Housing Value, Costs, and Measures of **Physical Adequacy**

Paul Emrath Heather Taylor National Association of Home Builders

Abstract

Part of the U.S. Department of Housing and Urban Development's (HUD's) mission is to create quality affordable homes for all. To accomplish this mission, HUD must define quality and must develop a method for detecting physically inadequate housing units. In the past, researchers have relied on summary indicators of inadequacy provided on the American Housing Survey (AHS) public use data file. These measures are designed by HUD and are used by HUD for HUD purposes. This article reexamines these standard indicators in a hedonic regression framework, using AHS data to develop models that estimate house values and rent. The hedonic models are then used to define a new indicator of physical inadequacy that has a statistically significant negative effect on house values and rents, in contrast to the traditional indicators that are not statistically significant and often have the wrong sign. The new indicator indentifies a substantially larger number of housing units in the United States as being physically inadequate, especially single-family units, suggesting that the need for housing assistance is more widespread than is generally recognized. Housing units identified as inadequate under this new criterion are concentrated in the older stock and are disproportionately occupied by households with children. The new criterion also identifies a substantial number of nonseasonal, vacant single-family housing units as being physically inadequate, implying that the inventory of existing homes on the market may be effectively overstated. The statistical models used to derive these results also illustrate the practical utility of a large number of variables in different sections of the AHS. Many neighborhood characteristics are shown to have a significant effect on home values, for example, which is information of potentially great value to homeowners and local governments.

Introduction

An important aspect of the U.S. Department of Housing and Urban Development's (HUD's) mission is its goal of creating quality affordable homes for all. To accomplish this goal effectively, it is necessary to consider affordability and quality in tandem. Affordability may be achieved by neglecting routine maintenance and allowing properties to deteriorate, or by failing to replace or renovate very old housing units to bring them more in line with modern building codes. Few people would consider these to be desirable outcomes.

Historically, to judge the quality of the U.S. housing stock, researchers have typically relied on standard criteria for classifying housing units as physically adequate or inadequate, using characteristics of the housing units collected in the American Housing Survey (AHS). The AHS, which is funded by HUD and conducted by the U.S. Census Bureau in odd-numbered years, collects information on a large number of housing characteristics. A number of these characteristics are combined to produce a variable that classifies housing units as adequate, moderately inadequate, or severely inadequate. This variable is included on the AHS public use file. These measures are designed by HUD and are used by HUD for HUD purposes. It is much more common to accept this traditional classification scheme uncritically than to consider an alternative specification, despite the richness of the AHS data set that would permit extensive experimentation with alternatives.

Relying on the standard AHS adequacy classification scheme produces a view of the U.S. housing market in which problems of high housing costs relative to income are considerably more widespread than problems of poor-quality housing. For example, HUD's latest report to Congress on "Worst Case Housing Needs" (HUD, 2011) states, "Of the two types of priority problems that qualify as worst case needs, severe rent burden appears far more frequently than severely inadequate housing." An implication is that the problem of physically inadequate housing in the United States and its interaction with affordability, although of interest from a theoretic perspective, can often be comfortably neglected in favor of a concentration on affordability problems.

This article investigates the issue of housing quality in a hedonic regression framework that estimates the effect of various housing characteristics and different definitions of inadequacy on the value of owner-occupied and cost (rent) of renter-occupied housing units. The underlying hypothesis is that inadequate housing units should have lower values (if they are owner occupied) or rents (if they are renter occupied), controlling for other characteristics.

The Baseline Regressions section presents baseline hedonic regression results for owner-occupied single-family housing units and renter-occupied multifamily units. The Physical Inadequacy section proposes new definitions of inadequacy and shows how the new definitions compare with the standard ones that have been used historically when added to the pertinent baseline regression. The next section, Characteristics of Inadequate Housing Units, addresses the number of housing units flagged as inadequate and the characteristics of the inadequate units and their occupants under the proposed new classification scheme. The final section summarizes the results and offers conclusions.

Baseline Regressions

The primary statistical technique used in this study is called hedonic regression. In practice, hedonic regression refers to a technique that estimates the price of a good based on its characteristics. Hedonic price estimation dates back at least to Waugh (1928), although Griliches (1961) and Rosen (1974) are usually credited for establishing it as a widely used technique. One general use of hedonic regression is to estimate a constant quality price index for a heterogeneous set of commodities. It is often used this way in housing markets, for example to derive constant quality price indices for new construction (U.S. Census Bureau, 2005). It is also used to estimate the marginal effect of a particular characteristic on the price of a commodity.

Hedonic Regression Specification

One central problem in estimating a hedonic regression is model specification. This problem can be particularly challenging in the case of housing because a housing unit is a complex commodity with an effectively infinite number of characteristics that can affect its market price. For this reason, any hedonic housing model will fail to capture some relevant characteristics and, therefore, must be misspecified to a certain extent. Moreover, some characteristics are likely to be collinear, tending to obscure the marginal effect of a particular attribute. The collinearity problem, in general, is less of an issue when the primary objective of the hedonic model is to drive an index that predicts the value of a housing unit.

Even if prediction is the primary intent, however, most hedonic regression estimates eventually end up being used to assess the marginal effects of particular features. Indeed, a substantial body of hedonic housing market research targets the marginal effect of a particular attribute as the principal objective. Targeting the marginal effect of a particular attribute typically occurs when hedonic models are used for policy analysis, as when estimating the value of environmental improvements. A recent example of using hedonic methods to analyze the marginal effect of environmental improvements is Carruthers, Clark, and Renner (2010).

Marginal effects from hedonic models also have applications for private business decisions. McDonald and McMillen (2007), for example, have argued that hedonic models are useful tools for real estate appraisers because the models provide information on the value of particular attributes. For cases in which marginal effects of particular characteristics are important, as is typically the case, omitted variable bias and collinearity are particularly important and should be taken into account when developing an empirical strategy.

The empirical strategy employed by the National Association of Home Builders (NAHB) in applying the hedonic regression method to AHS data has been developed over a period of several years. The first example, for single-family house prices, appeared in Emrath (1993) with a subsequent extension to multifamily housing in Emrath (1996). Results from the regressions have been made available to the public on the Internet in a form that can be used interactively to estimate house prices, as described in Emrath (2004). The key elements of the NAHB strategy include a substantial effort to clean the data; experimentation with a large number of independent variables, including neighborhood characteristics and interaction terms; and care taken to avoid deleting variables that materially alter estimated coefficients on any of the variables retained.

This strategy is designed to exploit several strengths of the publicly available AHS data set. Although the AHS allocates item nonresponses and truncates many variables to preserve respondent confidentiality, these procedures are well documented so that researchers are able to adjust for them. The longitudinal nature of the AHS—revisiting the same housing units year after year—allows researchers to investigate the way key variables change over time, providing an additional useful aid for screening suspect values. Perhaps most significantly, from the perspective of a researcher trying to estimate a single regression equation, the AHS includes a very large number of housing attributes as well as a large number of observations. The flattened version of the 2009 public use AHS file, for example, contains 2,776 variables and 73,222 observations. Although the public use file includes vacant housing units and noninterviews, the number of observations with usable data is nevertheless large enough to accommodate virtually any conceivable single-equation specification while reducing effects of collinearity.

The complete list of independent variables used in all the hedonic regression models presented in this article are displayed in exhibit 1.

Exhibit 1

List of Independent Variables Used in the Regression Models (1 of 2)

Location ind	icator
CENT_NE	Central city in the Northeast census region
CENT_MW	Central city in the Midwest census region
CENT_SO	Central city in the South census region
CENT_CA	Central city in large California metropolitan areas
CENT_WE	Central city in the remainder of the West census region
BURB_NE	Suburban area in the Northeast census region
BURB_MW	Suburban area in the Midwest census region
BURB_SO	Suburban area in the South census region
BURB_CA	Suburban area in large California metropolitan areas
BURB_WE	Suburban area in the remainder of the West census region
NMET_NE	Nonmetropolitan area in the Northeast census region
NMET_MW	Nonmetropolitan area in the Midwest census region
NMET_WE	Nonmetropolitan area in the West census region
The base loca	ation omitted from this list is nonmetropolitan area in the South census region.

Housing unit	characteristic
SIZE1900	Size of the unit in 1,000 sq ft if built before 1950; 0 otherwise
SIZE1950	Size of the unit in 1,000 sq ft if built from 1950 through 1979; 0 otherwise
SIZE1980	Size of the unit in 1,000 sq ft if built from 1980 through 1994; 0 otherwise
SIZE1995	Size of the unit in 1,000 sq ft if built from 1995 through 2006; 0 otherwise
SIZE2007	Size of the unit in 1,000 sq ft if built after 2006; 0 otherwise
SIZEBF85	Size of the unit in 1,000 sq ft if built before 1985; 0 otherwise
SIZE1985	Size of the unit in 1,000 sq ft if built from 1985 through 1989; 0 otherwise
SIZEPOST	Size of the unit in 1,000 sq ft if built after 1990; 0 otherwise
LOTSIZE	Size of the lot in acres for single-family units
BATHS	Number of full bathrooms in the unit
HALFB	Number of half bathrooms in the unit
BEDRMS	Number of bedrooms in the unit
DINING	Number of dining rooms in the unit

Number of family rooms in the unit

FAMRM

Exhibit 1

List of Independent Variables Used in the Regression Models (2 of 2)

Housing unit characteristic (continued)

OTHROOMS Number of rooms other than baths and bedrooms for multifamily units; other than baths,

bedrooms, dining rooms, and family rooms for single-family units

BA_MW Indicator for full or partial basement in the Midwest census region BA_SO Indicator for full or partial basement in the South census region

BA_CA Indicator for full or partial basement in large California metropolitan areas

GARAGEX Indicator for garage or carport included with unit

FIREPLAC Indicator for presence of a fireplace

AC_MWSO Indicator for central air conditioning in the Midwest and South census regions

UDISH Indicator for presence of a working dishwasher in the unit
UDRY Indicator for presence of a working clothes dryer in the unit

Neighborhood and multifamily building characteristic

GREEN Indicator for open spaces within one-half of a block
COMRECR Indicator for presence of community recreational facilities
COMGATE Indicator for a unit lying within a gated community

XWATER Indicator for property not on the waterfront, but with a body of water within one-half of a block

XWFPROP Indicator for waterfront property

WFP_NE Indicator for waterfront property in the Northeast census region
WFP_MW Indicator for waterfront property in the Midwest census region
WFP_SO Indicator for waterfront property in the South census region

WFP_CA Indicator for waterfront property in large California metropolitan areas
WFP_WE Indicator for waterfront property in the remainder of the West census region

XTRAN Indicator for neighborhood with satisfactory public transportation

XTR_MET Indicator for satisfactory neighborhood public transportation within metropolitan areas

XSHOP Indicator for neighborhood with satisfactory shopping

SHP_MET Indicator for satisfactory neighborhood shopping within metropolitan areas

UELEV Indicator for unit within a multifamily building, on a floor with access to an elevator

UACCESSB Indicator for entry system that restricts access to the building

FLOOR_3 Indicator for unit in a building with 3 floors
FLOOR_49 Indicator for unit in a building with 4 to 9 floors
FLOOR_10 Indicator for unit in a building with 10 or more floors

FIFTY_1 Indicator for unit in a building with 1 floor and 50 or more housing units

BARCL Indicator for buildings within one-half of a block of the unit with metal bars on their windows

ABAN Indicator for abandoned or vandalized buildings within one-half of a block of the unit

BADROADS Indicator for roads in need of repair within one-half of a block of the unit COMCRIME Indicator for serious crime in the neighborhood within the past year COMODOR Indicator for unit in a neighborhood with smoke, gas, or bad smells XSTNOISE Indicator for unit in a neighborhood with heavy street noise or traffic

JNK_MET Indicator for trash/litter/junk within one-half of a block of the unit in metropolitan areas
JNK_NM Indicator for trash/litter/junk within one-half of a block of the unit outside of metropolitan areas
COM1 Indicator for businesses or institutions within one-half of a block of the unit

COM2 Indicator for factories/industrial structures within one-half of a block of the unit

MB_MET Indicator for mobile homes within one-half of a block of the unit in metropolitan areas

MB_NM Indicator for mobile homes within one-half of a block of the unit outside of metropolitan areas

Indicator of physical inadequacy

NEW_INAD Unit inadequate according to proposed new criteria described in the text

AHS_MOD Unit either moderately or severely inadequate according to the traditional AHS criteria

AHS_SEV Unit severely inadequate according to the traditional AHS criteria

A potential disadvantage of the AHS, which is sometimes cited in an application like the models in exhibit 1, is its national nature and somewhat limited level of geographic detail. The information collected on neighborhood characteristics compensates for the limited geographic detail, to some extent. For some purposes, information on characteristics of the neighborhood, rather than very precise information on the location of the neighborhood, may be a strength rather than a weakness. Although location indicators identifying housing units to within a fine level of geography, such as a specific census tract or block group, could likely be used to improve the fit of a hedonic regression, this approach would not provide information on what aspect of a neighborhood is responsible for improving the fit. Yet, information about the effect of being located on the waterfront or in a neighborhood with abandoned buildings on a property's value would typically be of interest to homeowners and local policymakers. The NAHB strategy of trying a relatively large number of neighborhood characteristics in the hedonic specification seeks to take as much advantage as possible of this section of the AHS.

The location indicators used in this article intersect the four principal census regions (Northeast, Midwest, South, and West) with a metropolitan status measure that identifies if an area is in a central city, suburb, or nonmetropolitan area. This combination of region and metropolitan status, in general, is the most precise level of geographic detail available in the AHS. The AHS does identify certain metropolitan areas (based on the definitions and boundaries of metropolitan areas that prevailed in 1980), but there are, in general, too few observations in a metropolitan area to treat each of them separately in the model. For this article, a number of the large California metropolitan areas were carved out as a separate "region" distinct from the rest of the West, however, a procedure first employed in Emrath (1995). The metropolitan areas included in the California region are Bakersfield, Fresno, Los Angeles-Long Beach, Modesto, Oakland, Orange County, Riverside-San Bernardino, Sacramento, San Diego, San Francisco, San Jose, Santa Barbara-Santa Maria, Santa Rosa, Stockton-Lodi, Vallejo-Fairfield-Napa, and Ventura. It is well known that home values tend to be high in many of these areas, and being in a central city or suburban location within one of the California metropolitan areas has the strongest effect on house values of any of the location indicators. In theory, intersecting 5 regions (Northwest, Midwest, South, West, and California) with 3 metropolitan status categories (central city, suburb, and nonmetropolitan) produces a total of 15 geographically unique areas. Because the California region does not include nonmetropolitan counties, however, in practice the intersection results in only 14 unique areas. To avoid perfect multicollinearity in models with a constant term, the geographically unique area of the nonmetropolitan South is omitted.

An additional specification issue in hedonic regression is the choice of functional form. Although the term hedonic is often used simply to indicate a regression that estimates price of a good as a function of its characteristics, economic theory underlies the technique. Because production costs and utility of a particular good typically both change with changes in its characteristics, a question arises as to whether the hedonic regression captures a demand or supply relationship. The answer

¹ The term *suburb* is used as a convenient way to describe territory inside metropolitan areas but outside the central city. The AHS variable METRO3 is used to define the three metropolitan status categories as follows: central city if METRO3 = 1, suburb if METRO3 = 2 or 3, and nonmetropolitan if METRO3 = 4 or 5. This classification scheme is consistent with the one used in the AHS printed reports.

is essentially both, as equilibrium in a hedonic model represents a point of tangency between the offer curves of a buyer and seller. The relationship between price and a particular characteristic traces an envelope of these tangency points, a point emphasized by Rosen (1974). Although the first derivative of a desirable characteristic is, in general, positive, as increasing a desirable characteristic in a commodity usually coincides with a higher production cost, economic theory otherwise provides little guidance on choice of functional form. Although use of the semilogarithmic specification in exhibit 2 has been commonplace dating back as far as a paper by Court (1939), in the absence of a theoretic justification, some testing of alternative forms is desirable. To allow for some flexibility in functional form and provide a way to test alternatives, all models considered in this article were initially estimated by employing a Box-Cox functional form that transforms the dependent variable (house value or rent) according to the following formula:

$$f(\lambda) = (y^{\lambda} - 1) / \lambda$$

Exhibit 2

where λ is a free parameter to be estimated. This specification includes linear ($\lambda = 1$) and semilogarithmic ($\lambda = 0$) models as special cases. This functional form was first employed by NAHB to analyze AHS data in Emrath (2002) for single-family housing and in Emrath (2001) for multifamily housing.

House Value: Ordinary Least Squares (OLS) Regression Estimates (1 of 2)

Dependent variable: natural logarithm of value of owner-occupied, single-family detached housing units

	Model 1	Model 2	Model 3	Model 4	Model 5		
Constant	10.644* (0.421)	10.725* (0.423)	10.738* (0.424)	10.727* (0.424)	10.724* (0.423)		
Location indicator							
CENT_NE	0.777* (0.076)	0.637* (0.066)	0.627* (0.065)	0.635* (0.066)	0.636* (0.066)		
CENT_MW	0.016 (0.024)	- 0.083* (0.030)	- 0.078* (0.030)	- 0.084* (0.030)	- 0.083* (0.030)		
CENT_SO	0.123* (0.021)	0.035 (0.024)	0.036 (0.024)	0.035 (0.024)	0.035 (0.024)		
CENT_CA	1.156* (0.109)	1.039* (0.099)	1.034* (0.099)	1.037* (0.099)	1.039* (0.099)		
CENT_WE	0.693* (0.067)	0.567* (0.059)	0.562* (0.058)	0.565* (0.059)	0.567* (0.059)		
BURB_NE	0.866* (0.080)	0.731* (0.070)	0.724* (0.069)	0.730* (0.070)	0.731* (0.070)		
BURB_MW	0.144* (0.024)	0.045* (0.026)	0.046* (0.026)	0.044* (0.026)	0.045* (0.026)		
BURB_SO	0.194* (0.022)	0.119* (0.024)	0.120* (0.024)	0.118* (0.024)	0.119* (0.024)		
BURB_CA	1.293* (0.122)	1.163* (0.110)	1.156* (0.110)	1.161* (0.110)	1.163* (0.110)		
BURB_WE	0.753* (0.071)	0.625* (0.062)	0.619* (0.061)	0.623* (0.062)	0.625* (0.062)		
NMET_NE	0.435* (0.047)	0.389* (0.043)	0.385* (0.043)	0.387* (0.043)	0.389* (0.043)		
NMET_MW	- 0.088* (0.023)	- 0.111* (0.024)	- 0.112* (0.024)	- 0.112* (0.024)	- 0.111* (0.024)		
NMET_WE	0.596* (0.059)	0.580* (0.057)	0.573* (0.056)	0.578* (0.057)	0.580* (0.057)		
Housing unit c	haracteristic						
SIZE1900	0.055* (0.008)	0.056* (0.008)	0.058* (0.008)	0.056* (0.008)	0.056* (0.008)		
SIZE1950	0.072* (0.008)	0.067* (0.008)	0.067* (0.008)	0.067* (0.008)	0.067* (0.008)		
SIZE1980	0.100* (0.011)	0.094* (0.010)	0.094* (0.010)	0.094* (0.010)	0.094* (0.010)		
SIZE1995	0.111* (0.012)	0.107* (0.011)	0.106* (0.011)	0.107* (0.011)	0.107* (0.011)		
SIZE2007	0.129* (0.018)	0.124* (0.017)	0.123* (0.017)	0.124* (0.017)	0.124* (0.017)		
LOTSIZE	0.009* (0.002)	0.012* (0.002)	0.012* (0.002)	0.012* (0.002)	0.012* (0.002)		
BATHS	0.213* (0.021)	0.193* (0.019)	0.191* (0.019)	0.193* (0.019)	0.193* (0.019)		
HALFB	0.122* (0.013)	0.115* (0.013)	0.113* (0.012)	0.115* (0.013)	0.115* (0.013)		
BEDRMS	0.040* (0.006)	0.044* (0.006)	0.045* (0.006)	0.044* (0.006)	0.044* (0.006)		
DINING	0.068* (0.010)	0.067* (0.009)	0.067* (0.009)	0.067* (0.009)	0.066* (0.009)		

Exhibit 2

House Value: Ordinary Least Squares (OLS) Regression Estimates (2 of 2) Dependent variable: natural logarithm of value of owner-occupied, single-family detached housing units

	Model 1	Model 2	Model 3	Model 4	Model 5					
Housing unit cha	Housing unit characteristic (continued)									
FAMRM	0.083* (0.011)	0.074* (0.010)	0.074* (0.010)	0.074* (0.010)	0.074* (0.010)					
OTHROOMS	0.039* (0.006)	0.038* (0.006)	0.040* (0.006)	0.038* (0.006)	0.038* (0.006)					
BA_MW	0.088* (0.018)	0.085* (0.018)	0.085* (0.018)	0.085* (0.018)	0.085* (0.018)					
BA_SO	0.193* (0.023)	0.182* (0.022)	0.180* (0.022)	0.182* (0.022)	0.182* (0.022)					
BA_CA	0.442* (0.077)	0.452* (0.076)	0.449* (0.075)	0.452* (0.076)	0.452* (0.076)					
GARAGEX	0.052* (0.012)	0.039* (0.011)	0.036* (0.011)	0.039* (0.011)	0.039* (0.011)					
FIREPLAC	0.141* (0.015)	0.126* (0.014)	0.126* (0.013)	0.126* (0.014)	0.126* (0.014)					
AC_MWSO	0.148* (0.018)	0.125* (0.016)	0.118* (0.016)	0.124* (0.017)	0.125* (0.016)					
Neighborhood c	haracteristic									
GREEN		0.028* (0.008)	0.029* (0.008)	0.029* (0.008)	0.028* (0.008)					
COMRECR		0.025* (0.008)	0.026* (0.008)	0.025* (0.008)	0.025* (0.008)					
COMGATE		0.105* (0.019)	0.104* (0.019)	0.105* (0.019)	0.105* (0.019)					
XWATER		0.066* (0.012)	0.067* (0.012)	0.066* (0.012)	0.066* (0.012)					
WFP_NE		0.170* (0.051)	0.166* (0.051)	0.170* (0.051)	0.170* (0.051)					
WFP_MW		0.304* (0.046)	0.301* (0.046)	0.304* (0.046)	0.304* (0.046)					
WFP_SO		0.365* (0.046)	0.363* (0.045)	0.365* (0.046)	0.365* (0.046)					
WFP_CA		0.324* (0.187)	0.316* (0.187)	0.324* (0.187)	0.325* (0.187)					
WFP_WE		0.436* (0.084)	0.429* (0.083)	0.436* (0.084)	0.436* (0.084)					
XTR_MET		0.121* (0.014)	0.120* (0.014)	0.121* (0.014)	0.121* (0.014)					
SHP_MET		0.046* (0.018)	0.044* (0.018)	0.046* (0.018)	0.046* (0.018)					
BARCL		- 0.096* (0.018)	- 0.095* (0.018)	- 0.096* (0.018)	- 0.096* (0.018)					
ABAN		- 0.154* (0.022)	- 0.149* (0.021)	- 0.154* (0.022)	- 0.154* (0.022)					
BADROADS		- 0.031* (0.008)		- 0.031* (0.008)	- 0.031* (0.008)					
COMCRIME		- 0.029* (0.010)	- 0.026* (0.010)	- 0.029* (0.010)	- 0.029* (0.010)					
COMODOR		- 0.046* (0.018)	- 0.043* (0.018)	- 0.046* (0.018)	- 0.046* (0.018)					
XSTNOISE		- 0.037* (0.010)	, ,	- 0.037* (0.010)	- 0.037* (0.010)					
JNK_MET		- 0.100* (0.020)	- 0.094* (0.020)	- 0.100* (0.020)	- 0.100* (0.020)					
JNK_NM		- 0.018* (0.031)	-0.010 (0.031)	- 0.017 (0.031)	- 0.018 (0.031)					
COM1		- 0.038* (0.010)	- 0.037* (0.010)	- 0.038* (0.010)	- 0.038* (0.010)					
COM2		- 0.030 (0.021)		- 0.030 (0.021)	- 0.030 (0.021)					
MB_MET			- 0.162* (0.024)	- 0.165* (0.024)	- 0.165* (0.024)					
MB_NM		- 0.092* (0.020)	- 0.091* (0.020)	- 0.092* (0.020)	- 0.093* (0.020)					
Alternative indic	ators of physical	inadequacy								
NEW_INAD			- 0.112* (0.017)							
AHS_MOD				- 0.014 (0.022)						
AHS_SEV					0.036 (0.036)					
Adj. R ²	0.5044	0.5320	0.5335	0.5320	0.5320					

Standard errors in parentheses. * Indicates a coefficient significant at the 0.1 level.

Single-Family Regressions

Exhibit 2 shows the results of estimating several specifications of a hedonic model using the 2009 national AHS, where the dependent variable is the natural logarithm of the value of a single-family detached, owner-occupied home. The five models were estimated using NLOGIT 3.0 (the software previously named LIMDEP). Manipulation of the AHS public use file and preparation of a data set for input into NLOGIT was accomplished with SAS.

One feature of the AHS that may be considered a disadvantage in a hedonic model like this is that the dependent variable, the logarithm of the current market value of an owner-occupied housing unit, is based on the owner's estimate of the home's value, rather than an independent appraisal or transaction price reported on real estate records. Researchers have raised questions about the accuracy of owner's self-reported valuations, which were systematically investigated by Kiel and Zabel (1999). They found that recent buyers on average reported home values 8.4 percent above the stated sales price, while owners who had been in their homes for a longer period of time tended to overvalue their homes by 3.3 percent. The average across all owners was a tendency to overvalue their homes by 5.1 percent. However, they also found that, with the exception of length of tenure in the current residence, owners' valuations were not systematically related to characteristics of the owner, housing unit, or neighborhood. If this finding holds, it suggests that coefficients in a correctly specified hedonic model could accurately identify independent variables that have a significant effect on the value of the home, although the coefficients on the independent variables may be inflated slightly.

Before the models were estimated, efforts were undertaken to clean the data, which consisted primarily of removing observations for which the dependent or key independent variables have been allocated, top coded, or appear unreasonable. Deleted were observations for manufactured housing units; cases in which value or size of the unit is allocated or top coded, in which the number of bathrooms or year built is allocated, or in which the house value is \$29,000 or less; and, taking advantage of the longitudinal nature of the AHS, in which the housing unit's value has fallen to less than one-third of the value reported for the same unit in 2007. After these deletions, a total of 20,340 observations from the 2009 AHS were used to generate the ordinary least squares (OLS) estimates shown in exhibit 2.

Exhibit 3 shows likelihood ratio tests for each semilogarithmic model shown in exhibit 2 against an alternative of the more general Box-Cox specification. In each case, the estimated value of λ is relatively close to zero—0.009 or smaller in absolute value—and not statistically significant even at, for example, a 0.1 significance level. In other words, the likelihood ratio tests fail to reject any of the semilogarithmic specifications in exhibit 2 in favor of the more general form. Moreover, even if the Box-Cox version of the model were to be used, with an estimated value of λ so close to zero, it would make little practical difference. Hence, the semilogarithmic specification was chosen for the models shown in exhibit 2. Because of the logarithmic transformation, coefficients in exhibit 2 have the interpretation of percentage change in home value that is attributable to a particular independent variable.

Exhibit 3

Likelihood Ratio Tests of Alternatives to House Value Models in Exhibit 2 Dependent variable: natural logarithm of value of owner-occupied, single-family detached housing units

	Model 1	Model 2	Model 3	Model 4	Model 5
	Location indicator				
	Housing unit characteristic				
		Neighborhood characteristic	Neighborhood characteristic	Neighborhood characteristic	Neighborhood characteristic
			NEW_INAD	AHS_MOD	AHS_SEV
Test against alte	ernative of a Box	-Cox form with p	arameter λ		
estimate of λ	0.0089	0.0089	0.0078	0.0088	0.0090
χ² statistic	1.386	1.437	1.096	1.403	1.456
d.f.	1	1	1	1	1
p-value	0.2391	0.2306	0.2952	0.2362	0.2275
Test against a n	nodel with fewer	independent vari	ables		
null hypothesis	constant only	Model 1	Model 2	Model 2	Model 2
χ² statistic	14,307.826	1,188.738	68.215	0.401	1.032
d.f.	31	23	1	1	1
p-value	< 0.0001	< 0.0001	< 0.0001	0.5265	0.3097

d.f. = degrees of freedom.

Model 1 in exhibit 2 shows the result of regressing the logarithm of home value on locational indicators and housing unit characteristics. Because the geographically unique area of the nonmetropolitan South is omitted to avoid collinearity, the coefficients on the location indicators in exhibit 2 are estimated percentage changes in value relative to a location in the nonmetropolitan South.

Model 1 excludes an important set of characteristics—those that describe neighborhood conditions. Researchers broadly agree that neighborhood conditions can affect property values, so model 1 is included not as a likely candidate for the final model, but to illustrate the sensitivity of coefficients on certain variables, particularly the location indicators, to a significant change in the model's specification. Model 2 includes all the variables in model 1 plus 23 neighborhood characteristics, most of which are statistically significant. The adjusted R² statistic, a conventional measure of the fit of the model, is higher for model 2 than for model 1.

Exhibit 3 also shows likelihood ratio tests of model 1 against a model with a constant term as the only regressor, and of model 2 against model 1. The likelihood ratio test is applicable under relatively general conditions that can be used to evaluate all models considered in this article, including those that are nonlinear and estimated with a technique other than OLS. The likelihood ratio tests in exhibit 3 reject a constant-only model in favor of model 1, and model 1 in favor of model 2.

² Also, HUD is currently considering deleting many of the neighborhood characteristics the next time the AHS is redesigned. In the future, researchers may have to work with models that do not include these variables.

Coefficients on the independent variable common to models 1 and 2 are relatively stable across specifications, especially the housing unit characteristics. Among the location indicators, the smallest coefficient in model 1 (for CENT_MW) changes sign in model 2 and the second smallest (for CENT_SO) becomes insignificant at the 0.1 level.

Housing characteristics employed in the models include size of the home crossed with the year it was built, because of a systematic tendency for a square foot of living space to be worth more for homes that were built more recently. This tendency may reflect lower maintenance or operating costs that are capitalized into the value of the home, or more stringent building codes that increase perceived safety of the homes and their construction costs. The age of the home may also be acting in part as a proxy for floor plans that change over time in response to consumer preferences or for other features not captured in the data set.

Interactions between location and many of the housing unit characteristics and all of the neighborhood characteristics were tried in preliminary versions of the models and retained in the specifications shown in exhibit 2, where they made a noticeable difference. Among the housing unit characteristics, the strongest effects on home values are associated with the presence of a basement if the presence of a basement occurs in the California metropolitan region, followed by the addition of a full bathroom. The difference in the cost of excavating and constructing a basement relative to the alternatives can vary substantially with location, depending on soil conditions and code requirements for a foundation in the absence of a basement.

All of the housing unit characteristics included in exhibit 2 have a positive effect on home values that is significant at the 0.1 level or better. The 23 neighborhood characteristics in model 2 include those with both positive and negative effects on value. The strongest positive estimated effects on value among the neighborhood characteristics are associated with a location directly on the waterfront. The strongest negative estimated effects come from the presence of abandoned buildings in the neighborhood. Coefficients on two of the neighborhood variables, JNK_NM and COM2, are not significant at the 0.1 level, but the sign of the coefficients and their magnitude relative to other coefficients estimated in the models in exhibit 2 are plausible.

Multifamily Regressions

Exhibit 4 shows estimated results for several specifications of a model in which the dependent variable is a Box-Cox transformed version of monthly rent for multifamily units, defined as units in structures with five or more units. Manufactured housing is excluded. The measure of rent is based on gross rent for units in which no fuel (electricity, gas, or oil) is included in the rental payment. Gross rent includes rent paid to the property owner and the cost of nontelecommunication utilities, regardless of who pays for them. To remove ambiguity and achieve a consistent measure across observations, it is necessary to use gross rent rather than simply the payment made to the property owner, due to the substantial variation that exists in practices for including utility costs in the rental payment. Some ambiguity remains due to differential treatment of items such as water and sewer payments, but these payments are typically small relative to the payment for energy.

Efforts to clean the multifamily rental data from the 2009 national AHS included deleting cases in which the tenant is occupying the unit without paying cash rent; the occupant reports that the unit

Exhibit 4

Multifamily Rent: Box-Cox Regression Estimates

Dependent variable: rent paid by tenants in structures with five or more units, transformed by Box-Cox parameter $\boldsymbol{\lambda}$

BOX-COX pa	Model 1	Model 2	Model 3	Model 4	Model 5			
λ -	- 0.4291*(0.0344)	- 0.3887*(0.0334)	- 0.3872*(0.0334)	- 0.3887*(0.0334) -	- 0.3885* (0.0334)			
Constant	2.1586*(0.1368)	2.3230*(0.1491)		, ,	2.3238* (0.1491)			
Location indi	, ,	,	,	,	,			
CENT_NE	0.0328*(0.0079)	0.0308*(0.0076)	0.0312*(0.0076)	0.0308*(0.0076)	0.0308* (0.0075)			
CENT_MW	0.0134*(0.0037)	0.0091*(0.0034)	0.0092*(0.0034)		0.0091* (0.0034)			
CENT_SO	0.0127*(0.0035)	0.0116*(0.0036)	0.0118*(0.0036)		0.0116* (0.0036)			
CENT_CA	0.0314*(0.0076)	0.0365*(0.0087)	0.0369*(0.0088)		0.0364* (0.0087)			
CENT_WE	0.0160*(0.0042)	0.0157*(0.0045)	0.0158*(0.0045)		0.0157* (0.0045)			
BURB_NE	0.0309*(0.0075)	0.0342*(0.0082)	0.0346*(0.0083)	, ,	0.0342* (0.0082)			
BURB MW	0.0134*(0.0037)	0.0106*(0.0035)	0.0107*(0.0036)		0.0106* (0.0035)			
BURB_SO	0.0163*(0.0042)	0.0164*(0.0044)			0.0164* (0.0044)			
BURB_CA	0.0329*(0.0079)	0.0409*(0.0097)	0.0413*(0.0098)		0.0409* (0.0097)			
BURB_WE	0.0159*(0.0042)	0.0168*(0.0047)	0.0170*(0.0047)		0.0169* (0.0047)			
NMET_NE	0.0210*(0.0065)	0.0208*(0.0072)	0.0214*(0.0073)		0.0209* (0.0072)			
				- 0.0045 (0.0035) -				
NMET_WE	0.0044 (0.0035)	0.0049 (0.0044)	0.0047 (0.0044)		0.0047 (0.0044)			
Housing unit characteristic								
SIZEBF85	0.0003 (0.0005)	0.0009 (0.0007)	0.0008 (0.0007)	0.0009 (0.0007)	0.0008 (0.0007)			
SIZE1985	0.0010 (0.0009)	0.0013 (0.0011)	0.0013 (0.0011)		0.0014 (0.0011)			
SIZEPOST	0.0034*(0.0011)	0.0036*(0.0012)	0.0036*(0.0012)		0.0036* (0.0012)			
BATHS	0.0069*(0.0019)	0.0076*(0.0021)	0.0077*(0.0021)		0.0078* (0.0021)			
HALFB	0.0015 (0.0008)	0.0023*(0.0011)	0.0023*(0.0011)		0.0023* (0.0011)			
BEDRMS	0.0042*(0.0011)	0.0071*(0.0017)	0.0072*(0.0017)		0.0071* (0.0017)			
OTHROOMS	0.0003 (0.0006)	0.0006 (0.0007)	0.0006 (0.0008)		0.0006 (0.0007)			
GARAGEX	0.0050*(0.0014)	0.0043*(0.0014)			0.0042* (0.0014)			
FIREPLAC	0.0018 (0.0010)	0.0038*(0.0015)	0.0039*(0.0015)		0.0038* (0.0015)			
UDISH	0.0057*(0.0015)	0.0056*(0.0016)	0.0056*(0.0016)	, ,	0.0057* (0.0016)			
UDRY	0.0049*(0.0014)	0.0059*(0.0016)		, ,	0.0058* (0.0016)			
Building/neig	hborhood chara		,	, ,	, ,			
UELEV	inbornood ondra	0.0043*(0.0019)	0.0044*(0.0019)	0.0043*(0.0019)	0.0044* (0.0019)			
UACCESSB		0.0026*(0.0011)	0.0026*(0.0011)	, ,	0.0026* (0.0011)			
FLOOR_3		0.0038*(0.0012)	0.0039*(0.0012)	, ,	0.0038* (0.0012)			
FLOOR_49		0.0104*(0.0029)	0.0106* (0.0029)		0.0104* (0.0029)			
FLOOR_10		0.0247*(0.0063)	0.0250*(0.0064)	, ,	0.0248* (0.0063)			
FIFTY_1		0.0395*(0.0120)	0.0397*(0.0121)		0.0397* (0.0120)			
COMRECR		0.0024*(0.0010)	0.0024*(0.0010)		0.0024* (0.0010)			
XWATER		0.0032*(0.0013)			0.0032* (0.0013)			
XWFPROP		0.0048*(0.0027)	0.0048*(0.0028)		0.0048* (0.0027)			
XTRAN		0.0028*(0.0012)	0.0028*(0.0012)		0.0027* (0.0012)			
XSHOP		0.0050*(0.0023)	0.0050*(0.0023)	0.0050*(0.0023)	0.0050* (0.0023)			
JNK_MET		, ,	, ,	- 0.0013 (0.0012) -	, ,			
JNK_NM				- 0.0110*(0.0049) -				
	hysical inadequa	` ,	,	,	,			
NEW_INAD	, oroar madeque	,	- 0.0030*(0.0017)					
AHS_MOD			0.0000 (0.0017)	0.0002 (0.0014)				
AHS_SEV				0.0002 (0.0014)	0.0044 (0.0029)			
Amemiya's	0.6170	0.5524	0.5507	0.5525	,			
prediction criteria	0.6179	0.5531	0.5527	0.5535	0.5530			
Standard errors	in parentheses. * Indi	icates a coefficient sigi	nificant at the 0.1 level.					

is public housing, that the government subsidizes the rent or limits the rent through rent control or stabilization, or that the rent is adjusted because the tenant is related to the owner; the size of the unit is allocated or top coded; the rent or the number of bathrooms is allocated; or the rent has either tripled or fallen to less than one-third of value reported for the same unit in 2007. After these deletions, a total of 2,645 observations from the 2009 AHS were used to generate the regression estimates shown in exhibit 4.

The regressions in exhibit 4 are a Box-Cox functional form with λ , an additional parameter that is estimated with the rest through a maximum-likelihood procedure. Because of the use of this functional form, the adjusted R² statistic is not available. As a substitute, the Prediction Criterion (PC) introduced in Amemiya (1980) is provided as an alternate goodness-of-fit measure, where a smaller PC indicates a better fit.

Model 1 in exhibit 4 shows the result of regressing the transformed version of rent on locational indicators and housing unit characteristics. Model 2 includes all the variables in model 1 plus a number of characteristics that pertain to the overall building or community surrounding it. According to PC, model 2 is preferred to model 1. Exhibit 5 shows likelihood ratio tests for the models in exhibit 4. These tests reject a constant-only model in favor of model 1, and model 1 in favor of model 2 against a model with a constant term as the only regressor, and of model 2 against model 1. Both the PC and likelihood ratio tests indicate that a model that includes building and neighborhood characteristics is appropriate.

Exhibit 5

Likelihood Ratio Tests of Alternatives to the Multifamily Rent Models in Exhibit 4 Dependent variable: rent paid by tenants in structures with five or more units, transformed by Box-Cox parameter λ

	Model 1	Model 2	Model 3	Model 4	Model 5
	Location indicator				
	Housing unit characteristic				
		Neighborhood characteristic	Neighborhood characteristic	Neighborhood characteristic	Neighborhood characteristic
			NEW_INAD	AHS_MOD	AHS_SEV
Test against null	of a linear mode	el (λ = 1)			
χ² statistic	2,210.051	2,228.667	2,226.896	2,228.496	2,230.078
d.f.	1	1	1	1	1
p-value	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Test against null	of a semilogarit	thmic model (λ =	0)		
χ² statistic	167.435	145.452	144.488	145.448	145.470
d.f.	1	1	1	1	1
p-value	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Test against a m	odel with fewer	independent vari	ables		
null hypothesis	constant only	Model 1	Model 2	Model 2	Model 2
χ² statistic	1,323.449	317.750	373.347	0.0002	2.553
d.f.	24	13	1	1	1
p-value	< 0.0001	< 0.0001	< 0.0001	0.9902	0.1101

d.f. = degrees of freedom.

For each of the models considered in exhibit 4, exhibit 5 also includes likelihood ratio tests for linear and semilogarithmic models against the alternative of the Box-Cox specification. In each case, both the linear and semilogarithmic forms are rejected in favor of the more general alternative of the Box-Cox regression that appears in exhibit 4.

The estimated value of λ varies somewhat across the exhibit, but remains in the neighborhood of -0.4. Analogous to the results shown for the single-family regressions, coefficients on the independent variable common to models 1 and 2 are relatively stable across specifications in exhibit 4, although coefficients on two of the housing unit characteristics (HALFB and FIREPLAC) become significant with the addition of building and neighborhood characteristics to the model.

Because of the Box-Cox functional form, coefficients associated with independent variables no longer have the interpretation of a percentage change in the dependent variable. As an aid to interpretation, a few examples of marginal effects are included in the exposition. If a housing unit with characteristics that produced an estimated monthly rent of \$1,000 in the nonmetropolitan South were located instead in a southern central city, the coefficient in model 2 in exhibit 4 implies that its monthly rent would increase by \$190. If the original unit were located in a suburb in the "California metropolitan" region rather than the nonmetropolitan South, its rent would increase by \$962.

Housing characteristics considered in the rent models include size of the unit crossed with the year in which the structure was built. Although exhibit 4 shows a systematic tendency for a square foot of living space to be worth more in structures that were built more recently, this tendency is shown in considerably less detail than in the single-family model, with vintage differences becoming undetectable among units built before 1985. This lack of detail may be a symptom of the smaller number of observations used to generate exhibit 4, but renters may also be less knowledgeable than owners about the vintage of the building they are living in. Neither the coefficient on SIZEBF85 or SIZE1985 is significant at the 0.1 level, but both are retained in the models under the argument that the hypothesis that rents increase along with square footage of the unit remains reasonable, and the relative magnitude of the coefficients is consistent with the hypothesis that the value of a square foot of living space is higher in newer units.

The coefficient on OTHROOMS is also relatively small and not significant at the 0.1 level. This coefficient is estimated in a model that controls for a number of other characteristics, including square footage of the unit, bathrooms, and bedrooms. It is reasonable to suppose that an extra wall partition in an apartment that does not result in a larger unit, or an extra bed or bathroom, would add relatively little to the rent that can be charged. For a hypothetical unit that rents for \$1,000 per month, the coefficient in model 2 in exhibit 4 implies that an extra room without a corresponding increase in square footage would increase the estimated monthly rent by \$8.

A number of building characteristics have a significant effect on rents in model 2, particularly those related to the number of floors and units in the building. Tall buildings and buildings containing a large number of units tend to be more common in places where land is expensive, and the indicators for these characteristics are acting as a proxy for this tendency in the model. Some neighborhood characteristics have a significant effect, but there are fewer significant neighborhood effects in the multifamily models than in the single-family models shown in exhibit 2. Again, this difference between single-family and multifamily models may partly reflect the smaller number of

observations used to estimate the multifamily models, but it may also be that renters in multifamily structures feel more isolated from conditions in the surrounding neighborhood and therefore do not attach the same value to these conditions as do single-family homeowners. For a hypothetic unit that rents for \$1,000, the coefficient in model 2 in exhibit 4 implies that a waterfront location increases rent for the unit by \$73.

Physical Inadequacy

Because an important part of HUD's mission is to create quality housing for all, it is not surprising that a survey such as the AHS, which is funded by HUD to track the condition of housing in the United States, collects considerable information on housing unit quality. Combining this information into a single indicator is far from a trivial exercise, however. The multidimensional nature of housing means that a meaningful analysis of housing quality cannot be reduced to one or two simple characteristics, a point emphasized by Weicher (1979). One implication of this inability to measure quality by using one or two simple characteristics is that data sets that contain information on only a limited number of housing characteristics, such as the American Community Survey, are not useful for detecting physical inadequacy. In fact, it is difficult to imagine a data set other than the AHS that could be effectively used for this purpose.

The AHS had traditionally provided two standards for housing inadequacy—moderately inadequate and severely inadequate. A housing unit is classified as severely inadequate in the AHS if it has any one of the following conditions:

- Fewer than two full bathrooms without hot and cold running water, or without bathtub or shower, or without a flush toilet, or with shared plumbing.
- Respondent who reports being cold for 24 hours or more and at least two breakdowns of heating equipment lasting longer than 6 hours.
- Respondent reporting that the household does not use electricity.
- Exposed wiring, plus a lack of electrical outlets in every room, plus fuses that have blown more than twice.
- At least five of the following conditions:
 - · Outside water leaks.
 - · Inside water leaks.
 - · Holes in the floor
 - · Open cracks in the inside walls or ceilings.
 - · An area of peeling paint larger than 8 by 11 inches.
 - Respondent who reports seeing rats recently.

A housing unit that is not severely inadequate is moderately inadequate in the AHS if it has any one of the following conditions:

- At least three of the conditions listed in the previous list.
- More than two toilet breakdowns lasting 6 hours or more.
- Main heating equipment consisting of unvented room heaters.
- Lack of complete kitchen facilities.

Although these criteria are relatively complex, they exploit only a fraction of the data collected in the AHS that could be used as a basis to classify housing units as adequate or inadequate.

The basic structure of the AHS summary inadequacy definitions has been in place for decades, and it appears to have undergone relatively little scrutiny during that time. The changes that have taken place seem to be the result of attempts to streamline the AHS data set rather than to refine or improve the adequacy classification scheme. For example, the definitions of severe and moderate inadequacy were simplified slightly in 2007, removing additional criteria based on problems in common areas of multifamily structures (such as lack of lighting, broken stairways, and loose or no rails on stairs), when questions on some of these characteristics were deleted from the AHS questionnaire.

Moreover, when the concepts of moderate and severe inadequacy were originally defined, although considerable thought went into the process, it was nevertheless done in a relatively ad hoc fashion, without trying alternate specifications in an economic model, according to Crowe (2011), who was involved in the development of the AHS in the 1970s.

This section of the article proposes new summary criteria for determining the physical adequacy of housing units that use somewhat different AHS variables. The new definition of inadequacy presented in this article, identified in the exhibits as NEW_INAD, flags a single-family structure as physically inadequate if it has any one of the following conditions:

- · Missing siding.
- · Broken windows.
- Holes, cracks, or crumbling in the foundation.
- Sagging roof.
- Holes in the floor.

These characteristics were chosen by running a number of regressions with each possible indicator of inadequacy entered one at a time to identify those that individually tended to lower single-family home values, controlling for other factors. Building conditions, such as the ones used to construct NEW_INAD, would generally cost a significant amount of money, time, and effort to repair, so it is theoretically plausible that they would have a depressing effect on the reported value of single-family housing units.

To illustrate the effect of NEW_INAD on the reported value of the home, model 3 in exhibit 2 adds NEW_INAD to model 2. The coefficient is significant and indicates that, controlling for the other variables in the model, NEW_INAD reduces the value of the single-family unit by roughly 11 percent.

Models 4 and 5 introduce the standard AHS measures of inadequacy, AHS_MOD and AHS_SEV, into the hedonic specification. Although it would be unfair to expect these traditional inadequacy indicators to perform as well as NEW_INAD, which was constructed to work well in the model, the weakness of the traditional indicators in the hedonic specification is perhaps surprising. The estimated coefficients indicate that AHS_MOD reduces house value by 1.4 percent, while AHS_SEV increases house value by 3.6 percent. Neither effect is significant at the 0.1 level. Among the three alternative indicators of inadequacy in exhibit 2, only NEW_INAD improves the fit of the model as indicated by the adjusted R².

The hypothesis tests in exhibit 3 also reject model 2 in favor of model 3, but they fail to reject model 2 in favor of either model 4 or model 5.

An advantage of the conditions used to define NEW_INAD for single-family housing units is that they are based on questions asked in all AHS interviews, including interviews conducted for vacant units, allowing investigation into the condition of vacant as well as occupied housing units.

A drawback to the conditions used in the single-family definition of NEW_INAD is that they do not allow for a consistent definition across structure type—for example, the conditions, in general, are not applicable to multifamily structures. Most of the conditions used in single-family NEW_INAD were cited by very few to zero occupants of multifamily rental units in the 2009 AHS. This finding is probably not surprising, given the nature of the characteristics, which refer to the condition of the building rather than a particular unit in a building. Depending partly on the location of their unit within the building, occupants of multifamily structures may be unaware of the condition of, for example, the foundation or roof. Indeed, the occupant may not consider deficiencies in these parts of the building to be problems if the deteriorated condition is located in a part of the structure that is remote from his or her own unit.

Hence, the proposed definition for NEW_INAD for multifamily units is based on a different set of characteristics. NEW_INAD indicates that a multifamily housing unit is inadequate if it has any of the following conditions:

- · Lack of a kitchen sink.
- · Lack of a bathroom sink.
- Open cracks in the inside walls or ceilings.
- A breakdown of the sewage system since the previous interview.
- Lack of built-in equipment designed to distribute heat throughout the unit in climates with 4,000 or more heating degree days (HDDs).

The individual characteristics used to define the multifamily version of NEW_INAD were also chosen by running a number of regressions with each possible indicator of inadequacy entered one at a time to identify those that individually tended to have a depressing effect on the dependent variable in question—in this case, monthly rent. The characteristics in the multifamily version of NEW_INAD, in general, are observable by occupants of the multifamily unit. Except for a breakdown of the sewage system, the characteristics are also based on information that the AHS collects for both vacant and occupied housing units.

A housing unit is identified as having a lack of built-in heating equipment designed to distribute heat throughout the unit if it has no main heating equipment, or if the main heating equipment is one of the following:

- Vented room heaters burning kerosene gas or oil.
- Unvented room heaters burning kerosene gas or oil.
- Portable electric heaters.
- A cooking stove.

In the AHS, the only one of these heating deficiencies that is used in the definition of inadequate housing is unvented room heaters. Yet, all types of room heaters, portable electric heaters, and gas or electric cooking stoves are often cited as safety hazards by organizations such as the U.S. Consumer Product Safety Commission (2009, 2011).

The multifamily version of NEW_INAD assumes that reliance on any heating equipment that requires so much care on the part of tenants to operate safely is a reasonable indicator of the physical inadequacy of the unit, as is a total lack of heating equipment in an area where heating equipment is needed. The cutoff used to identify areas where heating is needed is a climate with, on average, at least 4.000 HDDs.

The information on climate available in the AHS public use file is somewhat restricted. Climate data are collapsed into six zones, based on both heating and cooling degree days, and sometimes are suppressed for confidentiality reasons. The first three zones (coldest, cold, and cool) are characterized by at least 4,000 HDDs. The standard reference for a picture of the area captured by this degree day requirement is the set of maps produced by the National Oceanic and Atmospheric Administration (2009). In the central part of the country, 4,000-plus HDDs roughly coincide with the area north of Missouri's southern border. Toward the east coast, the line of demarcation drifts upward and the northern part of Virginia is the only part of that state that has 4,000-plus HDDs. On the west coast, the 4,000-plus HDD zone starts well north of San Francisco. In mountainous areas, the cutoff is determined by elevation as much as by latitude. Around the Sierra Nevada and Rocky Mountain ranges, the 4,000-plus HDD zone extends well to the south, but much of this area is sparsely populated.

When considering the depressing effect that a lack of built-in heating in colder climates has on rent, it is useful to recall that the sample used to estimate the models in exhibit 4 excludes cases in which utility payments are included in rents, so the estimated effects on rents are effects on rents exclusive of utility costs.

Analogous to the treatment of inadequacy in single-family regressions, competing definitions of inadequacy are added to model 2 one at a time in exhibit 4. Model 3 adds the proposed new definition of physical inadequacy. The coefficient on this variable is statistically significant and implies that, for a hypothetical housing unit that would otherwise rent for \$1,000, inadequacy under this new definition reduces rent by \$42.

In contrast, in models 4 and 5, coefficients on the standard AHS measures of inadequacy, AHS_MOD and AHS_SEV, are insignificant and have the wrong signs. For the hypothetical

\$1,000-per-month rental apartment, the coefficients in these models imply that AHS_MOD would increase rent by 25 cents and AHS_SEV would increase rent by \$67. The PC statistics favor model 3 over any of the alternatives in exhibit 4. The hypothesis tests in exhibit 5 also reject model 2 in favor of model 3, but they fail to reject model 2 in favor of either model 4 or model 5.

Of the three indicators of physical inadequacy considered, the statistics presented in exhibits 2 through 5 consistently favor the new definition proposed in this article over the traditional summary measures provided on the AHS data file in the hedonic models that explain house value and rent levels.

Characteristics of Inadequate Housing Units

The first question that arises in evaluating a proposed definition of inadequacy is the number of housing units that it classifies as inadequate. Exhibit 6 shows the number of housing units classified as inadequate under both of the standard summary criteria in the AHS and under the new criterion proposed in this article.

The standard criteria in the AHS tend to classify a small share of single-family units as inadequate relative to multifamily units. According to these criteria, 1.3 percent of occupied single-family units and 2.9 percent of occupied multifamily units are severely inadequate and 3.5 percent of occupied single-family units are at least moderately inadequate compared with 10.1 percent of the occupied multifamily units. Because a large proportion of the U.S. housing stock consists of single-family housing, however, the standard AHS measures fairly equal numbers of occupied inadequate single-family and multifamily units—2.7 and 2.6 million, respectively, in the case of the moderately or severely inadequate category.

In contrast, the new inadequacy criterion proposed in this article, although based on somewhat different characteristics depending on structure type, captures near-equal shares of occupied single-family and multifamily units, classifying 8.5 percent of single-family units and 8.3 percent of multifamily units as inadequate. These percentages translate into a larger number of inadequate occupied housing units—8.8 million (6.7 million single-family and 2.1 million multifamily units) compared with only 5.3 million for the most inclusive of the traditional AHS inadequacy measures.

Exhibit 6

Number of Housing Units Classified as Inadequate Under Alternative Definitions

	Occupied		Nonseasor	nal Vacant	Total
	Single-Family	Multifamily	Single-Family	Multifamily	Iotai
AHS severely inadequate	991,358	744,606	0	0	1,735,965
	1.3%	2.9%	0.0%	0.0%	1.5%
AHS moderately or severely inadequate	2,727,494	2,607,392	0	0	5,334,886
	3.5%	10.1%	0.0%	0.0%	4.6%
Inadequate under new definition	6,733,007	2,153,890	1,104,633	397,619	10,389,149
	8.5%	8.3%	19.4%	8.9%	9.0%
Total housing units	79,133,307	25,920,344	5,707,567	4,449,398	115,210,615

AHS = American Housing Survey.

The tables published on the Census Bureau's website, based on the 2009 national AHS, show 5.7 million moderately or severely inadequate occupied units rather than 5.3 million, because the Census-published tables include 0.4 million inadequate manufactured housing units, and the tabulations in this article exclude manufactured housing.

Because the proposed inadequacy criterion relies primarily on data that are collected during AHS interviews conducted for both vacant and occupied housing units, it can be applied to vacant housing units. Exhibit 6 also shows estimates of the number of inadequate nonseasonal vacant units.³ The share of nonseasonal vacant multifamily units that are inadequate is only slightly higher than the share for occupied units, but the inadequate share of nonseasonal single-family units is more than 19 percent. The total number of nonseasonal vacant homes that are now defined as inadequate is 1.5 million. Knowledge of this statistic could significantly alter the way industry observers evaluate the inventory of existing single-family units on the market that are for sale or rent.

Exhibit 7 shows housing units by the year they were built, both for all units and for those that are inadequate under the definition of inadequacy proposed in this article. The inadequate units tend to be relatively old. Roughly one-third of inadequate units were built before 1940 compared with 16 percent of all occupied and nonseasonal vacant units. These results are not surprising because older units have had more time to undergo wear and tear and more time for problems of neglected maintenance to accumulate.

Exhibit 8 partitions the housing stock by geography rather than by vintage. Compared with housing units in general, the inadequate units are found less often in suburbs and more often in central cities and nonmetropolitan areas. A partial exception to this general rule is that, for nonseasonal vacant multifamily housing, the inadequate units are more concentrated in central cities but not in nonmetropolitan areas. Because the housing stock tends to be older in central cities and outside metropolitan areas, the results in exhibit 8, in general, are consistent with those shown in exhibit 7.

Exhibit 9 details the housing cost burden for households in units classified as inadequate under the new inadequacy criterion proposed in this article. Of the more than 8.8 million households living in inadequate housing units in the United States as of 2009, nearly 5.2 million are owners and 3.7 million are renters (exhibit 9). Fewer than 0.5 million owners and renters living in inadequate units are spending 30 percent or more of their incomes on housing. This relatively small overlap means that most occupants of inadequate units are not counted among the cost-burdened (using the traditional HUD definition of an occupant spending at least 30 percent of his or her income on housing) and represent a net addition to the number of American households with housing problems. All households that are both cost-burdened and living in inadequate units earn less than 50 percent of Area Median Income (AMI).

Renters living in inadequate housing units are particularly concentrated at the lower end of the income distribution. Nearly 40 percent of those renters earn less than 30 percent of AMI. Homeowners living in inadequate housing are more evenly spread across the income distribution.

³ Nonseasonal vacant housing units exclude vacant units that are coded as seasonal, migratory, or held for occasional use. These units are excluded from the tabulations because adequacy standards may be different for housing units not intended for year-round occupation.

Exhibit 7

Housing Units by Year Built							
·	Оссі	ıpied	Nonseasor	nal Vacant	T. 1 . 1		
	Single-Family	Multifamily	Single-Family	Multifamily	Total		
Units inadequate under	new definition						
Built before 1940	2,073,311	711,765	477,613	185,514	3,448,204		
	30.8%	33.1%	43.2%	46.7%	33.2%		
Built 1940 to 1949	688,909	150,805	127,836	28,666	996,217		
	10.2%	7.0%	11.6%	7.2%	9.6%		
Built 1950 to 1959	935,887	157,986	165,810	49,159	1,308,842		
	13.9%	7.3%	15.0%	12.4%	12.6%		
Built 1960 to 1969	803,426	270,471	99,785	27,372	1,201,054		
	11.9%	12.6%	9.0%	6.9%	11.6%		
Built 1970 to 1979	1,069,371	521,169	142,209	63,983	1,796,733		
	15.9%	24.2%	12.9%	16.1%	17.3%		
Built 1980 to 1989	516,269	178,954	39,629	19,637	754,489		
	7.7%	8.3%	3.6%	4.9%	7.3%		
Built 1990 to 1999	380,763	90,624	15,802	10,095	497,284		
	5.7%	4.2%	1.4%	2.5%	4.8%		
Built 2000 to 2004	178,932	47,066	15,290	2,771	244,059		
	2.7%	2.2%	1.4%	0.7%	2.3%		
Built 2005 or later	86,137	25,050	20,659	10,422	142,268		
	1.3%	1.2%	1.9%	2.6%	1.4%		
Total	6,733,007	2,153,890	1,104,633	397,619	10,389,149		
All units							
Built before 1940	12,078,056	4,756,872	1,295,838	957,960	19,088,726		
	15.3%	18.4%	22.7%	21.5%	16.6%		
Built 1940 to 1949	5,410,738	1,284,757	521,825	221,246	7,438,567		
	6.8%	5.0%	9.1%	5.0%	6.5%		
Built 1950 to 1959	10,045,797	1,678,732	686,503	206,774	12,617,806		
	12.7%	6.5%	12.0%	4.7%	11.0%		
Built 1960 to 1969	9,705,514	3,231,039	572,304	456,795	13,965,651		
	12.3%	12.5%	10.0%	10.3%	12.1%		
Built 1970 to 1979	13,150,016	6,619,579	742,667	1,052,086	21,564,348		
	16.6%	25.5%	13.0%	23.7%	18.7%		
Built 1980 to 1989	8,642,609	4,034,224	405,514	674,840	13,757,187		
	10.9%	15.6%	7.1%	15.2%	11.9%		
Built 1990 to 1999	9,449,565	2,112,683	478,721	346,680	12,387,649		
	11.9%	8.2%	8.4%	7.8%	10.8%		
Built 2000 to 2004	6,044,946	1,296,530	350,801	182,769	7,875,046		
	7.6%	5.0%	6.2%	4.1%	6.8%		
Built 2005 or later	4,606,067	905,928	653,394	350,246	6,515,635		
	5.8%	3.5%	11.5%	7.9%	5.7%		
Total	79,133,307	25,920,344	5,707,567	4,449,398	115,210,615		

Exhibit 8

Housing Units by Geography

	Оссі	upied	Nonseasor	nal Vacant	Total
	Single-Family	Multifamily	Single-Family	Multifamily	Total
Units inadequate unde	r new definition				
Central city	1,930,735	1,315,266	376,062	267,035	3,889,098
	28.7%	61.1%	34.0%	67.2%	37.4%
Urban suburb	1,860,964	526,729	205,822	78,455	2,671,969
	27.6%	24.5%	18.6%	19.7%	25.7%
Rural suburb	932,336	55,413	143,889	9,666	1,141,304
	13.9%	2.6%	13.0%	2.4%	11.0%
Urban nonmetropolitan	673,566	187,674	126,685	30,723	1,018,647
	10.0%	8.7%	11.5%	7.7%	9.8%
Rural nonmetropolitan	1,335,406	68,809	252,176	11,740	1,668,131
	19.8%	3.2%	22.8%	3.0%	16.1%
Total	6,733,007	2,153,890	1,104,633	397,619	10,389,149
All units					
Central city	18,976,616	13,079,571	1,521,548	2,436,767	36,014,503
	24.0%	50.5%	26.7%	54.8%	31.3%
Urban suburb	28,624,659	8,933,583	1,686,688	1,290,695	40,535,626
	36.2%	34.5%	29.6%	29.0%	35.2%
Rural suburb	12,914,932	1,118,746	822,732	211,028	15,067,438
	16.3%	4.3%	14.4%	4.7%	13.1%
Urban nonmetropolitan	6,041,805	2,015,375	554,617	364,156	8,975,953
	7.6%	7.8%	9.7%	8.2%	7.8%
Rural nonmetropolitan	12,575,294	773,070	1,121,982	146,750	14,617,096
	15.9%	3.0%	19.7%	3.3%	12.7%
Total	79,133,307	25,920,344	5,707,567	4,449,398	115,210,615

Exhibit 9

Housing Cost Burden for Households in Units Classified as Inadequate Under the New Definition

	Ow	Owner Occupied			Renter Occupied			
Household Income	Under 30% of Income	30–50% of Income	50% of Income or More	Under 30% of Income	30–50% of Income	50% of Income or More	Total	
Under 30% AMI	769,197	38,605	106,118	1,165,377	55,057	231,845	2,366,200	
30 to 50% AMI	762,509	2,508	0	805,814	0	0	1,570,831	
50 to 80% AMI	976,511	0	0	752,663	0	0	1,729,174	
80 to 120% AMI	1,100,113	0	0	428,960	0	0	1,529,074	
120% AMI or more	1,421,420	0	0	270,198	0	0	1,691,618	
Total	5,029,751	41,113	106,118	3,423,013	55,057	231,845	8,886,897	

AMI = Area Median Income.

More than one-fourth of those homeowners earn at least 120 percent of AMI. It seems reasonable to suppose that many owners of inadequate units in this income range have resources to upgrade their existing units or move to new ones if they so choose. These homeowners may therefore not be appropriate targets for housing assistance in the conventional sense, but an educational effort that provides information about property repair and maintenance could be worthwhile.

Exhibit 10 shows the number of housing units captured under the new inadequacy criterion proposed in this article by race and ethnicity of the household head. The problems of physically inadequate housing persist across major racial and ethnic categories. More than 76 percent of homeowners living in inadequate units are non-Hispanic White, and this statistic is only about 2 percent less than the incidence of 78 percent of non-Hispanic Whites among all homeowners. In fact, among the groups shown in exhibit 10, only non-Hispanic Black households are overrepresented among the homeowners who live in inadequate units. The racial and ethnic breakdown of renters living in inadequate units is very similar to the breakdown for all renters. In exhibit 10, Hispanic renters are the only group that is overrepresented in the inadequate column compared with its share among all renters, but the difference is less than 1 percentage point.

In contrast to the relatively egalitarian distribution of physically inadequate housing across racial and the ethnic lines, exhibit 11 shows that particular categories of households are disproportionately affected by problems of inadequate housing, and these are families with children. The effect is strongest for single-parent households and other households that are not headed by a married couple but nevertheless contain children under age 18. Although these nonmarried households with children account for 6.7 percent of all homeowners, they account for 11.6 percent of owners living in inadequate housing. Nonmarried households with children, which account for 19.7 percent of renters, represent 26.1 percent of renters in inadequate housing units.

Exhibit 10

Race/Ethnicity of Household Head in Units Classified as Inadequate Under the New Definition

	Owners	s	Renter	s
	In Inadequate Units	All Owners	In Inadequate Units	All Renters
Non-Hispanic White	3,958,482	55,669,648	2,040,241	18,456,583
	76.5%	78.3%	55.0%	54.3%
Non-Hispanic Black	582,333	5,941,607	768,626	7,067,630
	11.3%	8.4%	20.7%	20.8%
Hispanic	424,317	5,952,846	692,640	6,067,802
	8.2%	8.4%	18.7%	17.9%
Other	211,850	3,514,626	208,407	2,382,910
	4.1%	4.9%	5.6%	7.0%
Total	5,176,982	71,078,727	3,709,915	33,974,924

Exhibit 11

Time of Household in Heite	Classified as Inadequate Under the New Definition
Type of Household in Offics	Classified as Inadequate Under the New Definition

	Owners		Renters	
	In Inadequate Units	All Owners	In Inadequate Units	All Renters
Married couple with children	1,433,851	19,536,247	592,948	4,827,748
	27.7%	27.5%	16.0%	14.2%
Other with children	600,449	4,746,176	967,907	6,701,880
	11.6%	6.7%	26.1%	19.7%
65 + householder	851,531	16,514,343	273,081	4,336,733
with no children	16.5%	23.2%	7.4%	12.8%
Other without children	2,291,152	30,281,960	1,875,980	18,108,562
	44.3%	42.6%	50.6%	53.3%
Total	5,176,982	71,078,727	3,709,915	33,974,924

Summary and Conclusion

This article presents baseline hedonic regression models that estimate house values for owner-occupied housing units and rents for rental apartments, building on models developed by NAHB during the past two decades and based on data from the 2009 national AHS. Distinguishing features of these models include extensive use of the allocation flags and other features of the AHS public use file to clean the data before estimation (which includes exploiting the longitudinal nature of the survey and comparing the same unit across years to detect outliers), interaction terms that combine information on the size of the units with the year they were built, and explanatory variables drawn from many sections of the survey, including the section that collects information on neighborhood characteristics.

A relatively large number of neighborhood characteristics have economically and statistically significant effects on the dependent variables, particularly on the value of owner-occupied single-family housing,

This subset of the results, by itself, has a number of potential uses, especially at the local level. For example, homeowners associations or local governments could use the results to estimate how certain public policies (such as providing public transportation, or finding a use for abandoned buildings) are likely to affect home values in particular neighborhoods. Moreover, it is not obvious that these effects could be estimated from any alternative data source that currently exists.

This article combines the regression models with information from the housing quality section and other related sections of the AHS to develop a new summary indicator of physically inadequate housing. This is another line of research that would be difficult or impossible to pursue without the information contained in the AHS. The new inadequacy indicator proposed in this article is based primarily on conditions on the outside of the building (such as missing siding, holes in the roof, and broken windows) for single-family units and conditions that are more readily observed from inside the unit (such as lack of a bathroom or kitchen sink and a household's reliance on heating equipment that poses a risk, or a home with no heating equipment inside the unit) for multifamily

units. This new indictor of inadequacy has a statistically significant and negative effect on house values and rents—in contrast to the traditional summary indicators of inadequacy that are provided in the public use AHS file, which are not significant and often have the wrong (positive) sign.

The new indicator also indentifies a substantially larger number of housing units in the United States as physically inadequate, especially single-family units. The inadequate units are strongly concentrated in the older housing stock and in geographic areas where the housing stock tends to be older, including both central cities and outlying nonmetropolitan areas.

Physical adequacy and affordability are two sides of the same coin, in that affordability may be achieved by neglecting maintenance and repairs, which leads to conditions such broken windows or holes in the roof—or failure to replace or upgrade older units that lack sinks or safe central heating equipment in colder climates—but achieving pure affordability through these means is not in general a desirable outcome.

Very few households identified by the new indicator as living in physically inadequate housing are also suffering from housing costs that are high relative to their incomes. Therefore, the larger number of households living in inadequate units represents primarily a net addition to the estimated number of U.S. households experiencing housing problems that need to be addressed in some fashion. Some households living in inadequate units are owners who appear to be relatively well off, with incomes of at least 120 percent of AMI and housing costs that are less than 30 percent of this income.

From a public policy perspective, the remedy for these cases may primarily be educational. Government organizations could provide homeowners with information about recommended repair and maintenance schedules, or how to cost-effectively upgrade older structures to more current standards. Even net of these relatively well-off cases, however, the estimates of inadequate housing presented in this article indicate that the need for housing assistance in the conventional sense is more widespread than is generally recognized. A disproportionate share of households suffering from inadequate housing are households with children.

The new indicator proposed in this article can also be applied to estimate the number of inadequate vacant housing units. The resulting estimate is more than 1 million inadequate nonseasonal vacant housing units in the United States, with a particularly high rate of inadequacy found among nonseasonal vacant single-family units. An implication of this relatively large number of inadequate vacant units is that the effective inventory of existing single-family units available on the market may be overstated, if it is assumed these units are ready to be sold to prospective full-time occupants without substantial repairs or upgrades.

Authors

Paul Emrath is the vice president for Survey and Housing Policy Research with the National Association of Home Builders.

Heather Taylor is a senior research associate with the National Association of Home Builders.

References

Amemiya, Takeshi. 1980. "Selection of Regressors," International Economic Review 21: 331–354.

Carruthers, John I., David E. Clark, and Robert N. Renner. 2010. The Benefits of Environmental Improvement: Estimates From Space-Time Analysis. Working Paper #REP 10-02. Washington, DC: U.S. Department of Housing and Urban Development. Available at http://www.huduser.org/portal/publications/pdf/jic_dec_rnr_rep_1002.pdf.

Court, Andrew. 1939. "Hedonic Price Indexes With Automotive Examples." In *Dynamics of Automobile Demand*. New York: General Motors Corporation: 99–117.

Crowe, David. 2011. Personal communication. Chief economist, National Association of Home Builders.

Emrath, Paul. 2004. "How Features of a Home Impact Its Price," HousingEconomics.com website: http://www.nahb.org/generic.aspx?sectionID=734&genericContentID=33077&channelID=311 (accessed February 18, 2011).

———. 2002. "Explaining House Prices," Housing Economics January: 9–13.
——. 2001. "Calculating the Value of Apartment Features and Amenities," NAHB Multifamily Market Outlook, October.
——. 1996. "Explaining Differences in Market Rents," Housing Economics July: 9–11.
. 1995. "Features That Influence Home Values," Housing Economics December: 5–8.
——. 1993. "Quality and Price for New and Existing Homes," Housing Economics August: 10–1

Griliches, Zvi. 1961. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Changes." In *The Price Statistics of the Federal Government.* Cambridge, MA: National Bureau of Economic Research: 173–196.

Kiel, Katherine A., and Jeffrey E. Zabel. 1999. "The Accuracy of Owner-Provided House Values: The 1978–1991 American Housing Survey," *Real Estate Economics* 27: 263–298.

McDonald, John F., and Daniel McMillen. 2007. *Urban Economics and Real Estate: Theory and Policy*. Malden, MA: Blackwell Publishing.

National Oceanic and Atmospheric Administration. 2009. "Climatography of the U.S. No. 81 - Supplement # 3 Maps of Annual 1961–1990 Normal Temperature, Precipitation and Degree Days." http://www.ncdc.noaa.gov/img/documentlibrary/clim81supp3/annualheatingDD_hires.jpg (accessed February 18, 2011).

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Market Product Differentiation in Pure Competition," *Journal of Political Economy* 82: 34–55.

U.S. Census Bureau. 2005. Construction Price Indexes. Washington, DC: U.S. Census Bureau, Manufacturing and Construction Division. Available at http://www.census.gov/const/www/priceindexregressionmodel.pdf.

U.S. Consumer Product Safety Commission. 2011. "Reducing Fire Hazards for Portable Electric Heaters." Safety alert. Washington, DC: U.S. Consumer Product Safety Commission. http://www.cpsc.gov/CPSCPUB/PUBS/098.pdf.

———. 2009. "CPSC Warns of Deadly Fire and Carbon Monoxide Hazards With Winter Home Heating." Press release. Washington, DC: U.S. Consumer Product Safety Commission. http://www.cpsc.gov/cpscpub/prerel/prhtml09/09109.html.

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

Waugh, Fredrick W. 1928. "Quality Factors Influencing Vegetable Prices," *Journal of Farm Economics* 10: 185–196.

Weicher, John C. 1979. "Urban Housing Policy." In *Current Issues in Urban Economics*, edited by Peter Mieszkowski and Mahlon Straszheim. Baltimore, MD: Johns Hopkins University Press: 501.

