
Basement leak	15	27	10	38	1976
Electric plug in every room	3	2	49	NA	1976
All wiring concealed	3	2	75	NA	1976
Blown fuses	10	51	5	100	1976
Roof leaked in last 3 months	5	29	2	42	1974
Roof leaked in last 3 months	5	28	2	51	1973

Nominal Response Variables (complete kitchen and plumbing)

	All (%)	Exclusive Use (%)	Shared (%)	No (%)	Survey Year
Complete kitchen	1	0.3	88	14	1978
Complete kitchen	1	0.2	NA	26	1977
Complete kitchen	1	0.3	89	11	1975
Complete plumbing	1	0.2	33	19	1977
Complete plumbing	1	1.0	46	23	1974

AHS = American Housing Survey. NA = insufficient responses in original survey.

Source: Adapted from Chakrabarty (1996)

[3] For complete plumbing and kitchen, the exhibit shows further detail on shared versus exclusive use.

5 percent of responses about roof leaks in the past 3 months were inconsistent between the initial interview and the followup, but a greater percentage of initial reports of a leaking roof changed to no leak (29 percent) than the opposite (2 percent). The test-retest results suggest that one-fourth or more of dwelling units classified as inadequate in the AHS may not be. Two caveats about these estimates are in order. First, although Census Bureau interviewers attempted to reinterview the same respondent as the initial interview, they did not always succeed, although the share of the same versus different respondents is not known (Cole, 2011). In addition, the 4-week time lag between the initial interview and the followup could be sufficient for some conditions to be resolved (and other conditions to emerge). Although less precise than would be ideal, the test-retest results are a cautionary note for analysts, particularly those focusing on inadequate housing.

Internal Consistency

Another test of the reliability of the AHS inadequacy index as a measure of housing quality is its internal consistency—the degree to which the individual items in the index are intercorrelated and, therefore, more likely to all be indicators of the same latent characteristic. The most frequent statistical test of internal consistency is Cronbach's alpha, which is the ratio of the sum of each item's variance to the variance of the entire scale (Bland and Altman, 1997; Cronbach, 1951):

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_T^2}\right), \quad (1)$$

where k is the total number of measures, s^2_i is the variance of each individual measure, and s^2_T is the variance of the overall index (the sum of the individual measures). Like a bivariate Pearson correlation coefficient, Cronbach's alpha ranges between 0 and 1. Higher alphas represent better internal consistency, and an alpha of 0.70 or higher is generally considered acceptable (Spector, 1991).

The Cronbach's alpha for the AHS inadequacy index is 0.37 for the total housing stock, 0.44 for rental units, and 0.30 for owned units, falling well short of the 0.70 threshold. Even restricting the sample to a homogenous subgroup of dwelling units (central city, attached single-family—typically a rowhouse—two-bedroom rentals occupied by households with incomes less than the median) produces an alpha of 0.58, still less than the acceptable level. These results suggest that all the items in the index are not tapping the single construct of housing quality.

Latent Trait Analysis

Another test of internal consistency shifts the emphasis from the degree to which the items in the index "hang together," which is the focus of Chronbach's alpha, to the properties of each item in the index and the contribution it makes to the underlying construct. Latent Trait Analysis (LTA) assumes that each item is a measure of the underlying variable, in this case, housing quality, and that lower quality houses are more likely to have each inadequacy included in the index. LTA estimates two parameters for each item in the index: its difficulty and its discrimination. *Difficulty* in this context means how deteriorated or inadequate the dwelling needs to be before any particular item in the index is present.[4] More difficult items (that is, those denoting serious inadequacy

[4] LTA was originally developed to evaluate standardized educational tests, which explains why the word "difficulty" is used.

problems) are likely to be found only in the worst quality units. A measure's *discrimination* is its ability to distinguish among types of units. For example, the index might include two items of equal difficulty, such as a collapsed wall and a bat infestation, and both inadequacies might characterize less than 1 percent of the housing stock. The collapsed wall may be more discriminating, however, because it does a better job distinguishing units in terms of their housing quality. By contrast, because bat infestations most often arise in particular geographic locations, special environmental conditions, and even in high-quality dwelling units, they do not constitute a strong predictor of housing quality. If the AHS index reflects the full range of housing quality from best to worst, LTA difficulty values should range from negative to positive. If the index ranks highly on its discrimination ability, LTA discrimination values should be large.

Using a logit/normal response function, LTA models the probability (P) of having a particular inadequacy item (i) by the following:

$$P_i(\theta_n) = \frac{e^{\lambda_i(\theta_n - \delta_i)}}{1 + e^{\lambda_i(\theta_n - \delta_i)}}, \qquad (2)$$

where λ_i is the discrimination of an item and δ_i its difficulty. To estimate these parameters for each measure, this article uses a Birnbaum 2-parameter model in STATA's GLLAMM package, using maximum likelihood estimation (Zheng and Rabe-Hesketh, 2007).

Exhibit 4 presents results for the full AHS sample, for rental units, and for owned units, and exhibit 5 displays the characteristic curves for each item in the index for the full sample. In the graph, an item's difficulty is indicated by its position along the latent variable axis and its discrimination is measured by its slope at the inflection point. As shown in exhibit 4, the measures follow mostly the same pattern in each sample.[5] The difficulty of all items exceeds 2.00, indicating that they are limited to housing units with a level of disrepair that is more than two standard deviations greater than the mean. The least difficult items are signs of minor disrepair such as peeling paint, cracks in walls, and leaks, and the most difficult items are the presence of unvented room heaters, incomplete kitchens, and incomplete plumbing. As their steep slopes indicate, the most discriminating measures are cracks in walls, holes in floors, and peeling paint, followed by frequent toilet and heating breakdowns. By and large, the remaining 10 items are not discriminating. This is particularly the case for incomplete plumbing, incomplete kitchen, and unvented room heaters.

Consistent with the prevalence of each inadequacy shown in exhibit 1, the LTA results suggest that the AHS inadequacy items differentiate a small fraction of units with multiple physical inadequacies. LTA provides the additional insight that the items do not differentiate among the most inadequate units, because the high-difficulty items are also nondiscriminating. The results further confirm that none of the individual items distinguishes among units categorized as "adequate," which constitute most of the housing stock. In LTA parlance, the items in the inadequacy index are too difficult to discriminate among most units.

[5] Some distinctions from the overall pattern include the greater likelihood that measures of severe inadequacy (for example, incomplete plumbing) characterize rented units and that measures of moderate inadequacy (for example, peeling paint) characterize owned units.

Exhibit 4

Latent Trait Analysis: "Difficulty" and "Discrimination" of AHS Inadequacy Index Items

	Full Sample		Rental Units		Owned Units	
	Difficulty	Discrimination	Difficulty	Discrimination	Difficulty	Discrimination
Incomplete plumbing	7.71	0.62	11.28	0.36	7.44	0.70
Incomplete kitchen	10.41	0.42	24.76	0.13	9.49	0.62
Exterior leak	2.95	0.82	2.66	1.03	2.95	0.80
Interior leak	3.31	0.84	2.50	1.00	4.28	0.67
Cracks in walls	2.49	1.77	2.13	1.86	2.76	1.64
Holes in floor	3.16	2.34	2.98	2.12	3.20	2.69
Peeling paint or plaster	2.81	2.22	1.83	3.00	3.05	2.12
Rats in unit	5.30	1.06	4.66	1.13	6.59	0.86
Unvented room heaters	10.34	0.45	12.32	0.37	10.98	0.43
Frequent toilet breakdowns	5.30	1.39	5.23	1.23	5.54	1.46
Frequent heating breakdowns	4.06	1.15	4.61	1.00	4.02	1.17
Not all rooms have outlets	5.35	0.91	5.13	0.87	5.43	0.93
Frequent blown fuses	5.17	0.82	4.53	0.91	5.87	0.73
Exposed wiring	6.75	0.74	7.27	0.63	6.68	0.77

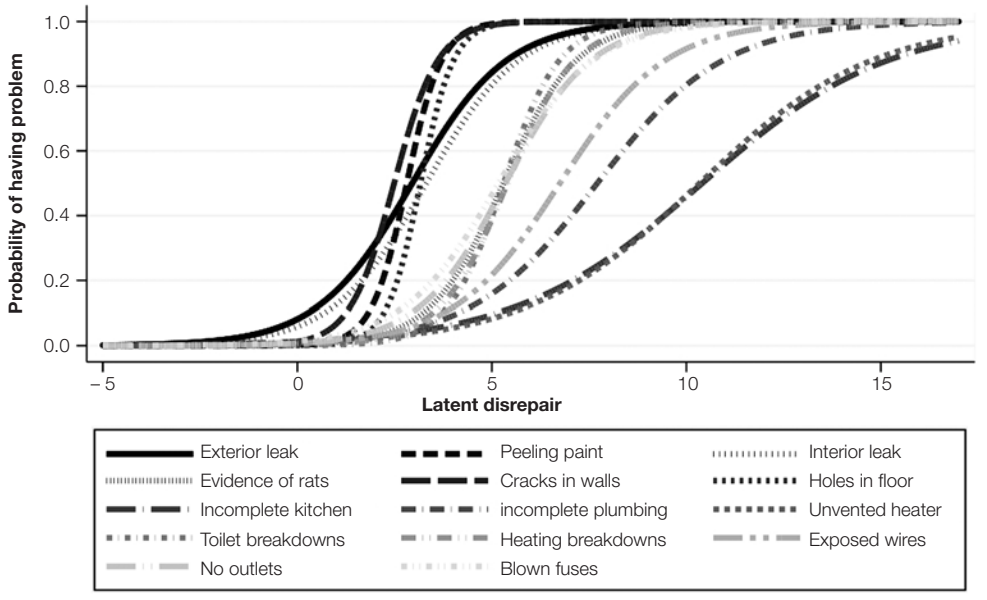
AHS = American Housing Survey.

Notes: All values are statistically significant. The "no electricity" item is excluded because of its incredibly low prevalence. Listwise deletion of missing data.

Source: 2007 AHS

Exhibit 5

Latent Trait Analysis: Item Characteristic Curves



Validity

A valid index is one that accurately measures the latent construct, housing quality, and not some other feature of the unit. To this end, it should correlate with measures typically associated with housing quality and predict outcomes known to be affected by unit quality (Carmines and Zeller, 1979).

Convergent Validity

If the AHS inadequacy index is a valid measure of the underlying construct of interest, housing quality, then it should correlate highly with other measures that also purport to tap this underlying construct. Indicators that should yield strong correlations include housing cost (for example, Emrath and Taylor, 2012; Kain and Quigley, 1970; Thibodeau, 1995), age of structure (for example, Malpezzi, Ozanne, and Thibodeau, 1987; Rubin, 1993), the tenant's satisfaction with the housing unit, and residential mobility (Lu, 1999; Newman and Duncan, 1979; Speare, 1974).[6] We measure housing cost as the log of house value reported by owners and the log of gross rent reported by renters. "Age of structure" is the estimated age reported by the respondent. "Satisfaction with the housing unit" is the respondent's rating on a 10-point scale (10 is "best" and 1 is "worst"). "Residential mobility" reflects whether a different household occupied the unit in the 2009 AHS than in the 2007 AHS.

Exhibit 6 lists the Pearson *r* correlations between each of the three points on the inadequacy index and each of the five measures theoretically associated with physical inadequacies in the dwelling. The correlations have the expected signs but are small, indicating little convergent validity. The greatest correlations are between adequacy and house value (0.10), structure age (-0.10), and housing satisfaction (0.15).

Predictive Validity

Another test of the extent to which the AHS inadequacy index is tapping the housing quality construct is to determine whether the index predicts outcomes to which it is theoretically associated. Four of the five measures used to estimate convergent validity meet this criterion: house value, rent, satisfaction, and residential mobility. To predict value and rent, we use a hedonic framework. Because hedonic models are highly sensitive to specification, we replicate, to the extent possible, models developed by two well-regarded housing economists (Coulson and Li, 2011; Thibodeau, 1995). We model housing satisfaction using an ordered logit and model the likelihood that the household moved between 2007 and 2009 using a logistic function.[7]

Exhibit 7 presents the results. Although the Coulson and Li (2011) and Thibodeau (1995) specifications are somewhat different, our replication using 2007 data produces nearly identical coefficients for the inadequacy index. For owners, the coefficient on "moderately inadequate" is negative and

[6] The predominant reasons AHS respondents give for moving focus on attributes of the housing unit (for example, Holupka and Newman, 2011).

[7] Models using contract rent and gross rent produce similar estimates. We report estimates from gross rent models in the exhibit.

Exhibit 6

Correlations Between AHS Inadequacy Index Components and Items With Housing Cost, Age, Satisfaction, and Moves

	Log Rent (renters)	Log Value (owners)	Age (all)	Satisfaction (all)	Move (all)
Adequate	0.05*	0.10*	– 0.10*	0.15*	– 0.07*
Moderately inadequate	– 0.05*	– 0.09*	0.09*	– 0.12*	0.05*
Severely inadequate	– 0.02**	– 0.03*	0.05*	– 0.09*	0.04*
Incomplete plumbing	– 0.01	– 0.03*	0.04*	– 0.05*	0.02*
Incomplete kitchen	– 0.04*	– 0.01	0.03*	– 0.05*	0.06*
Exterior water leak	0.02	– 0.03*	0.13*	– 0.09*	– 0.01*
Interior leak	0.00	– 0.01*	0.03*	– 0.12*	0.04*
Cracks in walls	– 0.03*	– 0.07*	0.10*	– 0.18*	0.05*
Holes in floor	– 0.01	– 0.08*	0.04*	– 0.11*	0.02*
Peeling paint or plaster	– 0.02	– 0.05*	0.09*	– 0.13*	0.03*
Rats in unit	– 0.01	– 0.01**	0.04	– 0.07*	0.01
Unvented room heaters	– 0.05*	– 0.10*	0.06*	– 0.04*	0.01
Frequent toilet breakdowns	– 0.02**	– 0.02*	0.01**	– 0.05*	0.02*
Frequent heating breakdowns	– 0.01	0.00	0.04*	– 0.07*	0.02*
Not all rooms have outlets	– 0.01	– 0.01	0.03*	– 0.05*	0.02*
Frequent blown fuses	0.03*	0.01	0.03*	– 0.09	0.03*
Exposed wiring	– 0.02**	0.01	0.01*	– 0.01	0.02*
No electricity	NA	0.00	0.00	0.00	0.00

AHS = American Housing Survey. NA = insufficient responses in original survey.

**p < 0.05. **p < 0.10.*

Notes: Weighted data. Listwise deletion of missing data.

Sources: 2007 and 2009 AHS

statistically significant. It suggests that moving from an adequate to a moderately inadequate unit is associated with a 20- to 22-percent decline in housing value.[8] None of the other coefficients reach statistical significance at the 0.10 level, and all are very small. The lack of significance of the severe inadequacy variable confirms the lack of discrimination at the extreme end of poor housing quality. Emrath and Taylor's (2012) recent *Cityscape* article also tested several different hedonic specifications that included the AHS inadequacy index. They report statistically insignificant coefficients on both the moderately and severely inadequate variables in predictions of both house value and rent.

The results for housing satisfaction are decidedly different. The coefficients on both moderate and severe inadequacy are large, of similar size, and statistically significant in all models. For example, moving from an adequate to a moderately inadequate unit reduces by 19 percentage points the likelihood that respondents ranked their housing satisfaction a 9 or 10. By contrast, moderate or severe inadequacy in the dwelling does not appear to be closely associated with making a residential move. Living in a severely inadequate unit, for example, increases the likelihood of a household moving by 2 percentage points. The only result that reaches significance ($p < 0.10$) occurs for homeowners, but even here, the size of the effect is small, increasing the probability of moving by about 3 percentage points.

[8] These marginal effects are calculated by exponentiating the coefficients.

Exhibit 7

Multivariate Analysis of AHS Inadequacy Index

		AHS (moderately inadequate)	AHS (severely inadequate)
Coulson and Li (2011) hedonic[a]	Renters (log rent)	– 0.01	0.03
		(0.02)	(0.04)
	Owners (log value)	– 0.25*	0.01
		(0.05)	(0.07)
Thibodeau (1995) hedonic[b]	Renters (log rent)	0.00	0.03
		(0.02)	(0.03)
	Owners (log value)	– 0.22*	0.01
		(0.04)	(0.06)
Housing satisfaction[c]	All	– 0.79*	– 0.87*
		(0.06)	(0.08)
	Renters	– 0.68*	– 0.82*
		(0.08)	(0.11)
	Owners	– 0.88*	– 0.87*
		(0.09)	(0.13)
Moving in 2 years[c]	All	0.12	0.16
		(0.07)	(0.10)
	Renters	0.10	0.14
		(0.09)	(0.13)
	Owners	0.10	0.32**
		(0.14)	(0.19)

AHS = American Housing Survey.

**p < 0.05. **p < 0.10.*

[a] Includes household income, race, school adequacy, shopping adequacy, the presence of public transportation, the number of bathrooms, the age of the housing unit, the presence of a garage, the presence of central air conditioning, the type of heat, region, whether the unit was in a central city, the lot square footage, and the interior square footage.

[b] Includes slightly different specifications for renters and owners and controls for number of bathrooms, number of bedrooms, number of other rooms, structure type, age of property, garage, basement, heating system, air conditioning system, assessment of neighborhood, abandoned properties nearby, litter in neighborhood, neighborhood crime, neighborhood noise, head of household Black, head of household Hispanic, people per room, lot size, head of household moved before 1949, and utility inclusion in rent.

[c] Include log household income, head of household race, head of household education, head of household gender, head of household age, urban/suburban, and region.

Notes: Standard errors in parentheses. Hedonic models are estimated using ordinary least squares. Housing satisfaction, measured on a scale of 1 to 10 (10 being best), is estimated using an ordered logit linking function. Moving, measured by a change in household between the 2007 and 2009 surveys, is estimated using a logit linking function. Coefficients (log odds) are reported. Listwise deletion of missing data.

Sources: 2007 and 2009 AHS

Conclusions

The AHS inadequacy index identifies a small share of dwelling units with multiple inadequacies. Whether a dwelling unit is consistently characterized by the same inadequacies when the initial survey is administered again within 1 month is indeterminate with the available data, but evidence at least suggests that data users should exercise caution, particularly when focusing on units categorized as inadequate by the index. Although we could not find any documentation about how the

AHS inadequacy index was formed, the AHS (and users) view it as a measure of housing quality. Tests of internal consistency suggest that the 15 items in the index do not tap the same underlying construct of housing quality, and that no differentiation is made among the items that characterize the worst quality units. Tests of convergent and predictive validity also raise questions about whether the index taps the housing quality construct.

These results are reminiscent of the many past failed efforts, primarily in the 1970s, to develop a single measure of housing quality (for example, Goedert and Goodman, 1976; Goodman, 1978; Kain and Quigley, 1970). With increased interdisciplinary interest in housing and greater analytic sophistication, now may be the time to revisit the conceptualization and measurement of the elusive concept of housing quality.

Acknowledgments

This research was conducted as part of the MacArthur Foundation's Research Network on Housing and Families With Children.

Author

Sandra J. Newman is a professor of policy studies in the Institute for Policy Studies at Johns Hopkins University.

Philip M.E. Garboden is a doctoral candidate in the Department of Sociology at Johns Hopkins University.

References

Bland, J. Martin, and Douglas G. Altman. 1997. "Statistics Notes: Cronbach's Alpha," *BMJ* 314 (7080): 572.

Carmines, Edward G., and Richard A. Zeller. 1979. *Reliability and Validity Assessment*. Thousand Oaks, CA: SAGE Publications.

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

Chakrabarty, Rameswar P. 1996. *American Housing Survey: A Quality Profile*. Washington, DC: U.S. Department of Housing and Urban Development.

Cole, Tamara. 2011 (February 8). Personal communication (e-mail). Chief, American Housing Survey Branch, U.S. Census Bureau.

Coulson, Edward, and Herman Li. 2011. "Measuring the External Benefits of Homeownership." Paper presented at the 46th Annual AREUEA Conference, Denver, January 7–9.

Cronbach, Lee J. 1951. "Cronbach's Alpha Internal Consistency: Coefficient Alpha and the Internal Structure of Tests," *Psychometrika* 16: 297–334.

Emrath, Paul, and Heather Taylor. 2012. "Housing Value, Costs, and Measures of Physical Adequacy," *Cityscape* 14 (1): 99–126.

Friedman, Samantha, and Emily Rosenbaum. 2004. "Nativity Status and Racial/Ethnic Differences in Access to Quality Housing: Does Homeownership Bring Greater Parity?" *Housing Policy Debate* 14 (4): 865–901.

Goedert, Jeanne, and John Goodman. 1976. Indicators of Housing Quality: An Exploration of the Annual Housing Survey. Working Paper #249-2. Washington, DC: Urban Institute.

Goodman, John. 1978. "Causes and Indicators of Housing Quality," *Social Indicators Research* 5: 195–210.

Holupka, Scott, and Sandra Newman. 2011. "The Housing and Neighborhood Conditions of America's Children: Patterns and Trends Over Four Decades," *Housing Policy Debate* 21 (2): 215–245.

Joint Center for Housing Studies of Harvard University (JCHS). 2010. *The State of the Nation's Housing, 2010.* Cambridge, MA: Joint Center for Housing Studies. Also available at http://www.jchs.harvard.edu/research/publications/state-nations-housing-2010.

Kain, John F., and John M. 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): 21–41.

Lu, Max. 1999. "Determinants of Residential Satisfaction: Ordered Logit vs. Regression Models," *Growth and Change* 30 (2): 264–287.

Malpezzi, Stephen, Larry Ozanne, and Thomas G. Thibodeau. 1987. "Microeconomic Estimates of Housing Depreciation," *Land Economics* 63 (4): 372–385.

Newman, Sandra J., and Greg J. Duncan. 1979. "Residential Problems, Dissatisfaction, and Mobility," *Journal of the American Planning Association* 45 (2): 154–166.

Nunnaly, Jum C. 1978. *Psychometric Theory*. New York: McGraw-Hill.

Reichman, Nancy, Julien Teitler, Irwin Garfinkel, and Sara McLanahan. 2001. "Fragile Families: Sample and Design," *Children and Youth Services Review* 23 (4–5): 303–326.

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.

Rubin, Geoffrey M. 1993. "Is Housing Age a Commodity? Hedonic Price Estimates of Unit Age," *Journal of Housing Research* 4 (1): 165–184.

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

Speare, Alden. 1974. "Residential Satisfaction As an Intervening Variable in Residential Mobility," *Demography* 11 (2): 173–188.

Spector, Paul. 1991. *Summated Rating Scale Construction: An Introduction*. Thousand Oaks, CA: SAGE Publications.

Thibodeau, Thomas G. 1995. "Housing Price Indices From the 1984–1992 MSA American Housing Surveys," *Journal of Housing Research* 6 (3): 439–479.

U.S. Department of Housing and Urban Development (HUD). 2011a. "Codebook for the American Housing Survey, Public Use File: 1997 and Later." Available at http://www.huduser.org/portal/datasets/ahs/ahs_codebook.html (accessed February 15, 2013).

———. 2011b. *Worst Case Housing Needs 2009: A Report to Congress*. Washington, DC: U.S. Department of Housing and Urban Development. Also available at http://www.huduser.org/portal/publications/affhsg/wc_HsgNeeds09.html.

U.S. Millennial Housing Commission. 2002. *Meeting Our Nation's Housing Challenges*. Washington, DC: Government Printing Office.

Winston, Pamela W., Ronald J. Angel, Linda M. Burton, P. Lindsay Chase-Landsdale, Andrew J. Cherlin, Robert A. Moffitt, and William J. Wilson. 1999. *Welfare, Children, and Families: Overview and Design*. Baltimore: Johns Hopkins University.

Zheng, Xiaohui, and Sophia Rabe-Hesketh. 2007. "Estimating Parameters of Dichotomous and Ordinal Item Response Models With GLLAMM," *The Stata Journal* 7 (3): 313–333.