Is the Behavior of Sellers with Expected Gains and Losses Relevant to Cycles in House Prices?*

Tingyu Zhou, Florida State University, tzhou@business.fsu.edu**

John M Clapp, University of Connecticut, John.Clapp@uconn.edu

Ran Lu-Andrews, California Lutheran University, rluandrews@callutheran.edu

December 20, 2019

Abstract

We examine anchoring to the price paid at purchase during an important cycle, 2000-2017. Our repeat sales model, corrected for endogeneity and unobserved heterogeneity, provides robust estimates of negotiated price premiums (discounts) of sellers with expected losses (expected gains). I.e., it estimates the influence of anchoring to price paid by individual sellers. This paper extends the anchoring literature which has focused on individual behavior, and it supports new stylized facts associated with housing market cycles.

We associate individual anchoring with the aggregate price cycle using a model that multiplies negotiated premiums and discounts by the magnitudes of those quantities and by the proportion of sales with expected losses and gains. Results suggest that anchoring was associated with reductions in observed changes in house prices during the boom (2004-2006) as sellers with gains dominate with their price discounts, and with reduced price declines during the bust (2007-2012) when the behavior of those with losses becomes important.

Additional results that make minimal model assumptions suggest that loss behavior is statistically significant and important at turning points, i.e., during the transition from a boom to a bust. Double mean differences suggest that those with losses used the mild recovery as an opportunity to realize losses at reduced premiums after long delays.

Results are robust to substituting asking price for sales price, to restricting the sample to repeat sales, to including loan-to-value ratio, to alternative definitions of cyclical phases and to alternative method of correcting for unobserved quality.

Keywords: Housing Cycles, Anchoring, Reference Dependence, Price-volume Relationship, Cyclical Turning Points

JEL Codes: E00, G11, N2, R3

*Acknowledgements: We thank seminar participants at the Amsterdam University, Baruch College, University of Cambridge, Concordia University, University of Florida, Florida State University, Free University, University of Groningen, University of Reading, the Urban Economics Association (UEA), the American Real Estate and Urban Economics Association (AREUEA) and the Weimer School. We particularly thank David Genesove, Katya Salavei Bardos, Casey Dougal, Marc Francke, Ed Glaeser, Christian Hilber, Lu Han, John Leahy, Zhenguo Lin, David Ling, Thies Lindenthal, Jordy Meekes, Jan Rouwendal, Mike Sklar, Eva Steiner, Will Strange, Geoffrey Turnbull, Paul Willen, Yildiray Yildiram and Kelvin Wong for insightful comments.

** Corresponding author

1. Introduction

The growing literature relevant to micro foundations of housing market cycles has developed a significant body of theory and empirical results. Several authors have documented stylized facts for the house price cycle: 1) too much volatility in prices relative to changes in fundamental value (i.e., house values based on income, employment, financing and construction costs); 2) short term positive serial correlation; 3) longer term mean reversion; 4) positive correlation between house prices and the volume of transactions. Structural models explaining housing market cycles have been designed to fit these facts.[1]

A substantial literature shows that the price paid to purchase a house influences the asking price set by sellers at the end of their holding periods, and that this influences negotiated sales prices (Stein, 1995; Yavas and Yang, 1995; Genesove and Mayer, 2001; Einiö *et al.*, 2008; Han and Strange, 2016; Andersen *et al.*, 2019). Anchoring to the asset purchase does not have to reflect psychological loss aversion or irrational behavior.[2] It might be due to financial or tax frictions or limited ability to obtain and process information as suggested by Stein (1995), Barberis and Xiong (2009), Ben-David and Hirshleifer (2012), Anenberg (2016) and Andersen *et al.* (2019). When sellers anchor their decisions to the price paid for their asset, they consider the expected gain or loss over their holding period when deciding whether to sell and at what reservation price. Evidence from the directed search model (Han and Strange, 2016) shows that sellers' asking prices influence prices paid by buyers.

In this paper, we establish new stylized facts relating anchoring behavior to aggregate house price movements. We find that anchoring to the price paid to purchase a house is associated with substantial reductions in the observed change in house prices during the boom as sellers with gains dominate with their price discounts and with substantially reduced price declines during the bust when the price premiums negotiated by those with losses become important.

[1] The roots of this literature are in Case and Shiller (1990, 2003) and Stein (1995). A review of the literature on microeconomic behavior contributing to housing cycles is provided by Davis and van Nieuwerbergh (2015). Studies most influential here include Han and Strange (2016), Anenberg (2016), Carrillo (2013), Favara and Imbs (2015), Campbell et al. (2011), Coulson and Zabel (2013); Glaeser and Nathanson (2014, 2017), Ortalo-Magné and Rady (2006), Favilukus and van Nieuwerburgh (2018).

[2] The literature attempting to demonstrate psychological loss aversion in real estate markets derives from the seminal work of Genesove and Mayer (GM, 2001). Bokhari and Geltner (2011) examine loss aversion using commercial real estate transactions. Zhou, Clapp and Lu-Andrews (2019) provide detailed explanation of our extensions of GM's anchoring model which deal with omitted variable bias. This paper uses the bias-corrected anchoring model to study the aggregate housing cycle.

We estimate our models with over 540,000 housing transactions and over 90,000 repeat sales pairs in Connecticut. We evaluate a full cycle in house prices from 2000 through 2017 when a repeat sales index increased by more than 40% (ending in 2006), then declined to a level that was 10% lower than 2002, followed by a mild recovery. In our model, the premiums and discounts associated with expected losses and gains (the expected coefficients from the anchoring regressions are positive for losses, negative for gains) are multiplied by the magnitudes of those quantities and by the proportion of sales with expected losses and gains. This produces reduced-form house price indices that adjust house price indices for observed anchoring behavior: we refer to these as “contrast relative indices.”[3]

Using the contrast-relative estimates from our multiplicative model we examine the association between anchoring behavior and each phase of an important cycle in house prices, 2000-2017. We show the mix of sellers with expected gain/loss, a mix that changes dramatically over the house price cycle, together with the response of sellers to gains and losses suggests new stylized facts: i.e., the behavior of these sellers is relevant at the macro level. This contrasts with previous literature on anchoring which has shown substantial effects at the individual transactions level but has not shown an association between this behavior and the aggregate level of house prices.[4] If we had found little or no reduced-form association, then there would be no reason to add anchoring to the stylized facts listed above.

We address two potential threats to our empirical analysis. First, the anchoring effect may not be identified separately from omitted variables such as time on the market (TOM). For example, in a booming market, sellers may plausibly expect high opportunity cost of waiting. If so, they will set low asking prices to lower TOM. To mitigate this concern, we assert the testable

[3] Maldonado and Greenland (2002) rigorously define causal contrast relative statements for epidemiological studies. That paper clarifies that our comparison of change in indices adjusted for anchoring behavior to changes in unadjusted indices is a contrast relative association, a reduced form relationship which reflects some unknown causal process. For further discussion of contrast relative associations see Menzies (2017).

[4] Bracke and Tenreyro (2016) use over twenty million transactions in England and Wales to produce some evidence that aggregate fluctuations in sales volume are influenced by anchoring behavior. Andersen *et al*. (2019) use a rich database to explore the association between losses and gains, listing premia, down payment constraints and time on the market. Some of the most persuasive evidence in these two papers is based on minimal theory, a strategy also employed by us. These papers analyze seller decisions at the micro level whereas we focus on the relationship between expected loss/gain and the aggregate house price cycle. Our model is related to Bokhari and Geltner (2011) who explore the effects of loss aversion behavior on aggregate market cycles in commercial real estate prices. They find little association between loss aversion and the market cycle. Our findings differ because we examine loss and gain (anchoring) behavior, not loss aversion, in the housing market.

assumption that all sellers are responding simultaneously to any time-dependent variables such as demand shocks or changes in financial constraints when they set TOM or make listing or withdrawal decisions, whereas each seller has her own anchor which influences selling decisions.[5] To test this hypothesis we include town-year fixed effects. Towns in Connecticut are small municipalities that govern local public schools, other services and property taxes, providing unusually well-designed controls. Town-year interacted fixed effects capture many sources of time-varying unobserved town-level heterogeneity. The resulting estimates, and experiments with a range of county and year fixed effects, support our hypothesis.

The second threat we examine is that sellers, and buyers who inspect the property, observe characteristics not known to the econometrician.[6] These omitted quality characteristics bias estimated losses and gains to be larger than true losses and gains causing loss and gain parameters to be biased upward in absolute value. To mitigate this concern, we follow and extend Zhou *et al*. (2019) and use property tax assessed values at the time of sale to control the effects of unobserved quality on sales prices expected by sellers and on their expected loss or gain, and we find extensive evidence that the assessor does control unobserved heterogeneity. If unobserved quality can be controlled then coefficients on loss and gain explanatory variables based on assessed values will be reduced, and that is what we find.[7] Due to the large differences between the estimated results following previous literature and quality-adjusted results using assessed value, we rely on the latter to draw our conclusions.

Despite the reduced coefficients on loss and gain explanatory variables (i.e., premiums on expected losses and discounts on gains are both smaller in absolute value) based on assessed values, we still find that these coefficients are statistically significant and that the loss and gain effects are large. This is because, even with small loss and gain coefficients, the imbalance of

---

[5] Seller specific anchors will influence their decisions about when to sell, list prices, withdraw a listing or wait for a buyer as shown by Andersen *et al.* (2019). Anchoring influences these decisions over the cycle which implies that anchoring is relevant to the cycle. We simply claim that our reduced form results support a significant influence of anchoring behavior on the housing market cycle independently of the interaction between anchoring and other variables. Here, we include original loan amount as one of the anchoring variables.

[6] Examples of quality typically unobserved by the econometrician include an above average kitchen or attractive landscaping as well as a view or location on a busy street. We reject the more common term, "unobservables" because it misses the fact that unobserved qualities are problematical when they are reflected in market prices (or other variables) but not known to the econometrician.

[7] Clapp and Zhou (2019) also shows an order of magnitude upward bias in estimated coefficients on expected loss, where the upwardly biased coefficients are based on Genesove and Mayer's (2001) model.

proportion between loss and gain could lead to a large impact on aggregate changes in transactions prices. As long as the magnitudes and proportions are large, we do not need a large coefficient to get the big effects on the cycle.

During the normal and boom periods when 95% of sales have quality-adjusted expected gains, contrast relative changes in house prices were between 8% and 24% higher than those observed, suggesting that loss/gain behavior dampened price increases. During the bust, when about 48% of sales experienced expected losses, contrast relative price changes are over 30% lower than observed. Loss behavior dominated during the mild recovery when it was associated with increases in the observed price change.

Results using asking price supports an influential role for seller behavior. Han and Strange (2016), Carrillo (2013) and others argue that asking price conveys an important signal to the market. Information in the market (e.g., inquiries by potential buyers) is strongly influenced by asking price (Carrillo, 2013). Our results based on sales prices are robust to substituting asking prices.

We report reduced form patterns with limited implications for causation. Premiums and discounts associated with anchoring, magnitudes, and number of transactions with expected gains and losses are determined simultaneously with house prices. Causal relationships require structural models; our point is that these models need to be consistent with the new stylized facts documented here, whereas this would not be necessary if we had found no association. I.e., in a world without loss and gain behavior we should observe little effect from our contrast relative associations.

The rest of the paper is organized as follows. The next section summarizes our bias-correction addition to the literature on loss/gain behavior. Section 3 develops our contrast-relative analysis. Sections 4 and 5 present data and empirical results. Section 6 concludes.

2. Models of Loss/Gain Effects over a Housing Cycle

2.1 The Genesove and Mayer (2001) Model

Genesove and Mayer (GM, 2001) model the effect of reference dependence on house prices, providing a well-established basis for our model. In a repeat sales framework, GM use a standard hedonic model in order to find the expected sale price at the time of the second sale. We modify GM's model by including town-year dummies to control time-varying spatial heterogeneity:

$$P_{ilt} = \beta_0 + \beta X_{il} + FE_{lt} + \varepsilon_{ilt} \quad (1)$$

where P_{ilt} is the natural log of sales price of property *i* at location *l* in time *t*, X_{il} is a vector of property and locational variables and ε_{ilt} is a zero-mean disturbance term. We control for time and spatial effects, notably for variation in local public services and taxes, with a dummy for each year in each town (i.e., town-year fixed effects, FE_{lt}). Equation (1) is the first stage of a two-stage model. The price predicted from equation (1) defines "market value": $\hat{P}_{ilt}$ is the most likely sales price based on information typically available in the market.

In the second stage, GM test for anchoring with repeat sales (a subset of the sales used in the first stage, equation (1)) to model the second sales price:[8]

$$P_{ils} = \beta_0 + \beta_s \hat{P}_{ils} + \alpha_l \left(P_{ilp} - \hat{P}_{ils}\right)^{+} + \alpha_g \left(P_{ilp} - \hat{P}_{ils}\right)^{-}$$
$$+\alpha_m M_{s-p} + \alpha_q \hat{\varepsilon}_{ilp} + FE_{ls} + \varepsilon_{ils} \quad (2)$$

Here, *p, s* indexes the first (*p*) and second (*s*) price of a repeat pair. I.e., equation (2) is estimated for repeat pairs, whereas equation (1) is estimated from all sales including one-only.[9] $(P_{ilp} - \hat{P}_{ils})^{+}$ is the expected loss, calculated as the positive part of the difference between purchase price, P_{ils} and expected sales price, $\hat{P}_{ils}$. $(P_{ilp} - \hat{P}_{ils})^{-}$ is the absolute value of the negative part, the expected gain. These variables are entered because sellers may anchor to the price paid, P_{ilp}. M_{s-p} is number of months between first and second sale. GM include $\hat{\varepsilon}_{ilp}$, the estimated residual from equation (1) for the first transaction, as a noisy proxy for unobserved characteristics such as a good view or a particularly attractive kitchen.

Anchoring to the first sales price, P_{ilp}, is measured by non-zero coefficients for α_l and α_g. If the first price is above (below) the market value of the second sale, $\hat{P}_{ils}$, then reference dependence would say that the seller insists on a randomly high (buyers can negotiate a low) price. I.e., the α's are not equal to zero because of anchoring, a testable assumption strongly supported by previous literature. We take no position on whether this is due to loss aversion or to constraints due to limited information, liquidity or mortgage repayment; our purpose is to focus on the association between an aggregate housing cycle and loss and gain behavior.

---

[8] This model is used in numerous studies on anchoring effect, including Beggs and Graddy (2009), Bokhari and Geltner (BG, 2011), Anenberg (2011) and others.

[9] We re-ran our analysis by restricting the first stage estimation to repeat sales sample only and find similar results. See Appendix 9.

2.2 Identifying the Anchoring Effect and Controlling Unobserved Heterogeneity

Empirical studies of anchoring typically use time and price of the first sale and information on the mortgage (or loan-to-value ratio), all variables unique to each purchase. This setting assume variables that occur at a later point in time such as loan availability, supply or demand shocks or unanticipated changes in the business cycle are independent of historical purchase information.[10] To identify anchoring separately from unanticipated events occurring later, we flood our model with town-and-year interacted fixed effects, the FE_{ls} in equation (2). Since Connecticut is a small state divided into 169 small towns controlling local services and taxes, this gives us a unique opportunity to test the ability of historical data on the first sale to control time-and-space dependent omitted variables. The resulting estimates control time-varying location-specific drivers of unobserved heterogeneity such as motivation to sale and time-on-market. Identification derives from within town-year variation in individual seller anchors.

Zhou *et al*. (2019) show that, when the econometrician estimates an expected loss (purchase price is above expected sales price), all or part of the loss may be due to high unobserved quality: i.e., the true expected price is higher than the measured one and this might explain the premium attributed to anchoring, and this holds in reverse for gains. Moreover, the inclusion of the first residual, $\hat{\varepsilon}_{ilp}$ in equation (2) may not fully capture unobserved heterogeneity.

To deal with this, we follow and extend Zhou *et al*. (2019) who use assessed value to mitigate unobserved quality. Appendix 1 provides substantial evidence on three characteristics of assessed value (e.g., Clapp and Giaccotto, 1992; Han and Strange, 2016; The Appraisal Institute, 2013). First, the property tax assessor observes many property characteristics unobserved by the econometrician.[11] Second, a unique feature of our data is that assessed value is the one in place as of the date of the sale: the value is predetermined by a 2.5 year average lag for most properties,

[10] Anchoring behavior likely interacts with time-dependent omitted variables as in Andersen *et al.* (2019), and this implies that anchoring is potentially important to the housing cycle. Our claim is to have separated anchor effect from any independent effect of time-dependent variables, and we control loan-to-value ratio.

[11] A typical story might be that unobserved neighborhood characteristics cause correlation between sales prices and gains and losses. The assessor takes most neighborhood characteristics into account as explained in Appendix 1. In Connecticut, the assessor asks to enter the house to verify interior condition, but the homeowner is not required to grant access. This is one reason the assessor imperfectly controls quality. But the assessor considers variables including landscaping, house maintenance, construction quality, heating type and garage type typically omitted by the econometrician: see Appendix 1 for detail.

reducing endogeneity concerns.[12] Third, relative to the seller, the tax assessor is less likely to suffer from psychological biases and to anchor to price paid by any particular seller.[13] It is considered unprofessional for tax assessors to use *any* personal information about sellers or buyers; to do so would fuel the already abundant law suits alleging discrimination. This precludes using the price paid by the seller as a variable determining valuation.

Based on the above discussion, we construct another measure of expected loss, $losslnAV_{ils}$, by substituting $lnAV\_norm_{ils}$ for expected second price, $\hat{P}_{ils}$. With this substitution, equation (2) becomes:

$$P_{ils} = \beta_0 + \beta_s^{av} lnAV\_norm_{ils} + \alpha_l^{av}(P_{ilp} - lnAV\_norm_{ils})^+ + \alpha_g^{av}(P_{ilp} - lnAV\_norm_{ils})^- + \alpha_m^{av} M_{s-p} + \alpha_q^{av}\hat{\varepsilon}_{ilp} + FE_s + \varepsilon_{ils}^{av} \quad (3)$$

where $lnAV\_norm_{ils}$ is normalized assessed value (NAV). The difference between equation (2) and (3) is that we replace $\hat{P}_{ils}$ with $lnAV\_norm_{ils}$. $lnAV\_norm_{ils}$ is calculated using the algorithm described in Appendix 2.

The normalization adjusts for the fact that each town has a potentially different 5-year assessment cycle. All the within town-year variation in $lnAV\_norm_{ils}$ comes from $lnAV_{ilt}$, as required by our identification strategy, and its level is normalized to the town-year average level of sales prices, as required to calculate expected losses based on $lnAV\_norm_{ils}$.

Can tax assessors use the extra information they observe to measure property value more accurately than the econometrician using equation (1)? To investigate this issue, we construct a leave-one-out (LOO) validation exercise comparing predicted values from equation (1) to normalized assessed value. Results of this horse race, reported in Appendix 2, strongly support the superior performance of assessed value. The average out-of-sample mean square error (MSE) of NAV is 0.095 compared to 0.127 for expected price (PV) estimated using the hedonic model, equation (1). The superior performance of NAV varied over the cycle. For example, the average MSE of NAV (hedonic model) is 0.098 (0.123), 0.087 (0.108), 0.095 (0.143) and 0.012 (0.166) in

[12] Each town in Connecticut is required to revalue all property once every 5 years. The 5-year cycle is potentially different for each town, producing an average lag of 2.5 years between the date of revaluation and the sale. A few revaluation companies are certified to do the statistical analysis and rigorous sales ratio studies evaluate the predictive accuracy out-of-sample.

[13] It is possible that the assessor considers price paid as a comparable sale, but only if price paid is within 3 years of the sale. Moreover, a comparable sale is very different than seller anchoring (and bargaining hard) based on financial constraints, loss aversion or information advantages: assessors do not use models such as equation (2).

normal, boom, bust and recovery, respectively. Interestingly, assessed value outperformed the hedonic model most during the bust and recovery when sales with losses were a large percentage of all sales.

3. Our Analytical Framework for Aggregate Market Effects

In this section, we develop Bokhari and Geltner's (BG, 2011) method in a way that allows us to analyze aggregate market effects in the broader context of all loss/gain behavior. Equations (1)–(3) produce reduced form patterns, whereas our analysis produces contrast-relative associations (Menzies, 2017): in a world without loss and gain behavior, our contrast-relative analysis should find little association between this behavior and the aggregate market cycle.

3.1 Contrast-relative Analysis: Market Prices with Loss and Gain Effects Removed

We identify four sub-periods indexed by *j*: normal, boom, bust and recovery. The three variables (i.e., six quantities in total) are first, the premium per dollar of expected loss (discount for gains) given by α_{lj} (α_{gj}) from equation (3) estimated using observations from period *j*; second, ML_t (MG_t), the average amount of the expected loss (gain) conditional on loss (gain) in the market. This is the average of $\left(P_{ilp} - \hat{P}_{ils}\right)^{+}$ ($\left(P_{ilp} - \hat{P}_{ils}\right)^{-}$) during any year *t* of individual seller with expected loss (gain); and third, $\%L_t$ ($(1 - \%L_t)$), the percentage of sales with expected loss (gain).

BG examine the association between loss/gain behavior and the aggregate price movement in pre- and post-crisis for commercial real estate. BG propose a five-step method designed "to produce a loss-aversion-adjusted price index (BG, 2011, p. 664)." For each time period (pre-2007, 2007 and 2008-2009) they estimate α_{lj} (α_{gj}) from the GM model, i.e., our equation (2), and they calculate average values for ML_j (MG_j) and $\%L_j$.

Our method is different from BG in several ways. First, we construct our contrast-relative index adjusted for omitted variable bias. We show that these indices with and without this adjustment are radically different. Second, we adjust for any anchoring, including but not limited to loss aversion. For each regime, BG take the difference of the loss and gain coefficient (to reflect the pure behavioral loss aversion), multiply this difference by the magnitude of loss and then multiply by the proportion of loss. In contrast, we let *both* the loss and gain coefficients interact with their magnitudes and proportions respectively. We multiply sales price premiums or discounts by magnitudes of these quantities and their proportions in the aggregate market to construct contrast-relative loss-gain-free indices for comparison with the observed repeat sales index. Third, we draw conclusions on changes by calculating the changes in the adjustment factor. Lastly, we fit out model over a full cycle of the US housing market from 2001 to 2017 whereas BG used commercial real estate data over a shorter period around the financial crisis.

BG do not find much aggregate effect of loss aversion: the maximum effect is during the bust of 2007, and then it increased the market-wide index by only 1.2%. They conclude that loss aversion has a substantial effect at the individual property level but not at the aggregate level. In contrast, we establish the relevance of loss/gain behavior to the housing cycle.

The intuition behind our multiplicative model is that loss and gain behavior in any time period can be used to produce a contrast-relative index that removes the three variables associated with that behavior. The relationship between a standard repeat sales index, computed using the same method as for the widely applied Case-Shiller index (AI_t= actual or observed index in year *t*), and the contrast-relative index ($CFAI_t$ = contrast-relative factor adjusted index in year *t*) is as follows:

$$CFAI_t = AI_t - (LAF_t + GAF_t) = AI_t - AF_t \qquad (4)$$

where the LAF_t (GAF_t) represents the loss (gain) adjustment factor. The multiplicative model is: $LAF_t = \alpha_{l,j} ML_t \%L_t$ and $GAF_t = \alpha_{g,j} MG_t (1 - \%L_t)$. We define the α, *ML/MG* and *%L* variables (three variables and six statistics for each time period) as described earlier in this section. Ideally, we would allow all the three variables (coefficients, magnitudes, and weights) to vary by each year. Due to insufficient data, we are only able to estimate loss/gain coefficients for each sub-period *j* (normal, boom, bust and recovery). For weights and magnitudes, we calculate the average over all sellers with expected losses (gain) in the market in year *t*. We define $LAF_t + GAF_t$ as the total adjustment factor (i.e. $AF_t = LAF_t + GAF_t$).

In general, the *CFAI* adjusts for the association between expected gain or loss and the observed repeat sales index. We focus on both the *level* and the *change* of *CFAI*. The *level* of *CFAI* provides a contrast-relative term for what market prices would have been had the three factors been zero: it estimates an index with loss/gain behavior set to zero.[14] If *CFAI* lies above (below)

[14] For coefficients, levels of contrast-relative estimates implicitly assume that marginal behavior (coefficients from equation (2)) applies infra-marginally. I.e., our contrast-relative estimate indicates possible outcomes, not predictions. An important reason for levels estimates is to test the null hypothesis of no loss/gain effect on aggregate price indices. Moreover, changes over cyclical phases show the association between changes in loss/gain behavior and the cycle.

AI, so that $LAF + GAF < 0$ and $AF = AI – CFAI < 0$ ($LAF + GAF > 0$ and $AF = AI – CFAI > 0$). As the *CFAI* shows how much lower (higher) the house price would achieve when the loss factor (gain factor) dominates, we will test the significance of any difference between *AI* and *CFAI* (i.e., *AF*). More importantly, we can test the significance of *LAF* and *GAF* magnitudes separately, allowing us to see which type of behavior is more statistically important in each phase of the cycle.

We compare the *change* of *AI* and that of *CFAI* from the previous period in order to estimate contrast-relative fluctuations in house price indices. By construction, $\Delta AI – \Delta CFAI = \Delta AF$. For example, if the increase of *AI* from normal to boom is greater (less) than the change of *CFAI* (e.g., $\Delta AI > 0$, and $\Delta AI – \Delta CFAI = \Delta AF > 0$), we would conclude that the increase in the house price index without loss/gain behavior would have been less (more) in the boom. In other words, the presence of loss/gain behavior increases (dampens) the boom. Changes in contrast-relative variables don't predict behavior, but they do measure the relative importance of changes in loss *versus* gain behavior over different parts of the housing cycle. Most importantly, our reduced-form estimates show the direction of the association between each type of behavior and each phase of the cycle.

We summarize our analytical framework with a diagram in Appendix 3. The interpretation of changes in our contrast-relative is as follows:

	Interpretation of $AF_j = AI_j – CFAI_j$ as Differences in Levels	
	$AF_j > 0$	**$AF_j < 0$**
$\Delta AI_j > 0$ or $\Delta AI_j < 0$	The presence of L/G behavior *increases* the price index	The presence of L/G behavior *lowers* the price index
	Interpretation of $\Delta AF_j = \Delta AI_j – \Delta CFAI_j$ as Differences in Changes	
	$\Delta AF_j > 0$	**$\Delta AF_j < 0$**
$\Delta AI_j > 0$	The presence of L/G behavior *increases* the cycle phase	The presence of L/G behavior *dampens* the cycle phase
$\Delta AI_j < 0$	The presence of L/G behavior *dampens* the cycle phase	The presence of L/G behavior *increases* the cycle phase

4. Data

Our initial dataset includes 1,409,127 individual residential transactions in 169 towns in Connecticut between 1994 and 2017. Our data are comprehensive because the town hall records contain all residential property transactions, including for-sale-by-owner (FSBO) and sales financed privately.[15] In addition, data were collected on a monthly basis. The immediacy of data collection is important because it allows us to eliminate repeat pairs with a significant change in characteristics or sales by someone other than the purchaser.

We restrict our sample to single-family residential properties with warranty deeds, with sale price over $40,000, with interior footage over 300 square feet and lot size less than 500,000 square feet, with at least one bedroom, with at least a half bathroom, with structures built between 1799 and 2018 and with non-missing transaction date. We require at least 10 sales in each town year because the fixed effects in our regressions imply identification from variation in individual seller anchors within each town-year. After applying these filters and deleting observations with missing or incorrectly coded data (for example, year built is less than year sold), we end up with 548,568 observations.

To calculate sellers' expected gains or losses we further restrict our sample to repeat sale pairs in which we could identify a buyer at the first sale having the same name as the seller at the second sale in order to determine the extent to which the second sale is associated with seller anchoring to the price she paid. We keep repeat sale pairs with the minimum holding period of 12 months in order to remove flips.[16] We control for observable quality changes between sales by deleting observations with changes of interior size between sales greater than 5%. These requirements further reduce our sample size to 90,345 repeat pairs.[17] Appendix 4 summarizes the sample construction procedure.

We use both the expected 2nd sale price (PV) from the hedonic model and normalized assessed value (NAV) as proxies for expected sale price; as explained in Section 2. The hedonic estimation is reported in Appendix 5. NAV is important to our empirical strategy, primarily because the assessor observed many characteristics unobserved by the econometrician.

---

[15] It is noted that our sample includes all transactions such as homes financed with risky non-agency loans and seller financing. Compared with transactions financed with conforming loans, these properties saw more appreciation during the boom, and larger price declines afterward. This explains the discrepancy between our repeat sales index and the FHFA index.

[16] Results are similar when we keep repeat sale pairs with the minimum holding period of 24 months.

[17] We compare the differences of the first-stage hedonic estimation between one-only and repeat sales. Unreported results suggest that houses in the repeat sales sample are smaller, older and have comparable sale price. Given these findings, we re-ran our results by restricting the 1st stage to be repeat sales only: we use 249,497 repeats (more than 90,342 x 2 because changed repeats are included) instead of 548,568 observations in the 1st stage estimation (see Appendix 9). Results are highly robust to using 249,497 transactions.

We divide our sample period based on the recent housing cycle. In a broad context, we divide our sample into two regimes, *Pre-2007* and *Post-2007* (including 2007), because the year 2007 was a transition year when the US and Connecticut housing markets started to decline. We further divide our sample into four distinct periods: *Normal* (2000-2003), *Boom* (2004-2006), *Bust* (2007-2012) and *Recovery* (2013-2017).[18] The housing market in Connecticut experienced stable growth in the period from 2000 to 2003 then a boom in the period from 2004-2006. Following the bubble bursting in 2007, the market continued to fall with the collapse of the subprime mortgage industry and an increase in foreclosure activities. The market started to recuperate from its trough in 2013: see the repeat sales index (AI) in Appendix 6.

5. Results

5.1 Descriptive Statistics

Variable definitions are in Table 1. Table 2 provides a summary of the repeat sale sample and Appendix 7 provides details on the distribution (i.e., quantiles) for each variable. Statistics using NAV are adjusted for unobserved quality.

An average second sale in our sample was sold at $378,263 (the geometric mean is exp(12.537)=$278,452). The price ratio given loss is much higher than the ratio given gain (1.22 versus 1.02 using PV; 1.09 versus 1.01 using NAV), strongly suggesting that loss behavior is important. Mean PV-based conditional losses (gains) are .25 (.38) compared to NAV-based of .20 (.37). Moreover, NAV has a larger impact on loss (.250 versus .204) compared with its impact on gain (.378 versus .374), shown over time in Figure 1 Panel A and B suggesting more selection on quality by those with losses.

The sub-period results in Table 2 and Figure 1 support several findings. First, the magnitude of loss and gain variables shows a large variation across different periods. In Panel A and B of Figure 1, magnitudes of expected losses go down in the boom and up during the bust; gains go down during the bust. During the boom, expected gains were quite large.

Second, we find that the mixture (as a % of all sales) of sellers with gain/loss varies as expected over the housing cycle: see Panel C and D of Figure 1 and *Exp. Loss Dummy (NAV)* in

[18] The FHFA Connecticut Housing Index depicts an early boom starting in 2002 and an early bust starting in 2006Q3. Our results are robust to the alternative definition of the housing cycle of normal (2000-2001), boom (2002-2005), bust (2006-2012) and recovery (2013-2017). Results are summarized in Appendix 9. We thank Zhenguo Lin for his helpful comment on this point.

Table 2. In the normal and boom period, only 5-7% of the transactions had an expected quality-adjusted loss. This number rose to 48% in the bust period and 57% in the recovery period. Quality adjustment using NAV is working as expected: magnitudes become smaller on average, especially for loss. Panels C and D of Figure 1 show that an economically significant proportion of transactions classified as losses (gain) during the normal, boom and recovery periods are reclassified as gains (loss) after quality adjustment.

Third, loss and gain behavior is associated with transaction volume. There are more transactions per quarter in the boom and recovery period, compared with the bust period reflecting the well-documented positive relationship between price change and volume of transactions (Stein, 1995). The ratio of expected gain to loss transactions varies dramatically over the cycle: roughly 1.3x, 1.7x, 1x and .8x in the normal, boom, bust and recovery respectively. These results suggest the possibility that the behavior of those with gains and losses (e.g., deciding when to realize a gain or loss) contributed importantly to the cycle in number of transactions as predicted by realization utility (Barberis and Xiong, 2012).

The variation of loss and gain magnitudes and proportions is consistent with the substantial variation in *2^{nd} Sale Price / Exp. 2^{nd} Price (NAV)* over the cycle, suggesting the important role for loss/gain behavior. For example, the premium fetched by conditional losses decreases from a range around .37 in the normal and boom periods to .1 in the bust and recovery.[19] Together with the very large percentages of sales with losses in recovery (57%), this leads us to conjecture a throw-in-the-towel effect where many of those with losses decide not to defer sales further: they sell at reduced premiums during the recovery. We can't prove this conjecture, but it supports further investigation into the role of loss and gain behavior as an explanation of the positive association between prices and volume of transactions, an association that has been widely debated (e.g., Stein, 1995; Glaeser and Nathanson, 2014; 2017).

In Panel A of Table 3, we track the differences between two subsequent periods to measure incremental changes as the market moves through the stages of the cycle. The comparison between row (1) and (2) highlights the importance of quality adjustment: NAV-based loss suggests we have less amount of loss in boom, more in bust, and less again in recovery, while PV-based loss suggests

[19] There are few losses in the normal and boom period, so we average .395 and .336. Note that the premium on gains given gains stayed roughly constant over the cycle at about .09; here we averaged the .015 and .002 numbers for bust and recovery periods.

no change in boom and a small positive change in recovery. The quality-adjusted conditional expected gain is .091 higher in boom but .134 (.054) lower in bust (recovery) compared with the previous period. This suggests the possibility that those with gains, like those with losses, decide to sell at prices that are even lower during the mild recovery than the bust. In bust and recovery, sellers with both loss and gain appear to wait longer to sell: the conditional month variables increase significantly from boom to bust, as well as from bust to recovery.

One of our important findings is that the recovery period exhibits a high percentage of sales with losses and these had a high average expected loss. The quality-adjusted loss dummy continues to increase from bust to recovery (*Exp. Loss Dummy (NAV)* = +.089 in column (3)). The mild recovery may not have been perceived primarily as an opportunity to cut losses, just to realize them. In "Recovery-Normal", column (5), the difference of quality-adjusted conditional expected loss is positive (+.029) and statistically significant and this increases to +.062 when recovery is compared to the boom, column (6), further supporting the conjecture that many chose to realize larger losses during the recovery than during the boom and normal periods, consistent with the possibility of throw-in-the-towel behavior. The big differences between gains and losses motivate a further investigation using double mean differences in the next section.

5.2 Double Mean Differences (DD) Analysis over the Cycle

Panel B of Table 3 presents the significance of the differences between gains and losses as well as differences in the mean differences ("double mean differences", or DD) over the cycle. These bivariate statistics give intuition for loss/gain behavior over the cycle based on minimal model assumptions. We interpret the results based on NAV.

The DD estimates of *2^{nd} Sale Price / Exp. 2^{nd} Price (NAV)* show that loss behaviors change significantly in *Boom – Normal* (boom minus normal). The DD is driven by the change in loss: sellers with a potential loss were associated with a lower price premium in the boom compared with normal, whereas gains changed little. However, the percentage of sales that were losses is small (between 5 and 7 percent in Table 2) suggesting that the overall influence of losses is overstated by DD statistics. From boom to bust, sellers with expected losses were associated with significantly smaller premiums in the bust than in the boom whereas those with expected gains did not change much and percentage of sales with losses was nearly equal to gains. This implies a large statistically significant DD estimate of -.242, suggesting that loss behavior was strongly associated with house price changes during the bust.

The direction of the association is to reduce observed price declines by reducing price premiums on losses. From bust to recovery, the DD estimate is not statistically significant because the two conditional changes (-.013 for gain; -.035 for loss) are small and in the same direction. This is consistent with the previous evidence that the mild recovery may have been perceived as an opportunity to continue realizing losses.

In all the periods, the first difference between PV- and NAV-based statistics, column (6) compared to column (3), suggests more conservative estimates from NAV-based calculation.[20] We have data on asking price for the period ending in 2013. The DD results of *Asking Price / Exp. 2nd Price* in Appendix 8 support the explanations we provide for Table 3 since asking price is set several months before negotiations determine sales prices, on average.

The possibility of a strong role for loss/gain behavior during the boom and recovery periods is suggested by the *Excess Months* variables, a measure of the holding period for repeat sales relative to the average of each period. Findings are generally consistent with the hypothesis that sellers were reluctant to sell after substantial price declines, and those with losses held back more. However, losses sold relatively quickly in the recovery compared to the bust.[21]

We claim to have results that make minimal model assumptions because the results in Table 3 do not require interpretation of model parameters. Together, the price ratios and month variables suggest a more complex story than the univariate and bivariate statistics can tell on their own. Next, we use model estimates to further investigate these patterns.

5.3 Estimates of the GM Model

Table 4 presents our main results on the relation between anchoring and the transactions price of the second sale, equation (3). Variables in Panel A (Panel B) are calculated based on PV (NAV). We control for the *Months* following GM (2001), Beggs and Graddy (2009), and BG (2011). All the specifications flood the data with town-year fixed effects to control spatially and

[20] To explain the difference between PV- and NAV-based statistics, take normal and boom as an example. First, using NAV helps mitigate misclassification of sales which appears to have a large loss but simply have high values for unobserved quality. Second, in both normal and boom, Exp. 2nd Price (NAV) is higher than Exp. 2nd Price (PV). These two ratios explain the difference between column (3) and (6): quality adjustment using NAV shrinks the first difference, from .493 to .379 in normal; from .467 to .326 in boom. Similar relationships hold in all the periods.

[21] The last six rows of Table 3, Panel B are presented for logical completeness, supporting conclusions above.

time-varying characteristics. Given that the loss and gain variables are constructed from a separate regression, the estimates of the asymptotic covariance matrix are corrected using the bootstrap strategy.

The results suggest anchoring effects on transactions prices: the coefficient estimates of the *Exp. Loss* variable are positive in all model specifications and statistically significant in most cases. We conclude that sellers facing potential losses obtain transactions prices higher than the expected selling prices. Results in Panel B column (1) suggest a .20 expected loss (based on the mean value using NAV) is associated with a .016 (=.079x.2) increase in log sale price.

Comparing the loss coefficients between Panel A and B (Table 4), we find quality adjustment shrinks the coefficient estimates of *Exp. Loss* in all the periods except the bust when the two coefficient magnitudes are similar. Another noticeable difference is that, in Post-07 column (3), PV-based results suggest a positive and statistically significant effect of .123 while NAV-based results indicate much smaller loss effect of .038. Breaking the Post-07 period into bust and recovery provides further support for this difference, supporting the effectiveness of using NAV to control omitted variable bias.

Quality-adjustment has an even bigger effect on *Exp. Gain* coefficients, especially in post-crisis. In the whole sample period, column (1), PV-based coefficients suggest an anomalous positive gain effect while NAV-based coefficients suggest the opposite. The negative NAV gain effect is consistent with the predictions from loss aversion models which are based on a kink in the utility function at zero, implying that marginal gains are valued much less than equivalent losses at least over an interval near the zero point. In view of the superior LOO performance of NAV values (Appendix 2), we think the anomalous PV results are likely due to unobserved quality. We conclude that omitted variable bias is likely present in Table 4, Panel A.

Our confidence in NAV-based results is increased by examining coefficients on the *Residual* variable. Recall that this variable is the residual from the hedonic model for the first sale. In the GM literature, this is interpreted as a noisy proxy for unobserved quality which persists to the second sale. The coefficient of *Residual* is .557 in column (1) Panel A compared with only .082 in Panel B, suggesting the assessors are able to account for many, but not all, characteristics influencing sales price but unobserved by the econometrician. The coefficient of *Residual* increases substantially in the bust and recovery, suggesting that higher quality properties were trading relatively frequently.

5.4 Contrast-relative Analysis: Market Prices with Loss and Gain Effects Controlled

This section uses house price indices to further analyze the association between loss/gain behavior and aggregate repeat sales price cycles. There are six quantities (three variables) from the loss/gain behavior all of which we adjust for quality: the coefficients on gain and loss; the average magnitudes of gains and losses; and the proportion of transactions that are gains and losses. Table 5 summarizes our empirical results for these six quantities each year.[22]

We analyze the association between the quality-adjusted loss/gain factor adjusted index (*CFAI* (*NAV*)) and the repeat sales index using Figure 2 and Table 6. Contrast-relative loss/gain behavior differs substantially from the observed repeat sales index, decreasing the observed change during normal and boom periods and increasing the observed change during the bust. This is important because, if we had found small differences between *CFAI* and *AI* then causal models would be free to ignore the stylized contrast-relative reported here. Figure 2 also shows the economic significance of quality adjustment as the difference between the two CFAI estimates: *CFAI* (*NAV*) deviates the most from *CFAI* (*PV*) in boom and bust when the misclassification of loss/gain is more likely to happen.

Table 6 columns (6)-(8) show the quality-adjusted comparison between the *level* of *AI* and *CFAI* based on our analytical framework in Section 3. We test the null hypothesis that the adjustment factors equal to zero. Numbers in bold denote for *p*-value of *F*-statistics significant at 5%. Both *LAF* (column (6)) and GAF (column (7)) are statistically significant during the normal and boom periods (2000-2006) but *GAF* is 10 to 20 times larger, suggesting that gain behavior dominates, and reduces the level of the observed index. This is expected because the expected losses during this period are 4 to 7% of all sales on a quality-adjusted basis. During the bust (2007-2012), the gain adjustment factor remains important but the magnitude became smaller than in the boom. The magnitude of loss adjustment factor increased relative to the declining GAF throughout the years in the bust, ending at about two-thirds of GAF's level. Together LAF and GAF

[22] Note that the coefficients are taken from table 4. Also note that the proportions of loss in column (8) and the proportion of gain in column (11) do not add up to one in the first few years due to missing values of NAV. This is because we define loss/gain dummy as one if loss/gain (NAV) is greater than one and zero otherwise. Therefore, when NAV is missing, loss dummy and gain dummy are both zero. Our results are robust if we define gain dummy as one minus loss dummy or if we define loss dummy as one minus gain dummy.

significantly decreased the level of the observed index (column (8)) contributing substantially to changes in the contrast relative variables. In the recovery (2013-2017), the effect of *GAF* was minimal (statistically insignificant) after 2012 while *LAF* significantly increased the level of the observed repeat sales index.

Our most important results focus on the comparison between the *change* of *AI* and *CFAI* after quality adjustment, columns (9)-(11). Column (9) shows that changes in *AF* (= ΔLAF + ΔGAF)) are negative during the normal period (2001-2003), meaning that loss/gain behavior is associated with reduced changes in the observed repeat sales index. The average reduction during the three-year period is about 24% (=-.020/.082, where -.020 is the average of ΔAF and .082 is the average of ΔAI, both from 2001-2003) of changes in the observed.

Changes in AF are negative in the first two years (2004-2005) during the boom period, meaning that loss/gain behavior decreased observed changes in the actual index, as it did in the earlier period. Based on this, and DD results, we conclude that anchoring was associated with reductions of between 8% and 24% in observed price changes during the normal and boom periods.

In 2006, ΔAF became positive and it remained positive for eight years. We interpret 2006 as a year of transition to the bust. In fact, quarterly house price indices in Connecticut peaked in the middle of 2006.[23] Therefore, we re-ran our analysis using the alternative house price cycle. The results in Appendix 9 suggest our findings are highly consistent when we use alternative timing of the housing cycle in Connecticut: anchoring was associated with reductions in observed changes in house prices during the boom (2002-2005) and with reduced price declines during the bust (2006-2012). These results suggest the possibility that structural models of loss and gain behavior might be able to explain turning points in the housing cycle.

Changes in AF are positive during the bust period (2007-2012), meaning that loss/gain behavior increased observed changes in the repeat sales index (i.e., they are less negative).[24] The average increase during the six-year period is 33% (=.023/-.070) of the observed changes. Most of this is concentrated in 2007 because we can measure the coefficient only for each period, making

[23] Zhou *et al*. (2018) uses the Case-Shiller gap measure to diagnose the housing bubble in Connecticut and find the gap declines in 2006. By using a pattern-recognition method, the Harding and Pagan (2002) algorithm, to determine the turning points of house prices, we find the repeat sales index peaked in 2006Q3 and this is consistent with the Federal Housing Finance Agency (FHFA) Purchase-Only Indices.

[24] From boom to bust, the presence of loss/gain behavior dampens the bust as the decrease of AI is less than the decrease of CFAI (i.e., $\Delta CFAI < \Delta AI < 0$. Therefore $\Delta AI - \Delta CFAI = \Delta AF > 0$.) because the sign of the first two terms are negative.

ΔAF substantially larger in 2007. Our interpretation is that the changes in the underlying coefficients were likely spread out over time.

In recovery (2013-2017), changes in AF was large and positive in 2013 but diminished in later years, averaging .0028 per year. Changes in the repeat sales index were small too (the recovery in Connecticut was slow), averaging .0054, implying that loss/gain behavior was associated with about 50% increased growth in the observed index. Changes during the recovery are small, but the pattern (loss/gain behavior increasing changes in the index) is an interesting reversal of the boom period, a reversal that is driven by loss behavior, providing additional preliminary evidence of throw-in-the-towel behavior (see discussion above).

5.6 Robustness Tests

Anchoring variables might be correlated with any variable available at the time of purchase, such as loan amount or expectations of price changes, but they will not be correlated with later information such as shocks to productivity or unanticipated changes in the business cycle. But anchors vary across individual sellers and they are historical facts that do not vary over time. This characteristic of anchors controls many omitted variables or feedback mechanisms that might explain the associations we find with the housing market cycle. For example, when the economy is expanding (contracting) then owners may have many (few) attractive alternative investments in housing and other assets and so be quick (slow) to sell at reduced (increased) prices. If this hypothesis is incorrect, then these time dependent variables should change the loss and gain coefficients reported in Table 4.

We test this hypothesis in Appendix 10.1 with various combinations of spatial and temporal fixed effects available in our unique Connecticut data. If the documented loss/gain effects were biased by any omitted time (or location)-dependent variables in a consistent way, we should expect big differences between the coefficient estimates with and without the fixed effects. However, we find the town-dummy effect is relatively small and in inconsistent directions. Furthermore, omission of time dummies has much less effect than omitted town dummies. A model with no spatial or temporal fixed effects produces coefficients reasonably close to a model which has town-year fixed effects. This finding supports identification based on the special characteristics of anchoring variables in many previous studies including GM (2001), BG (2011), Anenberg (2016) and Andersen *et al*. (2019).

The literature has shown that anchoring behavior influences the market through the asking prices set by sellers and that these prices influence buyers (Han and Strange, 2016; Carrillo, 2013; BG, 2011; Anenberg, 2011; GM, 2001). Appendix 10.2 confirm results similar to Table 4 for the period where we have data on asking price. Empirical results reported in previous literature strongly support significance for equation (2) and (3) coefficients with asking price as the dependent variable. E.g., GM (2001) show that sellers with expected losses set an asking price that exceeds the asking price of other sellers by between 25 and 35 percent of their loss. We conclude our asking price (substituted for 2nd sales price) results are consistent with the literature and with our results.

We also re-ran Table 4 by restricting to the 1st stage to be repeat sales only: we use 249,497 repeats (more than 90,342 x 2 because changed repeats are included) instead of 548,568 observations in the 1st stage estimation. Appendix 10.3 shows that our results are still robust.

GM add loan-to-value (LTV) ratio and conclude that it is not strongly binding in their market.[25] As a robustness check, in Appendix 10.4 we follow the literature (e.g., GM, 2001; Engelhardt, 2003; Anenberg, 2011) and include LTV as the difference between the loan-to-value ratio and 80%, truncated from below at zero.[26] The loss and gain coefficients are highly consistent with those in Table 4.

In Appendix 11, we also follow Clapp and Zhou (2019)'s simulation model to correct for unobserved quality, an entirely different method than NAV. Changes in contrast-relative are consistent with the assessed value (NAV) results reported here, establishing that the NAV method is robust. This is important because it suggests that those modeling loss/gain behaviors can correct for unobserved quality without having access to our special assessed value data.

In Appendix 12, we isolate the contrast-relative effect to the changes in loss/gain coefficients by holding loss/gain coefficients constant at the normal period. Results suggest that negotiated prices, holding constant proportions and magnitudes, produce contrast-relative estimations associated with dampening each phase of the cycle. Without negotiated premiums and

---

[25] GM find that LTV greater than .8 has a small positive effect (.04 for LTV of 100%) on asking prices for unsold properties and a slightly larger effect (.06) for sold property. Similarly, Anenberg (2011), Bracke and Tenreyro (2016) and Einiö et al. (2008) control loan to value, but find that it has modest influence on the main results.

[26] Initial value (i.e., the denominator of LTV) was updated to time of sale with town-level house price index and the original loan amount was amortized using the 30-year fixed mortgage rate prevailing at time of origination. This approximation and data errors result in some erroneous LTV's; we therefore winsorize LTV at the 99.9 percentile; variation in this cutoff does not change results. Results without winsorization are highly consistent.

discounts contrast-relative estimations suggest that the boom and recovery would have been larger, and the bust would have been more severe, the same conclusion as we get from Table 6. This implies that causal models based on search and bargaining might be able to explain the stylized facts documented here.

6. Conclusions and Discussions

Results fall into several categories. First, we extend Genesove and Mayer (2001) and Bokhari and Geltner (2011): we multiply sales price premiums or discounts by magnitudes of losses and gains and by their proportions in the aggregate market to construct contrast-relative loss-gain-free indices for comparison with the observed repeat sales index. When gains dominate during the normal and boom periods, negotiated prices are discounted; contrast-relative estimates reduce the observed aggregate increase in house prices by between 8% and 24%. When losses become important during the bust, sellers with losses negotiate a price premium; contrast-relative estimates reduce the observed aggregate decline in house prices. The two behaviors worked in the same direction during the bust: contrast-relative estimate declines in log house prices during the bust were more than 30% less than actual declines. During the mild recovery, when price changes averaged only about .5% per year loss/gain contrast-relative behavior was associated with most of this change. Loss behavior dominated during this period reversing the dampening pattern during the gain-dominated boom, leading us to speculate that many sellers were realizing their losses after years of delaying sale, adding to the recovery. Of course, this is a tentative hypothesis requiring further investigation.

Second, we correct losses and gains for unobserved quality. We use normalized assessed value, a unique feature of our data to construct quality-adjusted expected second sales prices to correct coefficients on losses and gains for omitted variable bias. Both univariate and multivariate results support the effectiveness of this adjustment.

Third, we report additional results that make minimal model assumptions in the sense that they do not depend on interpreting parameters from a regression model. For example, our univariate results suggest that loss behavior is statistically significant and important at turning points, i.e., during the transition from a boom to a bust. Double mean differences show little association between loss/gain behavior and the transition to a recovery, when the two behaviors roughly cancel each other.

Results are robust to substituting asking price for sales price, to restricting the sample to repeat sales, to including loan-to-value ratio, to alternative definitions of cyclical phases and to alternative method of correcting for unobserved quality.

These reduced-form results are not conclusive absent structural models of anchoring, search and bargaining. Also, these results are based on a single housing market cycle in Connecticut – the important cycle from 2000 through 2017. However, Connecticut's cycle was typical of many other states. Moreover, the Connecticut data allow precise control for local public policy (e.g., taxes, schools, and other services) by flooding the data with town-year fixed effects, and major quality change between sales can be controlled. The data allow identification of parameters associated with loss and gain behavior from substantial variation across sellers within town-years, and after correction for unobserved quality.

Zhou et al. (2019) argue the loss coefficients (i.e., negotiated premiums given an expected loss) documented in the literature are largely biased upward. In our analysis, the role of loss/gain is determined by the direction of adjustment which comes from the coefficients, the magnitude and the proportion. As long as the magnitude and proportion is large, we do not need a large coefficient to get the big effects found here. This is because, even with small loss and gain coefficients, the imbalance of proportion between loss and gain could lead to a large impact on final transaction price.

By differencing loss minus gain holding periods (i.e., months between sales) we find the possibility of a strong role for selection of when to sell. Losses were selling relatively quickly in the boom. Holding periods lengthened in the bust and recovery for both losses and gains, suggesting further study of how loss/gain behaviors influence the timing of sales and therefore the volume of transactions. Evidence for further research is provided by the association between transactions volume and the ratio of sales prices to expected prices conditional on loss and gains.[27]

Our work provides stylized facts relevant to structural models of housing market cycles with heterogeneous agents (Davis and van Nieuwerburgh, 2015) where losses and gains could be introduced as endogenous changes in constraints on sellers. In overlapping generations models (e.g., Ortalo-Magne and Rady, 2006) loss and gain are largely determined by when the homeowner

[27] Changes in contrast-relative estimates provide additional very preliminary evidence suggesting that turning points in the housing cycle might be better understood with models including loss and gain percentages and magnitudes: their first and second moments respond to the cycle in an obvious way.

enters the market. In the life cycle model of Favilukus and van Nieuwerburgh (2018) endogenous variation in the magnitudes and percentages of losses and gains are good candidates for influencing housing wealth and the down payment available for purchase.

References

Andersen, S., C. Badarinza, L. Liu, J. Marx and T. Ramadorai. 2019. Reference Dependence in the Housing Market. Available at SSRN: https://ssrn.com/abstract=3396506 or http://dx.doi.org/10.2139/ssrn.3396506

Anenberg, E. 2011. Loss Aversion, Equity Constraints and Seller Behavior in the Real Estate Market. *Regional Science and Urban Economics* 41 (1): 67-76.

Anenberg, E. 2016. Information Frictions and Housing Market Dynamics. *International Economic Review*. 57(4): 1449-1479.

Barberis, N. and W. Xiong. 2012. Realization Utility. *Journal of Financial Economics* 104: 251-271.

Beggs, A. and K. Graddy. 2009. Anchoring Effects: Evidence from Art Auctions. *American Economic Review* 99 (3): 1027-1039.

Ben-David, I. and D. Hirshleifer. 2012. Are Investors Really Reluctant to Realize their Losses? Trading Responses to Past Returns and the Disposition Effect. *Review of Financial Studies* 25 (8): 2485-2532.

Bokhari, S. and D. Geltner. 2011. Loss Aversion and Anchoring in Commercial Real Estate Pricing: Empirical Evidence and Price Index Implications. *Real Estate Economics* 39 (4): 635-670.

Bracke, P and S. Tenreyro. 2016. History Dependence in the housing market. Bank of England Staff Working Paper No. 630.

Campbell, J. Y., S. Giglio and P. Pathak. 2011. Forced Sales and House Prices. *American Economic Review* 101 (5), August 2011: 2108-31.

Carrillo, P.E. 2013. To Sell or Not to Sell: Measuring the Heat of the Housing Market. *Real Estate Economics* 41 (2): 310-346.

Case, K. E. and R. J. Shiller. 1990. Forecasting Prices and Excess Returns in the Housing Market. *Real Estate Economics* 18 (3): 253-273.

Case, K. E. and R. J. Shiller. 2003. Is there a Bubble in the Housing Market? *Brookings Papers on Economic Activity* (2): 299-342.

Clapp, J. and C. Giaccotto. 1992. Estimating Price Indices for Residential Property: A Comparison of Repeat Sales and Assessed Values Methods. *Journal of the American Statistical Association* 87(418), 300-306.

Clapp, J. M, R. Lu-Andrews and T. Zhou. 2018. Anchoring to Purchase Price and Fundamentals: Application of Salience Theory to Housing Cycle Diagnosis. *Real Estate Economics* forthcoming.

Clapp, J. M. and T. Zhou. 2019. Controlling Unobserved Heterogeneity in Repeat Sales Models: Application to Anchoring to Purchase Price. Available at SSRN: https://ssrn.com/abstract=3358401 or http://dx.doi.org/10.2139/ssrn.3358401

Coulson, N. E. and J. E. Zabel. 2013. What can we Learn from Hedonic Models when Housing Markets Dominated by Foreclosures? *Annual Review of Resource Economics* 5, (1): 261-279.

Davis, M.A. and S. van Nieuwerburgh. 2015. Housing, Finance, and the Macroeconomy, 1st ed, *Handbook of Regional and Urban Economics*. Elsevier B.V.

Engelhardt, G.V. 2003. Nominal Loss Aversion, Housing Equity Constraints, and Household Mobility: Evidence from the United States. *Journal of Urban Economics* 53 (1): 171-195.

Einiö, M., M. Kaustia, and V. Puttonen. 2008. Price Setting and the Reluctance to Realize Losses in Apartment Markets. *Journal of Economic Psychology* 29 (1): 19-34.

Favara, G. and J. Imbs. 2015. Credit Supply and the Price of Housing. *American Economic Review* 105(3): 958-992.

Favilukis, J.Y and S. van Nieuwerburgh. 2018. Out-of-Town Home Buyers and City Welfare. Working Paper.

Genesove, D. and C. Mayer. 2001. Loss Aversion and Seller Behavior: Evidence from the Housing Market. *Quarterly Journal of Economics* 116 (4): 1233-1260.

Glaeser, E. L. and C.G. Nathanson. 2014. Housing Bubbles. NBER Working Paper No. 20426. http://www.nber.org/papers/w20426

Glaeser, E. L. and C.G. Nathanson. 2017. An Extrapolative Model of House Price Dynamics. *Journal of Financial Economics* 126 (1): 147-170.

Han, L., W.C. Strange. 2016. What is the Role of the Asking Price for a House? *Journal of Urban Economics* 93: 115-130.

Harding, D., A. Pagan. 2002. Dissecting the Cycle: A Methodological Investigation. *Journal of Monetary Economics* 49 (2): 365–381.

Maldonado, G. and S. Greenland. 2002. Estimating causal effects. *International Journal of Epidemiology*. 31(2):422-9.

Menzies, Peter. 2017. Counterfactual Theories of Causation. *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.).

Ortalo-Magné, F. and R. Sven. 2006. Housing market dynamics: On the contribution of income shocks and credit constraints. *Review of Economic Studies*. 73, 459–485.

Stein, J. C. 1995. Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects. *Quarterly Journal of Economics* 110 (May): 371-40.

The Appraisal Institute. *The Appraisal of Real Estate*. 14th Ed. Chicago: The Appraisal Institute. 2013.

Yavas, A. and S. Yang. 1995. The Strategic Role of List Pricing in Marketing Real Estate: Theory and Evidence. *Real Estate Economics* 23 (3): 347-36.

Zhou T., J.M. Clapp and R. Lu-Andrews. 2019. Do Sellers with Expected Losses Obtain a Premium Price? Really? Working paper.

Table 1: Variable Definition

Variable	Definitions
Variable	**Definitions**
Log of Second Sale Price	Log of sale price of the second sale.
Log of Exp. 2nd Price (PV)	Log of predicted price in the year of second sale estimated by the 1st stage hedonic model. This is the expected 2nd Price estimated by the econometrician.
Log of Assessed Value (AV)	Log of assessed value at the time of second sale. Assessed value is estimated by the property tax assessor every 5 years using the assessors' observables.
Log of Normalized Assessed Value (NAV)	Log of Assessed Value (AV) normalized for the 5-year assessment cycle. Normalized AV is the assessors' estimate of expected 2nd Price.
Anchor (PV)	Log of first sale price minus the log of the Exp. 2nd price (PV).
Anchor (NAV)	Log of first sale price minus the log of the normalized assessed value of the second sale.
Exp. Loss Dummy (PV)	A dummy equal to one if the first sale price is greater than the Exp. 2nd price (PV) and zero otherwise.
Exp. Loss Dummy (NAV)	A dummy equal to one if the first sale price is greater than the normalized assessed value and zero otherwise.
Exp. Loss (PV)	Anchor (PV) if anchor (PV) > 0, 0 otherwise. It is the expected loss based on the econometrician's estimate of expected 2nd price.
Exp. Loss (NAV)	Anchor (NAV) if anchor (NAV) > 0, 0 otherwise. It is the expected loss based on the assessors' estimate of expected 2nd price.
Exp. Gain (PV)	Absolute value of Anchor (PV) if anchor (PV) < 0, 0 otherwise. It is the expected gain based on the econometrician's estimate of expected 2nd price.
Exp. Gain (NAV)	Absolute value of Anchor (NAV) if anchor (NAV) < 0, 0 otherwise. It is the expected gain based on the assessors' estimate of expected 2nd price.
Exp. Loss, Given Loss (PV)	Conditional Exp. Loss (PV) if Exp. Loss Dummy (PV) = 1. Zero values are set to missing.
Exp. Loss, Given Loss (NAV)	Conditional Exp. Loss (NAV) if Exp. Loss Dummy (NAV) = 1. Zero values are set to missing.
Exp. Gain, Given Gain (PV)	Conditional Exp. Gain (PV) if Exp. Loss Dummy (PV) = 0. Zero values are set to missing.
Exp. Gain, Given Gain (NAV)	Conditional Exp. Gain (NAV) if Exp. Loss Dummy (NAV) = 0. Zero values are set to missing.
2nd Sale Price / Exp. 2nd Price (PV)	Ratio of 2nd Sale Price (in dollars) to Exp. 2nd price (PV) (in dollars)
2nd Sale Price / Exp. 2nd Price (NAV)	Ratio of 2nd Sale Price (in dollars) to normalized assessed value at the second sale (in dollars)
2nd Sale Price / Exp. 2nd Price, Given Loss (PV)	2nd Sale Price / Exp. 2nd price (PV) if Exp. Loss Dummy (PV) = 1. Zero values are set to missing.
2nd Sale Price / Exp. 2nd Price, Given Loss (NAV)	2nd Sale Price / normalized assessed value of the second sale if Exp. Loss Dummy (NAV) = 1. Zero values are set to missing.
2nd Sale Price / Exp. 2nd Price, Given Gain (PV)	2nd Sale Price / Exp. 2nd price (PV) of the second sale if Exp. Gain Dummy (PV) = 0. Zero values are set to missing.
2nd Sale Price / Exp. 2nd Price, Given Gain (NAV)	2nd Sale Price / normalized assessed value if Exp. Gain Dummy (NAV) = 0. Zero values are set to missing.
Months	Number of months between the first and second sale divided by 100.
Excess Months	Months minus the mean of Months in each period.
Excess Months, Given Loss (PV)	Excess Months if Exp. Loss Dummy (PV) = 1. Zero values are set to missing.
Excess Months, Given Loss (NAV)	Excess Months if Exp. Loss Dummy (NAV) = 1. Zero values are set to missing.
Excess Months, Given Gain (PV)	Excess Months if Exp. Gain Dummy (PV) = 0. Zero values are set to missing.
Excess Months, Given Gain (NAV)	Excess Months if Exp. Gain Dummy (NAV) = 0. Zero values are set to missing.
Residual	The residual from the 1st stage hedonic regression for the first sale.
Normal	If the date of the second sale is between 2000 and 2003 (including end years).
Boom	If the date of the second sale is between 2004 and 2006, inclusive.
Bust	If the date of the second sale is between 2007 and 2012, inclusive.
Recovery	If the date of the second sale is between 2013 and 2017, inclusive.
Pre2007	If the date of the second sale is before 2007.
Post2007	If the date of the second sale is on or after 2007.

Table 2: Summary Statistics

This table reports summary statistics of variables based on a sample of repeat sales transactions in all years ("All") from 2000 to 2017, and in four sub-periods, "normal" period (2000-2003), "boom" period (2004-2006), "bust" period (2007-2012) and "recovery" period (2013-2017). Column (1) shows the number of observations in the whole sample period. Columns (2)-(5) show means of the variables. Columns (6)-(10) show the number of observations per quarter. Months is the number of months between the first and second sale divided by 100.Table 1 summarizes variable definitions. See Appendix 5 for more detailed statistics including standard deviation, lower quartile, median and upper quartile.

	Obs.	Mean					Obs./Qtr			
	(1) All	(2) All	(3) Normal	(4) Boom	(5) Bust	(6) Recovery	(7) Normal	(8) Boom	(9) Bust	(10) Recovery
Log of Second Sale Price	90,345	12.537	12.441	12.653	12.547	12.502	742	1,477	1,166	1,638
Log of Exp. 2nd Price (PV)	90,345	12.494	12.400	12.616	12.519	12.440	742	1,477	1,166	1,638
Log of Assessed Value (AV)	89,909	12.040	11.763	11.814	12.170	12.148	718	1,473	1,166	1,638
Log of Normalized Assessed Value (NAV)	89,909	12.540	12.457	12.654	12.547	12.502	718	1,473	1,166	1,638
2nd Sale Price / Exp. 2nd Price (PV)	90,345	1.105	1.101	1.078	1.104	1.122	742	1,477	1,166	1,638
2nd Sale Price / Exp. 2nd Price (NAV)	89,909	1.042	1.042	1.027	1.056	1.038	718	1,473	1,166	1,638
2nd Sale Price / Exp. 2nd Price, Given Loss (PV)	38,056	1.220	1.526	1.501	1.194	1.191	102	139	588	1,032
2nd Sale Price / Exp. 2nd Price, Given Loss (NAV)	33,819	1.094	1.395	1.336	1.100	1.065	50	79	560	932
2nd Sale Price / Exp. 2nd Price, Given Gain (PV)	52,288	1.021	1.034	1.034	1.012	1.005	640	1,339	578	606
2nd Sale Price / Exp. 2nd Price, Given Gain (NAV)	56,090	1.010	1.016	1.010	1.015	1.002	669	1,394	606	706
Anchor (PV)	90,345	-0.113	-0.299	-0.404	-0.041	0.050	742	1,477	1,166	1,638
Anchor (NAV)	89,909	-0.156	-0.344	-0.440	-0.069	-0.012	718	1,473	1,166	1,638
Exp. Loss Dummy (PV)	90,345	0.421	0.137	0.094	0.504	0.630	742	1,477	1,166	1,638
Exp. Loss Dummy (NAV)	89,909	0.376	0.069	0.054	0.480	0.569	718	1,473	1,166	1,638
Exp. Loss (PV)	90,345	0.105	0.024	0.016	0.128	0.164	742	1,477	1,166	1,638
Exp. Loss (NAV)	89,909	0.077	0.012	0.007	0.108	0.111	718	1,473	1,166	1,638
Exp. Gain (PV)	90,345	0.219	0.323	0.420	0.169	0.114	742	1,477	1,166	1,638
Exp. Gain (NAV)	89,909	0.233	0.345	0.446	0.176	0.123	718	1,473	1,166	1,638
Exp. Loss, Given Loss (PV)	38,056	0.250	0.172	0.175	0.253	0.260	102	139	588	1,032
Exp. Loss, Given Loss (NAV)	33,819	0.204	0.166	0.133	0.225	0.195	50	79	560	932
Exp. Gain, Given Gain (PV)	52,288	0.378	0.374	0.463	0.341	0.309	640	1,339	578	606
Exp. Gain, Given Gain (NAV)	56,086	0.374	0.382	0.473	0.339	0.285	669	1,394	606	706
Months (in 00's)	90,345	0.673	0.427	0.476	0.634	0.901	742	1,477	1,166	1,638
Excess Months, Given Loss (PV)	38,056	0.020	-0.118	-0.211	-0.105	-0.032	102	139	588	1,032
Excess Months, Given Loss (NAV)	34,255	0.033	-0.154	-0.253	-0.113	-0.020	50	79	560	932
Excess Months, Given Gain (PV)	52,288	-0.015	0.019	0.022	0.107	0.054	640	1,339	578	606
Excess Months, Given Gain (NAV)	56,522	-0.018	0.011	0.014	0.104	0.026	669	1,394	606	706

Table 3: Test of Significance of Differences

This table summarizes tests of mean differences in variables between and across four sub-periods, normal period (2000-2003), boom period (2004-2006), bust period (2007-2012) and recovery period (2013- 2017). In Panel A, each column labeled X-Y subtracts the Table 2 statistic for period Y from the corresponding statistic for period X. Panel B summarizes results based on double mean differences (DD) tests of statistics in Table 2. ***, **, * denote for 1%, 5% and 10% significance, respectively.

Panel A: Univariate Tests

	Boom-Normal	Bust-Boom	Recovery-Bust	Bust-Normal	Recovery-Normal	Recovery-Boom
	(1)	(2)	(3)	(4)	(5)	(6)
(1) Exp. Loss, Given Loss (PV)	0.003	0.079***	0.007**	0.082***	0.089***	0.086***
(2) Exp. Loss, Given Loss (NAV)	-0.033**	0.092***	-0.029***	0.058***	0.029***	0.062***
(3) Exp. Gain, Given Gain (PV)	0.089***	-0.123***	-0.031***	-0.034***	-0.065***	-0.154***
(4) Exp. Gain, Given Gain (NAV)	0.091***	-0.134***	-0.054***	-0.043***	-0.097***	-0.188***
(5) 2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (PV)	-0.025	-0.307***	-0.003	-0.332***	-0.335***	-0.310***
(6) 2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (NAV)	-0.058***	-0.237***	-0.035**	-0.295***	-0.330***	-0.272***
(7) 2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (PV)	0.001	-0.022**	-0.008	-0.021*	-0.029***	-0.030***
(8) 2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (NAV)	-0.006	0.006	-0.013	-0.000	-0.014	-0.008
(9) Exp. Loss Dummy (PV)	-0.043***	0.411***	0.126***	0.367***	0.493***	0.536***
(10) Exp. Loss Dummy (NAV)	-0.015***	0.426***	0.089***	0.411***	0.500***	0.515***
(11) Excess Months, Given Loss (PV)	-0.093***	0.106***	0.073***	0.013***	0.086***	0.179***
(12) Excess Months, Given Loss (NAV)	-0.100***	0.140***	0.093***	0.041***	0.134***	0.234***
(13) Excess Months, Given Gain (PV)	0.003	0.085***	-0.053***	0.088***	0.035***	0.032***
(14) Excess Months, Given Gain (NAV)	0.051***	0.249***	0.188***	0.300***	0.488***	0.437***
(15) Log of Second Sale Price	0.212***	-0.107***	-0.045***	0.105***	0.061***	-0.152***

Panel B: Significance of Double Mean Differences (DD)

		PV			***NAV***		
		Gain	Loss	Col. (2) – (1)	Gain	Loss	Col. (5) – (4)
		(1)	(2)	(3)	(4)	(5)	(6)
Boom-Normal							
2nd Sale Price / Exp. 2nd Price	(1) Normal	1.034	1.526	0.493***	1.016	1.395	0.379***
	(2) Boom	1.034	1.501	0.467***	1.010	1.336	0.326***
	Row (2) – (1)	0.001	-0.025		-0.006	-0.058***	
	DD			**-0.026**			**-0.052****
Excess Months	(3) Normal	0.019	-0.118	-0.137***	0.011	-0.154	-0.165***
	(4) Boom	0.022	-0.211	-0.233***	0.014	-0.253	-0.267***
	Row (4) – (3)	0.003	0.009***		0.002	-0.100***	
	DD			**-0.096*****			**-0.102*****
Bust-Boom							
2nd Sale Price / Exp. 2nd Price	(5) Boom	1.034	1.501	0.467***	1.010	1.336	0.326***
	(6) Bust	1.012	1.194	0.182***	1.015	1.100	0.084***
	Row (6) – (5)	-0.022*	-0.307***		0.006	-0.237***	
	DD			**-0.285*****			**-0.242*****
Excess Months	(7) Boom	0.022	-0.211	-0.233***	0.014	-0.253	-0.267***
	(8) Bust	0.107	-0.105	-0.211***	0.104	-0.113	-0.217***
	Row (8) – (7)	0.085***	0.106***		0.090***	0.140***	
	DD			**0.021*****			**0.050*****
Recovery-Bust							
2nd Sale Price / Exp. 2nd Price	(9) Bust	1.012	1.194	0.182***	1.015	1.100	0.084***
	(10) Recovery	1.005	1.191	0.187***	1.100	1.065	0.063***
	Row (10) – (9)	-0.088	-0.003		-0.013	-0.035**	
	DD			**0.005**			**-0.022**
Excess Months	(11) Bust	0.107	-0.105	-0.211***	0.104	-0.113	-0.217***
	(12) Recovery	0.054	-0.032	-0.085***	0.026	-0.020	-0.045***
	Row (12) – (11)	-0.053***	0.073***		-0.078***	0.094***	
	DD			**0.126**			**0.172**
Bust-Normal							
2nd Sale Price / Exp. 2nd Price	DD			**-0.311*****			**-0.294*****
Excess Months	DD			**-0.075*****			**-0.052*****
Recovery-Normal							
2nd Sale Price / Exp. 2nd Price	DD			**-0.306*****			**-0.316*****
Excess Months	DD			**0.051*****			**0.120*****
Recovery-Boom							
2nd Sale Price / Exp. 2nd Price	DD			**-0.280*****			**-0.264*****
Excess Months	DD			**0.147*****			**0.222*****

Table 4: Regression Model with Anchoring

This table reports the regression results based on a sample of repeat sales transactions from 2000 to 2017. Dependent variable is log price of the second sale. *Exp. Loss* (*Exp. Gain*) is the difference between the first sale price and the expected 2nd price truncated above (below, then take absolute value) at zero. In Panel A, the expected 2nd price is the predicted value (PV) estimated by a standard hedonic model, the 1st stage regression: see Appendix 3 for regression details. In Panel B, the expected 2nd price is the normalized assessed value (NAV): see Section 2.2 for the NAV calculation. Residual is the residual from the 1st stage hedonic regression for the first sale. Months is number of months since the first sale divided by 100. Results in Column (1) are based on all observations from 2000 to 2017. Results in Column (2)-(3), Pre-2007 and Post-2007, use observations prior to 2007 and after 2007, respectively. Results in Column (4)-(7), Normal, Boom, Bust and Recovery, use repeat sale transactions in the period of 2000-2003, 2004-2006, 2007-2012 and 2013-2017, respectively. All price variables are in logs. All the specifications include town-year fixed effects. Bootstrap standard errors are reported. ***, **, * denote for 1%, 5% and 10% significance, respectively.

Panel A: PV. Dependent is log of 2nd sales price.

	All	Pre-07	Post-07	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exp. Loss (PV)	0.128***	0.859***	0.123***	0.921***	0.828***	0.046**	0.149***
	(0.011)	(0.047)	(0.013)	(0.078)	(0.060)	(0.021)	(0.019)
Exp. Gain (PV)	0.211***	-0.326***	0.319***	-0.453***	-0.277***	0.330***	0.360***
	(0.008)	(0.018)	(0.011)	(0.037)	(0.024)	(0.019)	(0.018)
Months	0.577***	0.136***	0.590***	0.082**	0.143***	0.677***	0.567***
	(0.008)	(0.020)	(0.010)	(0.040)	(0.026)	(0.019)	(0.014)
Residual	-0.122***	0.182***	-0.129***	0.209***	0.160***	-0.166***	-0.122***
	(0.003)	(0.013)	(0.003)	(0.021)	(0.018)	(0.010)	(0.003)
Exp. 2nd Price (PV)	0.945***	0.897***	0.970***	0.919***	0.881***	0.951***	0.985***
	(0.004)	(0.007)	(0.005)	(0.011)	(0.008)	(0.007)	(0.006)
Constant	0.750***	1.359***	0.460***	1.088***	1.575***	0.683***	0.289***
	(0.047)	(0.090)	(0.066)	(0.140)	(0.104)	(0.089)	(0.077)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	90,345	29,592	60,753	11,866	17,726	27,993	32,760
R-squared	0.848	0.887	0.834	0.890	0.881	0.820	0.847

Panel B: NAV. Dependent is log of 2nd sales price.

	All	Pre-07	Post-07	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exp. Loss (NAV)	0.079***	0.715***	0.038***	0.780***	0.584***	0.063***	0.067***
	(0.016)	(0.060)	(0.015)	(0.076)	(0.086)	(0.020)	(0.015)
Exp. Gain (NAV)	-0.116***	-0.352***	-0.039***	-0.362***	-0.347***	-0.153***	0.021*
	(0.006)	(0.011)	(0.007)	(0.021)	(0.015)	(0.015)	(0.012)
Months	-0.058***	0.159***	-0.077***	0.140***	0.163***	0.004	-0.092***
	(0.002)	(0.010)	(0.003)	(0.014)	(0.013)	(0.007)	(0.003)
Residual	0.082***	-0.057***	0.115***	0.007	-0.095***	0.073***	0.124***
	(0.005)	(0.011)	(0.007)	(0.018)	(0.013)	(0.013)	(0.011)
Exp. 2nd Price (NAV)	0.875***	0.872***	0.904***	0.903***	0.849***	0.894***	0.928***
	(0.003)	(0.007)	(0.004)	(0.010)	(0.008)	(0.007)	(0.005)
Constant	1.624***	1.680***	1.259***	1.262***	1.987***	1.349***	0.973***
	(0.042)	(0.091)	(0.052)	(0.129)	(0.095)	(0.087)	(0.064)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	89,909	29,176	60,733	11,495	17,681	27,976	32,757
R-squared	0.863	0.898	0.851	0.901	0.894	0.834	0.867

Table 5: Summary of Variables Used for Contrast-Relative Analysis

This table summarizes the three variables (six quantities in each time period) used for contrast-relative analyses by year. Quantities in columns (1)-(5) are calculated based on Exp. 2[nd] Price (PV) and those in columns (7)-(12) are based on normalized assessed values (NAV). "Coefficient of Loss (α_l)" equals coefficient estimates on the loss variables in Table 4, where coefficients are constant within subperiods. "Mean Exp. Loss, given loss (ML)" equals the mean expected loss variable, conditional on the sold property facing a loss. "Proportion of Loss (%L)" equals the proportion of transactions facing a loss. "Coefficient of Gain (α_g)" equals the coefficient estimates on the gain variables in Table 4. "Mean Exp. Gain, given gain (MG)" equals the mean expected gain variable, conditional on the sold property facing a gain. "Proportion of Gain (%G)" equals the proportion of transactions facing a gain.

	PV						**NAV**						
	ML	%L	α_l	MG	%G	α_g	ML	%L	α_l	MG	%G	α_g	Actual Log Price Index (AI) (Year 2000 = 0)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
2000	0.164	0.216	0.921	0.313	0.784	-0.453	0.148	0.116	0.78	0.316	0.824	-0.362	0.000
2001	0.174	0.160	0.921	0.331	0.840	-0.453	0.193	0.068	0.78	0.338	0.887	-0.362	0.069
2002	0.169	0.121	0.921	0.376	0.879	-0.453	0.176	0.071	0.78	0.392	0.897	-0.362	0.156
2003	0.179	0.102	0.921	0.421	0.898	-0.453	0.150	0.050	0.78	0.426	0.947	-0.362	0.245
2004	0.170	0.088	0.828	0.463	0.912	-0.277	0.133	0.045	0.584	0.471	0.952	-0.347	0.348
2005	0.182	0.080	0.828	0.473	0.920	-0.277	0.126	0.042	0.584	0.483	0.958	-0.347	0.452
2006	0.173	0.112	0.828	0.454	0.888	-0.277	0.137	0.073	0.584	0.465	0.923	-0.347	0.494
2007	0.168	0.209	0.046	0.393	0.791	0.330	0.144	0.158	0.063	0.398	0.840	-0.153	0.433
2008	0.199	0.426	0.046	0.323	0.574	0.330	0.216	0.384	0.063	0.327	0.616	-0.153	0.297
2009	0.262	0.583	0.046	0.307	0.417	0.330	0.244	0.576	0.063	0.295	0.423	-0.153	0.183
2010	0.257	0.604	0.046	0.304	0.396	0.330	0.217	0.591	0.063	0.296	0.408	-0.153	0.159
2011	0.280	0.640	0.046	0.315	0.360	0.330	0.226	0.635	0.063	0.314	0.365	-0.153	0.110
2012	0.290	0.686	0.046	0.314	0.314	0.330	0.243	0.669	0.063	0.303	0.331	-0.153	0.076
2013	0.281	0.652	0.149	0.296	0.348	0.360	0.225	0.617	0.067	0.290	0.383	0.021	0.083
2014	0.271	0.666	0.149	0.300	0.334	0.360	0.217	0.620	0.067	0.275	0.380	0.021	0.049
2015	0.268	0.653	0.149	0.309	0.347	0.360	0.201	0.589	0.067	0.286	0.411	0.021	0.055
2016	0.258	0.628	0.149	0.322	0.372	0.360	0.185	0.559	0.067	0.301	0.441	0.021	0.078
2017	0.236	0.583	0.149	0.311	0.417	0.360	0.164	0.506	0.067	0.274	0.494	0.021	0.103
Average	0.221	0.401		0.351	0.599		0.183	0.375		0.347	0.621		0.188

Table 6: Tests of Significance: Actual Housing Price Index minus Actual Factor Adjusted Index (CFAI)

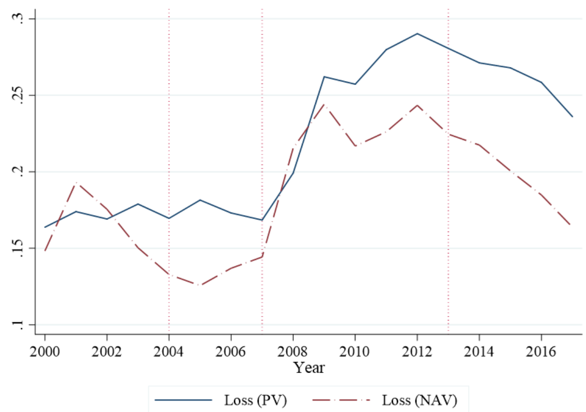
This table summarizes results based on tests of significance on the difference between the actual housing price index (AI) and the contrast-relative loss-gain-factor-adjusted index (CFAI). Loss and gain factors in columns (1)-(5) are based on Exp. 2nd price (PV) and those in columns (6)-(10) are based on normalized assessed value (NAV). The total adjusted factor (AF) (Column (3) or (8)) consists of the loss adjusted factor (LAF, in column (1) or (6)) plus the gain adjusted factor (GAF, in column (2) or (7)): AF = LAF + GAF and the contrast-relative adjusted index is CFAI = AI – AF. Column (4) and (9) shows the changes of total adjustment factors (ΔAF) (i.e. ΔLAF + ΔGAF). Column (5) and (10) shows the changes of the contrast-relative loss-gain-factor adjusted factor index (ΔCFAI). Column (11) shows the changes of actual housing price index (ΔAI). By construction, ΔAI – ΔCFAI = ΔAF and columns (4)-(5) and (9)-(10) are recorded as missing in 2000. Numbers in bold denote for *p*-value of *F* statistics significant at 5%. We cannot calculate *F* statistics for any change variables since they are based on only two numbers.

	PV ---					**NAV** --					
	LAF	GAF	AF	ΔAF	ΔCFAI	LAF	GAF	AF	ΔAF	ΔCFAI	ΔAI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
2000	**0.033**	**-0.111**	**-0.079**			**0.013**	**-0.094**	**-0.081**			
2001	**0.026**	**-0.126**	**-0.100**	-0.021	0.090	**0.010**	**-0.109**	**-0.098**	-0.017	0.086	0.069
2002	**0.019**	**-0.150**	**-0.131**	-0.031	0.118	**0.010**	**-0.127**	**-0.118**	-0.019	0.106	0.087
2003	**0.017**	**-0.171**	**-0.155**	-0.024	0.113	**0.006**	**-0.146**	**-0.140**	-0.023	0.112	0.089
2004	**0.012**	**-0.117**	**-0.105**	0.050	0.052	**0.003**	**-0.156**	**-0.152**	-0.012	0.114	0.102
2005	**0.012**	**-0.120**	**-0.108**	-0.004	0.108	**0.003**	**-0.161**	**-0.157**	-0.005	0.109	0.104
2006	**0.016**	**-0.112**	**-0.096**	0.013	0.029	**0.006**	**-0.149**	**-0.143**	0.014	0.027	0.042
2007	**0.002**	**0.103**	**0.104**	0.200	-0.260	**0.001**	**-0.051**	**-0.050**	0.093	-0.154	-0.060
2008	**0.004**	**0.061**	**0.065**	-0.039	-0.097	**0.005**	**-0.031**	**-0.026**	0.024	-0.161	-0.137
2009	**0.007**	**0.042**	**0.049**	-0.016	-0.098	**0.009**	**-0.019**	**-0.010**	0.015	-0.129	-0.113
2010	**0.007**	**0.040**	**0.047**	-0.002	-0.022	**0.008**	**-0.018**	**-0.010**	0.000	-0.024	-0.024
2011	**0.008**	**0.037**	**0.046**	-0.001	-0.047	**0.009**	**-0.018**	**-0.008**	0.002	-0.050	-0.048
2012	**0.009**	**0.032**	**0.042**	-0.004	-0.031	**0.010**	**-0.015**	-0.005	0.003	-0.038	-0.035
2013	**0.027**	**0.037**	**0.064**	0.023	-0.015	**0.009**	0.002	**0.012**	0.017	-0.009	0.008
2014	**0.027**	**0.036**	**0.063**	-0.001	-0.033	**0.009**	0.002	**0.011**	0.000	-0.034	-0.035
2015	**0.026**	**0.039**	**0.065**	0.002	0.005	**0.008**	0.002	**0.010**	-0.001	0.007	0.006
2016	**0.024**	**0.043**	**0.067**	0.003	0.020	**0.007**	0.003	**0.010**	-0.001	0.023	0.023
2017	**0.021**	**0.047**	**0.067**	0.000	0.025	**0.006**	0.003	**0.008**	-0.001	0.026	0.025

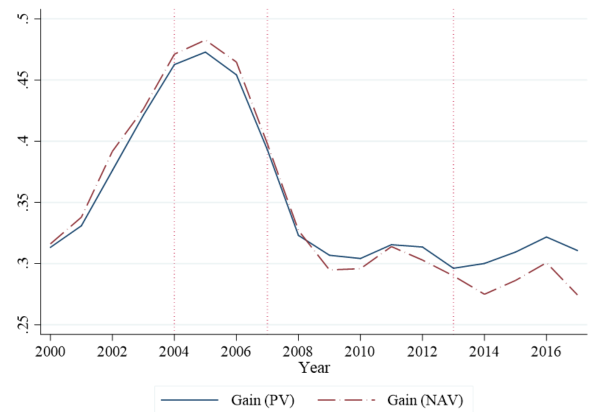
Figure 1: Magnitude of Loss/Gain and Proportion of Loss/Gain

This figure shows (1) magnitude of loss (i.e., Exp. Loss, Given Loss) in Panel A, (2) magnitude of gain (i.e., Exp. Gain, Given Gain) in Panel B, (3) proportion of loss in Panel C and (4) proportion of gain in Panel D. In Panel A-D, solid lines indicate estimated loss/gain based on *Exp. 2nd price (PV)* and dash lines indicate loss/gain based on normalized assessed value (NAV).

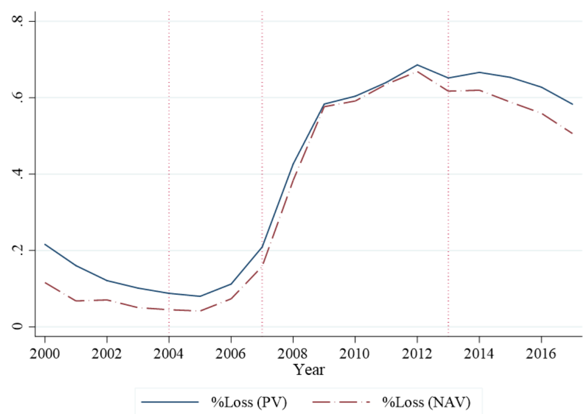
Panel A: Magnitude of Loss



Panel B: Magnitude of Gain



Panel C: Proportion of Loss



Panel D: Proportion of Gain

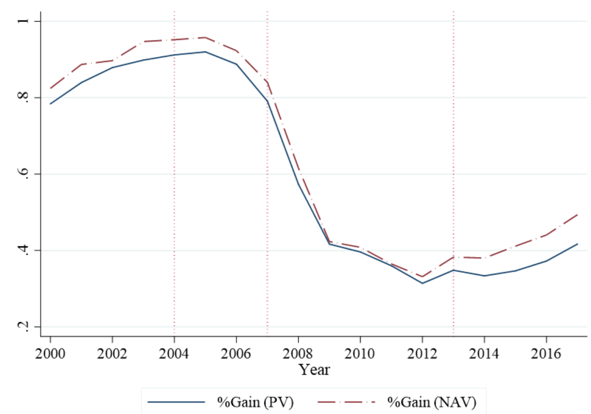


Figure 2: Repeat Sale Index and the Loss-Gain-Factor Adjusted Index (CFAI)

This figure shows the repeat sales index (AI) and the loss-gain-factor adjusted index (CFAI) based on a sample of repeat transactions from 2000 to 2017. Both indices are exponents of log price indices. The repeat sales index (AI) is normalized to 1 in 2000. The CFAI is calculated as the actual repeat sales index (AI) minus the loss (gain) adjustment factors, LAF (GAF). LAF (GAF) is the loss (gain) variable for a given period (i.e. normal, boom, bust and recovery) multiplied by the mean magnitude of loss (gain) among sold properties that were facing an expected loss (gain) and the proportion of sellers facing an potential loss (gain) in that period. The AI minus CFAI suggests the potential magnitude of loss/gain behavior. I.e., If the difference is negative (positive), then the behavior is associated with dampened (accentuated) price movements if the actual change is positive in that part of the cycle, $\Delta AI>0$. This logic reverses if $\Delta AI<0$. "CFAI (PV)" indicates CFAI is calculated based on *Exp. 2nd Price (PV).* "CFAI (NAV)" indicates CFAI is based on normalized assessed value (NAV).

Is the Behavior of Sellers with Expected Gains and Losses Relevant to Cycles in House Prices?

Online Appendix

Appendix 1 provides support for using assessed value to mitigate unobserved quality.
Appendix 2 describes the calculation of normalized assessed value (NAV) and summarizes results of leave-one-out (LOO) cross-validation.
Appendix 3 summarizes the analytical framework.
Appendix 4 displays our sample construction process.
Appendix 5 reports the first-stage hedonic regression results.
Appendix 6 shows the comparison between repeat sales index (based on repeat sales pairs), loss index and gain index.
Appendix 7 provides additional summary statistics that supplement Table 2.
Appendix 8 summarizes double results using asking price.
Appendix 9 summarizes results using an alternative house price cycle.
Appendix 10 reports additional robustness tests.

- A10.1 report robustness tests with various combinations of spatial and temporal fixed effects.
- A10.2 report robustness tests by restricting to the repeat sale sample only.
- A10.3 report results using asking price as the dependent variable.
- A10.4 reports robustness tests by adding the loan-to-value ratio.

Appendix 11 summarizes Clapp and Zhou (2019)'s method for correcting unobserved quality and report results using their method.
Appendix 12 presents and discusses contrast-relative analyses holding loss/gain coefficients constant.

Appendix 1: Assessed Valuation Methods

The purpose of this appendix is to support our claim that, compared with the estimated value of a hedonic model, assessed value has better information relevant to the true value and that the assessor does not anchor to price paid.

A1.1 Academic studies supporting assessment value to control for unobserved characteristics

Genesove and Han (2011) estimate parameters to capture market thinness. Their reduced-form estimation includes three inter-related equations on the list price, number of bidders and the final price. As the list price is composed of a quality component which is unobserved by the econometrician, regressing the number of bidders on list price is subject to an error-in-variable (EIV) problem. In a follow-up study, Han and Strange (2016) estimate the role of asking price on number of bidders for a house. One of the concerns in their reduced-form estimates of the coefficient of asking price is that "the econometrician is unlikely to observe all housing characteristics that are observed by buyers and sellers. To the extent that houses with nice but unobserved features are both listed at a higher price and attract more bidders, this will introduce a bias into the estimated effect of the asking price." (page 125).

Importantly, Genesove and Han argue assessed value contains part of unobserved house characteristics because "assessed value is typically based not only on housing attributes reported in the MLS database but also on the assessor's actual visit of the house and the neighborhood." Because the unobserved characteristics would bias the estimates in the manner of an EIV bias, they show (in their Table 5 and 8) that the estimated coefficients of asking price to become larger after controlling for property tax assessment.[1]

A1.2 Appraisal practices:

A) Do assessors know more than econometricians?

The assessor has an enormously difficult job of periodically valuing every real property parcel regardless of whether similar properties in similar locations are trading frequently. This must be done while meeting rigorous statewide tests of the predictive accuracy and while minimizing challenges from property owners who might feel unfairly treated. To balance these competing interests, the assessor collects a large amount of data and considers many details unobserved by the econometrician who is estimating hedonic valuation models based on a standard hedonic database. See Table A1 for details on variables available in a typical assessor database.

Equity in property taxation is the subject of a large literature seeking to uncover biases for or against various groups, typically based on income or racial/ethnic characteristics (Clapp, 1990). This transparency – assessment data are publicly available in Connecticut and most US states – increases confidence that the assessment is a good faith estimate of market value.

[1] The value added by assessment practices is documented in Clapp and Giaccotto (1992) and Clapp and O'Connor (2008).

One of the most relevant characteristics considered by the assessor and not by the econometrician is neighborhood, with potentially different land values for each neighborhood. The assessor separates land value from structure value in order to increase equity and improve predictive performance, whereas the econometrician does not do this. Assessors have detailed maps with neighborhood boundaries based on primary arteries and secondary streets, traffic, view, topography, proximity to points of interest and the characteristics of the structures. In one Connecticut town, there are 17 Census tracts and over 40 neighborhoods defined by the assessor. Important characteristics of the property include whether the lot was split from another or combined with other lots and whether sales prices are influenced by personal property or other factors unrelated to valuation. The assessor considers many characteristics of the property including landscaping, garage space, type of heating, quality of construction, updates to kitchen and bathrooms and finish of basement and attic space.

According to the Appraisal Institute, "In the valuation process, the appraiser gathers much of the information needed to describe and analyze the improvements by personally visiting the site of the real estate. Careless or inadequate inspection of the physical characteristics and features of the subject and comparable properties can create difficulties for an appraiser in later phases of the appraisal. For example, if a structural problem is overlooked, the conclusions of the three approaches to value could be meaningless. The goal of the site visit is identifying the site and building characteristics that create value." (page 219, the Appraisal Institute 2013). And "Failure to disclose defects in an improvement (because those defects were missed during the site visit) or to verify information gathered through other means are flaws of an appraisal report that can result in litigation against an appraiser." (page 221, the Appraisal Institute 2013).

B) Does the assessor anchor to the price paid by any particular seller?

A potential criticism is that the assessor might anchor to price paid since we show that this is relevant to the sales price at the time of the second sale. To the best of our knowledge, we have never heard or read that assessors anchor to price paid. In Connecticut, if the previous price is within 3 years, assessors will use it as a comparable sale. But too much reliance on recent prices is called "sales chasing" by assessors. i.e., assessors make a clear distinction between house value which is based on fundamentals and endures for the 5 years of the assessment and prices which are more volatile (Clapp, Giaccotto and Richo, 1994 & 1996).

In fact, assessors should not use or collect any information on individual characteristics or on groups of buyers or sellers. To do so would open the assessor to charges of discrimination or bias in the assessment process. For example, in 2017, the Brighton Park Neighborhood Council and Logan Square Neighborhood Association have filed a lawsuit in circuit court alleging that the office of Cook County Assessor Joseph Berrios conducts assessments that systematically and illegally shift residential property tax burdens from Whites to Hispanics and African-Americans and from the rich to the poor.[2] One defense is that the assessor does not maintain any information on the personal characteristics of the owners.

[2] The link to the lawsuit (last access on November 8, 2019):

References A1:

Clapp, J. and O'Connor, P. 2008. Best Practice Automated Valuation Models of Time and Space. *Journal of Property Tax Assessment and Administration* 5(2) 57-67.

Clapp, J., Giaccotto, C. and Richo, G. 1996. Estimating Time Adjustments with Sales Prices and Assessed Values. *The Appraisal Journal* 64(3), 319-327.

Clapp, J., Giaccotto, C. and Richo, G. 1994. Adjusting for Time Using Sales-Ratio Data. *The Assessment Journal* 1(1), 56-62.

Clapp, J. and Giaccotto, C. 1992. Estimating Price Indices for Residential Property: A Comparison of Repeat Sales and Assessed Values Methods. *Journal of the American Statistical Association* 87(418), 300-306.

Clapp, J. 1990. A New Test for Equitable Real Estate Tax Assessment. *The Journal of Real Estate Finance and Economics* 3, 233-249.

Genesove, D., Han, L., 2011. Measuring the Thinness of Real Estate Markets. Working Paper.

Han, L. and Strange, W.C., 2016. What is the role of the asking price for a house? *Journal of Urban Economics* 93, 115–130.

The Appraisal Institute. *The Appraisal of Real Estate*. 14th Ed. Chicago: The Appraisal Institute. 2013.

https://static1.squarespace.com/static/5871061e6b8f5b2a8ede8ff5/t/5a329fcaec212d09e4001553/1513267154250/2017.12.14+complaint+and+supporting+documents.pdf

Table A1 Selected Characteristics Considered in a Typical Assessor Database

Variable	Description
parcelid	A unique identifier. Format: First 3 numbers= Book, Next 2 numbers = Map, Last 3 numbers = Lot, and split (if applies) denoted by a letter (A-Z)
prcl_st	Parcel status: X denotes parcel has since been canceled through either a split or combine.
market	Residential Market Area
nbhd	Residential Neighborhood: If 5 digits long, the first 2 characters are the residential market area. Otherwise first character denotes residential market area.
sprice	Sale price as recorded on the Affidavit of Sale
vl_perpr	Value of personal property if denoted on the Affidavit of Sale
smonth	Sale Month
syear	Sale Year
deedtype	Type of Deed (Warranty Deed)
multprop	Y/N- if multiple parcels involved in the sale
proptype	Property Type as designated on the Affidavit of Sale. A= Vacant Land, B=Single Family Residential, C= Condo/Townhouse, D= 2-4 Plex, E= Apartment Building, F=Commercial/Industrial Use, G= Agricultural, H= Mobile or Manufactured Home, I= Other
fintype	Financing Type
own_cc	Only for residential properties- indicates if the buyer intends to use property as a primary residence
per_prop	Y/N- if the personal property was involved in sale. A similar variable for sale of partial interests.
sale_solar_indc	Y/N- solar involved in sale
landsqft	the total amount of land (square feet) in parcel
std_zne	Assessor's standardized zoning code
corner	The parcel is located on a corner
culdesac	The parcel is located in a culdesac
gated	The parcel is located in a gated community (similar variables for golf course, greenbelt, lake or other water features)
premium	The parcel has a premium view
adj_apt	The parcel is located adjacent to an apartment/multi-family complex
adj_cm	The parcel is located adjacent to commercial/industrial property
trans_ln	The parcel is located adjacent to a transmission line
waterway	The parcel is located adjacent to a waterway
paved	The parcel is accessible via a paved road
ut_none	The parcel has no utilities
ut_elec	The parcel has electricity
ut_water	The parcel has water
ut_well	The parcel has a well
ut_gas	The parcel is connected to gas lines (similar entries are available for sewer and septic)
fld_plan	The parcel is in a flood plain
flt_sub	The parcel is in a substantial noise flight path
r_totimpsqft	The residential square foot of living area in an economic unit. Similar variable for finished basement.
r_iclass	Residential quality class. The scale is 0-7 with 3 being average, 7 being highest
r_wtdyrblt	Residential weighted construction year: a weighted calculation which accounts for the age and square footage of livable additions
r_stories	Residential number of stories (maximum is 4- a basement + three floors)
r_addqual_att	Residential attached addition quality (In comparison to main quality (R_ICLASS), 1=below, 2= comparable, 3= above)
r_carport_att_sqft	Residential attached carport square feet (similar variables for detached carport and for garage space)
r_pool	Residential pool (square feet) or spa.
r_sport_court	Residential sport court (square feet)

Note: In Connecticut "property cards" and GIS systems contain similar detail about property characteristics and its location. For example, see https://westhartfordct.mapgeo.io/datasets/properties?abuttersDistance=300&latlng=41.7626%2C-72.756789&panel=themes&zoom=12 and http://gis.vgsi.com/westhartfordct/ (last accessed 11/09/19).

Appendix 2: Calculation and Validation of Normalized assessed value (NAV)

A2.1 Calculation of NAV

$lnAV\_norm_{ils}$ is calculated using the following algorithm:

1. Calculate town-year averages of the second sales price, P_{ils}. I.e., Regress P_{ils} on town-year fixed effects and calculated the expected P_{ils}, $\hat{P}^{a}_{ils}$.
2. Calculate deviations of the log of assessed value, $lnAV_{ilt}$, from town-year averages. I.e., Regress $lnAV_{ilt}$ on town-year fixed effects and obtain the residual, ε^{lnav}_{ils}.
3. Calculate $lnAV\_norm_{ils}$ as $\hat{P}^{a}_{ils} + \varepsilon^{lnav}_{ils}$. I.e., add the results from step 1 to step 2. This works with our identification strategy which comes from variation within town-years, so the only information added is to make the level of $lnAV\_norm_{ils}$ comparable to the average level of the second sales price, as required in order to calculate loss.

A2.2 Leave-one-out (LOO) Cross-Validation

We construct a leave-one-out (LOO) validation exercise comparing predicted values from the hedonic regression, equation (1), to normalized assessed value. Note that assessed value is at a disadvantage because it lags the date of sale by up to 5 years: any changes to property characteristics or to neighborhood values during the time from assessment to sale will be ignored by assessed value but included in the hedonic which uses all the information up to and including the time of the left-out observation. On the other hand, assessors inspect the property and include many neighborhood and property characteristics unavailable to the econometrician. This implies that our cross-validation exercise is relevant to determining the balance of advantages and disadvantages of the two methods for estimating property value.

To simplify the exercise and reduce computational time, we conduct the validation by town. For hedonic regression we run the following equation using all the observations from the beginning of our sample to through year *t* by leaving out one observation, *i*, in year *t*:[3]

$$P^{(-i)}_{it} = \beta X^{(-i)}_{i} + \emptyset^{(-i)}_{t} + \varepsilon^{(-i)}_{it} \qquad \text{(A2.1)}$$

where $\emptyset_t$ are year dummies and we omit town *l* notation because we do each step for a given town. As our analysis is performed by town, we only include year dummies. Then we calculate, for the left-out observation *i*, the predicted value $\hat{P}_{it}^{(-i)}$ and its mean square error

$$MSE^{Hedonic}_{it} = (P_{it} - \hat{P}_{it}^{(-i)})^2. \qquad \text{(A2.2)}$$

We repeat the calculation N times for all the observations in the given town in year *t*. We calculate the cross-validation test statistic (CV) by averaging the N results from equation (A2.2):

$$CV^{Hedonic}_{t,(N)} = \frac{1}{N}\sum_{i=1}^{N} MSE^{Hedonic}_{it}. \qquad \text{(A2.3)}$$

[3] Note that we do not include observations after the year *t* of the cross-validation exercise, but we do use all information before *t*, as does the assessor; but the assessor lags year *t* due to the 5-year assessment cycle.

The calculation of normalized assessed value is discussed in A2.1. For observation *i* in year *t* in the given town we first calculate the expected second sales price by leaving out that observation. i.e., Regress P_{ils} on year fixed effects and calculated the predicted value $\widehat{P}_{it}^{(-i)}$.

$$P_{it}^{(-i)} = \emptyset_t^{(-i)} + e_{it}^{(-i)} \qquad \text{(A2.4)}$$

where $\emptyset_t$ are year dummies. The predicted values are the average price for every sale in year *t* except the left-out sale.

Next, we calculate deviations of log of assessed value, $lnAV_{it}$, from year averages. I.e., Regress $lnAV_{it}$ on town-year fixed effects and obtain the residual, $\hat{\varepsilon}_{it}^{lnav}$.

$$lnAV_{it} = \emptyset_t + \varepsilon_{it}^{lnav}. \qquad \text{(A2.5)}$$

Lastly, we calculate

$$MSE_{ilt}^{NAV} = (P_{ilt} - \widehat{lnAV\_norm}_{ilt}^{(-i)})^2 \qquad \text{(A2.6)}$$

where $\widehat{lnAV\_norm}_{ilt}^{(-i)} = \widehat{P}_{ilt}^{(-i)} + \hat{\varepsilon}_{it}^{lnav}$.

The test statistic, $CV_{t,(N)}^{NAV}$, is estimated by averaging the N resulting MSE's,

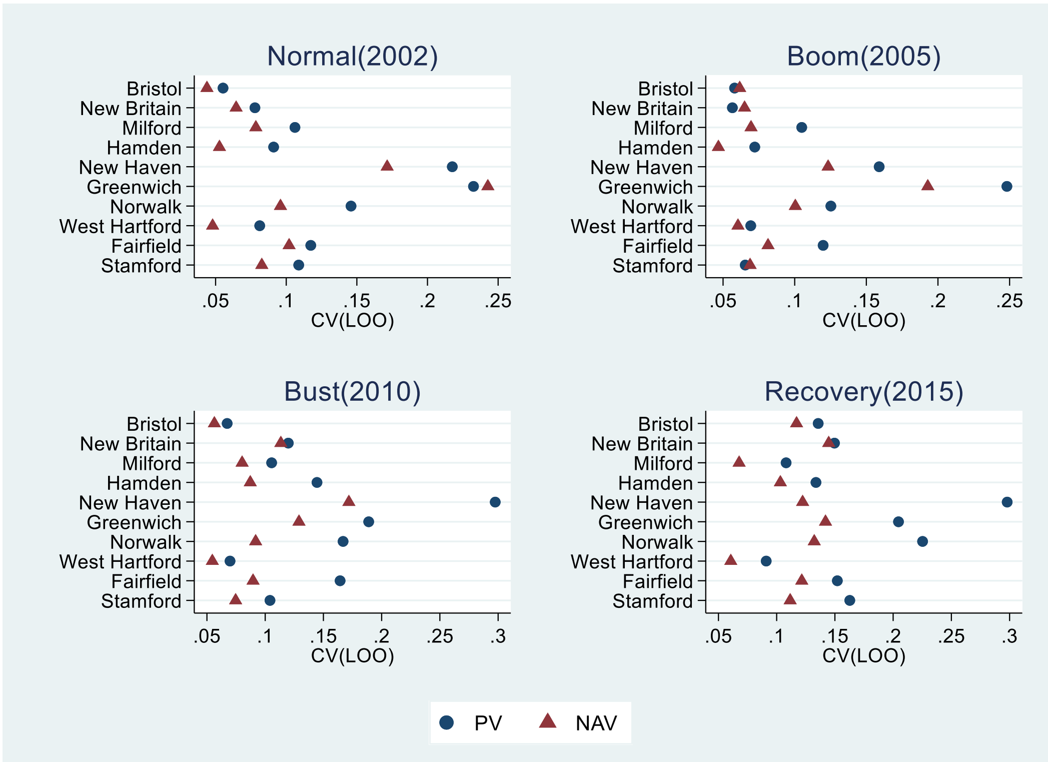
$$CV_{t,(N)}^{NAV} = \frac{1}{N}\sum_{i=1}^{N} MSE_{it}^{NAV}. \qquad \text{(A2.7)}$$

We choose top ten towns (out of 169 towns) based on their numbers of transactions during our sample period. Transactions in these towns represent 24% of our sample. For each town, we choose one year in the middle of each cycle: 2002 for normal, 2005 for boom, 2010 for bust and 2015 for recovery. In sum, we calculate 40 MSE pairs, four years for each of the top ten towns, to compare predicted values from hedonic regression and normalized assessed value.

The results of this horse race are summarized in Figure A2 below. In most town-years, the MSEs using normalized assessed value are consistently smaller than those using hedonic regressions. The results strongly support the superior performance of assessed value.

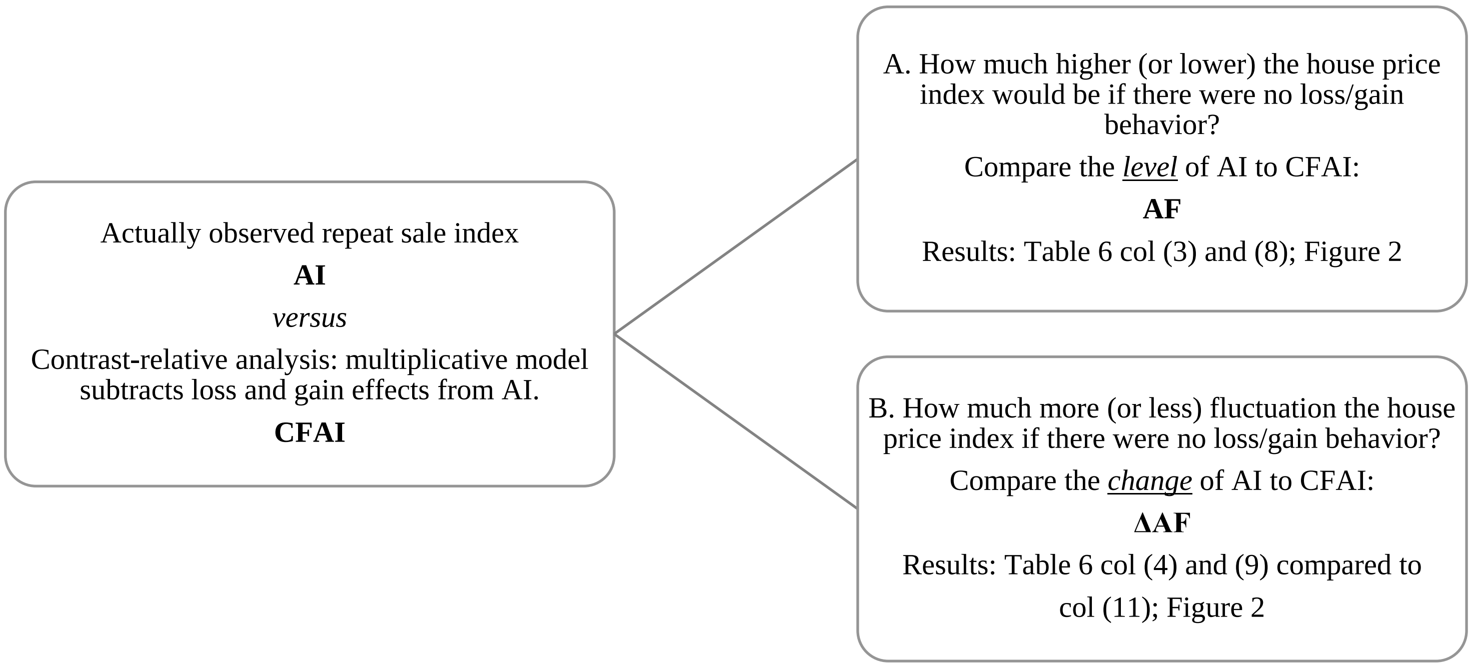
Figure A2. Leave-one-out (LOO) Cross-Validation

This figure shows the results of leave-one-out (LOO) validation exercise comparing predicted values from hedonic regression, marked "PV" in blue dot, to normalized assessed value, marketed "NAV" in red triangle. The validation is performed by town for the top ten towns ranked based on numbers of transitions during our sample period. The horizontal axis is cross-validation statistic of leave-one-out cross-validation CV(LOO). The vertical axis lists top ten towns in our sample.



Appendix 3 Summary of Analytical Framework

This appendix summarizes the analytical framework of contrast-relative analysis.



Appendix 4: Sample Construction

This table displays the sample construction process.

	Observations
Individual residential transactions between 1994 and 2017	1,409,127
Transactions with missing dates	(63)
Transactions with lot size less than 500,000 square feet	(12,764)
Transactions with sale price less than $40,000	(109,625)
Transactions with interior footage less than 300 square feet	(42,874)
Transactions with less than one bedroom	(395,000)
Transactions with structures built earlier than 1799 or after 2018	(66,585)
Transactions with less than 0.5 bathrooms	(9,480)
Transactions with year built later than year sold	(708)
Transactions with property types that are not single-family residential	(150,417)
Transactions without warranty deeds	(67,093)
Transactions with bought and sold on the same date	(5,950)
Final Sample used in the hedonic estimation	**548,568**
Non-repeat Sale	(351,005)
Repeat sales before 2000	(75,469)
Repeat sales in town-year with less than 10 observations	(663)
Repeat sales with holding period less than 12 months	(18,715)
Repeat sales with non-matched buyer and seller	(12,371)
Final Sample of repeat sales	**90,345**

Appendix 5: Hedonic Price Estimation

This table reports the hedonic regression results based on a sample of individual transactions from 1994 to 2017. The following hedonic characteristics are used: interior size, interior size squared, lot size of the property, lot size squared, age of the property, age of the property squared, a dummy variable that is equal to 1 if the number of bathrooms is between 2 and 3 and 0 otherwise, a dummy variable that is equal to 1 if the property has more than 3 bathrooms and 0 otherwise, number of bathrooms. The dependent variable is log transaction price. The inclusion of town-year fixed effects in the hedonic model is essential for calculating expected gains and losses on a consistent basis: see discussion of equations (1) and (2) in the text for details. Robust standard errors are clustered at town level. ***, **, * denote for 1%, 5% and 10% significance, respectively.

Dependent Variable = log of transaction price	
Interior Size	.0003029***
	(1.22e-06)
Interior Size Squared	-1.24e-08***
	(1.33e-10)
Lot Size	1.94e-06***
	(2.75e-08)
Lot Size Squared	-3.74e-12***
	(8.28e-14)
2-3 Bathrooms	.0675572***
	(.001179)
> 3 Bathrooms	.1868796***
	(.0024935)
Age	-.0061597***
	(.0000444)
Age Squared	.0000223***
	(2.80e-07)
Constant	12.05676***
	(.0024308)
Town-Year FE	Yes
Observations	548,568
Adj. R-squared	0.7876

Appendix 6: Repeat Sales Index, Loss Index, and Gain Index

This figure shows the actual repeat sales index (AI, based on repeat sales pairs), loss index, and gain index. The loss (gain) index is calculated as AI plus a loss (gain) adjustment factor. The loss (gain) adjustment factor equals the loss (gain) coefficient, calculated by each period, multiply by magnitude of loss (gain) in year *t*. The loss (gain) index assumes 100% loss (gain). If the loss (gain) coefficient is greater (less) than zero, the loss (gain) index lies above (below) AI. Panel A uses loss and gain based on *Exp. 2nd Price* (PV). Panel B uses loss and gain based on normalized assessed value (NAV). The indices are constructed by year.

Panel A: RSI, Estimated Loss Index, and Estimated Gain Index

Panel B: RSI, Quality-adjusted Loss and Gain Indexes using Assessed Values

Appendix 7: Detailed Summary Statistics

This table supplements Table 2 and reports detailed summary statistics of variables based on a sample of repeat sales transactions from 2000 to 2017. Panel A shows results based on the whole sample period. Panel B shows results based on four sub-periods, normal period (2000-2003), boom period (2004-2006), bust period (2007-2012) and recovery period (2013-2017). Table 1 in the text summarizes variable definitions.

Panel A: Full Sample

	N	Mean	Std. Dev	Q1	Median	Q3
	(1)	(2)	(3)	(4)	(5)	(6)
Log of Second Sale Price	90,345	12.537	0.705	12.100	12.449	12.899
Log of Exp. 2nd Price (PV)	90,345	12.494	0.627	12.051	12.378	12.828
Log of Assessed Value (AV)	89,909	12.040	0.674	11.587	11.952	12.393
Log of Normalized Assessed Value (NAV)	89,909	12.540	0.671	12.077	12.424	12.886
2^{nd} Sale Price / Exp. 2^{nd} Price (PV)	90,345	1.105	0.846	0.916	1.068	1.221
2^{nd} Sale Price / Exp. 2^{nd} Price (NAV)	89,909	1.042	0.784	0.904	1.024	1.143
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (PV)	38,056	1.220	1.085	1.014	1.165	1.324
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (NAV)	33,819	1.094	1.030	0.954	1.083	1.200
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (PV)	52,288	1.021	0.603	0.875	1.008	1.135
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (NAV)	56,090	1.010	0.585	0.887	0.994	1.102
Anchor (PV)	90,345	-0.113	0.410	-0.365	-0.079	0.168
Anchor (NAV)	89,909	-0.156	0.382	-0.390	-0.116	0.108
Exp. Loss Dummy (PV)	90,345	0.421	0.494	0.000	0.000	1.000
Exp. Loss Dummy (NAV)	89,909	0.376	0.484	0.000	0.000	1.000
Exp. Loss (PV)	90,345	0.105	0.184	0.000	0.000	0.168
Exp. Loss (NAV)	89,909	0.077	0.148	0.000	0.000	0.108
Exp. Gain (PV)	90,345	0.219	0.297	0.000	0.079	0.365
Exp. Gain (NAV)	89,909	0.233	0.297	0.000	0.116	0.390
Exp. Loss, Given Loss (PV)	38,056	0.250	0.209	0.102	0.208	0.343
Exp. Loss, Given Loss (NAV)	33,819	0.204	0.178	0.080	0.166	0.285
Exp. Gain, Given Gain (PV)	52,288	0.378	0.303	0.147	0.311	0.535
Exp. Gain, Given Gain (NAV)	56,086	0.374	0.297	0.148	0.314	0.529
Months	90,345	0.673	0.450	0.329	0.559	0.910
Excess Months, Given Loss (PV)	38,056	0.020	0.397	-0.308	-0.052	0.294
Excess Months, Given Loss (NAV)	33,819	0.033	0.384	-0.288	-0.021	0.306
Excess Months, Given Gain (PV)	52,288	-0.015	0.485	-0.367	-0.160	0.182
Excess Months, Given Gain (NAV)	56,090	-0.018	0.486	-0.372	-0.167	0.176

Panel B: By Four Periods

Normal (2000-2003, 16 Qtr) N=9,288	N/Qtr	Mean	Std. Dev	Q1	Median	Q3
Log of Second Sale Price	742	12.441	0.747	11.918	12.324	12.858
Log of Exp. 2nd Price (PV)	742	12.400	0.678	11.898	12.255	12.784
Log of Assessed Value (AV)	718	11.763	0.644	11.313	11.627	12.101
Log of Normalized Assessed Value (NAV)	718	12.457	0.726	11.919	12.307	12.864
2^{nd} Sale Price / Exp. 2^{nd} Price (PV)	742	1.101	0.620	0.903	1.047	1.201
2^{nd} Sale Price / Exp. 2^{nd} Price (NAV)	718	1.042	0.663	0.897	1.005	1.124
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (PV)	102	1.526	0.820	1.194	1.333	1.586
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (NAV)	50	1.395	0.862	1.106	1.261	1.465
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (PV)	640	1.034	0.552	0.881	1.015	1.143
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (NAV)	669	1.016	0.638	0.891	0.996	1.103
Anchor (PV)	742	-0.299	0.320	-0.471	-0.279	-0.106
Anchor (NAV)	718	-0.344	0.289	-0.491	-0.322	-0.174
Exp. Loss Dummy (PV)	742	0.137	0.344	0.000	0.000	0.000
Exp. Loss Dummy (NAV)	718	0.069	0.254	0.000	0.000	0.000
Exp. Loss (PV)	742	0.024	0.098	0.000	0.000	0.000
Exp. Loss (NAV)	718	0.012	0.081	0.000	0.000	0.000
Exp. Gain (PV)	742	0.323	0.279	0.106	0.279	0.471
Exp. Gain (NAV)	718	0.345	0.265	0.157	0.313	0.484
Exp. Loss, Given Loss (PV)	102	0.172	0.210	0.044	0.105	0.217
Exp. Loss, Given Loss (NAV)	50	0.166	0.263	0.037	0.082	0.188
Exp. Gain, Given Gain (PV)	640	0.374	0.266	0.181	0.326	0.510
Exp. Gain, Given Gain (NAV)	669	0.382	0.252	0.207	0.344	0.506
Months	742	0.427	0.216	0.254	0.387	0.565
Excess Months, Given Loss (PV)	102	-0.118	0.165	-0.248	-0.167	-0.039
Excess Months, Given Loss (NAV)	50	-0.154	0.147	-0.261	-0.195	-0.093
Excess Months, Given Gain (PV)	640	0.019	0.217	-0.155	-0.016	0.157
Excess Months, Given Gain (NAV)	669	0.011	0.216	-0.162	-0.025	0.148

Boom (2004-2006, 12 Qtr) N=13,766	N/Qtr	Mean	Std. Dev	Q1	Median	Q3
Log of Second Sale Price	1,477	12.653	0.652	12.211	12.521	12.994
Log of Exp. 2nd Price (PV)	1,477	12.616	0.597	12.187	12.468	12.949
Log of Assessed Value (AV)	1,473	11.814	0.662	11.340	11.669	12.175
Log of Normalized Assessed Value (NAV)	1,473	12.654	0.653	12.190	12.509	13.001
2^{nd} Sale Price / Exp. 2^{nd} Price (PV)	1,477	1.078	0.345	0.918	1.041	1.174
2^{nd} Sale Price / Exp. 2^{nd} Price (NAV)	1,473	1.027	0.257	0.898	1.004	1.122
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (PV)	139	1.501	0.643	1.185	1.310	1.563
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (NAV)	79	1.336	0.431	1.139	1.262	1.432
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (PV)	1,339	1.034	0.261	0.905	1.020	1.138
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (NAV)	1,394	1.010	0.231	0.892	0.995	1.104
Anchor (PV)	1,477	-0.404	0.357	-0.611	-0.373	-0.169
Anchor (NAV)	1,473	-0.440	0.324	-0.624	-0.408	-0.223
Exp. Loss Dummy (PV)	1,477	0.094	0.292	0.000	0.000	0.000
Exp. Loss Dummy (NAV)	1,473	0.054	0.226	0.000	0.000	0.000
Exp. Loss (PV)	1,477	0.016	0.083	0.000	0.000	0.000
Exp. Loss (NAV)	1,473	0.007	0.053	0.000	0.000	0.000
Exp. Gain (PV)	1,477	0.420	0.327	0.169	0.373	0.611
Exp. Gain (NAV)	1,473	0.446	0.310	0.221	0.407	0.623
Exp. Loss, Given Loss (PV)	139	0.175	0.215	0.044	0.106	0.214
Exp. Loss, Given Loss (NAV)	79	0.133	0.186	0.030	0.081	0.163
Exp. Gain, Given Gain (PV)	1,339	0.463	0.312	0.228	0.410	0.636
Exp. Gain, Given Gain (NAV)	1,394	0.473	0.298	0.254	0.429	0.638
Months	1,477	0.476	0.285	0.250	0.400	0.644
Excess Months, Given Loss (PV)	139	-0.211	0.145	-0.310	-0.246	-0.160
Excess Months, Given Loss (NAV)	79	-0.253	0.096	-0.323	-0.276	-0.216
Excess Months, Given Gain (PV)	1,339	0.022	0.287	-0.207	-0.048	0.199
Excess Months, Given Gain (NAV)	1,394	0.014	0.285	-0.213	-0.058	0.185

Appendix 7 Panel B continued

Bust (2007-2012, 24 Qtr) N=24,689	N/Qtr	Mean	Std. Dev	Q1	Median	Q3
Log of Second Sale Price	1,166	12.547	0.707	12.128	12.468	12.899
Log of Exp. 2nd Price (PV)	1,166	12.519	0.611	12.091	12.401	12.836
Log of Assessed Value (AV)	1,166	12.170	0.673	11.750	12.060	12.512
Log of Normalized Assessed Value (NAV)	1,166	12.547	0.661	12.100	12.442	12.886
2^{nd} Sale Price / Exp. 2^{nd} Price (PV)	1,166	1.104	1.242	0.901	1.061	1.214
2^{nd} Sale Price / Exp. 2^{nd} Price (NAV)	1,166	1.056	1.117	0.896	1.028	1.156
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (PV)	588	1.194	1.502	0.972	1.132	1.291
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (NAV)	560	1.100	1.341	0.934	1.077	1.207
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (PV)	578	1.012	0.894	0.858	0.999	1.128
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (NAV)	606	1.015	0.859	0.875	0.991	1.105
Anchor (PV)	1,166	-0.041	0.393	-0.263	0.004	0.211
Anchor (NAV)	1,166	-0.069	0.377	-0.280	-0.016	0.175
Exp. Loss Dummy (PV)	1,166	0.504	0.500	0.000	1.000	1.000
Exp. Loss Dummy (NAV)	1,166	0.480	0.500	0.000	0.000	1.000
Exp. Loss (PV)	1,166	0.128	0.196	0.000	0.004	0.211
Exp. Loss (NAV)	1,166	0.108	0.171	0.000	0.000	0.175
Exp. Gain (PV)	1,166	0.169	0.270	0.000	0.000	0.263
Exp. Gain (NAV)	1,166	0.176	0.273	0.000	0.016	0.280
Exp. Loss, Given Loss (PV)	588	0.253	0.211	0.103	0.209	0.349
Exp. Loss, Given Loss (NAV)	560	0.225	0.187	0.091	0.183	0.310
Exp. Gain, Given Gain (PV)	578	0.341	0.297	0.117	0.266	0.481
Exp. Gain, Given Gain (NAV)	606	0.339	0.296	0.117	0.267	0.486
Months	1,166	0.634	0.372	0.354	0.559	0.835
Excess Months, Given Loss (PV)	588	-0.105	0.261	-0.305	-0.146	0.059
Excess Months, Given Loss (NAV)	560	-0.113	0.247	-0.304	-0.149	0.048
Excess Months, Given Gain (PV)	578	0.107	0.433	-0.243	0.035	0.397
Excess Months, Given Gain (NAV)	606	0.104	0.432	-0.248	0.033	0.397

Recovery (2013-2017, 20 Qtr) N=29,798	N/Qtr	Mean	Std. Dev	Q1	Median	Q3
Log of Second Sale Price	1,638	12.502	0.705	12.061	12.437	12.861
Log of Exp. 2nd Price (PV)	1,638	12.440	0.626	11.985	12.328	12.771
Log of Assessed Value (AV)	1,638	12.148	0.631	11.725	12.048	12.458
Log of Normalized Assessed Value (NAV)	1,638	12.502	0.660	12.040	12.394	12.827
2^{nd} Sale Price / Exp. 2^{nd} Price (PV)	1,638	1.122	0.673	0.935	1.099	1.257
2^{nd} Sale Price / Exp. 2^{nd} Price (NAV)	1,638	1.038	0.655	0.917	1.038	1.150
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (PV)	1,032	1.191	0.726	1.019	1.163	1.309
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Loss (NAV)	932	1.065	0.761	0.956	1.074	1.179
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (PV)	606	1.005	0.554	0.848	0.989	1.132
2^{nd} Sale Price / Exp. 2^{nd} Price, Given Gain (NAV)	706	1.002	0.478	0.888	0.994	1.096
Anchor (PV)	1,638	0.050	0.369	-0.123	0.099	0.270
Anchor (NAV)	1,638	-0.012	0.331	-0.145	0.039	0.186
Exp. Loss Dummy (PV)	1,638	0.630	0.483	0.000	1.000	1.000
Exp. Loss Dummy (NAV)	1,638	0.569	0.495	0.000	1.000	1.000
Exp. Loss (PV)	1,638	0.164	0.206	0.000	0.099	0.270
Exp. Loss (NAV)	1,638	0.111	0.158	0.000	0.039	0.186
Exp. Gain (PV)	1,638	0.114	0.236	0.000	0.000	0.123
Exp. Gain (NAV)	1,638	0.123	0.239	0.000	0.000	0.145
Exp. Loss, Given Loss (PV)	1,032	0.260	0.206	0.118	0.223	0.353
Exp. Loss, Given Loss (NAV)	932	0.195	0.165	0.080	0.162	0.274
Exp. Gain, Given Gain (PV)	606	0.309	0.301	0.090	0.213	0.442
Exp. Gain, Given Gain (NAV)	706	0.285	0.294	0.074	0.188	0.412
Months	1,638	0.901	0.533	0.465	0.838	1.241
Excess Months, Given Loss (PV)	1,032	-0.032	0.406	-0.361	-0.034	0.263
Excess Months, Given Loss (NAV)	932	-0.020	0.384	-0.314	-0.008	0.262
Excess Months, Given Gain (PV)	606	0.054	0.695	-0.549	-0.182	0.663
Excess Months, Given Gain (NAV)	706	0.026	0.681	-0.550	-0.226	0.623

Appendix 8: Double Mean Differences Results Using Asking Price

This table shows robustness tests for Table 3 Panel B of mean differences (DD) tests. The sample period is from 2000-2013. "Asking Price / Exp. 2nd Price" is the ratio of asking price to expected 2nd sale price. ***, **, * denote for 1%, 5% and 10% significance, respectively.

		PV			***NAV***		
		Gain	Loss	Col. (2) – (1)	Gain	Loss	Col. (5) – (4)
		(1)	(2)	(3)	(4)	(5)	(6)
Boom-Normal							
Asking Price / Exp. 2nd Price	(1) Normal	1.125	1.646	0.520***	1.088	1.522	0.434***
	(2) Boom	1.120	1.583	0.463***	1.062	1.365	0.303***
	Row (2) – (1)	-0.006	-0.063		-0.026	-0.157*	
	DD			**-0.057**			**-0.131****
Bust-Boom							
Asking Price / Exp. 2nd Price	(5) Boom	1.120	1.583	0.463***	1.062	1.365	0.303***
	(6) Bust	1.109	1.297	0.188***	1.098	1.190	0.092***
	Row (6) – (5)	-0.011	-0.286***		0.036***	-0.175***	
	DD			**-0.275*****			**-0.211*****
Recovery-Bust							
Asking Price / Exp. 2nd Price	(9) Bust	1.109	1.297	0.188***	1.098	1.190	0.092***
	(10) Recovery	1.089	1.269	0.181***	1.078	1.142	0.065***
	Row (10) – (9)	-0.020	-0.028		-0.020	-0.047**	
	DD			**-0.008**			**-0.028**
Bust-Normal							
Asking Price / Exp. 2nd Price	DD			**-0.322*****			**-0.342*****
Recovery-Normal							
Asking Price / Exp. 2nd Price	DD			**-0.339*****			**-0.370*****
Recovery-Boom							
Asking Price / Exp. 2nd Price	DD			**-0.282*****			**-0.238*****

Appendix 9: Robustness to Using Alternative Cycle Classification

According to the Federal Housing Finance Agency (FHFA) and our actual repeat sale index (AI) in Figure 2, Connecticut experienced an early boom, starting roughly from 2002, and an early bust. Housing prices in Connecticut started to fall starting from 2006.[4] Based on these facts, we redefine the four cycles as normal from 2000-2001, boom from 2002-2005, bust from 2006-2012, and recovery from 2013-2017. The intuition behind this robustness test is to classify the peak year, 2006, as a bust year instead of a boom year. The purpose of this appendix is to discuss our main results using this alternative cycle classification.

Table A9.1 reproduces the main results in Table 4. The dependent variable is log price of the second sale. In Column (1)-(4), the expected 2nd price is the predicted value (PV) estimated by a standard hedonic model, the 1st stage regression. In Column (5)-(8), the expected 2nd price is the normalized assessed value (NAV). The results suggest that coefficient estimates are robust to the alternative classification of the housing cycle.

Figure A9 and Table A9.2 are compared with Figure 2 and Table 6, respectively. Figure A9 summarizes Table A9.2 by comparing annual levels of the observed repeat sales index (i.e. actual index, *AI*) to the quality-adjusted loss/gain factor adjusted index (*CFAI (NAV)*) and the CFAI estimated without quality adjustment (*CFAI (PV)*). Consistent with Figure 2, contrast-relative loss/gain behavior differs substantially from the observed repeat sales index, showing that loss/gain behavior decreased the observed change during normal and boom periods and increased the observed change during the bust: i.e., holding loss/gain behavior constant suggests that the contrast-relative change would have been substantially greater than observed in normal and boom, less during the bust. Slightly different from Figure 2, Figure A9 shows *CFAI (NAV)* deviates the most from *CFAI (PV)* in normal and bust.

In Table A9.2, we focus on the NAV results and the comparison between the *change* of AI and CFAI after quality adjustment, columns (9)-(11). Column (9) shows that changes in AF (= ΔLAF + ΔGAF)) are negative during the normal period (2001), meaning that loss/gain behavior is associated with reduced changes in the observed repeat sales index. The average reduction during the three-year period is about 26% (=-.018/.069, where -.018 is ΔAF and .069 is ΔAI, both in 2001) of changes in the observed.

Changes in AF are negative during the boom period (2002-2005), meaning that loss/gain behavior decreased observed changes in the repeat sales index. Anchoring was associated with an average of 13% (=.012/-.096, where -.012 is the average of ΔAF and .096 is the average ΔAI, both in 2002-2005) in observed price changes during the boom periods.

Changes in AF are positive during the bust period (2006-2012), meaning that loss/gain behavior counterfactually increased observed changes in the repeat sales index (i.e., they are less negative). The average increase during the six-year period is 39% (=.021/-.054) of the observed changes. Our re-classification does not change in recovery (2013-2017). The average effect is very small, suggesting loss/gain behavior had little effect on changes in the repeat sales index.

---

[4] Specifically, the FHFA Purchase-Only Indexes (Estimated using Sales Price Data, https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index-Datasets.aspx#qpo) suggest the house prices in Connecticut peaked in 2006Q2. Our repeat sales index (estimated by quarter) suggests the peak in 2006Q3.

Overall, our findings are highly consistent when we use a more accurately defined housing cycle in the state of Connecticut. The results suggest that anchoring was associated with reductions in observed changes in house prices during the boom (2002-2005) and with reduced price declines during the bust (2006-2012).

Table A9.1: Results Using Alternative Cycle Classification

This table shows robustness tests using alternative cycle classification, normal from 2000-2001, boom from 2002-2005, bust from 2006-2012, and recovery from 2013-2017. Results in this table are comparable to Table 3 in the text. The dependent variable is log price of the second sale. In Column (1)-(4), the expected 2nd price is the predicted value (PV) estimated by a standard hedonic model, the 1st stage regression. In Column (5)-(8), the expected 2nd price is the normalized assessed value (NAV). Months is the number of month since the first sale. All price variables are in logs. All the specifications include town-year fixed effects. Bootstrapped standard errors are used. ***, **, * denote for 1%, 5% and 10% significance, respectively.

	PV				NAV			
	Normal	Boom	Bust	Recovery	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exp. Loss	1.075***	0.920***	0.095***	0.149***	0.920***	0.635***	0.103***	0.067***
	(0.129)	(0.068)	(0.019)	(0.020)	(0.099)	(0.070)	(0.027)	(0.015)
Exp. Gain	-0.544***	-0.375***	0.261***	0.360***	-0.368***	-0.327***	-0.203***	0.021*
	(0.074)	(0.022)	(0.019)	(0.013)	(0.048)	(0.015)	(0.013)	(0.011)
Months	0.151***	0.211***	-0.145***	-0.122***	0.079***	0.140***	0.034***	-0.092***
	(0.031)	(0.012)	(0.009)	(0.003)	(0.025)	(0.013)	(0.009)	(0.003)
Residual	-0.004	0.109***	0.616***	0.567***	0.025	-0.029**	0.025**	0.124***
	(0.073)	(0.023)	(0.017)	(0.012)	(0.034)	(0.014)	(0.012)	(0.008)
Exp. 2nd Price	0.960***	0.887***	0.938***	0.985***	0.948***	0.865***	0.878***	0.928***
	(0.019)	(0.008)	(0.006)	(0.005)	(0.018)	(0.008)	(0.006)	(0.005)
Constant	0.592**	1.496***	0.830***	0.289***	0.709***	1.770***	1.550***	0.973***
	(0.230)	(0.100)	(0.079)	(0.070)	(0.222)	(0.103)	(0.080)	(0.061)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	4,369	18,908	34,308	32,760	4,125	18,759	34,268	32,757
R-squared	0.899	0.886	0.827	0.847	0.911	0.899	0.839	0.867

Figure A9: Repeat Sale Index and the Loss-Gain-Factor Adjusted Index (CFAI) Using Alternative Cycle Classification

This figure shows robustness tests using alternative cycle classification, normal from 2000-2001, boom from 2002-2005, bust from 2006-2012, and recovery from 2013-2017. Results in this figure are comparable to Figure 2 in the text. The repeat sales index (AI) and the loss-gain-factor adjusted index (CFAI) are based on a sample of repeat transactions from 2000 to 2017. Both indices are calculated in the exponential values. The CFAI is calculated as the actual repeat sales index (AI) minus the loss (gain) adjustment factors, LAF (GAF). LAF (GAF) is the loss (gain) variable for a given period (i.e. normal, boom, bust and recovery) multiplied by the mean magnitude of loss (gain) among sold properties that were facing an expected loss (gain) and the proportion of sellers facing an potential loss (gain) in that period. The AI minus CFAI suggests the potential magnitude of loss/gain behavior. I.e., If the difference is negative (positive), then the behavior is associated with dampened (accentuated) price movements if the actual change is positive in that part of the cycle, $\Delta AI>0$. This logic reverses if $\Delta AI<0$. "CFAI (PV)" indicates CFAI is calculated based on *Exp. 2nd Price (PV).* "CFAI (NAV)" indicates CFAI is based on normalized assessed value.

Table A9.2: Tests of Significance - Actual Housing Price Index minus Actual Factor Adjusted Index (CFAI) Using Alternative Cycle Classification

This table shows robustness tests using alternative cycle classification, normal from 2000-2001, boom from 2002-2005, bust from 2006-2012, and recovery from 2013-2017. Results in this table are comparable to Table 6 in the text. This table summarizes results based on tests of significance on the difference between the actual housing price index (AI) and the contrast-relative loss-gain-factor-adjusted index (CFAI). Loss and gain factors in columns (1)-(5) are based on Exp. 2nd price (PV) and those in columns (6)-(10) are based on normalized assessed value (NAV). The total adjusted factor (AF) consists of the loss adjusted factor (LAF, in column (1) and (6)) and the gain adjusted factor (GAF, in column (2) and (7)). The total adjustment factor (AF) (Column (3) and (8)) is the sum of these two: AF = LAF + GAF and the contrast-relative adjusted index, CFAI = AI – AF. Column (4) and (9) show the changes of total adjustment factors (ΔAF) (i.e. ΔLAF + ΔGAF). Column (5) and (10) shows the changes of the contrast-relative loss-gain-factor adjusted factor index (ΔCFAI). Column (11) shows the changes of actual housing price index (ΔAI). By construction, ΔAI – ΔCFAI = ΔAF and columns (4)-(5) and (9)-(10) are recorded as missing in 2000. Numbers in bold denote for *p*-value of *F* statistics significant at 5%. We cannot calculate *F* statistics for any change variables since they are based on only two numbers.

	PV ---					**NAV** --					
	LAF	GAF	AF	ΔAF	ΔCFAI	LAF	GAF	AF	ΔAF	ΔCFAI	ΔAI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
2000	**0.038**	**-0.134**	**-0.096**			**0.016**	**-0.096**	**-0.080**			
2001	**0.030**	**-0.151**	**-0.121**	-0.026	0.094	**0.012**	**-0.110**	**-0.098**	-0.018	0.087	0.069
2002	**0.019**	**-0.124**	**-0.105**	0.016	0.071	**0.008**	**-0.115**	**-0.107**	-0.009	0.096	0.087
2003	**0.017**	**-0.142**	**-0.125**	-0.020	0.110	**0.005**	**-0.132**	**-0.127**	-0.020	0.110	0.089
2004	**0.014**	**-0.158**	**-0.145**	-0.019	0.122	**0.004**	**-0.147**	**-0.143**	-0.016	0.118	0.102
2005	**0.013**	**-0.163**	**-0.150**	-0.005	0.109	**0.003**	**-0.151**	**-0.148**	-0.005	0.109	0.104
2006	**0.002**	**0.105**	**0.107**	0.257	-0.215	**0.001**	**-0.087**	**-0.086**	0.062	-0.020	0.042
2007	**0.003**	**0.081**	**0.084**	-0.023	-0.038	**0.002**	**-0.068**	**-0.066**	0.021	-0.081	-0.060
2008	**0.008**	**0.048**	**0.056**	-0.028	-0.109	**0.009**	**-0.041**	**-0.032**	0.033	-0.170	-0.137
2009	**0.015**	**0.033**	**0.048**	-0.009	-0.105	**0.014**	**-0.025**	**-0.011**	0.021	-0.135	-0.113
2010	**0.015**	**0.031**	**0.046**	-0.002	-0.023	**0.013**	**-0.025**	**-0.011**	0.000	-0.024	-0.024
2011	**0.017**	**0.030**	**0.047**	0.000	-0.049	**0.015**	**-0.023**	**-0.008**	0.003	-0.051	-0.048
2012	**0.019**	**0.026**	**0.045**	-0.002	-0.033	0.017	**-0.020**	-0.004	0.005	-0.040	-0.035
2013	**0.027**	**0.037**	**0.064**	0.020	-0.012	**0.009**	0.002	**0.012**	0.015	-0.008	0.008
2014	**0.027**	**0.036**	**0.063**	-0.001	-0.033	**0.009**	0.002	**0.011**	0.000	-0.034	-0.035
2015	**0.026**	**0.039**	**0.065**	0.002	0.005	**0.008**	0.002	**0.010**	-0.001	0.007	0.006
2016	**0.024**	**0.043**	**0.067**	0.003	0.020	**0.007**	0.003	**0.010**	-0.001	0.023	0.023
2017	**0.021**	**0.047**	**0.067**	0.000	0.025	**0.006**	0.003	**0.008**	-0.001	0.026	0.025

Appendix 10.1: Baseline Results with Alternative Fixed Effects

This table shows robustness tests for our baseline results using the whole sample period (Model 1 in Panel B of Table 4). Result with town-year fixed effects in model 8 is compared with no fixed effect in column (1), county fixed effects in column (2), town fixed effects in column (3), year fixed effects in column (4), county fixed effects and year fixed effects in column (5), county-year fixed effects in column (6), and town fixed effects and year fixed effects in column (7). Dependent variable is log price of the second sale. *Exp. Loss* (*Exp. Gain*) is the difference between the first sale price and the expected 2^{nd} sale price truncated above (below) at zero. The expected 2^{nd} price is the normalized assessed value (NAV). All price variables are in logs. All the specifications include town-year fixed effects. Bootstrapped standard errors are used. ***, **, * denote for 1%, 5% and 10% significance, respectively.

	No FE	County FE	Town FE	Year FE	County FE Year FE	County-year FE	Town FE Year FE	Town-year FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exp. Loss (NAV)	0.056***	0.041***	0.032**	0.091***	0.069***	0.072***	0.046***	0.079***
	(0.016)	(0.011)	(0.013)	(0.013)	(0.015)	(0.014)	(0.014)	(0.016)
Exp. Gain (NAV)	-0.117***	-0.117***	-0.090***	-0.157***	-0.150***	-0.152***	-0.108***	-0.116***
	(0.005)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)
Months	-0.056***	-0.058***	-0.062***	-0.039***	-0.045***	-0.045***	-0.058***	-0.058***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Residual	0.037***	0.047***	0.089***	0.012***	0.028***	0.026***	0.084***	0.082***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Exp. 2nd Price (NAV)	0.976***	0.954***	0.893***	0.979***	0.956***	0.956***	0.882***	0.875***
	(0.001)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)	(0.004)	(0.003)
Constant	0.358***	0.638***	1.401***	0.320***	0.594***	0.616***	1.485***	1.624***
	(0.018)	(0.029)	(0.034)	(0.018)	(0.021)	(0.025)	(0.045)	(0.042)
Observations	89,909	89,909	89,909	89,909	89,909	89,909	89,909	89,909
R-squared	0.860	0.860	0.862	0.860	0.861	0.861	0.863	0.863

Appendix 10.2: Baseline Results Restricting 1st Stage to Repeat Sales Only

This table shows robustness tests for Table 4 by restricting the sample to repeat sales only. Dependent variable is log price of the second sale. *Exp. Loss* (*Exp. Gain*) is the difference between the first sale price and the expected 2nd sale price truncated above (below) at zero. In Panel A, the expected 2nd price is the predicted value (PV) estimated by a standard hedonic model, the 1st stage regression using repeat sales only. In Panel B, the expected 2nd price is the normalized assessed value (NAV). All price variables are in logs. All the specifications include town-year fixed effects. Bootstrapped standard errors are used. ***, **, * denote for 1%, 5% and 10% significance, respectively.

Panel A: PV. Dependent is log of 2nd sales price.

	All	Pre-07	Post-07	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exp. Loss (PV)	0.137***	0.868***	0.132***	0.934***	0.832***	0.053**	0.162***
	(0.013)	(0.055)	(0.013)	(0.077)	(0.053)	(0.025)	(0.016)
Exp. Gain (PV)	-0.189***	0.336***	-0.294***	0.466***	0.285***	-0.307***	-0.327***
	(0.008)	(0.021)	(0.012)	(0.033)	(0.021)	(0.019)	(0.018)
Months	0.562***	0.129***	0.576***	0.072**	0.139***	0.663***	0.548***
	(0.008)	(0.023)	(0.010)	(0.034)	(0.020)	(0.021)	(0.015)
Residual	-0.116***	0.185***	-0.123***	0.212***	0.162***	-0.157***	-0.116***
	(0.003)	(0.014)	(0.003)	(0.019)	(0.016)	(0.011)	(0.003)
Exp. 2nd Price (PV)	0.953***	0.904***	0.977***	0.924***	0.888***	0.958***	0.992***
	(0.004)	(0.006)	(0.004)	(0.013)	(0.009)	(0.009)	(0.007)
Constant	0.657***	1.271***	0.367***	1.018***	1.476***	0.602***	0.188**
	(0.052)	(0.081)	(0.052)	(0.164)	(0.108)	(0.113)	(0.084)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	90,345	29,592	60,753	11,866	17,726	27,993	32,760
R-squared	0.848	0.887	0.833	0.890	0.881	0.820	0.846

Panel B: NAV. Dependent is log of 2nd sales price.

	All	Pre-07	Post-07	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exp. Loss (NAV)	0.076***	0.716***	0.036***	0.780***	0.585***	0.058**	0.066***
	(0.015)	(0.059)	(0.013)	(0.071)	(0.078)	(0.029)	(0.017)
Exp. Gain (NAV)	-0.112***	-0.353***	-0.034***	-0.363***	-0.348***	-0.146***	0.025**
	(0.006)	(0.011)	(0.009)	(0.021)	(0.015)	(0.014)	(0.010)
Months	-0.058***	0.160***	-0.078***	0.140***	0.164***	0.000	-0.092***
	(0.003)	(0.007)	(0.003)	(0.014)	(0.011)	(0.007)	(0.003)
Residual	0.089***	-0.058***	0.123***	0.006	-0.096***	0.083***	0.131***
	(0.005)	(0.010)	(0.007)	(0.019)	(0.011)	(0.012)	(0.009)
Exp. 2nd Price (NAV)	0.872***	0.872***	0.901***	0.904***	0.850***	0.890***	0.925***
	(0.003)	(0.005)	(0.004)	(0.012)	(0.006)	(0.008)	(0.005)
Constant	1.664***	1.668***	1.302***	1.257***	1.971***	1.400***	1.012***
	(0.043)	(0.065)	(0.046)	(0.147)	(0.079)	(0.096)	(0.057)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	89,909	29,176	60,733	11,495	17,681	27,976	32,757
R-squared	0.863	0.898	0.851	0.901	0.894	0.834	0.867

Appendix 10.3: Baseline Results Using Asking Price as Dependent Variable

This table shows robustness tests using asking price as dependent variable. The sample period is from 2000-2013. Panel A and B show robustness tests for Table 4. The Dependent variable is log of initial asking price. ***, **, * denote for 1%, 5% and 10% significance, respectively.

Panel A: Dependent is Log of Asking Price (PV measures based on expected 2nd sales price from the hedonic model, equation (1))

	All	Pre-07	Post-07	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exp. Loss (PV)	0.234***	0.764***	0.149***	0.822***	0.725***	0.157***	0.132
	(0.032)	(0.094)	(0.036)	(0.134)	(0.138)	(0.043)	(0.099)
Exp. Gain (PV)	0.121***	-0.268***	0.296***	-0.384***	-0.189**	0.276***	0.410***
	(0.019)	(0.054)	(0.029)	(0.096)	(0.091)	(0.026)	(0.114)
Months	0.516***	0.185***	0.593***	0.104	0.242***	0.585***	0.630***
	(0.021)	(0.057)	(0.031)	(0.090)	(0.087)	(0.027)	(0.071)
Residual	-0.081***	0.144***	-0.140***	0.201***	0.091	-0.135***	-0.147***
	(0.015)	(0.044)	(0.015)	(0.045)	(0.071)	(0.015)	(0.032)
Exp. 2nd Price (PV)	0.984***	0.950***	1.011***	0.981***	0.927***	1.013***	1.000***
	(0.009)	(0.015)	(0.011)	(0.021)	(0.021)	(0.012)	(0.029)
Constant	0.301***	0.729***	0.011	0.334	1.016***	-0.025	0.165
	(0.113)	(0.192)	(0.144)	(0.263)	(0.262)	(0.156)	(0.381)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	31,838	1,949	15,889	6,543	9,406	13,461	2,428
R-squared	0.722	0.697	0.750	0.732	0.662	0.775	0.646

Panel B: Dependent is Log of Asking Price (Normalized Assessed Values, NAV measures expected 2nd sales price as described in Section 2.2)

	All	Pre-07	Post-07	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exp. Loss (NAV)	0.273***	0.802***	0.180***	0.926***	0.578***	0.208***	0.125*
	(0.035)	(0.092)	(0.039)	(0.125)	(0.133)	(0.052)	(0.068)
Exp. Gain (NAV)	-0.188***	-0.300***	-0.079***	-0.395***	-0.229***	-0.113***	0.031
	(0.018)	(0.029)	(0.025)	(0.046)	(0.045)	(0.021)	(0.069)
Months	0.033***	0.131***	-0.022	0.184***	0.079**	0.005	-0.082***
	(0.013)	(0.024)	(0.013)	(0.029)	(0.040)	(0.013)	(0.029)
Residual	0.011	-0.043	0.052**	-0.085*	-0.008	0.033	0.095*
	(0.016)	(0.027)	(0.021)	(0.047)	(0.046)	(0.022)	(0.051)
Exp. 2nd Price (NAV)	0.912***	0.894***	0.948***	0.949***	0.853***	0.953***	0.949***
	(0.009)	(0.013)	(0.011)	(0.020)	(0.017)	(0.011)	(0.023)
Constant	1.191***	1.423***	0.746***	0.728***	1.954***	0.681***	0.741**
	(0.115)	(0.168)	(0.142)	(0.248)	(0.222)	(0.138)	(0.295)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	31,631	15,745	15,886	6,350	9,395	13,458	2,428
R-squared	0.732	0.707	0.760	0.744	0.674	0.785	0.656

Appendix 10.4: Baseline Results Controlling for LTV

This table shows robustness tests for Table 4 including LTV, defined as the greater of the difference between the loan-to-value ratio and 0.8, and zero. Loan-to-value ratio is the mortgage balance at second sale divided by the initial purchase price inflated at a hedonic price index. Dependent variable is log price of the second sale. *Exp. Loss* (*Exp. Gain*) is the difference between the first sale price and the expected 2nd selling price truncated above (below) at zero. In Panel A, the expected 2nd price is the predicted value (PV) estimated by a standard hedonic model, the 1st stage regression: see Appendix 3. In Panel B, the expected 2nd price is the normalized assessed value (NAV). All price variables are in logs. All the specifications include town-year fixed effects. Bootstrapped standard errors are used. ***, **, * denote for 1%, 5% and 10% significance, respectively.

Panel A: PV. Dependent is log of 2nd sales price.

	All	Pre-07	Post-07	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exp. Loss (PV)	0.107***	0.812***	0.118***	0.848***	0.794***	0.029	0.154***
	(0.014)	(0.049)	(0.013)	(0.083)	(0.062)	(0.027)	(0.019)
Exp. Gain (PV)	0.218***	-0.330***	0.321***	-0.459***	-0.279***	0.334***	0.358***
	(0.009)	(0.018)	(0.011)	(0.038)	(0.020)	(0.021)	(0.013)
Months	0.595***	0.163***	0.595***	0.117***	0.166***	0.692***	0.562***
	(0.010)	(0.019)	(0.011)	(0.037)	(0.022)	(0.021)	(0.012)
Residual	-0.121***	0.191***	-0.129***	0.223***	0.166***	-0.164***	-0.122***
	(0.003)	(0.013)	(0.003)	(0.019)	(0.014)	(0.010)	(0.003)
Exp. 2nd Price (PV)	0.947***	0.897***	0.970***	0.919***	0.881***	0.952***	0.985***
	(0.004)	(0.007)	(0.004)	(0.011)	(0.007)	(0.008)	(0.005)
LTV	0.079***	0.286***	0.017	0.377***	0.240***	0.052**	-0.016
	(0.011)	(0.028)	(0.013)	(0.058)	(0.034)	(0.021)	(0.016)
Constant	0.726***	1.346***	0.453***	1.079***	1.560***	0.661***	0.295***
	(0.046)	(0.087)	(0.050)	(0.138)	(0.090)	(0.103)	(0.066)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	90,345	29,592	60,753	11,866	17,726	27,993	32,760
R-squared	0.848	0.890	0.834	0.893	0.884	0.820	0.847

Panel A: NAV. Dependent is log of 2nd sales price.

	All	Pre-07	Post-07	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exp. Loss (NAV)	0.077***	0.691***	0.047***	0.743***	0.569***	0.062**	0.080***
	(0.015)	(0.049)	(0.018)	(0.089)	(0.076)	(0.027)	(0.016)
Exp. Gain (NAV)	-0.116***	-0.374***	-0.038***	-0.392***	-0.364***	-0.153***	0.021*
	(0.007)	(0.011)	(0.009)	(0.024)	(0.014)	(0.011)	(0.013)
Months	-0.057***	0.179***	-0.079***	0.164***	0.179***	0.005	-0.094***
	(0.002)	(0.010)	(0.003)	(0.013)	(0.013)	(0.008)	(0.003)
Residual	0.083***	-0.048***	0.110***	0.016	-0.088***	0.073***	0.117***
	(0.007)	(0.008)	(0.008)	(0.018)	(0.012)	(0.010)	(0.011)
Exp. 2nd Price (NAV)	0.875***	0.871***	0.905***	0.903***	0.848***	0.894***	0.928***
	(0.004)	(0.005)	(0.004)	(0.010)	(0.007)	(0.007)	(0.005)
LTV	0.009	0.216***	-0.035***	0.302***	0.169***	0.005	-0.054***
	(0.009)	(0.024)	(0.012)	(0.052)	(0.026)	(0.016)	(0.015)
Constant	1.626***	1.681***	1.255***	1.259***	1.986***	1.348***	0.972***
	(0.045)	(0.062)	(0.048)	(0.119)	(0.093)	(0.088)	(0.065)
Town-Year FE	Y	Y	Y	Y	Y	Y	Y
Observations	89,909	29,176	60,733	11,495	17,681	27,976	32,757
R-squared	0.863	0.900	0.851	0.904	0.895	0.834	0.867

Appendix 11: Baseline Results Corrected for Unobserved Quality

This appendix summarizes Clapp and Zhou (CZ, 2019) who develop an algorithm for simulating unobserved quality for each sale.[5] Simulated values are used to estimate true values for expected sales price, loss and gain. The relevance of this Appendix is that contrast-relative estimates over the cycle using CZ's method are close to those using Table 3, Panel B coefficients: i.e., the stylized facts established in the body of this paper are robust to using an entirely different method to estimate unobserved quality. Table A11.1 contains results comparable to Table 3 Panel B, NAV in this paper. Table A11.2 contains results comparable to Table 6 in this paper. Figure A.9 plots the housing cycle using CZ's method side-by-side with the NAV method: although CZ's quality adjustment produces a more conservable result, the pattern shows that NAV results are robust to CZ quality adjustment.

CZ propose a new method for using the first residual to estimate unobserved quality beginning with equation (A11.1):

$$P_{ilt} = \beta_0 + \beta X_{il} + FE_{lt} + \varepsilon_{ilt} \qquad \text{(A11.1)}$$

where P_{ilt} is the natural log of sales price of property *i* at location *l* in time *t* and the subscript *t* = *p, s, o* indexes the first (*p*), second (*s*) or one-only (*o*) sale of property i, X_{il} is a vector of property and locational variables and ε_{ilt} is a zero-mean disturbance term. CZ control for time and spatial effects, notably for variation in local public services and taxes, with a dummy for each year in each town (i.e., town-year fixed effects, FE_{lt}). Equation (A11.1) is the first stage of the simulation algorithm. The price predicted from equation (A11.1) defines "market value": $\hat{P}_{ilt}$ is the most likely sales price based on information typically available in the market and to the econometrician.

The residual from equation (A11.1) contains qualities unobserved by the econometrician:

$$\varepsilon_{ilt} = v_{il} + w_{ilt}\ . \qquad \text{(A11.2)}$$

Here, w_{ilt} is an iid disturbance term and v_{il} is an unobserved component such as a view or busy street that persists between the two sales. The two are assumed to be independently and normally distributed. This is the logic behind including ε_{ilt} as a noisy proxy for unobserved quality in the Genesove and Mayer (2001)'s model. Here, CZ propose an entirely new way of modeling equation (A11.2).

Because of zero covariance in equation (A11.2) the variance of the first residual $(v_{il} + w_{ilp})$, a known quantity, is the sum of variances of unknown quantities:

$$var(\varepsilon_{ilp}) = var(v_{il}) + var(w_{ilp})\ . \qquad \text{(A11.3)}$$

Assuming joint normality, CZ grid $var(v_{il})$ from zero to $var(\hat{\varepsilon}_{ilp}) = var(\widehat{v_{il} + w_{ilp}})$ which is known from the first stage hedonic, equation (A11.1). The distribution of v_{il}, conditional on $v_{il} + \omega_{ilp}$ is

[5] The paper is available on SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3146943

(based on Bayes rule and probability density function of normal distribution) $v_{il}|(v_{il} + w_{ilp})\sim N(\frac{(v+w)\sigma_v^2}{(\sigma_v^2+\sigma_w^2)}, \frac{\sigma_v^2\sigma_w^2}{(\sigma_v^2+\sigma_w^2)})$.[6] Therefore, one can calculate w_{ilp}^s as observed 1st stage residual, $\hat{\varepsilon}_{ilp}$, minus simulated v_{il}, v_{il}^s. Here, the superscript indexes a simulation, draws from the conditional distribution.

Persistence in v_{il} is important because it implies that expected loss and gain – i.e., the variables that are intended to measure anchoring behavior – will be incorrectly estimated by the econometrician. CZ simulate the expected value of the second sales price, including a simulated value for the unobserved quality, at each grid point for mean $var(v_{il})$:

$$\mu_{ils}^s = \hat{\beta}X_{il} + FE_{ils} + v_{il}^s = \hat{P}_{ils} + v_{il}^s \ . \qquad \text{(A11.4)}$$

Here μ_{ils}^s is a simulated value for the true expected second sales price and v_{il}^s is a draw conditional on the known first residual for each data point as explained above. $\hat{P}_{ils}$ is estimated from equation (A11.1) (i.e., CZ substitute estimates of the coefficients in equation (A11.4)).

At each grid point:

$$P_{ils} = \gamma^s(\hat{P}_{ils} + v_{il}^s) + FE_{ils} + w_{ils} \ . \qquad \text{(A11.5)}$$

Since the term in parentheses is an estimate of the true predicted second price conditional on the grid point then the estimate of γ^s is conditionally unbiased. The grid point (or range of points) that maximize the percent of variance explained by (A11.5) identify the true unobserved mean variance $var(v_{il}^{s*})$ and percent of total variance where the superscript indicates simulations at maximizing grid points. Intuitively, variance in the persistent unobserved quality will be included in the second sales prices because buyers and sellers do observe these characteristics, so the R-squared for (A11.5) will be maximized at (or near, since these are estimated values) the true grid point's share in total variance.

Since the grid point that maximizes the variance explained by (A11.5) provides an estimate of the true mean $var(v_{il})$, the random draws from the true distribution can be plugged into the GM model to obtain unbiased estimates of the parameters of interest, α_l and α_g:

$$P_{ils} = \gamma^s\mu_{ils}^{s*} + \alpha_l\left(\widehat{exp\_loss}^*\right) + \alpha_g(\widehat{exp\_gain}^*) + FE_{ils} + \zeta_{ils} \qquad \text{(A11.6)}$$

where $\mu_{ils}^{s*} = \hat{P}_{ils} + v_{il}^{s*}$ and

$$\widehat{exp\_loss}^* = \left(P_{ilp} - \hat{P}_{ils} - v_{il}^{s*}\right)^+ = \left(FE_{ilp} - FE_{ils} + v_{il} + w_{ilp} - v_{il}^{s*}\right)^+ \quad \text{(A11.7)}$$

$$\widehat{exp\_gain}^* = |\left(P_{ilp} - \hat{P}_{ils} - v_{il}^{s*}\right)^-| = |\left(FE_{ilp} - FE_{ils} + v_{il} + w_{ilp} - v_{il}^{s*}\right)^-| \ \text{(A11.8)}$$

[6] GM (2001) first proposed using a grid and conditional distribution for the unknown v_{il} to estimate the bias in coefficients from their model 2. CZ use it in a very different way. CZ find that the coefficients from their model 2 are upwardly biased: i.e., the unobserved quality component is inadequately captured by the first residual.

CZ take the absolute value of simulated gains so that the sign in equation (A11.6) can take on the intuitive negative value, discounts on gains. However, it is not necessary to find negative values because those selling repeatedly during the sample period may obtain premiums relative to the one-only sales; this is true for expected gains as well as losses.

Summary of CZ algorithm for quality adjustment

Estimation equations are summarized here with an explanation of the iterative algorithm to achieve consistent estimators. Steps in estimating unobserved quality are:

1. Estimate equation (A11.1) using ordinary least squares and all sales to find $\hat{P}_{ils}$. It is known that coefficients are likely biased because $cov(v_{il}, X_{il}) \neq 0$.
2. Grid $mean(\sigma_{v,i})$ over its range which is [0, $sqrt(mean(\hat{\varepsilon}_{ilp}^2))$]. At each grid point:
   a. Draw unobserved quality, v_{il}^s (simulation at a grid point denoted with superscript *s*) randomly from $v_{il}^s|\hat{\varepsilon}_{ilp} \sim N(\frac{(\hat{\varepsilon}_{ilp})\sigma_{v,i}^{2s}}{(\sigma_{v,i}^{2s}+\sigma_w^2)}, \frac{\sigma_{v,i}^{2s}\sigma_w^2}{(\sigma_{v,i}^{2s}+\sigma_w^2)})$.
   b. Use these draws and $\hat{P}_{ils}$ from step 1 to estimate equation (A11.5).
   c. Find the grid point $s*$ that maximizes the R^2 for equation (A11.5). This is a Bayesian shrinkage estimator for $\sigma_{v,i}^{2*}$ as a percentage of var($\hat{\varepsilon}_{ilp}$).
3. Estimate v_{il}^{s*} for first and second sales: $v_{il}^{s*}|\hat{\varepsilon}_{ilt} \sim N(\frac{(\hat{\varepsilon}_{ilt})\sigma_{v,i}^{2*}}{(\sigma_{v,i}^{2*}+\sigma_w^2)}, \frac{\sigma_{v,i}^{2*}\sigma_w^2}{(\sigma_{v,i}^{2*}+\sigma_w^2)})$ where σ_w^2 is estimated from equation (A11.2) and *t*=*p*,*s*.
4. Estimate bias-corrected coefficients using equation (A11.6) – (A11.8).

Step 1 may produce biased estimates of coefficients of X_{il} because $Cov(v_{il}, X_{il}) \neq 0$, and these influence coefficients in steps 2 - 4. CZ address this with an iterative method developed by Oberhofer and Kmenta (1974) and summarized in Greene (2012). There are six assumptions required by Oberhofer and Kmenta, most importantly normal distributions for stochastic terms and independence between parameter estimates and the identifying assumptions for the variance-covariance matrix. Their identifying assumptions are joint normality of $v_{il}, \hat{\varepsilon}_{ilp}$ and that σ_w^2 is constant for all observations. By using first sales to estimation, CZ assert that $w_{ilt,t\neq p}$, is independent of v_{il} and so step 4 plausibly satisfies the independence assumption in Oberhofer and Kmenta (1974). It follows that, if the iterations converge then parameters converge to consistent maximum likelihood estimators. Convergence is not a given for every dataset.

Additional steps for iterating coefficients:

5. Estimate unobserved quality for one-only sales, $v_{il,oo}^{s*}$ using $v_{il,oo}^{s*} = \frac{(\hat{\varepsilon}_{il,o})\sigma_v^{2*}}{(\sigma_v^{2*}+\sigma_w^2)}$ if a one-only sale, otherwise zero.[7] Here σ_v^{2*} is the grid point from step 3.d., i.e., the mean optimizing variance. While

[7] The equation for $v_{il,oo}^{s*}$ holds at mean values if σ_v^{2*} estimated from repeat sales data is equal to the sigma-squared for one-only sales.

we could use a more complex method we find that the simple approach works well in our application.

6. Modify the regression using all sales, equation (A11.1) (i.e., step 1) as follows:

$$P_{ilt} = \beta^i X_{il} + \tau_{rs}^i v_{il,rs}^{s*} + \tau_{oo}^i v_{il,oo}^{s*} + FE_{lt}^i + \varepsilon_{ilt} \ . \qquad \text{(A11.1')}$$

Here the superscript *i* indexes the iteration of parameters and $v_{il,rs}^{s*} = v_{il}^{s*}$ from step 3 if the sale is a repeat, otherwise zero.

7. Compare estimated β^i to estimated β^{i-1}. The algorithm ends if the two vectors differ by economically insignificant amounts. In our application, we will iterate to step 4 to ensure that $FE_{lt}^i = FE_{lt}^{i-1}$ and therefore the loss and gain coefficients estimated from iterated equation (A11.6) are consistent estimators of the parameter of interest.

Table A11.1: Baseline Results Corrected for Unobserved Quality

This table shows robustness tests by correcting for unobserved quality. Results in this table are comparable to Table 3, Panel B (NAV) in the text. Dependent variable is log price of the second sale. Exp. Loss (adj.), Exp. Gain (adj.) and Exp. 2^{nd} Price (adj.) are quality-adjusted, constructed following Clapp and Zhou (2019). Months is number of month since the first sale. Results in Column (1)-(4), Normal, Boom, Bust and Recovery, use transactions in the period of 2000-2003, 2004-2006, 2007-2012 and 2013-2017, respectively. All price variables are in logs. All the specifications include town-year fixed effects. Bootstrapped standard errors are used. ***, **, * denote for 1%, 5% and 10% significance, respectively.

	Normal	Boom	Bust	Recovery
	(1)	(2)	(3)	(4)
Exp. Loss (adj.)	0.996***	0.918***	0.161***	0.231***
	(0.205)	(0.174)	(0.023)	(0.021)
Exp. Gain (adj.)	-0.370***	-0.202***	0.067***	0.358***
	(0.022)	(0.019)	(0.016)	(0.016)
Months	0.159***	0.094***	-0.057***	-0.137***
	(0.012)	(0.013)	(0.008)	(0.004)
Exp. 2^{nd} Price (adj.)	0.903***	0.858***	0.969***	0.984***
	(0.011)	(0.009)	(0.008)	(0.006)
Constant	1.291***	1.868***	0.431***	0.328***
	(0.131)	(0.109)	(0.099)	(0.073)
Town-Year FE	Y	Y	Y	Y
Observations	11,866	17,726	27,993	32,760
R-squared	0.888	0.878	0.813	0.842

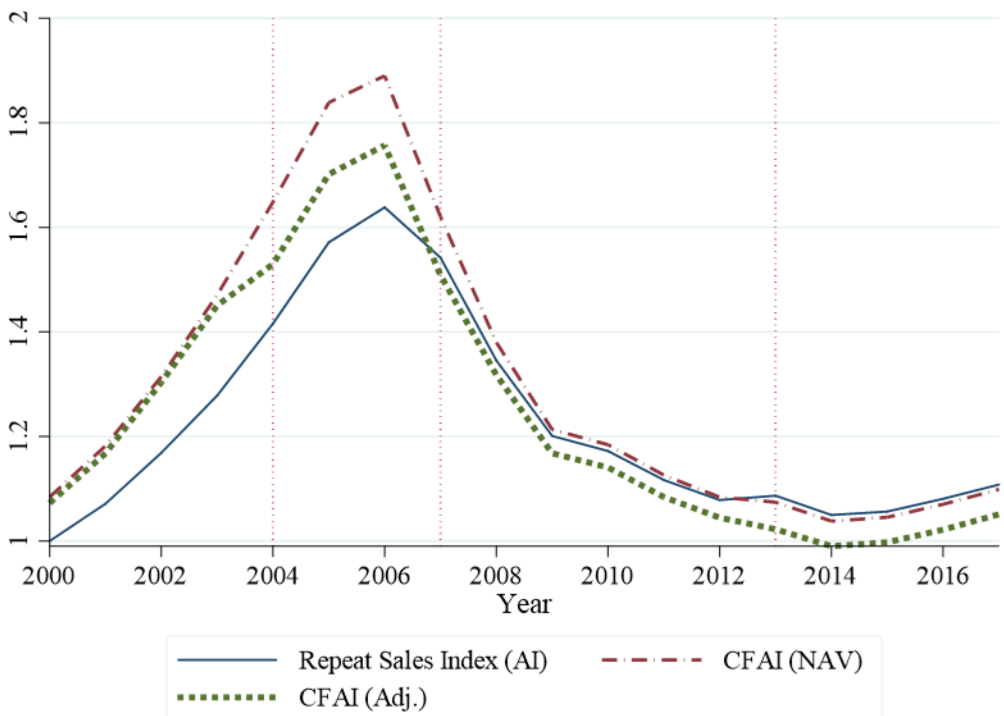
Table A11.2: Tests of Significance: Actual Housing Price Index minus Actual Factor Adjusted Index (CFAI) Corrected for Unobserved Quality

This table summarizes results based on tests of significance on the difference between the actual housing price index (AI) and the contrast-relative loss-gain-factor-adjusted index (CFAI). Loss and gain factors in columns (1)-(5) are the same as Table 6 and based on Exp. 2nd price (PV). Those in columns (6)-(10) are based on Exp. 2nd Price (adj.). The total adjusted factor (AF) consists of the loss adjusted factor (LAF, in column (1) and (6)) and the gain adjusted factor (GAF, in column (2) and (7)). The total adjustment factor (AF) (Column (3) and (8)) is the sum of these two: AF = LAF + GAF and the contrast-relative adjusted index, CFAI = AI – AF. Column (4) and (9) shows the changes of total adjustment factors (ΔAF) (i.e. ΔLAF + ΔGAF). Column (5) and (10) shows the changes of the contrast-relative loss-gain-factor adjusted factor index (ΔCFAI). Column (11) shows the changes of actual housing price index (ΔAI). By construction, ΔAI – ΔCFAI = ΔAF and columns (4)-(5) and (9)-(10) are recorded as missing in 2000. Numbers in bold denote for *p*-value of *F* statistics significant at 5%. We cannot calculate *F* statistics for any change variables since they are based on only two numbers.

	PV ---					**Exp. 2nd Price (adj.)**--					
	LAF	GAF	AF	ΔAF	ΔCFAI	LAF	GAF	AF	ΔAF	ΔCFAI	ΔAI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
2000	**0.033**	**-0.111**	**-0.079**			**0.011**	**-0.081**	**-0.070**			
2001	**0.026**	**-0.126**	**-0.100**	-0.021	0.090	**0.008**	**-0.094**	**-0.086**	-0.016	0.085	0.069
2002	**0.019**	**-0.150**	**-0.131**	-0.031	0.118	**0.005**	**-0.114**	**-0.109**	-0.023	0.110	0.087
2003	**0.017**	**-0.171**	**-0.155**	-0.024	0.113	**0.005**	**-0.132**	**-0.126**	-0.017	0.107	0.089
2004	**0.012**	**-0.117**	**-0.105**	0.050	0.052	**0.003**	**-0.081**	**-0.077**	0.049	0.053	0.102
2005	**0.012**	**-0.120**	**-0.108**	-0.004	0.108	**0.003**	**-0.083**	**-0.080**	-0.003	0.107	0.104
2006	**0.016**	**-0.112**	**-0.096**	0.013	0.029	**0.006**	**-0.076**	**-0.070**	0.010	0.032	0.042
2007	**0.002**	**0.103**	**0.104**	0.200	-0.260	**0.003**	**0.019**	**0.022**	0.092	-0.152	-0.060
2008	**0.004**	**0.061**	**0.065**	-0.039	-0.097	**0.010**	**0.011**	**0.021**	-0.001	-0.136	-0.137
2009	**0.007**	**0.042**	**0.049**	-0.016	-0.098	**0.021**	**0.007**	**0.028**	0.007	-0.121	-0.113
2010	**0.007**	**0.040**	**0.047**	-0.002	-0.022	**0.021**	**0.006**	**0.027**	-0.001	-0.023	-0.024
2011	**0.008**	**0.037**	**0.046**	-0.001	-0.047	**0.024**	**0.006**	**0.030**	0.003	-0.051	-0.048
2012	**0.009**	**0.032**	**0.042**	-0.004	-0.031	**0.027**	**0.005**	**0.032**	0.002	-0.037	-0.035
2013	**0.027**	**0.037**	**0.064**	0.023	-0.015	**0.034**	**0.027**	**0.061**	0.029	-0.022	0.008
2014	**0.027**	**0.036**	**0.063**	-0.001	-0.033	**0.034**	**0.025**	**0.059**	-0.002	-0.032	-0.035
2015	**0.026**	**0.039**	**0.065**	0.002	0.005	**0.032**	**0.025**	**0.057**	-0.001	0.007	0.006
2016	**0.024**	**0.043**	**0.067**	0.003	0.020	**0.028**	**0.028**	**0.056**	-0.001	0.024	0.023
2017	**0.021**	**0.047**	**0.067**	0.000	0.025	**0.021**	**0.031**	**0.053**	-0.003	0.028	0.025

Figure A11: Repeat Sale Index and the Loss-Gain-Factor Adjusted Index (CFAI) Corrected for Unobserved Quality

This figure shows the repeat sales index and the loss-gain-factor adjusted index (CFAI) corrected for unobserved quality following Clapp and Zhou (2019). Both indices are calculated in the exponential values. The CFAI is calculated as the actual repeat sales index minus the loss (gain) adjustment factors in which the latter is the loss (gain) variable for a given period (i.e. normal, boom, bust and recovery) multiplied by the mean magnitude of loss (gain) among sold properties that were facing an expected loss (gain) and the proportion of sellers facing an potential loss (gain) in that period. The repeat sales index minus CFAI suggests the potential magnitude of loss/gain behavior. I.e., If the difference is negative (positive), then the behavior is associated with dampened (accentuated) price movements if the actual change is positive in that part of the cycle, ΔAI>0. This logic reverses if ΔAI<0. "CFAI (NAV)" indicates CFAI is calculated based on normalized assessed value (NAV). "CFAI (Adj.)" indicates CFAI is based on Exp. 2nd Price (adj.).



Appendix 12: Contrast-relative Estimate Analysis Holding Loss/Gain Coefficients Constant

A12.1 Analytical Framework

This section develops the model summarized in figure A12.2. The purpose is to propose a method for isolating the changes in loss/gain coefficients from changes in magnitudes and proportions. The method developed here measures what market prices would have been holding coefficients constant at the normal period. This is important because coefficients are related to the bargaining process: i.e., search and bargaining models apply to coefficients. Results summarized in Section A12.2 suggest that negotiated prices, holding constant proportions and magnitudes, produce contrast-relative estimates associated with dampening each phase of the cycle. Without negotiated premiums and discounts contrast-relative estimates suggest that the boom and recovery would have been larger and the bust would have been more severe. This implies that causal models based on search and bargaining might be able to explain the stylized facts documented here.

We construct a contrast-relative estimate, $CFAI_1$, using the loss/gain coefficients from the normal period as the baseline.[8]

$$CFAI_1 = AI - (LAF_1 + GAF_1) \quad \text{(A12.1)}$$

where $LAF_1 = \alpha_{l,j=normal\,period} ML_t \%L_t$; $GAF_1 = \alpha_{g,j=normal\,period} MG_t (1 - \%L_t)$. We follow Section 3.1 and denote the adjustment factors of $CFAI_1$, $LAF_1 + GAF_1$, as AF_1.

The comparison between $CFAI_1$ and $CFAI$ helps analyze the importance of the change in loss/gain coefficients. As defined in Section 3.1, $CFAI = AI - (LAF + GAF)$. Therefore,

$$CFAI - CFAI_1 = (LAF_1 - LAF) + (GAF_1 - GAF) = AF_1 - AF \quad \text{(A12.2)}$$

where LAF is $\alpha_{l,j} ML_t \%L_t$, GAF is $\alpha_{g,j} MG_t (1 - \%L_t)$, as defined in Section 3. By definition the difference between $CFAI$ and $CFAI_1$ equals to $AF_1 - AF$.[9] To separate loss effect from gain, we define $LAF_1 - LAF$ as the effect of loss coefficient (EL) and $GAF_1 - GAF$ as the effect of gain coefficient (EG).

$$LAF_1 - LAF = EL = (\alpha_{l,j=normal\,period} - \alpha_{l,j}) ML_t \%L_t \quad \text{(A12.3)}$$

$$GAF_1 - GAF = EG = (\alpha_{g,j=normal\,period} - \alpha_{g,j}) MG_t (1 - \%L_t) \quad \text{(A12.4)}$$

Noted that we compare $CFAI_1$ with $CFAI$ instead of AI because $CFAI$ is loss/gain behavior exclusive while AI is composed exhaustively of those with losses and gains. Comparing $CFAI_1$ with $CFAI$ shows the effect of coefficient changes, while comparing $CFAI_1$ with AI shows the effect of the other two

[8] We construct the contrast-relative index by year. One could also construct contrast-relative estimate by sub-period. In this way, one could construct three contrast-relative estimates holding constant each of the three variables (coefficients, magnitude and proportion) at the normal period. For example, the four elements in $CFAI_1$ holds coefficient constant at its normal period value. Wherever we suppress the subscript j we refer to a vector with four elements, one for each subperiod. E.g., $GAF_j = \alpha_{g,j} MG_j (1 - \%L_j)$ can be modified to construct$GAF_{1,j}$, $\alpha_{g,j=normal\,period} MG_j (1 - \%L_j)$, which holds constant any changes in the gain coefficient after the normal period. Intuitively, subtracting changes in $GAF_{1,j}$ from GAF_j isolates the estimated influence of change in gain coefficients $\alpha_{g,j}$ on the total gain effect over a house price cycle.

[9] As $CFAI = AI - (LAF + GAF)$ and $CFAI_1 = AI_1 - (LAF_1 + GAF_1)$, $CFAI - CFAI_1 = (LAF_1 - LAF) + (GAF_1 - GAF)$. Since $AF = LAF + GAF$ and $AF_1 = LAF_1 + GAF_1$, $CFAI - CFAI_1 = AF_1 - AF$. Similarly, we have $\Delta CFAI - \Delta CFAI_1 = \Delta AF_1 - \Delta AF$.

variables combined (i.e., magnitude and proportion). In summary, the difference between *CFAI and* $CFAI_1$ (i.e. $CFAI - CFAI_1$) equals $EL + EG$. $CFAI - CFAI_1$ can also be written as $(LAF_1 - LAF) + (GAF_1 - GAF)$.

We will test the difference between *CFAI* and $CFAI_1$ by year. By construction, the difference is always equal to zero for the normal period (2000-2003), the baseline. If the loss/gain behavior associated with coefficients does not change over the housing price cycle relative to the normal period, then the difference between *CFAI* and $CFAI_1$ should be insignificantly different than zero in every year afterwards. Similar to Table 6, we will test the statistical significance of the loss part separately from gain.

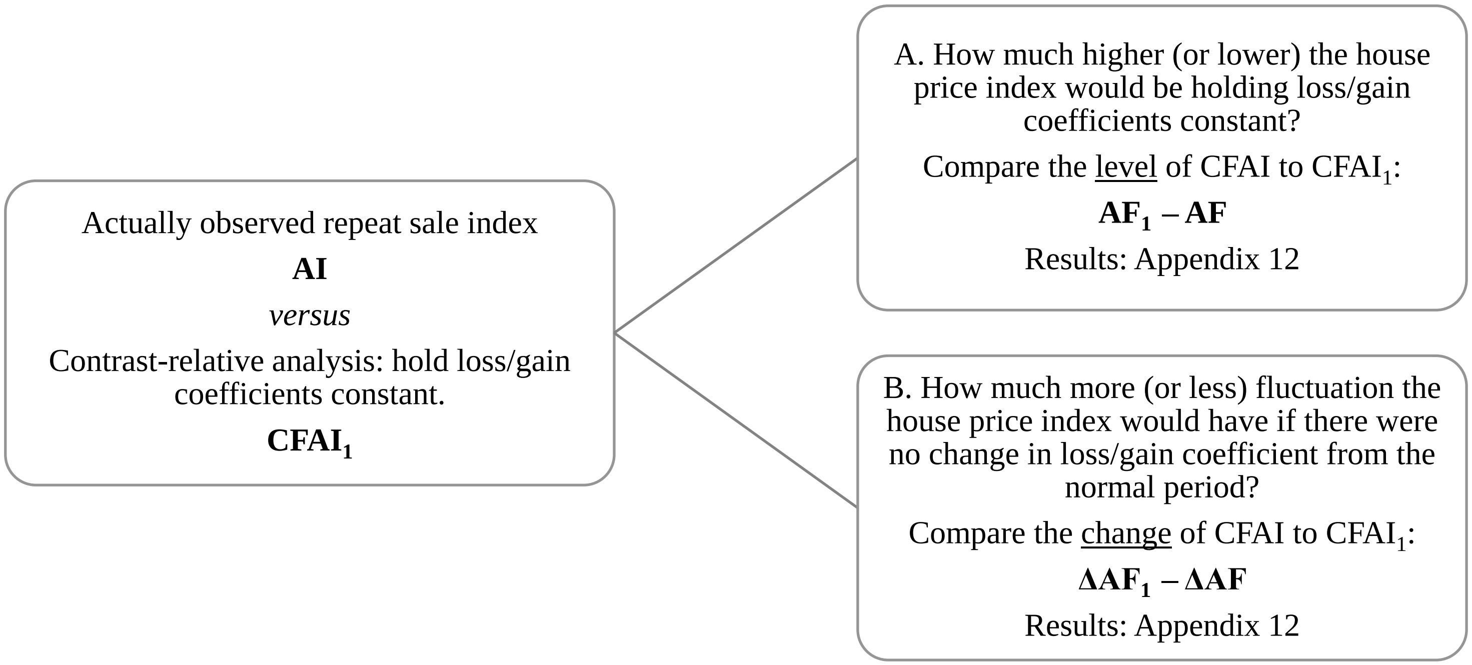
As the difference between *CFAI* and $CFAI_1$ is cumulative difference from the normal period, our focus is $\Delta CFAI - \Delta CFAI_1$.[10] For example, if coefficients change negatively from normal to boom, then the change in $CFAI_1$ is greater than the change in *CFAI* (i.e. $\Delta CFAI_1 > \Delta CFAI > 0$, or $\Delta AF > \Delta AF_1 > 0$). It means that the change in loss/gain coefficient dampened the boom: the intuition is that ΔAF allows all three variables (coefficients, magnitudes and weights) to change over the cycle whereas ΔAF_1 holds coefficients constant, so that variable is associated with a greater adjustment. Adjustments are subtracted from the repeat sales index when calculating *CFAI* so loss/gain coefficients dampened the boom in this example.

Moving from boom to bust, if $\Delta CFAI_1 < \Delta CFAI < 0$, or $\Delta AF - \Delta AF_1 < 0$, we would conclude that the change in loss/gain coefficients dampened the bust because the house prices would have decreased more without the coefficient change. In all cases we are looking at the sign of $\Delta CFAI$ and $\Delta CFAI_1$ and interpreting it depending on the phase of the cycle. We summarize the interpretation of changes in our contrast-relative estimates as follows:

	Interpretation of $AF_{1,t} - AF_t$ as Differences in Levels	
	$CFAI_t - CFAI_{1,t} > 0$ **$AF_{1,t} - AF_t > 0$**	**$CFAI_t - CFAI_{1,t} < 0$** **$AF_{1,t} - AF_t < 0$**
$\Delta AI_t > 0$ or $\Delta AI_t < 0$	(1) Change in L/G coefficients increases house prices	(2) Change in L/G coefficients decreases house prices
	Interpretation of $\Delta AF_{1,t} - \Delta AF_t$ as Differences in Changes	
	$\Delta CFAI_t - \Delta CFAI_{1,t} > 0$ **i.e. $\Delta AF_{1,t} - \Delta AF_t > 0$**	**$\Delta CFAI_t - \Delta CFAI_{1,t} < 0$** **i.e. $\Delta AF_{1,t} - \Delta AF_t < 0$**
$\Delta AI_t > 0$ **(in up markets)**	(3) Change in L/G coefficients increases cycle phase	(4) Change in L/G coefficients dampens cycle phase
$\Delta AI_t < 0$ **(in down markets)**	(5) Change in L/G coefficients dampens cycle phase	(6) Change in L/G coefficients increases cycle phase

[10] For example, because $CFAI - CFAI_1$ takes the value of zero in the normal period by construction, $\Delta CFAI - \Delta CFAI_1$ is equivalent to $CFAI - CFAI_1$ in the boom period.

A12.2 Summary of Analytical Framework for Isolating the Association between Coefficients and Contrast-relative Estimates



A12.3 Discussions of Our Results

We plot the contrast-relative index, $CFAI_1$, together with the actual factor adjusted index (*CFAI*), in Figure A12. The normal period is used as the baseline, so all the indices overlap 100% from the beginning of 2000 through the end of 2003. Our focus is $\Delta CFAI - \Delta CFAI_1$ because, by construction, the difference in levels is cumulative from the normal period. The table above provides the interpretation of results.

Figure A12 Panel A shows that the housing index would have increased more than the actual in the boom. The increase in $CFAI_1$ (which holds coefficients constant) is larger than that in *CFAI* without the coefficient change. In other words, the change in loss/gain coefficients was associated with dampening the boom. Compared estimated contrast-relative estimates with Panel A, quality-adjusted contrast-relative estimates in Panel B show very little difference. The small gap in Panel B is because "CFAI (NAV)" lies above "CFAI (PV)" while "CFAI-1 (NAV)" lies below "CFAI-1 (PV)", suggesting the effectiveness of quality adjustment using NAV. In Panel C, comparing AI with $CFAI_1$ shows the effect of the other two variables combined (i.e., magnitude and proportion). Holding coefficient constant, changes in magnitude and proportion combined generate a larger boom and a larger bust.

Table A12 provides further explanations. Results in Columns (1)-(3) and (7)-(9) help explain the difference between the level of *CFAI* and $CFAI_1$ and those in Column (4)-(6) and (10)-(12) help explain the difference between the change of *CFAI* and $CFAI_1$. Columns (1)-(6) show estimated results and columns (7)-(12) show quality-adjusted results.

In the boom period, the PV results suggest the effect of gain (EG) dominates the effect of loss (EL) while the NAV results suggest both effects are insignificant. In the bust, both PV and NAV results suggest EL dominates EG. Compared with EG, EL is larger in magnitude and has more statistical significance. Together, EL+EG is negative in most years.

In all the years, quality-adjusted EL in Column (7) is smaller than those unadjusted in Column (1). Similarly, quality-adjusted EG in Column (8) is smaller than those in Column (2). Comparing AF_1 (EL+EG in Column (9)) with *AF* (in Column (8) in Table 6), we find that AF_1 is smaller (more negative) than AF after 2008. This is consistent with our observation in Panel B of Figure A12 that $CFAI_1$ lie above *CFAI* after 2008 because $AF_1 < AF$ implies $CFAI_1 > CFAI$. Similarly, in Panel C "*CFAI-1(NAV)*" lies below "*CFAI-1(PV)*" after 2008 because EL+EG in Column (9) is less negative than those in Column (3).

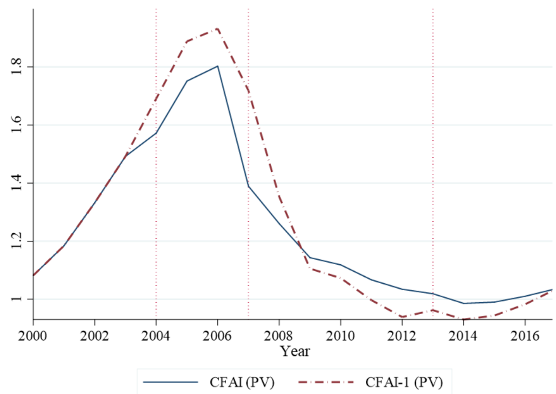
The difference between ΔAF_1 and ΔAF is shown in Column (6) and (12). The change in the total adjustment factor (ΔAF) is larger than the contrast-relative change (ΔAF_1) in 2004-2005, suggesting that the changes in loss/gain coefficients are associated with dampening the cycle because the increase in house prices would have been slightly higher if holding the loss and gain coefficients constant. This is scenario (4) as our summary of the interpretation of changes in our contrast-relative estimates.

In bust (except 2007), the differences between ΔAF_1 and ΔAF are strongly positive, suggesting changes in loss/gain coefficients dampened bust. This is scenario (5) as our summary of the interpretation of changes in our contrast-relative estimates. The recovery results are mostly large negative effects of holding loss/gain coefficients constant, a pattern consistent with scenario (4): negotiated coefficients dampen the recovery. Overall, the changes in loss/gain coefficients dampen each phase of the cycle.

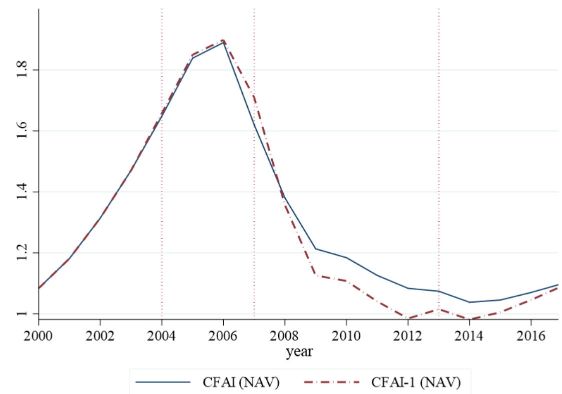
Figure A12: Loss-Gain-Factor Adjusted Index (*CFAI*) and the Contrast-relative Factor Adjusted Indices Holding Loss/Gain Coefficient Constant (*CFAI$_1$*)

This figure shows the loss-gain-factor adjusted index (CFAI) and a contrast-relative index holding loss/gain coefficients constant ($CFAI_1$) based on a sample of individual transactions from 2000 to 2017. "(PV)" is based on Exp. 2nd Price and "(NAV)" is based on normalized assessed value. Panel A compares "CFAI (PV)" with "CFAI-1 (PV)". Panel B compares "CFAI (NAV)" with "CFAI-1 (NAV)". Panel C plots actual "Repeat Sales Index (AI)", "CFAI-1 (PV)" and "CFAI-1 (NAV)". $CFAI_1$ minus CFAI suggests the potential magnitude of the loss/gain coefficients on the cycle. I.e., If the difference is negative (positive), then the coefficients dampened (accentuated) that part of the cycle.

Panel A:



Panel B:



Panel C:

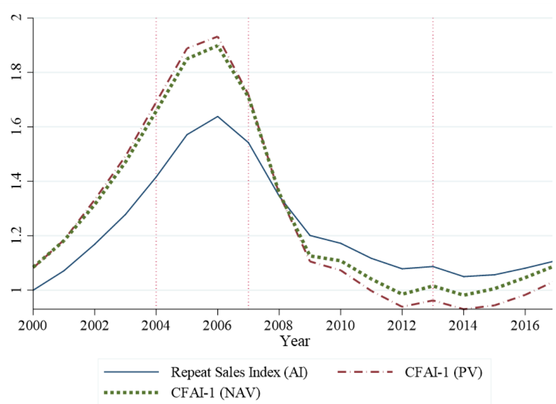


Table A12: Tests of Significance – Contrast-relative Index Holding Loss/Gain Coefficient Constant

This table summarizes results based on tests of significance on the difference between a contrast-relative index holding loss/gain coefficient constant ($CFAI_1$) and the loss-gain-factor-adjusted index (*CFAI*). Columns (1)-(6) are based on Exp. 2nd Price and Columns (7)-(11) are based on normalized assessed value. The total effect consists of the aggregate effect from loss (EL) and the aggregate effect from gain (EG). The total effect (Column (3) and (9)) is the sum of these two. Columns (4) and (10) show the changes of actual adjustment factors (ΔAF) (i.e. ΔLAF+ ΔGAF). Columns (5) and (11) show the changes of contrast-relative adjustment factors (ΔAF_1) (i.e. $\Delta LAF_1+\Delta GAF_1$). Columns (6) and (12) show the difference ΔAF and ΔAF_1. Numbers in bold denote for *p*-value of *F* statistics significant at 5%.

	PV						NAV					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	EL	EG	EL+EG	ΔAF_1	ΔAF	(4) – (5)	EL	EG	EL+EG	ΔAF_1	ΔAF	(10) – (11)
2004	-0.001	**0.074**	**0.073**	-0.023	0.050	-0.073	-0.001	0.007	0.006	-0.017	-0.012	-0.006
2005	-0.001	**0.077**	**0.075**	-0.006	-0.004	-0.002	-0.001	0.007	0.006	-0.006	-0.005	0.000
2006	-0.002	**0.071**	**0.069**	0.019	0.013	0.006	-0.002	0.006	0.004	0.016	0.014	0.001
2007	**-0.031**	**0.243**	**0.213**	0.056	0.200	-0.143	**-0.016**	**0.070**	**0.054**	0.044	0.093	-0.049
2008	**-0.074**	**0.145**	**0.071**	0.103	-0.039	0.142	**-0.059**	**0.042**	-0.017	0.095	0.024	0.071
2009	**-0.134**	**0.100**	**-0.034**	0.089	-0.016	0.104	**-0.101**	**0.026**	**-0.075**	0.073	0.015	0.057
2010	**-0.136**	**0.094**	**-0.042**	0.006	-0.002	0.008	**-0.092**	**0.025**	**-0.067**	-0.008	0.000	-0.008
2011	**-0.157**	**0.089**	**-0.068**	0.025	-0.001	0.026	**-0.103**	**0.024**	**-0.079**	0.014	0.002	0.012
2012	**-0.174**	**0.077**	**-0.097**	0.025	-0.004	0.029	**-0.117**	**0.021**	**-0.096**	0.020	0.003	0.017
2013	**-0.141**	**0.084**	**-0.057**	-0.017	0.023	-0.040	**-0.099**	**0.043**	**-0.056**	-0.022	0.017	-0.039
2014	**-0.140**	**0.081**	**-0.058**	-0.001	-0.001	0.001	**-0.096**	**0.040**	**-0.056**	-0.001	0.000	-0.001
2015	**-0.135**	**0.087**	**-0.048**	-0.008	0.002	-0.010	**-0.084**	**0.045**	**-0.039**	-0.017	-0.001	-0.017
2016	**-0.125**	**0.097**	**-0.028**	-0.018	0.003	-0.020	**-0.074**	**0.051**	**-0.023**	-0.017	-0.001	-0.016
2017	**-0.106**	**0.105**	-0.001	-0.027	0.000	-0.027	**-0.059**	**0.052**	-0.007	-0.017	-0.001	-0.016