The Destruction of Housing Capital: A Preliminary Exploration into Demolitions and Disasters

The research presented here was performed under the "AHS Analytical Support" Contract issued by the Office of Policy Development and Research at the U.S. Department of Housing and Urban Development (HUD).

The work was performed by ICF Consulting (http://www.icfconsulting.com) and Econometrica, Inc. (http://www.econometricainc.com). Gregory J. Watson of ICF Consulting and Frederick J. Eggers of Econometrica were the authors.

The authors would like to thank David A. Vandenbroucke and Ronald J. Sepanik of HUD for their oversight and comments, as well as the Census Bureau for clarification on certain data issues.

November 30, 2003

Background

Purpose

With the exception of fine art and jewelry, housing is the most durable of all consumer expenditures. According to the American Housing Survey (AHS), more than ten million units constructed prior to 1920 still survive in the United States. European cities and the surrounding countryside provide countless examples of structures built in the 17th, 18th, and 19th centuries that continue to furnish safe and comfortable habitats. Yet numerous housing units – both old and new -- are torn down or otherwise destroyed every year in this country. Between 1999 and 2001, 1.5 million housing units[1] disappeared permanently. Fires and natural disasters account for some of these losses but owners voluntarily demolished many other units.

We know little about this phenomenon – for example, how much capital is lost annually, which units are most susceptible to being loss, and what motivates owners to destroy housing capital? A better understanding of these issues could give us useful insights into important social questions. The price and tenure status of the units being lost affects the availability of affordability housing. The costs of regional economic dislocation include the impacts of declining population on the housing stock. Neighborhood transformations involve both people changes and structural changes. For these reasons, the Department of Housing and Urban Development (HUD) commissioned this exploratory study of housing loss.

This paper has two modest goals. First, we will examine how to use the AHS to study these questions. The AHS has features that make it well suited to an analysis of housing loss, particularly its large sample size, extensive information on the physical characteristics of the units, good neighborhood data, and the ability to track the same unit over time. However, researchers must first deal with a number of conceptual and data problems. Second, we will use the AHS to analyze, in a preliminary fashion, what units are destroyed and why.

Previous Research

Despite the apparent advantages of the AHS, researchers have devoted little attention specifically to the loss of residential capital. Most attention has been focused on two related areas – one broader and one narrower than this topic. Taking a broader perspective, researchers have looked at the various ways in which housing units appear in or disappear from the stock. Such analyses, which have been called *Components of Inventory Change* (CINCH), have relevancy mainly because they show that permanent losses are part of a wider set of changes that occur within the housing stock. Demolition is only one of the ways in which housing units can "leave" the stock. Housing units can be moved, merged with other residential units, or converted to non-residential use as offices, storage facilities, etc. The CINCH literature contributes to our efforts by

[1] Number based on calculations from the draft 1999-2001 Components of Inventory Change Report currently being produced by ICF Consulting.

suggesting paths that housing units might follow before being demolished. Unfortunately, the CINCH literature does not tell us why units leave the stock, but merely the characteristics of the units leaving the stock. The CINCH does also provide information about changes to the housing units, but that is not of concern for this research.

Many researchers have studied the phenomenon of abandonment, that is, why an owner would let a local government take possession of residential property in lieu of paying accrued real estate taxes. While some of this literature relates to specific locals, such as New York City, the literature contributes to our understanding of permanent loss by focusing on causation. An insightful example of the recent abandonment literature is *An Economic Analysis of Housing Abandonment* by Scafidi, Schill, Wachter and Culhane.[2] These authors model the abandonment decision as the exercise by an investor of a put option. Using data from New York City, they find that the influence of a put variable (a dummy variable equal to one if ratio of outstanding city liens to market value is greater than one) far exceeds that of any other variable in determining whether an owner will abandon a property. Other significant variables in their analysis related to the physical adequacy of the building and neighborhood conditions. Unfortunately, the abandonment literature deals with only one dimension of the permanent loss problem. It does not throw any light on the removal of existing structures to prepare for upgraded use of the land.

We did discover several studies that were directly relevant to our research objectives. In a 1981 article, Gleeson applied actuarial techniques to estimate mortality curves for housing.[3] Gleeson used AHS data to fix data on housing loss to Gompertz cuves. His analysis emphasized the non-linear relationship between age of a cohort of housing units and the rate at which units exit the cohort. Gleeson estimates suggest that the rate of loss increases with age.

The notion that, as housing ages, it will decline in value and ultimate drop out of the stock has a long history in the concept of filtration. Margolis relates filtering in the housing stock to underlying depreciation and replacement capital costs curves.[4] Malpezzi, Ozanne, and Thibodeau use AHS data to estimate depreciation rates for housing.[5] They found that “rents decrease at a nearly constant rate, values at a declining rate” but that some metropolitan areas “exhibit significant deviations from the average depreciation patterns.”

---

[2] Benjamin P. Scafidi, Michael H. Schill, Susan M. Wachter, and Dennis P. Culhane, “An Economic Analysis of Housing Abandonment,” in *Journal of Housing Economics*, Volume 7, Number 4, December 1998.
[3] Michael E. Gleeson, “Estimating Housing Mortality,” in *Journal of the American Planning Association*, volume 47, pp. 185-194, 1981.
[4] Stephen E. Margolis, “Depreciation and Maintenance of Houses,” in *Land Economics*, Volume 57, Number 1, February 1981, pp. 91-105.
[5] Stpehen Malpezzi, Larry Ozanne, and Thomas B. Thibodeau, “Microeconomic Estimates of Housing Depreciation, “ in *Land Economics*, Volume 63, Number 4, November 1987.

Baer examined permanent losses in a CINCH context.[6] Baer emphasized concentrated on demolitions and disasters because “a loss of [housing] space” occurs only with these outcomes. He also used the CINCH framework to consider the various paths by which unit could move to permanent removal. In particular, Baer talked about the recycling of units, that is, units that temporarily leave the stock only to return at a later date. Finally, he called attention to the importance of tenure and vintage, that is, he found that he rate of removal increases with age of the stock and units are more likely to be rental at the end of their existence than at the beginning.

Carliner has looked at the permanent losses in several different contexts.[7] In his 1991 paper, he distinguished between “depletion demolitions” and “displacement demolitions.” A depletion demolition results if “market or implicit rent falls below the cost of operation and maintenance.” Normal deterioration, physical damage from a disaster, or the failure of a major system could push operating costs higher than market rents. Alternatively high vacancy rates could push market rents below operating costs. In either case, the decision to demolish the structure would be, in Carliner’s view, independent of possible reuse of the land in a depletion demolition. Abandonment is consistent with depletion demolitions. In a displacement demolition, the value of the property with the structure falls below the value of the land alone. Carliner recognizes that these two categories of demolitions are not mutually exclusive. An investor who anticipates removing a structure to replace it with a more valuable structure may allow the original structure to deteriorate rather than incur maintenance costs.

Carliner used both CINCH data from the AHS and demolition permits to study demolition behavior. Limitations in the data restrict the range of conclusions that he can draw. The CINCH data indicated that “renter-occupied or vacant units were far more likely to be demolished than owner-occupied units, but among renter-occupied and vacant units single family detached units were more likely to be demolished than multifamily units.” Carliner observes a positive relationship between permits for retail buildings and permits for demolitions. He interprets this relationship as consistent with displacement demolitions.[8]

Mitchum employed another technique for studying removals from the stock.[9] He used decennial census data to measure the increase in the housing stock between 1990 and 2000. He subtracts this measure from Census Bureau data on housing completions to estimate “net removals”, which includes both physical removals and other CINCH types additions and subtractions from the housing stock. Over the decade of the 1990s, the housing stock increased by 13.6 million units. Since new construction provided 16.2 million new units, net removals had to be 2.6 million. Mitchum found that manufactured housing units accounted for 57 percent of all net removals. Mitchum estimated net

---

[6] William C. Baer , “Aging of the Housing Stock and Components of Inventory Change” in *Perspectives on Space, Time, and Housing Stock*, edited by Dowell Myers, pp. 249-273.
[7] See for example: Michael S. Carliner, “Housing Demolitions: An Analysis of Permit Data,” paper prepared for the AREUEA Midyear Meeting, May 28, 1991 and “Replacement Demand for Housing” in *Housing Economics*, December 1996.
[8] The quoted results are from Carliner’s AREUEA paper.
[9] Drew A. Mitchum, “Housing Removal Rates” in *Housing Economics*, February 2003.

removals for every State and attempted to explain variation in rates of net removal. Interestingly he was generally unsuccessful in explaining. High growth rates, age of the stock, share of the stock in manufacturing housing, and proportion of the stock without plumbing or kitchen facilities failed to correlate significantly with net removal rates. There was some correlation with the percentage increase in income and the percentage increase in the number of housing units. Mitchum speculates that a strong relationship between some of these variables and net removals exists only at lower levels of geography.

Research Issues

In this paper we will focus on permanent losses, specifically losses due to demolitions or disasters.[10] In this respect, we follow both Baer and Carliner in emphasizing the permanent loss of housing. Baer referred to the "loss of space;" we prefer to think of it being a loss of housing capital. From an economic perspective, the loss of capital appears to be the most policy relevant concern.

Concentrating on permanent losses creates some problems. Disaster and demolition losses are rare events. In any given AHS, usually less than ½ of one percent of the units have been destroyed by fires and natural disasters or demolished by owners since the last AHS. This problem is further complicated by the need to analyze disasters and demolitions separately to account for differences in the types of units affected and the underlying causes.

One could increase sample size by combining disasters and demolitions from several AHS's. Instead, we have chosen to analyze each AHS separately. This decision allows us to compare the characteristics of units loss through disasters and demolitions with the contemporaneous characteristics of units that have survived since the last AHS. If we simultaneously studied units that disappeared between 1985 and 1987 and units that disappeared between 1999 and 2001, we would not know how to choose a control group. We will need a control group to tell us whether a particular characteristic is over-represented among permanent losses.

Missing variables adds to the problems of relatively small samples. We have found that units that are loss to demolitions in one AHS often lack any data on the unit or household from the preceding AHS. This lack of information is as a result of the data design of the AHS. The absence of "baseline" data results from the status of the unit in the previous AHS. If the unit were vacant, uninhabitable, or used in a non-residential manner before being demolished, then the previous AHS may contain no information on the unit. We solved this problem by taking data from the most recent AHS for which data exists.[11]

---

[10] In the AHS, these units are coded as "30" under the variable that explains why there was no interview (NOINT). In AHS parlance, these are class of type C non-interviews. Other type C non-interviews include house or mobile home moved, unit eliminated in structural conversion, and permit abandoned.

[11] This approach was used for missing information on selected variables even when information on other variables was available in the previous AHS. The limit of how far back in time could be reached was the

The choice of variables also determines effective sample size. For example, certain variables have valid values only for a subsample of units, such as owners or renters. Using an owner-only variable would in effect remove all renter occupied properties.

We encountered some data problems. The AHS files contain all the units surveyed in a particular years. In some years, the Census Bureau includes units that are not surveyed in every year. The Census Bureau included a supplement rural sample in the 1987 and 1991 AHS's and a supplemental sample of units in New York, Los Angeles, Chicago, Philadelphia, Detroit, and Northern New Jersey in 1995 and 1999 AHS's. In a given survey, it is difficult to distinguish the supplemental units from the regular units. The approach we use is to count how many AHS's each unit appeared in. We eliminated units that appeared only in 1987 and 1991 and units that appeared only in 1995 and 1999.

The data used was from the National sample of the AHS. The units present in this sample had to be present in the data for the current year and all previous years, and <u>not</u> be a Type C – permanent loss in earlier years. The counts of units used in the analysis can be found in Appendix A. Metropolitan sample data was not used because of the changes in the metropolitan sample and the length of time between panels of the metropolitan AHS.

Reasons for Demolitions and Disasters

Disasters

We start by assuming that disaster losses are random. For this reason, we analyze demolitions separately from disasters. Disasters, such as fires and tornadoes, occur randomly but disaster losses may not be entirely random. If a disaster completely destroys a unit, then the loss is random. If, however, the disaster only damages a unit, then owner has to decide whether to repair the unit, tear the unit down and replace it with a new unit, or tear the unit down and do not replace the unit. Rational owners will base this decision on the extent of the damage and the value of the property. If the property has little value, then the owner may choose not to repair it.[12] For this reason, there may be significant correlations between a unit being disaster loss and its age, structure type, whether it was uninhabitable in the survey before the disaster loss, its physical condition in the survey year before the disaster loss, whether it was vacant in the survey year before disaster loss, tenure in the survey year before the disaster loss, and the trend in rent or value prior to disaster loss. Although this appears to be a reasonable conjecture, Baer found no clear-cut pattern between tenure and vintage for disaster losses.[13]

start of this sample in 1985. As a result, it is possible that certain information was from an AHS many years previously.

[12] In the AHS, a unit is considered to be a Type C demolition or disaster loss even if the owner replaces the unit with a new unit.

[13] Baer, page 265.

Demolitions

We explain demolition losses in terms of two different phenomena.

- Loss of economic value
- Land becoming more valuable than land plus existing structure

This is the same approach adopted by Carliner who distinguishes between depletion demolitions and displacement demolitions.[14]

Loss of Economic Value

Units can lose economic value for a variety of reasons. We believe that two possibilities encompass most of the reasons.

- **Physical Depreciation and Obsolescence**

Like any capital asset, a house will depreciate with age. Roofs need to be replaced periodically. Furnaces, hot water heaters, air conditioners also need to be replaced. Unless properly maintained, exterior surfaces will deteriorate and fail to provide adequate protection from the elements.

Obsolescence is also a problem. Older housing units frequently have inadequate wiring for the numerous electrical appliances used by most households. The presence of steam or hot water heating means that one needs to add air vents in order to retrofit air conditioning. Changing tastes affect the attractiveness of older units. For example, newly constructed homes typically have more and bigger bathrooms than older homes.

With proper maintenance a housing unit can last virtually forever. There are many 19th century and even a good number of 18th century homes still in service in this country. However, maintenance is expensive and may be more expensive for older homes. Moreover, obsolescence reduces the incentive to maintain. Overcoming obsolescence can often be very expensive.

- **Declining demand for residential structures**

This explanation focuses on the effects of declining demand caused by factors, such as, loss of population, shift of population and income from central cities to suburbs, growth in income, and general neighborhood decline.[15] Shifts in population structure can reduce demand for certain types of structures even when population is not declining. For example, an FHA-insured apartment complex in Brooklyn MN that was built in the early 1970's for the baby boomers just forming their own households. The apartment project contained mostly one-bedroom units and, for this reason, had high vacancy rates in the early 1990's when the baby boomers had moved beyond the new household status and

[14] Michael S. Carliner, "Housing Demolitions: An Analysis of Permit Data", paper prepared for the AREUEA midyear meetings, May 28, 1991, p. 1.

there were fewer new households. Declining demand for a property will lead to undermaintenance and ultimately to abandonment and demolition or to conversion to structures that are in higher demand.

Higher and better use

Increasing demand for land can also lead to demolition followed by replacement with higher density or more expensive properties, e.g., the tearing down of single-family units to build townhouses or high-rise apartments or the tearing down of old Cape Cod style homes to build "McMansions" or "starter castles." Increased demand for land can also result in the replacement of residential properties with commercial properties. Anticipating the replacement of a unit with more or better units, an owner might decide to undermaintain the existing unit.

Analysis

There are at least three different questions we could attempt to answer:

- What are the characteristics of the units that become permanent losses?
- What path do units follow in becoming permanent losses?
- What factors explain why a unit becomes a permanent loss?

The previous section posed different explanations for permanent losses from demolitions and disasters. One could use this discussion to derive conjectures about characteristics, path, and causes. For example, declining population would *cause* demolitions due to a lack of demand. In such circumstances, owners may cut variable costs (maintenance) and default on outstanding mortgages before abandoning the property. Undermaintenance, worsening physical condition, mortgage default, tax arrears, and abandonment may be steps along the *path* to abandonment. Value would be an important *characteristic* distinguishing permanent loss from lack of demand. Low-valued units would probably be the first to drop out of the stock because they have little room in which to "filter" down.

In this exploratory analysis, we test to see which of the approaches throws the most light on demolitions. The different approaches overlap in terms of the variables that they suggest.

The focus on characteristics suggests the following variables:

- **Region:** Slow growing or declining regions have lower demand for housing, and the age and quality of the housing stock may vary by region.

---

[15] Goodman notes that there has been "considerable substitution of metropolitan for central city housing in all three decades [1970 to 2000]." Allen C. Goodman, "Central Cities and Housing Supply", paper presented at the December 2002 meetings of the Allied Social Sciences Association.

We included three dummy variables for the census regions: *Midwest*, *South*, and *West*. *Northeast* was the excluded region.

- **Metropolitan status:** Declining population in rural areas should lead to a higher level of demolitions, and the rural housing stock may be older and less well maintained than the urban stock. In central cities, growth and decline may both produce demolitions. Gentrification or commercialization can lead to the replacement of existing restructures with new residential structures or commercial structures. Loss of central city population can result in the demolition of excess housing. The central city housing stock may older and less well maintained than the suburban stock
  We include two dummy variables for metropolitan status, *Non-metro* and *Central City*; *Suburban* is the omitted variable.
- **Structure type:** The costs of demolishing single family detached units should be lower than the costs of demolishing larger structures. Mobile homes are easily removed and may deteriorate faster than other structure types. Mitchum reports that 57 percent of the "removals" between 1990 and 2000 were manufactured homes.
  We include three structure type variables: *single-family attached*, *multifamily*, and *mobile home*. *Single-family detached* is the omitted variable.
- **Structure size:** The opportunity cost of demolishing a small unit is lower and smaller units may be less attractive in today's market.
  We experimented with several variables related to the size of a unit. The most successful variable was number of rooms, *ROOMS*. Less successful variables were *Unit Square Foot* and *Number of Bedrooms*.
- **Other factors:** It is possible that the demand for a unit, and therefore its value, is related to the configuration of a unit. We experimented with variables that might identify units that are less attractive in today's markets. These variables include lot size in square feet,[16] the ratio of unit size in square feet to lot size in square feet, and the number of baths and half baths. These variables were not significant in the test survey years. Note that we did not include variables that would only theoretically apply to a subsample of the universe and so bias the results. This is discussed in Appendix B: Data Notes.

We could have included age and condition of the unit under this characteristics category but we believe that they are more appropriately considered under causes and path respectively.

Analysis of path would lead us to consider the following variables:

[16] A large lot size might lead to demolition of a single unit followed by the construction of several units on the same land.

- **Tenure:** Single family units may move from owner-occupancy status to rental status prior to demolition. Baer found that rental units are more likely to be removed.
  We use a dummy variable, *Renter*; the omitted variable is *Owner-occupied*
- **Previous type B status:** A unit may reach one of the following stages prior to demolition: converted to non-residential use, converted to an institutional unit, unoccupied because the interior is exposed to the elements, or unoccupied because occupancy is prohibited.
  We added a dummy variable, *type_b*. The logic of a unit being a Type B being a predictor is that the unit has already been temporarily removed from the housing stock, and then the final loss is merely a continuation of a trend away from serving as housing. This variable is true if the unit is a Type B in the immediately prior AHS panel.
- **Vacancy status:** Being vacant is another possible stage prior to demolition.
  We include the dummy variable *vacant*. This variable is true if the unit is vacant in the immediately prior AHS panel.
- **Physical condition and changes in physical condition:** The likelihood of a unit having moderate or severe physical problems should be higher for units prior to demolition. In declining areas with excess housing, one would expect the lowest quality units to leave the stock first. Also one would expect owners to undermaintain units that are expected to be demolished in the future.
  We initially used two variables to measure the physical condition of a structure based on the AHS's definitions of moderate physical problems and severe physical problems. We found that having moderate physical problems was significantly related to being demolished but having severe physical problems was not statistically significant. Not wanting to use information in AHS measure, we combined these two variables in a single dummy variable, *inadeq*, which equals one if the unit has either moderate or severe physical problems. We also include the respondent's assessment of the unit, *HOWH*.
- **Maintenance effort:** In anticipation of demolition, owners should lower their maintenance efforts.
  We did not include any variables related to maintenance effort in the analysis because of the difficulty of finding comparable

In seeking causes, one might consider the following variables:

- **Age:** Older units may be most costly to maintain and may suffer from obsolescence. Baer, Carliner (1996), and Gleeson all find a relationship between age and permanent loss.
  We experimented with linear and non-linear measure of age. We found the most successful specification involved the following dummy variables: *Builtpre20*, *Built20_39*, *Built40_59*, and *Built60_80*. The omitted variable was *Builtpost80*.

The problem with a linear variable is that it would make the assumption of a single directional prediction – i.e. older is more likely to be removed than newer. We did not find that plausible logic necessarily because an older unit would have had to have "survived" to this point in time, and that means that poor quality units for example may have been weeded out. In addition, older homes may have characteristics which make them more attractive to buyers. As a result, using age dummies for the housing units make better theoretical sense.

- **Value and changes in value:** Units may decline in value or rent prior to demolition. The decline in values could result from a general market decline, worsening neighborhood conditions, or failure to maintain the unit adequately. Alternatively some owners may report rising values based on land prices alone.

  We experimented with several forms of the value variables, owner's estimate of value as a continuous variable, owner's estimate of value as a discrete variable, and value per square foot. In general, the results were disappointing. We dropped this variable from the analysis because of the lack of success and because including it required us to limit the analysis to owner-occupied units only. We also experimented with change in value between the most recent two AHS for which value information is available. This variable was somewhat successful in analysis involving disaster and demolitions losses combined. See the discussion in Appendix B about variables with limited universes.
- **Neighborhood quality and changes in neighborhood quality:** Changes in neighborhood quality or character could induce owners to demolish structures. The incentive could be negative (lack of value) or positive (higher or better use). These changes would probably be reflected in property values.

  We experimented with a limited number of variables related to neighborhood condition. Having abandoned structures in the neighborhood, *abandon*, having units with bars on the windows in the neighborhood, *bars*, having schools within one mile, and having shopping within one mile, were included in the analysis. However, we note that there was a change in collection in 1997 where instead of the interviewer collecting the data, it was asked of the respondent. This could lead to some inconsistency of results.

  We also experimented with the respondent's assessment of the neighborhood but it was generally insignificant.

Results

Table 1 presents the results for our preferred specification for all survey years. There are no results for 1985 because many of the variables for the demolished units are based on the preceding survey. (The 1985 AHS was the first survey to use the redesigned sample.)

The parameter estimates for some variables vary significantly over the years. This generally relates to the low statistical significance of these variables. But it also reflects the tenuousness of a fit in which fewer than one-half of one percent of the observations fall into the loss category. As noted in the table, the SAS output flagged the 1987 and 1997 runs with warnings that quasi-complete separation of data points was detected. Quasi-complete separation of the data points means that SAS was not able to determine a solution to the logistic regression at the appropriate level of precision. This could be because certain variables lack independence from one another.

Appendix A includes information on the number of observations used in each year as well as the number of losses.

Among the three approaches, analyzing the path that units might take to demolition seems the most promising. Units appear to pass through stages on the way to demolition. *Renter* was uniformly positive and highly significant in all years. *Type_b* and *vacant* were also uniformly positive and, except for 1987, were highly significant. *Inadequate* was uniformly positive and, except for 2001, was highly significant. The respondents assessment of his or her unit (*HOWH*) was uniformly negative and, except for 1993 and 1997, was highly significant.

We note about the Type B housing units, that some of them are vacant for many years prior to being returned to housing, becoming complete losses, or staying as Type B housing units. For our purposes, they are not complete losses from the stock yet. Future research should potentially address predicting the characteristics of a Type B unit, and how long it is likely to remain a Type B unit. However, that research is outside of the scope of what is presented here.

Although we tested fewer variables that might measure forces that lead to demolitions, the empirical results suggest some causal patterns. The age of a unit has an important influence on the probability of demolition and the effect appears to be non-linear. The three oldest age groups – *builtpre20*, *built20_39*, and *built40-59* – are uniformly positive and except for 1987 are significant in all years. Except for *built40_59* in 1993 (which is still significant at the 10% level), these variables are highly significant in all surveys after 1987. The coefficient of *builtpre20* is generally the largest among these variables and often the coefficients among the variables decline as the period of construction moves closer to the present. Among the neighborhood variables, *abandon* was the most successful. Having a near-by abandoned unitgenerally raised the probability of demolition. This variable was positive and significant in all years except 1991 and 1997 and was very significant in five of the eight years. Having a near-by unit with bars on the windows was significant in only 2001. This variable was generally negative, including in 2001.[17] Being conveniently located for shopping was significant and negative in two years. The accessibility to schools was significant and negative in three years.

[17] A negative sign would consistent if one interpreted bars on windows as a sign of gentrification and one assumed that gentrification is normally accomplished through substantial rehabilitation of units rather than the demolition and replacement of units. However, the generally low significance of this variable suggests not making too much of the sign.

Interestingly, in 1997 when convergence of the model was questionable, the magnitudes of these variables are markedly inconsistent with the results in other years.

After taking into account possible causes of demolitions and possible paths by which a unit might go to demolition, the characteristics variables were only somewhat helpful in explaining demolitions. The coefficients of the regional variables were significant in 14 out of 24 cases. When significant, they were always positive. This indicates that, relative to the northeast, demolitions are more likely in the other three regions. However, the signs of these variables were frequently negative when they were insignificant. The coefficients of the metropolitan status variables are significant in only 3 of 16 cases. If there is any legitimate interpretation of these results, it is that demolitions are more likely in non-metropolitan areas. The coefficients of the type of structure variables are statistically significant in only 5 of 24 cases, and was never significant for Mobile Homes. The signs of the variables indicate that, relative to single-family detached units, units in multifamily structures and single-family attached units are less likely to be demolished. Results for the mobile home variable were disappointing. The coefficients were never significant and the sign was frequently negative, contrary to expectations. Finally, the size of structure variable (*ROOMS*) had significant coefficients in only one year. Generally, but not always, this variable had the expected negative signs.

Table 1: Parameter Estimates and Log Likelihood Measures for Logit Regression of Demolitions on Variables Related to Unit Characteristics, the Path to Demolition, and Causes of Demolition.

		Characteristic Variables									*Path Variables*					*Cause Variables*								
Survey Year	*Intercept*	*Midwest*	*South*	*West*	*Central City*	*Non-Metro*	*Sng-Fam Attched*	*Multifamily*	*Mobile Home*	*Rooms*	*Renter*	*Type_b*	*Vacant*	*Inadequate*	*HOWH*	*Builtpre20*	*Built20_39*	*Built40_59*	*Built60_80*	*Abandon*	*Bars*	*Schools*	*Shopping*	*Ln Likelihood*
1987	-4.099	-0.045	0.529	0.762	-0.289	0.422	-1.305	-1.233	0.281	-0.320	1.575	-14.004	-14.077	0.725	-0.208	0.859	0.448	0.478	-0.391	0.680	-0.713	-0.644	-0.107	-520*
1989	-7.875	1.300	1.606	1.430	0.184	0.379	-0.426	-0.542	0.511	-0.004	0.672	2.521	1.170	0.532	-0.238	2.296	2.175	1.910	1.162	0.805	-0.434	-0.217	0.265	-735
1991	-8.200	1.086	1.214	1.238	0.316	0.721	-0.594	-0.440	-0.738	-0.102	0.695	2.929	1.277	0.897	-0.102	2.276	2.287	1.827	1.122	-0.090	-0.226	-0.560	0.149	-714
1993	-7.533	0.626	0.500	0.605	-0.494	-0.031	-0.936	-0.611	-0.210	-0.025	0.561	3.119	1.449	1.031	-0.063	1.675	1.617	1.282	0.987	0.733	-0.064	0.116	-0.373	-609
1995	-6.753	0.516	-0.328	-0.254	0.122	0.887	-0.198	-0.418	0.490	-0.072	0.806	2.832	1.170	1.064	-0.124	1.498	1.187	1.195	0.399	0.902	0.320	0.162	-0.085	-781
1997	-7.344	0.704	0.986	0.253	0.086	-0.382	0.130	-0.161	-0.809	0.049	0.620	2.940	1.029	1.123	-0.020	1.375	1.205	1.106	0.533	0.315	-0.295	-13.272	-14.858	-770*
1999	-6.947	1.115	0.809	0.451	0.281	0.175	-0.644	-0.236	0.310	0.008	0.940	2.532	0.915	0.510	-0.108	1.501	1.284	1.289	0.689	0.903	0.314	-1.351	-0.737	-1054
2001	-6.474	-0.223	0.372	-0.201	-0.065	0.175	0.257	-0.048	0.432	-0.049	0.613	2.842	1.682	0.256	-0.105	1.677	1.606	1.608	0.736	0.907	-0.439	-0.094	-0.420	-935

Parameter estimates highlighted in red are significant at the 10 percent level or better.
Parameter estimates highlighted in yellow are significant at the 5 percent level or better.
Parameter estimates highlighted in green are significant at the 1 percent level or better.
*The specification had difficult converging in 1987 and 1997.

Appendix A: Number of observations

The following table provides a count of:

1) Total number of observations in the data for that year
2) Number of observations used in the model, after missing values removed
3) Number of observations used in the model that were losses
4) Number of observations that were not losses

Year	Total number of observations in Data	# of observations after missing values removed	# of losses	# of not losses
1987	45,036	40,400	89	40,311
1989	44,548	41,069	133	40,936
1991	44,107	41,284	127	41,157
1993	43,626	41,208	111	41,097
1995	43,256	41,087	155	40,932
1997	42,322	40,511	180	40,331
1999	41,895	40,333	220	40,113
2001	41,422	40,049	181	39,868

The decline in number of observations is expected, and it is because of attrition from the sample.

The "improvement" in the ratio of number of observations used in the regressions compared to the total number of observations may be due to the improvements as a part of the changes starting with the 1997 data collection.

Appendix B: Data Notes

Variables with limited universes

The AHS includes variables that do not apply to all survey respondents. For example, renters are not asked about their mortgage payments because it does not apply to them. Variables that do not apply to all respondents are referred to as having "limited universes."

Inclusion of these limited universe variables into a logistic regression equation would lead to erroneous results because the SAS program we used would drop from the regression equation all observations to which these variables do not apply. The resulting predictors would only be valid for a subset of the universe. As a result, we did not include the following variables:

- Lot size –Not available for multifamily properties.
- Mortgage payment – Only applies to owner-occupied with mortgages.
- Rent – Only applies to rental properties.

Limitations on Data for Type C observations

Once a unit is lost from the AHS, there is a significant loss of data on that observation because certain information is no longer relevant. For example, if a unit is removed, it no longer has a "square footage". To deal with this problem, we took advantage of the time series nature of the AHS and used the last valid data. The next section describes our approach.

Filling in of missing data -- "Reach back"

Where data were missing for a housing unit – whether they were lost because the unit had been removed, the respondent refused to answer, or for other reasons, we went back and used the last valid data for that particular housing unit. We thought using dated information was more appropriate than randomly assigning values, particularly since many of the variables measure conditions that are unlikely to change.

We did not limit how far we "reached back" to gain valid information. There were cases where no valid information was available all the way back to 1985; in those cases the units were not included in the analysis.

Appendix A provides information on the sample sizes and the number of observations actually used in the regressions. The gap between the two is the number of observations where it was not possible to get appropriate information.

The variable that tended to have the most missing values, even after "reaching back" was the HOWH – Resident's satisfaction with unit, variable. This variable was the one that most frequently had missing data. As expected with that variable, the more satisfied a resident was with the unit, the less likely for the unit to be removed from the stock.

Reason for no regressions for 1985

As discussed in this section, the AHS records for units that have been removed from the housing stock are missing data for certain key variables. The technique to find substitutes for these missing data was to "reach back" to previous data. However, the Census Bureau drew a new AHS sample in 1985 and there are no previous surveys to reach back to. As a result, it would not be possible to construct a logistic regression equation that would be valid in 1985, consistent with later years, and provide useful results.

Shifts in 1997

As documented in other publications, such as "Documentation of Changes in the 1997 AHS", the Census Bureau changed how they collect information from housing units surveyed in the AHS. The data collection technique moved from a paper based survey to Computer Assisted Personal Interviewing (CAPI). As a consequence, certain data which had been collected through interviewer observation are now being collected through questions posed to respondents. This shift could lead to differences in answers to some questions, especially questions about the neighborhood , that are unrelated to changes in underlying conditions. In particular, there may be a shift in answers between the 1995 and 1997 surveys.

The regression results for 1997 have parameter estimates several orders of magnitude greater than every other year for the schools and shopping variables. The logistic regressions experienced difficulties in converging in 1997. An examination of the SAS logs of how the regression converged to a solution showed that those two variables were the reason there was not a convergence. The failure to converge could be related to a coincidence in the data, the 1997 change in data collection, or something completely different.

Sample limitations

We analyzed each AHS survey separately. The dependent variable equaled 1 if the unit was classified as demolished and equaled 0 otherwise. The sample consisted of all units present in the 1985 AHS and present in every wave of the AHS up to and including the year the unit was first classified as being demolished. After the unit was classified as a demolition, the unit was dropped from the sample used in the analysis of subsequent AHS surveys. If a unit became a permanent loss for reasons other than demolition, such as a disaster or being moved, the unit was dropped from the sample immediately.

We excluded from the analysis the rural oversample cases present in the 1987 and 1991 AHS's as well as the metropolitan oversample present in the 1995 and 1999 AHS's because we could not determine which period the loss occurred in. These cases are relatively few compared to the overall sample.

Units added to the AHS, mainly through new construction, were also not included in the analysis. Basically, we set a starting point in time and focused just on exiting from the sample. In addition, we thought the likelihood of a new unit both entering and permanently exiting the stock through voluntary demolition was relatively small.

Weighting

We made the decision not to weight the data for both theoretical and practical reasons. The first fundamental obstacle for weighting was the lack of an adjusted weight for observations that are permanent losses from the housing stock. There are techniques, such as applied in the Components of Inventory Change (CINCH) reports, to develop weights, but those techniques seemed excessive for three reasons. First, there were relatively few losses; second, we were only concerned with demolitions and not other changes to the housing stock; and third, we were also not comparing across years, but just inside of one year.

We did explore using the pure weight (PWT) variable which is present on every observation. However, the PWT for a unit varies across years if there is an oversample in one or more of the years. In addition, there are a few cases of changes in the pure weight for unknown reasons.

From the perspective of the logistic regression, use of the pure weight without many other adjustments also dramatically increased the apparent significance of results. This increase was an artifact and made it impossible to determine the real significance of variables.

For these reasons, the regressions all used unweighted data.

Other refinements or variables considered and rejected

In choosing independent variables, we considered what was theoretically valid and what was practically possible given the structure, characteristics, and quality of the data. This led us to explore many avenues that were not included in the final results. A brief discussion of some of these approaches follows:

Lagging variables to capture trends

We explored lagging variables to capture trends over time, to see if they were appropriate and useful for the modeling. For example, declining property value may indicate that the economic value of a property is being adversely affected by neighborhood effects or declining population. Similarly, trends in housing quality may also indicate a unit that is being run-down.

However, we had problems integrating trend variables with our "reach back" to fill missing information, and these variables exacerbated the limited universe problem. In addition, there was excessive sensitivity in some of the variables which could be attributed to allocations. Therefore, we did not include lagging of variables in the model.

Occupant characteristics

We did explore including occupant characteristics, such as age or race and ethnicity in the model. In the cases of demolished units, we used the characteristics of the last occupant. However, given the general lack of significance of these variables and the unclear theoretical underpinning, we decided not to include characteristics of the occupants.[18]

Neighborhood opinions on specific characteristics

The AHS contains a number of questions about the respondent's perspectives on certain elements of their neighborhood, such as crime, traffic, noise, etc. We explored using these variables but decided not to after examining the variables present through 1995 and in 1997 and later. The wording and approach to asking these questions changed significantly in the 1997 redesign, and the different approaches that we considered to conform the variables over time all had drawbacks. So as a proxy for these elements of neighborhood condition, we tried the single variable on the resident's opinion of their neighborhoods. We discovered that this variable was not significant and so it wasn't included. We were able to include variables dealing with having abandoned building near-by, having building with bars on the windows near-by, being near school, and being near shopping.

Decision of what variables to include

In the final analysis, we included variables based on what would make logical sense. We explored using various "data mining" techniques, such as Stepwise but discovered these techniques are very sensitive and do not always including the minimum appropriate variables. As a result, we returned to using theory, as discussed in the text, to choose variables.

[18] In the few experiments we conducted with these variables, the Black race variable was occasionally significant with a negative sign.