Federal Reserve Bank of New York
Staff Reports

Negative Equity and Housing Investment

Andrew Haughwout
Sarah Sutherland
Joseph Tracy

Staff Report No. 636
September 2013

Negative Equity and Housing Investment
Andrew Haughwout, Sarah Sutherland, and Joseph Tracy
Federal Reserve Bank of New York Staff Reports, no. 636
September 2013
JEL classification: R31, G21, D13

Abstract

Housing is a depreciating asset. The rate of depreciation depends on the degree to which households engage in housing investments. Housing investment expenditures economy-wide are sizable, averaging 45 percent of the value of new home construction over the past twenty years. The housing bust and recession coincided with a significant decline in housing investment. Using Consumer Expenditure Survey data from 2007 to 2012, we find that negative equity households reduce their housing investments by roughly 75 percent. The large increase in negative equity due to declining housing prices during the housing bust resulted in a cumulative decline of housing investment expenditures from 2006 to 2010 of $51.2 billion.

Key words: housing, negative equity

Haughwout, Sutherland, Tracy: Federal Reserve Bank of New York (e-mails: andrew.haughwout@ny.frb.org, sarah.sutherland@ny.frb.org, joseph.tracy@ny.frb.org). The authors thank Ravi Bhalla, Rachel Keller, and Ging Cee Ng for excellent research assistance and Joseph Gyourko for comments.

Housing is a long-lived asset that provides an ongoing flow of housing consumption. Absent active maintenance efforts by the owners, this flow of housing consumption will depreciate with time. In aggregate, these housing investments are sizeable averaging 45 percent of the magnitude of new residential construction between 1993 and 2013. Figure 1 shows the time series on aggregate residential investment expenditures.[1] These expenditures peaked at $145.6 billion in 2006 and then with the onset of the housing bust declined $33.6 billion by 2010. What explains this sharp decline in residential improvement expenditures? In particular, how important was the decline in incomes from the recession as compared to the decline in house prices from the housing bust?

An important aspect of housing investments is that they are tied to the house. Households in negative equity may behave more like renters than owners since any increases in the house value resulting from housing investments may go to the lender and not to the household.[2] Consequently, negative equity may lead households to under invest in maintenance and improvements. From their peak in April 2006 house prices then fell nationally by 33 percent, bottoming out in February 2012.[3] The sharp drop in house prices pushed a significant number of households with mortgages into negative equity. Data from CoreLogic data indicate that negative equity peaked in the fourth quarter of 2009 with an estimated 26 percent of households with mortgages in negative equity with the aggregate negative equity amounting to $861 billion. For those households in negative equity in 2009 Q4, the average amount was estimated to be over $70,000.

We quantify that negative equity has a significant adverse impact on housing investments. We use data from the Consumer Expenditure Survey (CES) since it provides detailed data on a wide range of housing maintenance and improvements. In addition, the CES contains detailed mortgage data covering first liens and any second liens as well as a self-reported house value. This information is required for determining whether the household perceives that it is in negative equity. Using the reported expenditures, households in our estimation sample on average invest $2,241 per year (0.95% of the self-reported house value) on maintenance and improvements. Adjusting for the labor supplied by the household to investment projects increases this average annual investment to $3,152 (1.4% of the self-reported house value). The results from our instrumental variables Tobit model

[1] Investment expenditures on improvements and maintenance of existing homes can be inferred using the Census Bureau's series for total private residential construction, new single-family construction, and new multi-family construction. We annualize the monthly data.
[2] See Haughwout, Peach and Tracy (2010).
[3] In the hardest hit states, house prices fell 50 percent from their peak values.

indicate that households in negative equity reduce their housing investment expenditures by an average of 74 percent. Our estimates indicate that from 2006 to 2010 negative equity contributed to a total decline in housing investment expenditures of $51.2billion. During the same period, total annual household income for homeowners did not fall below 2006 levels, ruling out income effects as a driving force behind the decline in housing investment expenditures.[4]

In the next section, we discuss in more detail the reasons why negative equity may impact a household's housing investment decisions, and we review the limited research to date on this issue. In the data and methodology section, we detail how we generate our estimation sample. In addition, we highlight three issues that are important for the empirical work. We then present the main findings as well as some robustness checks. In the conclusion we discuss some broader implications of the findings and some policy implications.

Motivation and literature review

We focus in our analysis on owner-occupied housing and do not consider housing held for investment purposes. The primary motive for an owner-occupied household to engage in costly housing investments is to maintain or improve the flow of housing services which the household enjoys. If a household expects to remain in the house for an extended period of time, then it may be able to fully capture the benefits of its investment decision through the flow of additional housing services over the useful life of the investment. However, since housing investment has a durable aspect to it, the household's expected tenure may in cases be less than the expected lifetime flow of housing services from the investment. If there was no way for the household to capture the value of the remaining service flow when it sells the house, then this would reduce the incentives for the household to make long-lived investments. However, if this remaining service flow can be verified and is valued by prospective buyers, then it may be capitalized into the sale price of the house.

Capitalization extends the horizon of the household when it is considering housing investment decisions. This helps to promote an efficient investment profile by the household regardless of its expected tenure in the house.[5] However, for capitalization to support efficient investment, the homeowner must expect to have a positive equity position in the house at the time of sale. That is, the homeowner needs to be fully exposed to any gains or losses in the value of the

[4] We use the CES's measure for household income before tax (as in our model) and aggregate over owner-occupied households to the US population using sample weights.
[5] See Fischel (2001).

house that are associated with these investment decisions.[6] In normal housing markets this is not a binding constraint since mortgages typically require a downpayment and nominal housing prices tend to rise over time.

Negative equity may lead to lower housing investment for a number of reasons. First, if a borrower in negative equity either has to default as the result of a negative income shock or chooses to default in order to move, the capitalization of any housing investments will first go to the lender. Consequently, negative equity creates an agency problem between the lender and the household. The lender would like the household to continue to make costly housing investments, but the lender has no direct control over these investment decisions until they take title to the property. Second, for those types of investments that can be delayed with little immediate impact on the flow of housing services, the household may choose to postpone these investments in order to create precautionary savings (see Carroll *et al* (2012)). Unemployment risk tends to be high in those housing markets suffering from extensive negative equity. As mentioned earlier, these savings could be used by the household to assist in making its mortgage payments during a period of lower household income, thereby reducing the risk of a costly default. Third, households may need to use the equity in their home to finance larger investment projects. Households in negative equity do not have access to this source of financing.

An additional potential constraint on a household's housing investment decision is whether the house is located in an area facing persistent weak housing demand. Gyourko and Saiz (2004) It may be rational for households with positive equity to forego costly housing investments if local house prices have fallen below replacement costs.[7] In this case, the relatively low demand for local housing implies that households may not be willing to pay for the cost of housing investments. For areas experiencing persistent weak housing demand, then, negative equity may have little to no incremental restraint on housing investment decisions.

The earlier empirical literature on housing investment typically does not control for the homeowner's equity, however Reschovsky (1992) includes a control for the household's equity in the house as a continuous variable and finds no significant effect on housing investment.[8] Melzer (2012) is the most relevant paper for our research. Melzer uses the CES data from 2006 Q1 to 2011

[6] We will discuss later whether making mortgage loans recourse mitigates the need for the borrower to expect to have a positive equity stake at the time of sale.
[7] See Gyourko and Saiz (2004).
[8] Reschovsky (1992) uses 1979 AHS data.

Q1. His key finding is that negative equity households spend on average 30 percent or $800 less on maintenance and improvements per year. Melzer controls for total household expenditures where we control for household income. Melzer also investigates the extent to which negative equity affects durable goods expenditures that are not tied to the house such as vehicles and other home furnishings and equipment. In contrast to his findings for home maintenance and improvement expenditures, Melzer does not find that negative equity significantly reduces these other durable expenditures. He also finds a significant negative equity effect on housing investments for households with higher incomes, more credit availability and with more financial assets. These results suggest that the negative equity variable is not acting as a proxy for liquidity constraints.

Sample construction and empirical methodology

Our primary data source is the Consumer Expenditure Survey. The CES is produced by the Bureau of Labor Statistics (BLS) and is used to generate expenditure weights for the Consumer Price Index (CPI). The sample consists of short panels for a rotating sample of households.[9] The data include detailed expenditure, income, household and individual characteristics. This section will describe the criteria we used to select our estimation sample, while the data appendix will provide more detailed information about variable construction, sampling considerations, and relevant material about using the CES.

Households are interviewed for 5 consecutive quarters. In the first quarter's interview, information is collected on demographics and family characteristics, and this information is updated at each subsequent interview. The second through fifth interviews collect expenditure information from the previous three months using a uniform questionnaire. Income information is collected in the second and fifth interviews and covers the previous 12 month periods. In the 2007 to 2012 surveys included in our estimation sample, quarterly surveys consist of an average of 7,000 households.

To focus on housing maintenance decisions in owner-occupied housing, we restrict our estimation sample to single family housing units identified as the household's primary residence.[10] The estimation sample includes owners of multiple homes. However, we focus on investment expenditures only for the primary residence. Consequently, there is only ever one housing unit per

[9] Each quarter, 20 percent of the sample are new households replacing those households who just completed their fifth interview.

[10] We exclude row or townhouses, duplexes, high-rise apartments, mobile homes and trailers even if they are listed as the primary residence.

household represented in the data, and all variables in the estimation that relate to the housing unit, including home renovations, are exclusively tied to the primary residence.[11]

We exclude some early years of the CES data due to missing values for a key variable. Prior to the second quarter of 2007, the question about property value, *propvalx*, was only asked in a household's second interview. Subsequently, the property value question was asked in each survey. As we will discuss, our estimation strategy uses negative equity indicators from the third, fourth, and fifth interviews. This requires that the estimation sample begin with the second quarter of 2007. Even without this restriction, if we wanted contiguous survey years we would only be able to additionally include 2006 in the estimation sample. In 2004 and 2005, only imputed values for the before tax income measure, *fincbtax*, were reported. In response to criticisms, the BLS reintroduced the actual measure in 2006.[12]

We adopt a number of selection criteria in creating our estimation sample. To improve the reliability of our negative equity indicator which we describe below, we exclude households where the initial (that is the 2nd interview) self-reported property value is either imputed or top-coded. We further restrict the data to include only those households with a reference person aged 20 to 59 inclusive in order to focus on housing investment decisions by households not yet approaching or in retirement. Properties with reported house values less than $30,000 or greater than $1,000,000 are dropped, along with properties with estimated loan-to-value ratios above 2.[13] To limit the degree to which the negative equity estimates may reflect liquidity constraints, households receiving unemployment insurance or reporting zero or negative total income were excluded from the sample. The data appendix details the number of observations excluded with each sample restriction. All nominal variables are deflated and expressed in real 2012 dollars using annual values of the Consumer Price Index for all Urban Consumers.

Our dependent variable, housing investment, is differentiated by household and by property. As noted earlier, for households with multiple properties, we restrict our attention to housing investment that is associated with the primary residence. Housing investment includes all reported

[11] Housing characteristic variables such as the number of rooms are only available for the primary housing unit.
[12] While the CES identifies the total income variable, *fincbtax*, as non-imputed, it does contain components that include some bracket based imputations for respondents who refused to answer the various income questions and were given the option to select a bracket instead. More details on the bracket based imputations are included in the appendix.
[13] Melzer (2012) uses similar sample restrictions with the exception of the $1 million cap.

expenditures related to home maintenance or improvements.[14] These expenditures are recorded in the CES data as project specific, classified by the survey respondents as "maintenance and repair", "addition", "alteration", "replacement", and "new construction". The expenditure data in the CES are reported at a monthly frequency with each interview including a full quarter's worth of expenditures.

Two important data issues relate to the measurement of housing expenditures. The first issue is whether the unit of analysis should be housing investment in a quarter or in a year. The data indicate that housing investment decisions tend to be lumpy. Even pooling housing investment expenditures across all projects in a quarter, 75 percent of households in our estimation sample report zero expenditures in a typical quarter. Pooling housing investment expenditures over a full year reduces the prevalence of zero expenditures to 40 percent. In addition, aggregating the housing investment expenditures across the four surveys provides a better alignment with the 12-month period covered by our income measure.[15] However, aggregating to annual data does require excluding any households with fewer than four interviews.[16] Our final estimation sample is 5,352 households.

Aggregating the reported housing investment expenditures over a full year mitigates but does not eliminate the concentration of zero values for our dependent variable. As a consequence, we will use a Tobit specification to analyze the data. As is well known, using a regression framework with this type of data would attenuate the impact of negative equity on housing investment. If, in addition, we estimated a regression using the data on a quarterly frequency, this attenuation bias would be further exacerbated.

The second data issue is that housing investment expenditures do not include the value of household members' time if they supplied labor to a project.[17] To the extent that households supply their own labor to projects, this breaks the connection between the reported expenditures and the actual investment. Fortunately, the CES contains information on whether the household supplied their own labor. For each project, the household indicates if it contracted out the entire project,

[14] Housing investment is constructed using 15 variables from the CRB expenditure dataset, *Construction, Repairs, Alterations and Maintenance of Owned and Rented Property*. Households are instructed to report their expenditures and any reimbursements they may have received from any source, and the BLS then reports expenditures net reimbursements.
[15] See Data Appendix for more details about the timing convention for the income and expenditure variables in the CES.
[16] This requirement results in dropping 9,904 households.
[17] This is also the case for housing investment expenditures reported in the AHS data.

supplied some of the labor, or supplied all of the labor.[18] For households that undertook a project, 37 percent of the projects were reported to have been completed with the household supplying all of the labor, and 3 percent with the household supplying some of the labor.

To get an overall estimate of the extent to which a household's decision to supply labor impacts the reported expenditures for a project, we pooled all of the project expenditures together and regressed the log of the reported project expenditures on the explanatory variables from our baseline housing investment model to be discussed below, a set of year/quarter fixed effects, a set of project fixed effects and indicators for whether the household supplied all or some of the labor. On average, for households that supplied all of the labor to a project, the reported expenditures are 76 percent lower than if the project was completed entirely by a contractor.[19] If we interpret the reported expenditures for projects where the household supplied all of the labor as comprising the material costs,[20] then this estimated discount indicates a high labor share in the total cost of these housing investment projects.

The possibility that households engage in home production of housing investments creates a potential problem for interpreting the estimation results using the reported housing investment expenditures. What we would like to estimate is the impact of our explanatory variables on the amount of housing investment as measured by the expenditures if the household contracted out the work. The expected reported expenditures for any project, however, will reflect a weighted average of the contractor-based costs and the material costs for the project when the household supplies the labor. Let E_{ijt}^{R} denote the reported housing investment expenditures by household *i* on a project of type *j* undertaken in period *t*. Similarly, let E_{ijt}^{C} denote the investment expenditures if the project is contracted out, and E_{ijt}^{S} denote the investment expenditures if the household supplied the labor. In addition, let p_{ijt} be the probability that the household supplies all of the labor on the project. Finally, let δ_j denote the material cost share for project type *j*; thus $E_{ijt}^{S} = \delta_j E_{ijt}^{C}$. The expected reported investment expenditures are given as follows.

[18] Households may list multiple expenditures for a given project. For each expenditure, they indicate if they supplied labor. However, in practice most households only list one expenditure per project.
[19] There is a small and statistically insignificant impact on expenditures for cases where the household indicates supplying some of the labor.
[20] Households may also include some capital costs if they purchased tools to do the work.

$$
\begin{aligned}
\mathrm{E}(E_{ijt}^{R}) &= (1-p_{ijt})E_{ijt}^{C} + p_{ijt}E_{ijt}^{S} \\
&= (1-p_{ijt})E_{ijt}^{C} + p_{ijt}\delta_{j}E_{ijt}^{C} \\
&= [1-(1-\delta_{j})p_{ijt}]E_{ijt}^{C} \\
\log[\mathrm{E}(E_{ijt}^{R})] &= \log(E_{ijt}^{C}) + \log[1-(1-\delta_{j})p_{ijt}]
\end{aligned}
\tag{1}
$$

The impact of a change in an explanatory variable on the log of the expected reported investment expenditures will reflect both the impact of the change in the contractor based expenditures (what we would like to measure) and the change in the probability that the household supplies the labor to the project.

$$
\frac{\partial \log \mathrm{E}(E_{ijt}^{R})}{\partial x} = \frac{\partial \log E_{ijt}^{C}}{\partial x} - \frac{(1-\delta_{j})}{1-(1-\delta_{j})p_{ijt}}\frac{\partial p_{ijt}}{\partial x} \tag{2}
$$

Without controlling for the likelihood that a household supplies its own labor to projects, our empirical estimates will not reflect the impact of the explanatory variables on housing investment as consistently defined by the contractor level of expenditures. For example, assume for simplicity that negative equity does not affect the propensity for households to engage in housing investment projects, but instead that households in negative equity are more likely to supply their own labor to a project. It would appear in the reported data as if negative equity reduces household investment. This, however, would be a spurious finding. Similar interpretation problems have arisen in other contexts where home production is involved such as comparing household consumption expenditures before and after retirement (see Aguiar and Hurst (2005)). We detail in the next section how we construct adjusted housing investment expenditures to attempt to correct for this issue.

Our key explanatory variable is an indicator for whether the household perceives itself to be in negative equity on its principal residence. We follow the convention used by Melzer (2012) to construct this negative equity indicator variable. First, an estimated loan-to-value for each property is generated based on the ratio of the total mortgage balance to the self-reported property value. Melzer defines the total mortgage balance as the sum of all outstanding mortgage balances, which includes mortgages, home equity loans, and home equity lines of credit (HELOC). The negative

equity indicator is then defined to be one if the loan-to-value is greater than one and zero otherwise. For each of the 4 interviews, we generate separate negative equity indicators for each household.[21]

To generate the total mortgage balance, we use property specific data from the CES expenditure files for mortgages, home equity loans, and HELOCs. Households report the following data related to their mortgages: origination balance, interest rate, term of the loan, month and year of first mortgage payment, monthly mortgage payment, and whether their interest rate is fixed or variable. The current mortgage and home equity installment loan balances are derived by the BLS using an amortization schedule that assumes the borrower is current on all required payments and has not made any additional principal payments. For HELOCs, borrowers report the loan balance and total payment made during the quarter, from which the CES imputes principal and interest assuming an interest rate of prime plus 1.5%. Although mortgage balances, home equity loans, and HELOCs are recorded in different expenditure datasets, it is possible to link them for the same household and property using the assigned property number in each of the files.[22] Combining all outstanding mortgages, home equity loans, and HELOCs, we construct the total mortgage balance variable for each household's primary residence.

The final component to estimating the loan to value (LTV) is the respondents' self-reported property value. Respondents are asked "About how much do you think this property would sell for on today's market?" This variable does not receive much editing from the BLS. There are flags to indicate whether the response is reported or imputed. Inspecting these self-reported house values reveals that some of the reported values are subject to significant but non-persistent measurement errors. A detailed review of this variable, including some examples of "problem observations," is included in the data appendix. To illustrate this measurement error, we regress the quarterly change in the log of the household's self-reported property value on its lag quarterly change and include location specific year-quarter fixed effects to control for the decline in house prices during this period. The coefficient is −0.47 (with a standard error of 0.01).[23] The negative coefficient is

[21] As noted earlier, we drop households if the calculated loan to value exceeds 2. When we investigated these cases, they typically reflected a reported house value that was significantly lower than the house values reported in the remaining interviews.

[22] The CES questionnaire asks survey respondents to list all their properties and the survey associate a "property number" to each of them. From there, questions are asked about mortgage and home equity loans on each of the properties. The expenditure files used here are MOR, HEL, and OPH, and the property number variable for each of these is *prop_nof*, *prop_nog*, and *prop_noh*, respectively.

[23] If we include two lags as explanatory variables the coefficient on the first lag is −0.63 (0.01) and on the second lag is −0.34 (0.01).

consistent with a mean reverting measurement error. A consequence is that this induces measurement error in the negative equity indicator. Left uncorrected, this measurement error will bias downwards the estimated impact of negative equity on housing investment.[24] We discuss our strategy for dealing with this measurement error in the next section.

Our baseline specification includes controls for household income and demographics including the reference person's age, race/ethnicity, and education. [25] The specification also controls for a number of house specific factors including the number of rooms and an indicator for a new house defined to be a house built in the prior three years. Since new construction may require less maintenance than other units, we explored different definitions for a new house and found that the discontinuity in reported housing investment expenditures was most pronounced over the first three years.[26] Empirical specifications in the literature also control for the duration that the household has lived in the house – housing tenure. We found higher housing investment activity during the first year of housing tenure. Controlling for this initial year effect, we did not find a significant additional effect of housing tenure on investment decisions. Descriptive statistics for all variables used in the estimation are provided in Table 1.

Estimation results

The first step in the estimation is to determine how to instrument for the measurement error in the negative equity indicator that is induced by the measurement error in the self-reported house value. Gyourko and Tracy (2006) identified a similar problem with the self-reported house values in the AHS. Their choice of instrument was a second negative equity indicator where the house value is estimated as the purchase price of the house updated to the time of the survey using a metropolitan repeat-sale house price index. Both the purchase date and price are also available in the CES. However, only 39 percent of our estimation sample has an identified metro area; for most of the rest only the state is identified, and in some cases even the state is suppressed. This makes it difficult to merge in an appropriate house price index.

[24]This downward bias would also impact the estimated impact of negative equity on other components on household expenditures.

[25] The CES defines the reference person as the person who responds yes to owning or renting the primary housing unit. We also controlled for family size but dropped this variable since it was not significant.

[26] The CES does not record the total square footage of the house or the lot size. Controlling for whether a house is new, we did not find a statistically significant or economically meaningful effect for the age of the house on housing investment.

An alternative instrumental variable approach is motivated by the feature that the measurement error in the self-reported house values typically takes the form of a non-persistent error that significantly increases or reduces the reported house value relative to the reported values from the other interviews. Less common were cases where it appeared that all of the reported house values were significantly over or under stated. This suggests that an alternative strategy is to use the negative equity indicators derived from the 3nd through 5th interviews as instruments for the negative equity indicator derived from the 2nd interview. This strategy will also tend to identify the effects on housing investment from those households who consistently perceive that they are in negative equity throughout the interview cycle. In our estimation sample, the three instruments are all highly significant in the first stage regression.

The next step in the estimation is to adjust the reported investment expenditures, E_{ijt}^{R}, for those projects where the household indicates that they supplied some or all of the labor. For each separate project type, we estimate the following regression.

$$\log(E_{ijt}^{R}) = X_{ijt}\gamma_j + \delta_j^S I_{ijt}^S + \delta_j^B I_{ijt}^B + \mu_{ijt} \tag{3}$$

The vector X contains all of our baseline control variables discussed earlier and a set of year/quarter fixed effects. The remaining two variables are indicators for a household supplying all (I_{ijt}^S) or some (I_{ijt}^B) of the labor. We use the following method for calculating the adjusted investment expenditures for a project, E_{ijt}^{A}, based on the reported expenditures.

$$E_{ijt}^{A} = \frac{E_{ijt}^{R}}{[(1-I_{ijt}^S)(1-I_{ijt}^B) + I_{ijt}^S e^{\hat{\delta}_j^S} + I_{ijt}^B e^{\hat{\delta}_j^B}]} \tag{4}$$

The adjusted project expenditure equals the reported expenditure when the project is entirely contracted out ($I_{ijt}^S = I_{ijt}^B = 0$).

The estimated project-specific discount factors are reported in Table 2. The only statistically significant discount factor associated with households supplying some of the labor was for inside painting. For households supplying all of the labor on a project, the estimated discount factors range from a low of 0.12 (implying a labor share of 0.88) for outside painting to a high of 0.41 (labor share of 0.59) for heating or air conditioning jobs.

Households engage in significant housing investment activity. For our estimation sample, the mean reported annual housing investment expenditure is $2,241, or 0.95 percent of the self-reported house value. Gyourko and Tracy (2006) using the AHS from 1985 to 1993 report an average annual expenditure level of $2,899 (in 2012 dollars), or 1.7 percent of the reported house value.[27] Similarly, the mean annual adjusted housing investment expenditure for our estimation sample is $3,152 or 1.4 percent of the house value. The reported housing investment expenditures understate the adjusted expenditures by 29 percent. Home production is an important aspect of housing investment.

A consequence is that we need to understand the determinants of when households supply their own labor to housing investment projects. As illustrated in equation (2), the difference between the impact of an explanatory variable on the reported and adjusted housing investment expenditures depends importantly on the degree to which the probability that the households uses its own labor on a project varies with that explanatory variable and the discount associated with that project.

We discussed earlier the estimated project specific discounts and we turn now to the decision by households to supply their own labor. Table 3 provides marginal effects for our baseline variables. We pool together all of the project data and include project fixed effects. Specification (1) is based on a standard Probit, while specification (2) uses an IV Probit where we instrument for the negative equity indicator using the three instruments discussed earlier. The results indicate that there is no significant relationship between negative equity and the likelihood that a household provides its own labor. Consistent with Becker's (1965) theory of home production, households with higher income and more education are less likely to provide their own labor to a project. Bogdon (1996) finds a similar result using AHS data. In addition, older household heads are more likely to rely on contractors. In terms of race and ethnicity differences, the data indicate that blacks, Hispanics and Asians are all significantly less likely to contribute labor to projects relative to white homeowners holding other variables constant. Bogdon (1996) also reports a similar finding for blacks. Households are less likely to supply their own labor to housing investment projects in their first year of tenure and more likely for larger houses.

[27] The decline in the size of maintenance expenditures relative to reported house values may reflect the significant appreciation of land values during the housing boom (Haughwout, Orr and Bedoll, 2008 and Haughwout, Peach, Sporn and Tracy, 2012). Land likely requires less on-going maintenance activity as compared to housing to maintain its service flow.

Our baseline housing investment results are provided in Table 4. We use Tobit specifications to account for the mass point of households with zero annual housing investment expenditures. For our main specification, we use the log of housing investment expenditures and the log of household income. In Appendix Table A1, we report the results using the levels of both variables. We report three specifications. Specification (1) provides results from a Tobit where we do not account for the measurement error in the negative equity indicator. Specification (2) switches to an IV Tobit where we instrument the negative equity variable with the 3rd to 5th interview negative equity indicators. Both specification (1) and (2) use as the dependent variable the log of the reported housing investment expenditures. Finally, in specification (3) we continue to use the IV Tobit, but switch to the log of the adjusted housing investment expenditures. In all three specifications, we report the unconditional marginal effects. For ease of interpretation, we also translate these marginal effects into their implied percent changes.

Before discussing our results for the household's perceived equity position, we summarize the other main findings. Housing investment has a strong relationship to household income with an estimated elasticity around 0.60. Gyourko and Tracy (2006) using the AHS data report an income elasticity of housing investment of 0.42 (OLS) and 0.47 (IV). Earlier work including Mendelsohn (1977) and Boehm and Ihlanfeldt (1986) also find positive income effects. Controlling for income, households where the head has a college degree or higher level of education spend significantly more each year on housing investment. Controlling for income and education, the data indicate that older household heads invest slightly more in their homes.[28] Turning to the race/ethnicity of the household head, the data indicate that Hispanics and Asians are significantly less likely to engage in housing investment, with the marginal effects indicating reductions of between 60 and 80 percent relative to white households. Given our earlier findings that they are also less likely to supply their own labor, the impacts are not mitigated when we move from the reported (specification (2)) to the adjusted (specification (3)) housing expenditures. The results for blacks are smaller in magnitude and less precisely estimated, with the impacts significant only at the 10 percent level for specification (3). Mendelsohn (1977) reports that non-whites are significantly less likely to report positive housing investment expenditures, but conditional on reporting positive expenditures they do not on average spend less than whites.

[28] Gyourko and Tracy (2006) find no significant age effect and no evidence of life-cycle nonlinearities in how age impacts housing investment decisions. They also limit their estimation sample to household heads aged 20 to 59.

Turning to characteristics of the house and the tenure of the household, the data indicate that a significant additional amount of housing investment takes place during the household's first year living in the house. Total expenditures on housing investments in the first year are more than 270 percent higher than otherwise. Gyourko and Tracy (2006) report an even higher first year effect of 422 percent. Consistent with the earlier empirical literature, maintenance expenditures are higher the larger the house as measured by total number of rooms. Finally, new houses receive significantly less maintenance, with the estimated reduction at around 75 percent.[29]

We move now to our central question of whether negative housing equity is associated with reduced housing investment expenditures. As expected, ignoring the concentration of households with zero housing investment expenditures leads to an attenuation of the estimated negative equity impact. Table 4's specification (1) which uses a Tobit indicates that households in negative equity on average spend 37 percent less annually on housing investment. From Table A1, using a Tobit with the dependent variable measured as the level of the annual housing investment expenditures, the unconditional marginal effect for negative equity is $1,072. In contrast, if we ignore the concentration of zero housing investment expenditures and estimate the levels specification using a linear regression, the estimated impact of negative equity is reduced to $731 – a decline of 32 percent relative to the Tobit specification. Switching from annual to quarterly data exacerbates the problem of the mass point of zero investment expenditures. Using a linear regression with quarterly data, the estimated annualized impact of negative equity declines further to $575 – a decline of 46 percent from the Tobit specification.

Measurement error in the self-reported house values has an important attenuation impact on the estimated degree to which negative equity is associated with lower housing investment activity. This can be seen by comparing specification (1) to specification (2) in Table 4. We use negative equity indicators derived from the 3rd through 5th survey as instruments for the negative equity indicator derived from the 2nd survey. Using an IV Tobit, the unconditional marginal effect from negative equity on housing investment doubles from 37 percent to 74 percent.[30] Similarly, switching

[29] Given the significant pace of new home construction during the housing boom (see Haughwout *et al* (2012)) and the fact that households that purchased these homes had less time for them to appreciate prior to the crash, there is a positive correlation between negative equity and living in a new home. Leaving the new home indicator out of the specification leads to an upwardly biased estimate of the negative equity effect on housing investment.
[30] As noted earlier, our instruments also identify the impact of persistent negative equity.

from the log to the level of housing investment expenditures (see Table A1), the unconditional marginal effect associated with negative equity increases from $1,072 to $2,610 when we instrument.

Home production also has a specific implication for measuring the impact of negative equity on housing investment.). Switching from the reported to the adjusted housing investment expenditures does not make a material impact on the estimated percent reduction in spending associated with negative equity. However, since adjusted expenditures exceed reported expenditures, this similar percent reduction translates into a significantly larger dollar amount. This can be seen by comparing specifications (2) and (3) in Table A1. The unconditional marginal effect associated with negative equity increases from $2,610 using reported expenditures to $3,892 using adjusted expenditures – an increase of 33 percent.

The data indicate a significant reduction in housing investment expenditures associated with households in negative equity. However, this does not identify if the reduction begins with the onset of negative equity, or only later when the household becomes delinquent on their mortgage.[31] If the drop off in housing investment only begins when the household enters delinquency or serious delinquency, then the impact of negative equity and delinquency will be larger than our estimate that conditions only on negative equity. The timing of the onset of the under maintenance is important for understanding the expected exposure time of a house to accelerated depreciation. Unfortunately, the CES does not identify if a household is current or delinquent on their mortgage payments. Consequently, this question will need to be addressed using a different data source.

Robustness Checks

An important issue to explore is whether our results are biased due to an important left-out variable that is correlated with housing investment as well as our negative equity indicator. Gyourko and Saiz (2004) using AHS data investigate the disincentive to make costly investments for households living in markets experiencing depressed housing demand. Analogous to a Tobin's Q for firms, Gyourko and Saiz calculate the house price relative to the replacement cost.[32] Similar to the impact of Tobin's Q on a firm's investment decisions, a house based Q that is below one may indicate low returns from a costly housing investment. Using AHS data, they report that 18% of households live in homes where the adjusted price is below the replacement cost. Controlling for

[31] A third possibility is that under maintenance may not take place until the household is seriously delinquent and there is a high probability that the household will go through foreclosure.
[32] See Gyourko and Saiz (2004) for details on the construction of their price to replacement cost indicator.

other factors, households with a price to replacement cost below one on average spent $310 less on housing investments (expressed in 2012 dollars). When they instrument the price to replacement cost using information from neighboring houses, the estimated effect indicates a decline of $1,184 – or 46% of their sample average housing investment expenditures.

We face two significant challenges in creating a similar variable for our CES sample. The first challenge is the lack of geographic information on the location for many households in the CES data. In our estimation sample 2,098 households (39% of total) have a reported metro area that has building cost data.[33] For 2,525 households (47% of total) we know their state of residence, but not the metro area within the state. In these cases, we assign a population weighted average of the RS Means building cost data for the metro areas in that state.[34] Finally, the CES suppresses the state identification for all observations from a small set of states. This impacts 792 households in our estimation sample. For these cases, we assign a population weighted average of the RS Means building cost data from the metro areas across these states.

The second challenge is that the house size is not reported in the CES. We estimate this size using a regression fit to AHS data from 1985 to 2009. This regression relates the square foot size of the house to the number of bedrooms, bathrooms and other rooms as well as when the house was built. We restrict the AHS data to houses between the 5th percentile (875 square feet) and the 90th percentile (3,800 square feet) of the sample size distribution.[35] The R-square from this regression is 0.37. We use the same AHS sample to estimate the average annual real depreciation rate on housing. The data indicate a value of 0.5 percent per year.[36] This is used to adjust older houses to a more comparable basis to new homes.

Our measure of the replacement cost will reflect measurement error both from the need to estimate the size of the house as well as the inability to merge in metro specific building cost data for 61% of our estimation sample. We find that 38 percent of our households are residing in housing

[33] The CES's metro area variable, *PSU*, is defined based on the Census' June 2003 definition of Core Based Statistical Areas (CBSAs), however it is not a direct match to the CBSAs. We map each *PSU* first to the CBSA to which it shares the most counties. We then match the CBSA to its largest Metropolitan Statistical Area (MSA) in order to merge with the MSA level building cost data.

[34] If any of these metro areas are reported by the CES (that is are used for the 2,098 households), we exclude these metro areas for the calculation of the state average. For 73 households there were no metro areas in the state with RS Means data. In these cases we computed the average using metro areas for neighboring states.

[35] The limited explanatory variables have a difficult time explaining each tail of the size distribution.

[36] We obtain this estimate of the depreciation rate by regressing the log of the self-reported house value per square foot on the age of the house in years, a set of indicators for the decade the house was built and a set of year effects.

with an adjusted self-reported price below the estimated replacement cost. Appendix Table A2 provides marginal effects from an IV Probit with the indicator for a Q below 1 as the dependent variable and the explanatory variables used in our housing investment Tobits. The results indicate that households in negative equity are 26 percentage point more likely to live in a house where the adjusted price is below replacement costs holding the other variables in our baseline specification constant. If this Q value has a significant impact on housing investments, then this would suggest that our earlier marginal effects associated with negative equity are biased upwards. Looking at our race/ethnicity indicators, blacks are 50 percentage points more likely and Asians 41 percentage points less likely to live in a house where the adjusted price is below the replacement cost.

Appendix Table A3 replicates Table A1 but adds the price relative to replacement cost indicator. We report the results from the levels specification so that we can compare the magnitudes to those reported in Gyourko and Saiz (2004). Our summary focuses on the IV Tobit estimates in specification (2), but the corresponding non-IV Tobit for reported investment and the IV Tobit for adjusted investment are reported in specification (1) and (3), respectively.

The data indicate that households with an adjusted price to replacement cost below one are associated with a lower housing investment of $822 (standard error of $283). This estimate is between the OLS and IV estimates (adjusted to 2012 dollars) presented in Gyourko and Saiz (2004). The degree of likely measurement error both from not being able to use metro area specific building costs for a large fraction of our data and from the estimation of the house size suggest that our estimate may suffer from significant attenuation bias. Controlling for the house specific Q, however, only slightly reduces the unconditional marginal effect associated with negative equity.[37] Including the Q measure does impact the unconditional marginal effect for blacks – changing the estimate from a decline of $866 to a decline of $712 (with a standard error of $576). The marginal effects for Asian households increase from a decline of $2,279 to a decline of $2,388. We leave to future research finding an explanation for the large estimated race/ethnic impacts on housing investment activity.

A related question to the impact of the price to replacement cost on housing investment is whether housing investment depends on the recent change in house prices controlling for whether the borrower is in negative equity. Rising house prices may encourage more housing investment due

[37] The IV unconditional marginal effect changes from -2,610 to -2,550.

to a higher perceived return on this investment.[38] In addition, rising prices may facilitate the financing of housing investment by increasing the borrower's equity in the house. An important question is whether the house price effect is asymmetric with respect to rising and falling house prices.

We face a similar challenge in including a measure of the change in house prices as we discussed for the price to replacement cost. We merge in the 12-month change in the CoreLogic overall metropolitan house price index for the subset of households where we know their metro area.[39] For households where we know their state but not their metro area, we use the corresponding 12-month change in the overall state house price index. Finally, for those households where their state is suppressed, we assume that they live in one of the five states where all household have a suppressed location and we assign the 12-month change in the population weighted average of these five state overall price indices.

The data indicate that the impact of house prices on housing investment is asymmetric. When we impose symmetry, the data indicate no significant association between recent house price changes on housing investment activity holding constant the other variables in our specification. This result holds both for our log and level specifications. When we allow for asymmetric effects, the data indicate that increases in house prices are associated with higher investment rates, but that decreases in house prices have a negligible impact. In logs, the data indicate that a one percent increase in house prices is associated with 9 percent higher investment. However, this is imprecisely estimated. In levels, a one percent increase in house prices is associated with an increase of housing investment of $320 (standard error of $110). Controlling for the recent change in house prices, however, only slightly reduces the measured impact of negative equity on housing investment.[40]

Legal remedies

Our finding of a significant reduction in housing investment expenditures for negative equity borrowers raises the important question of whether there are legal remedies that may help to protect the lender's interests. As noted earlier, in situations of negative equity an agency problem arises since the household retains the decision rights over any housing investment decisions so long as it

[38] Our house price measures are based on repeat-sales indices and are meant to reflect the change in prices of constant quality houses.
[39] We use the same matching process from the CES's metro area variable, *PSU*, to the CoreLogic's MSA variable as we did for the MSA level building cost data described in footnote 43.
[40] The IV Tobit marginal effect in logs declines from -1.334 to -1.303.

continues to hold title to the house, but the lender has the financial interest in the decisions. What is necessary is a mechanism to realign the interests of the borrower and the lender.

One possible legal remedy is to allow the lender financial recourse to the borrower in the event of a default. In the case of a recourse mortgage, a lender can file a deficiency judgment against the borrower for the difference between the balance on the mortgage and the proceeds from a foreclosure sale. Households are required to use any financial assets to satisfy the deficiency judgment. A deficiency judgment, like capitalization, may provide an additional financial incentive for the household to make costly housing investments. In practice, recourse is likely to only have a limited impact on housing investment decisions by negative equity households.

How might a recourse mortgage affect a household's housing investment decisions? For a household in negative equity, the impact likely depends on a number of factors. First, and importantly, the household must understand the implications of having a recourse mortgage. To the extent that households are not aware of whether their mortgages are recourse or not, this legal remedy will not impact their behavior.[41] Second, the impact of recourse will depend on the degree of the negative equity relative to the household's financial assets. If the extent of the negative equity exceeds the household's financial assets, then any additional decline in the house value due to decisions not to engage in costly housing investments will have no impact on the expected magnitude of any deficiency judgment. It is only in those cases where the extent of negative equity is less than the household's financial assets that the expected magnitude of any deficiency judgment depends on the housing investment decisions.[42]

We first explore the degree to which households are aware of the true recourse nature of their mortgage. We use data from the housing module of the 2013 *Survey of Consumer Expectations*.[43] Based on the household's state of residence and whether the mortgage was to purchase a house or to refinance an existing mortgage, we can determine if the mortgage is recourse. We infer the borrower's belief about their mortgage by their answer to the following question:

[41] However, even if under water households that are unaware that their mortgages are recourse choose to under maintain their homes, the consequences to the lender will be mitigated to the extent that the lender can collect a larger deficiency judgment.
[42] Even in these cases, if the household anticipates that it will use up these financial assets prior to a default, then recourse will not provide a financial incentive to housing investment decisions.
[43] See Fuster et al (2013).

> "If somebody with a mortgage like yours and living in your state went through foreclosure, do you think that their lender could legally go after their other assets (e.g. bank accounts, cars, other property, etc.) to cover the remaining amount they owe?"

Of the 582 respondents identified to have a recourse mortgage, only 324 or 56 percent answered the question correctly.[44] This accuracy rate declined in the confidence the respondent expressed in the answer, in the age of the respondent, and increased with the respondent's education. The accuracy did not improve if the respondent reported knowing one or more individuals who went through foreclosure. Households' frequently inaccurate understanding of the legal environment surrounding recourse suggests that these laws will have a limited impact on the maintenance behavior we observe.

Even when lenders have the legal right to pursue a deficiency judgment against a household that defaults, they may not always have a financial incentive to do so. The cost of obtaining a deficiency judgment varies by state and can amount to several thousand dollars. The lender also has imperfect information as to the borrower's ability to pay the deficiency judgment if one is granted. On a case by case analysis, the expected yield must at least cover the legal costs for it to make sense to the lender to pursue the borrower in the event of a default.[45] A consequence is that lenders are only likely to invest in a deficiency judgment if the size of the unpaid balance is fairly substantial.

The CES provides information both on the degree of a household's negative equity as well as their financial assets. This, in principle, would allow us to target those borrowers with recourse mortgages where we might expect that recourse would have the largest impact on their housing investment decisions. However, in practice, our sample of negative equity borrowers is too small to allow us to effectively do this targeting. We use the classification of states and types of mortgages in Ghent and Kudlyak (2011) to classify if a mortgage is recourse.[46] Table 5 provides tabulations of negative equity, recourse, financial assets, and deficiencies. In our estimation sample, while 79 percent of our mortgages are recourse, we have only 270 recourse mortgages where the household reports itself to be in negative equity. Restricting the focus to the subset of these borrowers who have a deficiency of at least ten thousand dollars decreases the number to 182 borrowers (67 percent

[44] For respondents identified to have non-recourse mortgages, 60 percent answered the question correctly.
[45] This ignores any possible value to the lender generated by having a reputation of being "tough" on defaulting borrowers.
[46] The seven non-recourse states represented in the CES are Alaska, Arizona, California, Minnesota, Oregon, Washington, and Wisconsin. Within California, we define homeowners with non-purchase mortgages as subject to recourse since deficiency judgments are prohibited only for purchase mortgages in the state. Some states are excluded or have recoded definitions in the CES, and a more detailed description of this is included in the appendix.

of the 270). Further restricting the focus to those borrowers who have both a deficiency of at least ten thousand dollars and sufficient financial assets to cover this deficiency decreases the number to 40 borrowers (15 percent of the 270).[47]

This suggests that the target set of borrowers who might be profitable for lenders to pursue for a deficiency judgment is likely to be a relatively small percentage of all negative equity borrowers with recourse mortgages. This is corroborated by recent experience by the GSEs. The FHFA Inspector General Report (2012) indicated that in 2011 the GSEs filed for deficiency judgments against 35,231 borrowers, or 10.3% of all foreclosures in that year. The reported yield was only 0.22 percent of the total deficiency.

Estimated Aggregate Impacts

We can use the model estimates to address the degree to which the sharp decline in housing investment expenditures over the period from 2006 to 2010 reflected income effects from the recession or negative equity from the housing bust. We create an annual aggregate income measure from the CES data restricting the sample to homeowners.[48] To calculate the negative equity effect we need an annual measure of the aggregate value of houses that are in negative equity. We use a random sample of active mortgages using servicing data from Lender Processing Services (LPS). For each mortgage, we create a negative equity indicator that takes a value of one if the current amortized loan balance exceeds the estimated current value of the house. We estimate this house value taking the appraised value at the mortgage origination date and updating using CoreLogic metro area repeat-sales house price indices. Our estimate of the annual maintenance for these properties is one percent of the last appraised value; we put this investment in real terms using the CPI. As a result, changes in our aggregate negative equity house value estimate will, like our model estimates, reflect primarily borrowers moving into or out of negative equity instead of changes in the degree of negative equity among negative equity borrowers.

Using the aggregate income changes and house values among negative equity borrowers, we can apply the model estimates to approximate the aggregate impacts. Figure 2 reproduces the time-

[47] For these borrowers with large deficiencies and significant financial assets, we cannot tell how much of their financial assets will be dissipated if they end up in foreclosure.

[48] The CES interview survey is designed to have nationally representative independent samples for each quarter, and we use sample weights to aggregate to the US population. The interview question asks for income over the last twelve months, so we use the fourth quarter totals of each year as our annual estimate.

series on aggregate housing investment expenditures with the estimated negative equity effects. Aggregate income for home owners never falls below the 2006 levels, therefore income changes do not help explain the drop in aggregate housing investment. In contrast, negative equity has a large impact with the cumulative effect from 2006 to 2010 amounting to $51.2 billion.

Conclusion, policy implications and further analysis

Households engage in significant housing investment. On average, we find that households report spending $2,241 or nearly 1 percent of the self-reported value of the house in various maintenance and improvement projects. Adjusting for the labor that some households supply themselves, the average annual expenditure rises to $3,152 or 1.4 percent of the house value. Capitalization of these expenditures in the value of the house can help support efficient investment decisions by allowing households to capture the benefits of their investment activities that persist beyond their expected remaining tenure in the house.

This useful role of capitalization may be largely shut down when the household is in negative equity. In this case, increases in the value of the house accrue to the lender/investor and not to the household. As a consequence, an agency problem may exist where the household retains the decision rights to the investment activity while the lender/investor holds the financial incentives to engage in these investments. We argue that legal recourse is likely only to be an effective safeguard for lenders[49]/investors in the few cases where the household has both a large deficiency and significant financial assets.

We investigate the extent to which negative equity adversely impacts housing investment decisions. Using Consumer Expenditure Survey data from 2007 to 2012, we find that on average negative equity households reduce their housing investments by around 75 percent. Our finding is robust to controlling in addition for whether the value of the house is below the replacement cost and for the recent changes in local house prices. The dramatic increase in negative equity resulting from the house price declines from the housing bust are estimated to have had a significant negative impact on aggregate housing investments. Over the period from 2006 to 2010, we estimate that the change in aggregate negative equity accounted for a cumulative decline in aggregate housing investment of $51.2 billion.

[49] This would be more typical for what has been described as "strategic" default as opposed to the traditional default resulting from an underwater borrower facing a job loss.

An important question that our data cannot answer is the persistence of the reduction in housing investment that we document. Housing units may pass from negative into positive equity when they change owners through short sales or the foreclosure process, or when house prices recover to such an extent that current owners' equity is restored. In these cases, how much of the maintenance foregone while the property was in negative equity is made up by the new owners, or the old owners with new equity? Is ex post catching up more expensive than steady investment would have been? Our data, with its short panels of owners, is not suited to answering these questions, but future work using a dataset like the American Housing Survey, with its long panel of housing units, could shed light on these important questions.

In this paper, we focused on decisions by negative equity borrowers to make costly investments in their house. More broadly, these same borrowers also must make decisions whether to make costly investments in their neighborhoods and communities. As in housing-specific investments, the household cannot take the neighborhood or community investments with them if they move out of the neighborhood or out of the community. While they may enjoy the service flow from these community investments while they remain in their house, these service flows may extend well beyond their expected remaining tenure in the house. Again, capitalization of the benefits of these neighborhood and community investments helps to promote efficient decisions by current homeowners. Negative equity, by turning down or off this capitalization, may discourage negative equity households from making these costly civic investments. Future research could examine to see if there is a connection between the percent of negative equity homeowners in a community and measures of civic investment activity.

References

Aguiar, Mark, and Erik Hurst. "Consumption vs Expenditure." *Journal of Political Economy* 113 (October 2005): 919-948.

Becker, Gary S. "A Theory of the Allocation of Time." *Economic Journal* 75 (September 1965): 493-517.

Boehm, Thomas P., and Keith R. Ihlanfeldt. "The Improvement Expenditures of Urban Homeowners: An Empirical Analysis." *AREUEA Journal* 14 (Spring 1986): 48-60.

Bogdon, Amy S. "Homeowner Renovation and Repair: The Decision to Hire Someone Else to Do the Project." *Journal of Housing Economics* 54 (1996): 323-350.

Dynan, Karen. "Is A Household Debt Overhang Holding Back Consumption?" *Brookings Papers on Economic Activity*, Spring 2012, 299-362.

Dynarski, Susan, and Jonathan Gruber. "Can Families Smooth Variable Earnings?" *Brookings Papers on Economic Activity* (1997).

Fischel, William A. *The Homevoter Hypothesis: How Home Values Influence Local Government Taxation, School Finance and Land-Use Policies*. Cambridge, Harvard University Press, 2001.

Fuster, Andreas, and Paul S. Willen. "Payment Size, Negative Equity, and Mortgage Default." Staff Report 582. Federal Reserve Bank of New York, November, 2012. http://www.newyorkfed.org/research/staff_reports/sr582.pdf.

Fuster, Andreas, Wilbert van der Klaauw, and Basit Zafar. "NY Fed Survey of Consumer Expectations Housing Module." 2013.

Gerardi, Kristhpher S., Eric Rosenblatt, Paul S. Willen, and Vincent W. Yao. "Foreclosure Externalities: Some New Evidence." Working Paper No. 12-5. Federal Reserve Bank of Boston, July, 2012.

Gyourko, Joseph, and Albert Saiz. "Reinvestment in the Housing Stock: The Role of Construction Costs and the Supply Side." *Journal of Urban Economics* 55 (March 2004): 238-256.

------, and Joseph Tracy. "Using Home Maintenance and Repairs to Smooth Variable Earnings." *Review of Economics and Statistics* 88 (November 2006): 736-747.

FHFA's Oversight of the Enterprises' Efforts to Recover Losses from Foreclosure Sales. Federal Housing Finance Agency, Office of Inspector General. Audit Report: AUD-2013-001, 2012.

Haughwout, Andrew, Ebiere Okah, and Joseph Tracy. "Second Chances: Subprime Mortgage Modification and Re-Default." Staff Report No. 417. Federal Reserve Bank of New York, August, 2010.

-------, James A. Orr and David Bedoll. "The Price of Land in the New York Metropolitan Area," *Current Issues in Economics and Finance* 14 (April/May 2008): 1-7.

------, Richard W. Peach, John Sporn, and Joseph Tracy. "The Supply Side of the Housing Boom and Bust of the 2000s." Staff Report 556. Federal Reserve Bank of New York, March, 2012. http://www.newyorkfed.org/research/staff_reports/sr556.pdf.

------, Richard Peach, and Joseph Tracy. "The Homeownership Gap." *Current Issues in Economics and Finance* 16 (May 2010): 1-11.

Melzer, Brian T. "Mortgage Debt Overhang: Reduced Investment by Homeowners with Negative Equity." Working Paper. Northwestern University, Kellogg School of Management, August, 2012.

Mendelsohn, Robert. "Empirical Evidence on Home Improvements." *Journal of Urban Economics* 4 (October 1977): 459-468.

Pence, Karen. "Foreclosing on Opportunity: State Laws and Mortgage Credit," *Review of Economics and Statistics* 88 (February 2006): 177-182.

Tracy, Joseph, and Joshua Wright. "Payment Changes and Default Risk." Staff Report 562. Federal Reserve Bank of New York, June, 2012. http://www.newyorkfed.org/research/staff_reports/sr562.pdf.
Zhu, Jun. "An Empirical Analysis of the Home Affordable Refinance Program's Impact on Default Rates." Working Paper. Freddie Mac, May, 2012.

Table 1: Summary Statistics

Dependent Variables:	Mean	Std	Min	Max
Reported Investment	2,241	6,605	0	133,937
Adjusted Investment	3,152	8,416	0	133,937
Reported Investment / Property Value	0.009	0.030	0	0.816
Reported Investment / Income	1.075	42.924	0	2,830
Adjusted Investment / Property Value	0.014	0.039	0	0.816
Covariates				
Negative equity	0.08	0.27	0	1
Household income	97,993	70,628	2	852,512
Log household income	11	1	1	14
Age of household head	45	9	20	59
Education of household head:				
High school graduate	0.20	0.40	0	1
Some college	0.19	0.40	0	1
College graduate	0.39	0.49	0	1
Graduate school	0.15	0.36	0	1
Race/ethnicity of household head:				
Black	0.06	0.24	0	1
Hispanic	0.03	0.18	0	1
Asian	0.10	0.30	0	1
Other	0.01	0.11	0	1
Number of rooms in home	7	2	1	99
First year in home	0.06	0.24	0	1
New home (≤3 years old)	0.05	0.22	0	1
Descriptive:				
Reported Property Value	269,754	168,099	30,621	968,905

Notes: Number of observations is 5,352. Investment, property value, and income are expressed in real 2012 dollars using annual values of the Consumer Price Index for all Urban Consumers.

Table 2: Estimated project discount factors

Project (code): Description	Percent of all positive expenditures (frequency)	Percent of all positive expenditures (magnitude)	Percent provided all labor	Average expenditure if positive	Discount factor: Provided All Labor
100 Dwellings under construction including a vacation or second home	1	5	34	6,374	0.231
110 Building an addition to the house or a new structure including porch, garage or new wing	2	10	45	8,227	0.323
120 Finishing a basement or an attic or enclosing a porch; Remodeling one or more rooms in house	9	21	46	3,870	0.221
140 Landscaping the ground or planting new shrubs or trees	11	6	49	899	0.266
150 Building outdoor patios, walks, fences, or other enclosures, driveways, permanent swimming pools	5	9	40	2,787	0.203
160 Repairing outdoor patios, walks, fences, driveways, or permanent swimming pools	3	2	35	1,086	0.271
170 Inside painting or papering	10	2	76	402	0.136
180 Outside painting	3	2	47	1,153	0.118
190 Plastering or paneling; Siding; Insulation	2	2	42	1,413	0.238
200 Plumbing or water heating installations and repairs	13	4	26	500	0.344
210 Electrical work	4	2	27	760	0.361
220 Heating or air-conditioning jobs	10	9	6	1,520	0.411
230 Flooring repair or replacement, including inlaid linoleum, vinyl tile; Replacing or installing carpeting	4	4	37	1,836	0.277
260 Roofing, gutters, or downspouts; Masonry, brick, or stucco work	5	8	17	2,588	0.155
280 Installation, repair, or replacement of window panes, screens, storm doors, awnings, etc.	6	7	25	2,124	0.182
300 Other improvements or repairs; Combined job codes	10	6	36	949	0.315

Notes: Calculations based on the 3,210 households in our sample that reported positive expenditures. Households can report multiple expenditures; our sample included 7,099 distinct expenditures in total. Discount factor is given by $e^{\hat{\delta}_j}$. Some related projects were combined; see section A3 of the appendix for details.

Table 3: Probability that Owner Supplied Some or All of Labor

	Probit MEs	IV Probit MEs
	(1)	(2)
Negative equity	-0.006	-0.014
	(.027)	(.042)
Log Household income	-0.008*	-0.008*
	(.004)	(.004)
Age of head (10 years)	-0.056**	-0.056**
	(.008)	(.008)
Education of household head:		
High school graduate	-0.044	-0.044
	(.038)	(.038)
Some college	-0.119**	-0.119**
	(.037)	(.037)
College graduate	-0.140**	-0.140**
	(.036)	(.036)
Graduate school	-0.194**	-0.194**
	(.038)	(.038)
Race/ethnicity of household head:		
Black	-0.105**	-0.105**
	(.031)	(.031)
Hispanic	-0.078**	-0.078**
	(.026)	(.026)
Asian	-0.147**	-0.147**
	(.048)	(.048)
Other	-0.011	-0.012
	(.07)	(.07)
First year in home	-0.016*	-0.016*
	(.004)	(.004)
Number of rooms in home	-0.054**	-0.054**
	(.03)	(.03)
New home (≤3 years old)	-0.012	-0.012
	(.034)	(.034)

Notes: Average marginal effects with standard errors reported in parentheses. Year/quarter and project fixed effects are included. Average probability is 0.4. Number of observations 7,090. Standard errors are calculated clustering on the household.

** significant at the 5 percent level * significant at the 10 percent level

Table 4: Determinants of Log Housing Investment

	Tobit		IV Tobit			
			Reported Investment		Adjusted Investment	
	(1)		(2)		(3)	
Negative equity	-0.463	*-37*	-1.334**	*-74*	-1.408**	*-76*
	(.308)		(.484)		(.514)	
Log Household income	0.581**		0.582**		0.609**	
	(.085)		(.085)		(.091)	
Age of household head	0.024**	*2*	0.020**	*2*	0.015	*1*
	(.009)		(.009)		(.01)	
Education of household head:						
High school graduate	0.199	*22*	0.197	*22*	0.171	*19*
	(.401)		(.401)		(.425)	
Some college	1.184**	*227*	1.185**	*227*	1.150**	*216*
	(.401)		(.402)		(.426)	
College graduate	1.612**	*401*	1.615**	*403*	1.591**	*391*
	(.386)		(.386)		(.409)	
Graduate school	2.102**	*718*	2.088**	*707*	2.006**	*643*
	(.42)		(.42)		(.446)	
Race/ethnicity of household head:						
Black	-0.591*	*-45*	-0.562*	*-43*	-0.686*	*-50*
	(.346)		(.346)		(.367)	
Hispanic	-0.869**	*-58*	-0.858**	*-58*	-0.945**	*-61*
	(.291)		(.291)		(.309)	
Asian	-1.538**	*-79*	-1.539**	*-79*	-1.727**	*-82*
	(.452)		(.452)		(.48)	
Other	0.312	*37*	0.263	*30*	0.159	*17*
	(.748)		(.748)		(.794)	
First year in home	1.320**	*274*	1.345**	*284*	1.398**	*305*
	(.342)		(.343)		(.364)	
Number of rooms in home	0.137**	*15*	0.134**	*14*	0.134**	*14*
	(.036)		(.036)		(.039)	
New home (≤3 years old)	-1.370**	*-75*	-1.339**	*-74*	-1.408**	*-76*
	(.388)		(.389)		(.412)	

Notes: Unconditional marginal effects with standard errors given in parenthesis. Percent changes are reported in bold italics and are calculated as $e^{\beta}-1$ where β is the unconditional marginal effect. Adjusted investment estimates the total project cost if contracted out for projects where owner supplied all or some of the labor. The left-out category consists of white household heads with less than a high school education living in a home built more than 3 years ago and with a current tenure in the house of two or more years. Number of observations is 5,352.

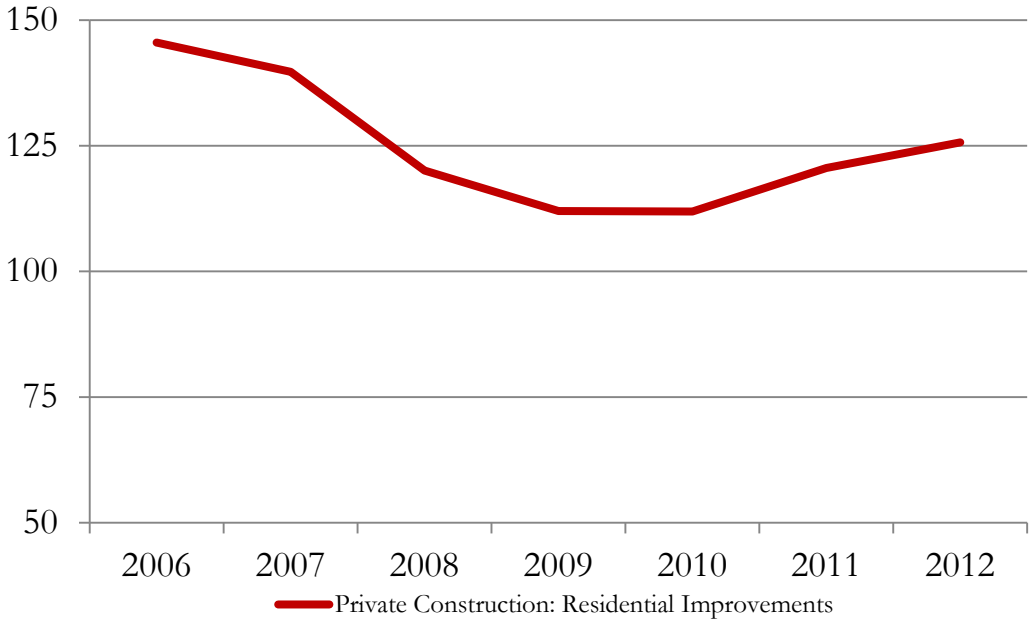
** significant at the 5 percent level * significant at the 10 percent level

Table 5: Tabulations of Negative Equity, Recourse, Financial Assets & Deficiency

		Recourse			
	4,623	0	1		
Negative	0	805	3,448	Financial Assets > $10k	
Equity	1	100	270	0	1
		Deficiency	0	71	17
		> $10k	1	142	40

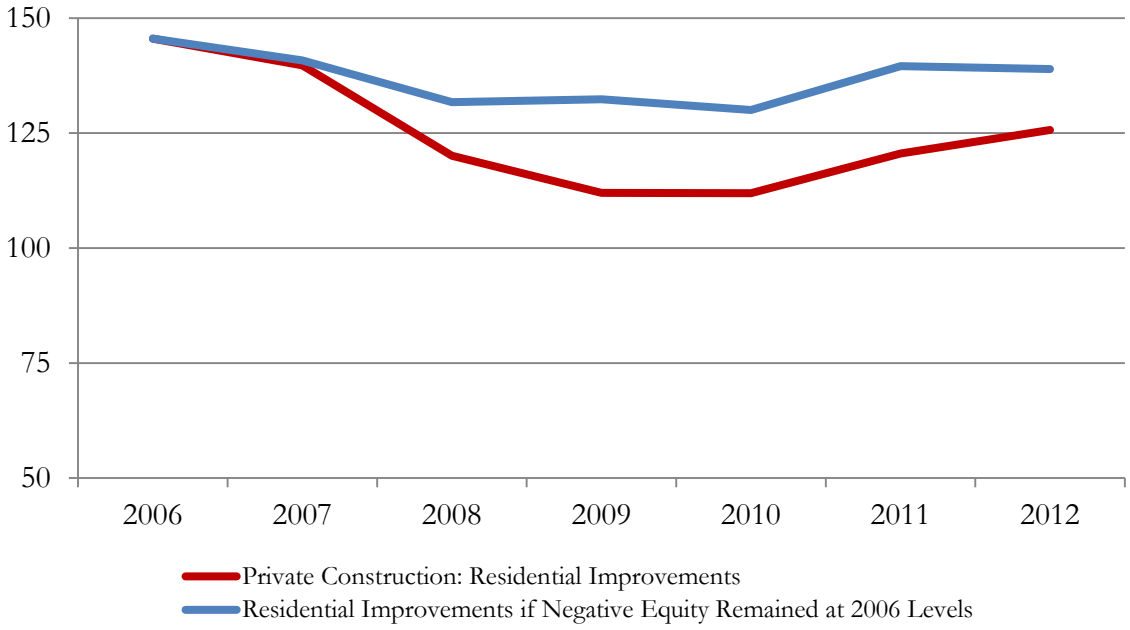
Notes: Sample size decreases by 729 observations where the state of residence is not identified.

Figure 1. Aggregate Residential Improvement Expenditures, $B



Source: Census Bureau, Haver

Figure 2. Impact of Negative Equity on Aggregate Housing Investments, $B



Source: Census Bureau, Core Logic, Consumer Expenditure Survey, and Authors' Calculations

Table A1: Determinants of Housing Investment

	Tobit	IV Tobit	
	Reported Investment		Adjusted Investment
	(1)	(2)	(3)
Negative equity	-1,072**	-2,610**	-3,892**
	(512)	(806)	(1,034)
Household income ($1000)	21**	21**	28**
	(2)	(2)	(3)
Age of household head	36**	29*	11
	(15)	(16)	(20)
Education of household head:			
High school graduate	-235	-238	-1,100
	(669)	(670)	(854)
Some college	978	981	463
	(669)	(669)	(854)
College graduate	1,483**	1,489**	1,018
	(645)	(646)	(824)
Graduate school	1,998**	1,976**	1,134
	(703)	(704)	(900)
Race/ethnicity of household head:			
Black	-915	-866	-1,179
	(573)	(574)	(736)
Hispanic	-1,087**	-1,071**	-1,785**
	(484)	(485)	(622)
Asian	-2,274**	-2,279**	-3,525**
	(752)	(753)	(969)
Other	973	888	1,706
	(1,231)	(1,233)	(1,576)
First year in home	2,535**	2,578**	3,487**
	(562)	(563)	(721)
Number of rooms in home	166**	161**	144*
	(60)	(60)	(78)
New home (≤3 years old)	-2,281**	-2,228**	-2,737**
	(648)	(649)	(831)

Notes: Adjusted investment estimates the total project cost if contracted out for projects where owner supplied all or some of the labor. Unconditional marginal effects are reported with standard errors given in parenthesis. The left-out category consists of white household heads with less than a high school education living in a home built more than 3 years ago and with a current tenure in the house of two or more years. Number of observations is 5,352.

** significant at the 5 percent level * significant at the 10 percent level

Table A2: Determinants of Price/Replacement Cost <1

	IV Probit Marginal Effect
Negative equity	0.264**
	(.104)
Household income ($1,000)	-0.003**
	(.000)
Age of head (10 years)	-0.028
	(.021)
Education of household head:	
High school graduate	0.127
	(.084)
Some college	0.115
	(.085)
College graduate	0.101
	(.082)
Graduate school	-0.061
	(.091)
Race/ethnicity of household head:	
Black	0.496**
	(.075)
Hispanic	0.102*
	(.061)
Asian	-0.409**
	(.107)
Other	0.092
	(.161)
First year in home	-0.015
	(.076)
Number of rooms in home	-0.014
	(.01)
New home (≤3 years old)	-0.231**
	(.086)

Notes: IV Probit marginal effects. Instrument for negative equity indicator using survey 3-5 negative equity indicators. Number of observations 5,352. Average incidence of P/C < 1 is 38%.

** significant at the 5% level, * significant at the 10% level

Table A3: Determinants of Housing Investment Controlling for Price Relative to Replacement Cost

	Tobit	IV Tobit	
	Reported Investment		Adjusted Investment
	(1)	(2)	(3)
Negative equity	-996**	-2,550**	-3,828**
	(512)	(806)	(1,035)
House based Q <1	-861**	-822**	-886**
	(282)	(283)	(363)
Household income ($1000)	20**	20**	27**
	(2)	(2)	(3)
Age of household head	35**	28*	10
	(15)	(15)	(20)
Education of household head:			
High school graduate	-210	-215	-1,074
	(668)	(668)	(853)
Some college	993	995	479
	(667)	(668)	(853)
College graduate	1,496**	1,501**	1,032
	(643)	(644)	(822)
Graduate school	1,962**	1,940**	1,096
	(702)	(703)	(898)
Race/ethnicity of household head:			
Black	-754	-712	-1,014
	(575)	(576)	(738)
Hispanic	-1,030**	-1,015*	-1,724**
	(483)	(484)	(621)
Asian	-2,389**	-2,388**	-3,644**
	(752)	(753)	(969)
Other	1,004	918	1,738
	(1,229)	(1,230)	(1,574)
First year in home	2,527**	2,570**	3,479**
	(561)	(562)	(720)
Number of rooms in home	162**	157**	139*
	(60)	(60)	(77)
New home (≤3 years old)	-2,350**	-2,293**	-2,808**
	(648)	(649)	(831)

Notes: Unconditional marginal effects are reported with standard errors given in parenthesis. The left-out category consists of white household heads with less than a high school education living in a home built more than 3 years ago and with a current tenure in the house of two or more years. Number of observations is 5,352.

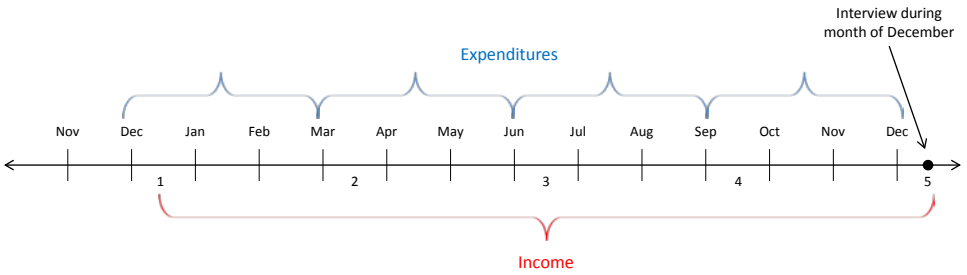
** significant at the 5 percent level * significant at the 10 percent level

Data Appendix

A.1 Using the CES

In our study we use several of the Consumer Expenditure Survey's Interview data files, including Family (FMLY) and several of the Expenditure (EXPN) files. The FMLY dataset is used to gather household-level information, including income, education, age, race, and family size, as well as housing characteristics. The EXPN files are used to collect information on household consumption, credit liability, mortgages and home equity loans, and property value.

One of the main benefits of the CES is its detail for both expenditure and income variables. One downside is that the questions for expenditure and income variables may produce inconsistent time periods. When a household is asked to report recent expenditures, the question specifies that the household consider all expenditures made since the first of the month three months prior to the interview month. For example, if a household is interviewed in December of 2009, they must report their expenditures for September, October, and November 2009. This reference period between September 1 and November 30 is the same regardless of the date of the interview during the month of December. When a household is asked to report income, however, the reference period is less clear, and only asks the respondent about income received during the past 12 months. Because people can be interviewed at different times within a month, the reference period for income can vary according to individual interpretation. This ambiguity in income reference period, and its imprecise relationship to that of expenditures, is illustrated in the timeline below.



A.2 Challenges with the self-reported property value

In our study we use self-reported property values and mortgage balances to construct negative equity indicators for households, but several issues with the CES's variable for property value, *propvalx,* should be noted. Respondents are asked to report property value quarterly, and when raw property value data is not

available or is deemed invalid, a rough imputation using "hot deck" procedure with broad-based income and regional measures (i.e. "North", "South", "East", "West") is used to fill in values. One basic measurement error arrives with the use of this "hot deck" imputation, but even in cases without imputation, implausible property values and inconsistencies across interviews exist for many household. Some examples are included below.

1. Example of the "hot deck" imputation procedure producing implausible property values across interviews.

newid	cuid	interi	propvalx	prop_alx	total mortgage balance
1838762	183876	2	70,000	D	15,267.33
1838763	183876	3	225,000	F	15,226.00
1838764	183876	4	109,000	D	15,183.67
1838765	183876	5	95,000	D	15,140.33

Cuid is the household identifier, *interi* is the interview number, *Newid* is *cuid* concatenated with the *interi*, *prop_alx* is a code for whether *propvalx* was imputed (D = not-imputed, F = imputed), *total mortgage balance* is the combination of all outstanding mortages, home equity lump sum loans, and HELOCs. In this example, it looks as if the property value tripled over the course of one quarter between the second and third interview. Part of the explanation in this example is that the second interview's property value was imputed using the "hot deck" procedure; however even the reported values of interview four and five seem unlikely. This is an example with multiple types of measurement error. To avoid the measurement error introduced by the hot deck approach, we require reported property values in the second interview to be non-imputed.

2. Example of a dramatic change in non-imputed property values across interviews.

newid	cuid	interi	propvalx	prop_alx	total mortgage balance
1871302	187130	2	300,000	D	95,510.34
1871303	187130	3	700,000	D	90,956.66
1871304	187130	4	500,000	D	186,844.70
1871305	187130	5	300,000	D	83,172.66

newid	cuid	interi	propvalx	prop_alx	total mortgage balance
1879732	187973	2	80,000	D	70,712.00
1879733	187973	3	225,000	D	65,719.00
1879734	187973	4	125,000	D	66,393.00
1879735	187973	5	115,000	D	45,143.33

In this first example, both the property variable and the mortgage balance vary greatly across interviews. Both households (187130 & 187973) provide examples of a common trend in the data: households tend to report higher property value in the third and fourth interviews, returning back to near interview two levels in the fifth interview. This widespread mean reversion promotes the use of third, fourth, and fifth interviews as instruments for the second interview.

3. Example of scaling issue in secondary home property value.

newid	propvalx	prop_almx	qyear	prop_noi
2022332	195000	D	20093	1
2022333	300000	D	20094	1
2022334	15000	D	20101	1
2022332	20000	D	20093	2
2022333	200000	F	20094	2

Note that *prop_noi* is the property number with a value of "1" indicating the home is the primary residence. Each of these values were reported by the same member of the same household (*cuid*, or household id, = 202233 from the 2009 and 2010 datasets). The data for this household seems to imply that they own one primary home and one vacation home, and the self-reported property values for each of those properties are incredibly volatile. This household would not be included in our estimation because we only include primary homes but is included here for descriptive purposes.

If we look at property number two, the imputed house value of $200K seems more reasonable than the self-reported $20K value for a house in 2009Q4. The error for property number two here could be a scaling issue, however the majority of problems are not resulting from scaling issues. Most of the measurement error in the property value variable takes the form of property one in this example (similar to the examples in #2 above), though there are a few examples of simple scaling error.

4. Survey to survey ratios of each consecutive interview's property value and mortgage balance compared to the first interview's values.

Survey to Survey Ratios	N	mean	sd	min	p1	p25	p50	p75	p99	max
2nd interview's property value / 1st interview's property value	16980	1.02	0.28	0.10	0.50	0.93	1.00	1.06	2.00	6.67
3rd interview's property value / 1st interview's property value	16980	1.01	0.30	0.13	0.48	0.91	1.00	1.06	2.00	7.50
4th interview's property value / 1st interview's property value	16980	1.00	0.28	0.11	0.49	0.89	1.00	1.06	1.93	6.36
2nd interview's mortgage balance / 1st interview's mortgage balance	14280	1.15	3.45	0.00	0.06	0.98	0.99	1.00	4.31	192.50
3rd interview's mortgage balance / 1st interview's mortgage balance	14280	1.26	6.93	0.00	0.00	0.96	0.99	0.99	4.79	336.00
4th interview's mortgage balance / 1st interview's mortgage balance	14280	1.47	17.47	0.00	0.00	0.95	0.98	0.99	4.42	1006.67

These ratios apply to our final estimation sample. Overall averages show both the property value and mortgage balance remaining fairly consistent over the course of the four interviews. While there are several problem examples similar to the ones discussed in this section, our use of the second, third, and fourth interview negative equity instruments helps to decrease the measurement error.

A.3 Topcoding and imputation of variables in our estimation

Many variables in the CES have been either topcoded, imputed, or both. Topcoding involves the replacement of data for instances where the value of the original data exceeds a defined critical value. Each observation that falls outside the critical value is replaced with a topcoded value that represents the mean of the subset of all outlying observations. All five quarters of data in the CES microdata release are used when calculating the critical value and topcode amounts. The replacement of these data is meant to protect users at the very high end of the income distribution from being identified by users.

Imputation in the CES refers to various methods depending on the variable. Information specific to the variables in our estimation is included below.

Household income: fincbtax

The variable for total household income before taxes (*fincbtax)* is comprised of many subcategories of income variables. Each of these components is subject to its own topcoding and imputation process. Additionally, some of these components are comprised of individual level income variables that have been aggregated to the household level. At times, the imputation or topcoding is applied to these individual level variables prior to aggregation. The CES interview user documentation provides detailed information on this. To give an example, let's consider the largest component of household income, household wage and salary (*fsalaryx*). In our estimation sample, 2.5% of households have at least one member of the household's wage and salary topcoded. The member level wage and salary variable (*salaryx*) is top-coded if over or equal to $150,000 and replaced with $279,006.[50] Since total income is the sum of other variables, it's possible for some of the components to be top-coded while others are not (eg, the husband's but not the wife's wage and salary is topcoded).

The series of income variables in the CES that end in "x" as opposed to "m" (*fincbtax, fsalaryx, fnonfarmx*) are dubbed non-imputed by the BLS, while the corresponding variables ending in "m" (*fincbtxm*, *fsalaryxm, fnonfrmm*) were introduced in 2004 and involve a multiple imputation method that allows income values to be estimated when respondents decline to report. In 2004 and 2005, the BLS removed the non-imputed versions of these variables in public data, only reporting the imputed series that end in "m". The non-imputed versions can be mostly reconstructed using the member level data in these years, though replicating identically is impossible due to topcoding.

In our estimation, we only use the non-imputed version of total household income. While the BLS reports it as non-imputed, it does contain components that include some bracket based imputations for respondents who refused to answer the various income questions and were given the option to select a bracket instead. Below is a list of the bracket options for the member level wage and salary variable (*salaryb*) from the 2010 data dictionary.

[50] Topcode critical values vary each year, and these numbers apply to 2010.

Bracket Options
01 $0 - $4,999
02 $5,000 - $9,999
03 $10,000 - $14,999
04 $15,000 - $19,999
05 $20,000 - $29,999
06 $30,000 - $39,999
07 $40,000 - $49,999
08 $50,000 - $69,999
09 $70,000 - $89,999
90,000 - $119,999
$120,000 and over

The method for converting the bracketed data into a single data entry is to take the pool of actual, non-bracketed responses from the last 3 years that fall into the bracket and take the median. This process is rolling. There are similar bracket options for other components of total household income as well, and this rolling median process is also applied to them.

Reported Property Value: propvalx

We use the reported property value variable (*propvalx*) to generate our negative equity indicator. In addition to the "hotdeck" imputation process described in section A.2, this variable is also subject to topcoding. In our benchmark specification, we impose the restriction that the initial property value reported in the 2nd survey must not be topcoded or imputed. The property value in the 3rd, 4th, and 5th interviews' property value could have been topcoded. Though worth mentioning, the imputation of this variable is not very prevalent: 90% of our benchmark estimation sample have all four of the propvalx non-imputed or topcoded; 97% of our estimation sample have at least 2 of the remaining 3 interviews with a non-imputed or topcoded propvalx; 99% of our estimation sample have at least 1 of the remaining 3 interviews with a non-imputed or topcoded propvalx.

Total Mortgage Balance

Our constructed total mortgage balance is also used to generate our negative equity indicator**.** This variable is constructed from 7 types of MOR/HEL/HELOC variables, 6 of which could have been individually topcoded if the reported value was above the prescribed critical value for that specific expenditure. Since the mortgage balance data was reasonable, we have never excluded observations based on imputed or topcoded mortgage balance components.

Housing Investment

This variable is constructed from 15 types of expenditure variables in the *Construction, Repairs, Alterations and Maintenance of Owned or Rented Property* (CRB) expenditure dataset. Of these 15 variables, 9 could have been individually topcoded if the reported value was above the prescribed critical value for that specific expenditure. Of note, each of these 9 expenditure categories has a different prescribed critical value depending on the project type associated with the expenditure.

Project Type

The variable *crmcodeb* in the *Construction, Repairs, Alterations and Maintenance of Owned or Rented Property* (CRB) expenditure dataset was used to define the type of projects for which households invested. Some related projects were combined when estimating the project discount factors presented in Table 2 and used to calculate the adjusted investment analyzed in Table 4, A1, and A3. The project code for siding (270), insulation (240), and plastering or paneling (190) were combined. Roofing, gutters, or downspouts (260) were combined with masonry, brick or stucco work (290). Remodeling one or more rooms in the house (130) and finishing a basement or an attic or enclosing a porch (120) were considered the same project type. Flooring repairs or replacement (230), wall to wall carpet installation and replacement (231 and 232, respectively) were also combined. Finally, other improvements (300) and combined job codes (310) were categorized together.

State

The value of the variable *state* is suppressed for some observations in order to meet the Census Disclosure Review Board's criterion regarding identity protection. The 2010 data dictionary reports that approximately 17 percent of the records in the family level datasets are left blank. In addition, approximately 4% of the state records have been completely recoded as a different state, and beginning in 2005, this recoding was not flagged. Thus, any analysis reliant on state identification should be interpreted with this measurement error in mind.

A.4 Restricting the Sample

As detailed in the data section, a few integral variables lacked sufficient characteristics in early years of the CE, and our estimation sample was selected around this. Our estimation sample begins with the second quarter of 2007 and concludes with the first quarter of 2012.[51] The sample includes primary housing units owned by a member of the CU. At this stage, there are 29,806 potential households.

We first enforce restrictions for the self-reported property value and the mortgage balance. Our estimation includes homes valued between $30,000 and $1,000,000 with loan-to-value ratios above 2. This takes our sample down to 28,110 households. To avoid the measurement error introduced by the hot deck imputation of the self-reported property value variable, we require reported values in the second interview to be non-imputed or top-coded. This eliminates 4,547 households. We also restrict the sample to single family detached homes, bringing the total down to 19,897 households.

To increase the number of non-zero reported housing investment, we aggregate expenditure for each household across all four quarters. Aggregating to annual data excludes any household with fewer than four interviews and constitutes our largest sample restriction. At this stage, the sample includes 9,993 households.

Finally, we restrict the sample to exclude those receiving unemployment insurance, including only those with reference person age 20 to 59 and with positive total household income. Accounting for missing responses to some right hand side variables, our final estimation sample includes 5,352.

[51] As of this writing, the 2011 microdata are the most recently available. Each year's worth of data includes 5 quarters, 4 for the year associated with the data disc and 1 for the subsequent year. In all years aside from 2012, we use all four quarters from the same year's data disc because BLS makes coding rules (e.g. topcoding) consistent within discs but subject to variance across discs. The first quarter of 2012 from the 2011 data disc is included in our sample.

Table A4: Restricting the Sample

Sample Restrictions	Number of Households
1. Our estimation sample begins with all primary housing units owned by a member of the CU. The sample is also restricted to only include homes with no more than 3 mortages, 3 home equity lump sum loans, or 3 home equity lines of credit. We include interviews from the following quarters: Q2-2007 through Q1-2012.	29,806
2. We exclude properties with house values less than $30,000 or greater than $1,000,000, along with properties with loan-to-value ratios above 2. This also excludes anyone with missing propvalx or mortgage balance (if they report a mortgage).	28,110
3. Include only households with a non-imputed property value in the first interview.	23,563
4. Include only single family detached homes.	19,897
5. Include only households who had all 4 interviews conducted in the following time frame: Q2-2007 through Q1-2012.	9,993
6. We further restrict the data to include only those with reference person age 20 to 59 inclusive.	6,582
7. Exclude observations for which total income <=0.	5,956
8. Exclude households receiving unemployment insurance.	5,607
9. Exclude households with missing covariates.	5,352

Additional Notes:

2. If a household is listed as an owner but does not have a mortgage, they are included in the sample as someone with positive equity.

5. Maintenance expenditure includes all 4 quarters for which the HH responded. Note: expenditure questions cover the 3 months prior to the interview month (regardless of where within the month the question falls) – this means some observations cover expenditure in Q1-2007. All RHS variables are the reported values in the household's first interview. This means the time frame for these responses is from Q2-2007 through Q2-2011. Since income questions cover the 12 months prior to the interview month, this means observations include income from Q2-2006 through Q2-2010.