An Alternative Approach to Estimating Foreclosure and Short Sale Discounts*

James N. Conklin,† N. Edward Coulson,‡ Moussa Diop,§ and Nuno Mota¶

February 13, 2023

Abstract

Current research documents astonishingly large price discounts for foreclosures and short sales. However, such outsized estimates may largely be due to omitted variables bias. We propose an innovative methodology relying on appraisers' ability to match properties along both observable and unobservable attributes when performing appraisals. Our empirical approach, which relies on the use of appraisal fixed effects, produces foreclosure and short sale discounts of approximately 5% after controlling for a rich set of characteristics, including quality and condition, attributable mostly to the stigma associated with distress ifself. We show that these lower estimates are not due to appraisers selecting high-price distressed properties as comps and are robust across a wide variety of subsamples and under alternative estimation methods.

*We thank Gilles Duranton, Mike Eriksen, Kris Gerardi, Michael LaCour-Little, Lauren Lambie-Hanson, Ruchi Singh, Patrick Smith, Paul Willen, and two anonymous referees for valuable comments. The views presented in this document are those of the authors and do not represent those of their respective institutions.

†Terry College of Business, University of Georgia, Athens, GA

‡Paul Merage School of Business, University of California, Irvine, Irvine, CA

§Sol Price School of Public Policy, University of Southern California, Los Angeles, CA

¶Fannie Mae, Washington, DC

1. Introduction

The size of price discounts associated with forced sales of real estate assets remains an unsettled question.[1] In their widely-cited study of forced sales of Massachusetts properties, Campbell, Giglio, and Pathak (2011) document astonishingly large foreclosure discounts of 27% on average, with much larger discounts recorded for low-price houses and in low-price neighborhoods. Discounts of similar magnitudes have been documented in other U.S. contexts, as well as Sweden, Germany, the Netherlands, and Italy and Poland.[2] Many studies find price discounts of 20% or more, especially for foreclosure sales. Even though there is wide agreement that forced property sales likely cause non-trivial price discounts, there is no consensus on the exact magnitude of these discounts and even more modest estimates found in the literature seem implausibly large.[3]

Hedonic log sales price regressions are generally used to identify forced sale discounts by including a forced sale indicator as well as a large number of covariates that attempt to control for key property and neighborhood characteristics. A primary concern, of course, is whether characteristics that are unobservable to the econometrician covary with both forced sales and prices. Indeed, this concern speaks to the definition of the forced sale discount itself. Generally, one can think of four broad reasons why the coefficient of the forced sale indicator would be negative in

[1]We use the terms "forced sale" and "distressed sale" interchangeably to refer collectively to foreclosure and short sales. Consistent with most of the literature, we define a foreclosure as a property sold by a lender as real estate owned (REO) and a short sale as a lender-approved pre-foreclosure sale producing less than the mortgage balance. Caution is required when comparing studies because the definition of a distressed sale is not universal across studies and early work has not used a consistent definition or considered different types of distressed sales (Doerner and Leventis, 2015). Most studies (including our own) use post-foreclosure REO sales, however, Chinloy, Hardin, and Wu (2017) explore discounts on foreclosure auctions. Other studies make no distinction between REO sales and short sales and thus estimate an average of the two discounts (Conklin, Coulson, and Diop, 2021). Furthermore, earlier studies may have had difficulty identifying distressed sales due to data limitations (Doerner and Leventis, 2015).

[2]Doerner and Leventis (2015); Donner (2020); Donner, Song, and Wilhelmsson (2016); Forgey, Rutherford, and VanBuskirk (1994); Just et al. (2019); Mehrotra, Nowak, and Smith (2021); Mocking and Overvest (2017); Renigier-Biłozor et al. (2018); William and Marvin (1996)

[3]Clauretie and Daneshvary (2009) and Carroll, Clauretie, and Neill (1997) find smaller discounts likely due to sample selection. Harding, Rosenblatt, and Yao (2012) find that the majority of real estate owned (REO) purchases do not earn economically significant excess returns, but this says little about the level effect of foreclosures, with which the research is generally concerned.

a hedonic regression. First, there are the observable and unobservable prior differences between distressed and non-distressed properties (Frame, 2010); second, there are differences in condition caused by the reduction in maintenance by distressed homeowners (Lambie-Hanson, 2015); third, distressed sellers (often financial institutions) have greater urgency and the briefer time on the market is correlated with lower price (Clauretie and Daneshvary, 2009); and fourth, there is the stigma associated with distressed sales, the outcome of asymmetric information endemic to real estate transactions (Lopez, 2021; Stroebel, 2016). This stigma effect refers to a price discount on distressed sales for no other reason than the distressed status of the property, with no differences in observable and non-observable characteristics between the distressed property and non-distressed properties, nor any differences in seller urgency.

Any given estimate of the distressed discount will have to reckon with whether said discount includes any or all of the above contributors to the conditional price differential between distressed and non-distressed properties, and whether such contributors are a causal impact of distress.[4] Evidently, the stigma effect identified above is a causal impact of distress on price. One may also argue that the reduction in maintenance by distressed homeowners is a causal impact of distress on price. Conversely, observable and unobservable differences in the property or neighborhood characteristics, and differences in seller urgency should not be thought of as causal impacts of distress on price. Our empirical approach and subsequent robustness checks allow us to identify the stigma effect of distress on price.

We provide an alternative methodology to estimating distress discounts, which while related to traditional regression methods, adds a novel matching technique of distressed and non-distressed properties to these traditional procedures. In doing so, we shed light on the sources and size of the contributors to distressed sale discounts. We employ a data set from a large secondary market

[4]In addition to differences in controls, there would seem to be significant geographic heterogeneity in the discount estimates. The sample of Campbell, Giglio, and Pathak (2011) is limited to the Boston area while lower-end estimates of Carroll, Clauretie, and Neill (1997) and Clauretie and Daneshvary (2009) are from Las Vegas. Note as well the significant geographic variation estimated in Zhou et al. (2015).

purchaser of mortgages, which contains not only data on property transactions but also on the professional appraisal typically required by the lender. These appraisals use recent transactions of comparable properties – comparables, or “comps” – to provide information on the market value of the subject property. The key to our estimation procedure is that these comps, which include distressed sales, are selected precisely because they are good matches to the subject property on a large number of dimensions. Properly matching comparables to subject properties is a core function of appraisers, and local market expertise is required to ensure properties are not only similar along observable dimensions, but also other dimensions that may be difficult for an outsider to observe.[5] Thus, within a given appraisal, the properties are not only a good match to the subject property, but also to each other. In our procedure, we will, in essence, compare transaction prices of arm’s length comps with forced sale comps *within the same appraisal*, allowing us to estimate forced sale discounts which are far less likely to be a function of differences in omitted property and location characteristics than estimates from traditional regression methods. Our procedure is thus akin to a matching estimator, where the matching is accomplished not by a statistical procedure but by a local expert.

The data set consists of residential property appraisals conducted in all 50 states from 2013 to 2017 associated with 7.2 million *arm’s length* home purchase loan applications. The data link each appraisal to the comparables used by the appraiser. There are 27.3 million comparable property sales used in these appraisals, including both forced and arm’s length sales. The comps data contain the sales price, characteristics, and location of comparables, and, unlike most available data, capture the condition and quality of these properties. Furthermore, the data distinguish foreclosures from short sales in the comparable transactions, allowing us to compare the price impacts of these two main channels used to dispose of distressed real estate properties.[6]

[5]The most common appraisal forms (1004, 1073, 1004C, 1025, and 1004D) require the appraiser to certify that “I selected and used comparable sales that are locationally, physically, and functionally the most similar to the subject property” (Fannie Mae, 2020; FHA, 2020).

[6]Information on data availability and code will be provided on request.

Using only the comps data, we estimate a sequence of regression models that provide increasingly stringent controls, in order to mirror distressed discounts estimated from previous traditional models, and then ending with our appraisal-centered matching model. The foreclosure and short sale discount estimates significantly drop from 36% and 19%, respectively, to about 16% after progressively accounting for time-invariant neighborhood attributes with location (census tract) fixed effects; time-variant national trends with year-quarter fixed effects; property characteristics; and property conditions, quality, and location and view desirability. These lower figures match some of the most conservative estimates found in the literature. By using appraisal fixed effects, which further control for unobservable differences in property location, characteristics, quality, and condition that tended to be present in distress discount estimates documented in previous studies, we substantially lower the foreclosure and short sale price discount estimates to roughly 5-6%. We note that the size of the short sale discount is similar to the foreclosure discount, which stands in contrast with findings in the existing literature.

We then conduct tests to rule out competing explanations. Importantly, we find that our results do not depend on market liquidity, as the estimate of our key parameter does not vary across markets with higher or lower numbers of comps. As we explain below, this would seem to rule out both seller urgency and selection of "good" foreclosures as comps as an explanation of our results. Ruling out these explanations leaves "foreclosure stigma" as the primary cause of the discount. Additional analysis shows that our results are also remarkably robust to sample splits by transaction year, neighborhood racial composition, sand vs. non-sand states and state foreclosure procedures, suggesting that stigma is somewhat constant across housing markets.

The rest of the paper is organized as follows. The next section presents a short accounting of both econometric and appraiser estimates of the forced sale discount. This is followed by a short description of the appraisal process. Section 4 describes the data used in this study. Next, Section 5 presents our empirical findings. Section 6 concludes with a discussion of the importance of this result for the study of residential markets.

2. Estimating forced sale discounts

Estimates of the foreclosure discount, such as those discussed in the introduction, are from regression models that include a distressed indicator, typically specified as:

$$P_i = X_i\beta + \gamma D_i + u_i + e_i \tag{1}$$

where P_i is the log sale price of property *i*, X_i is a vector of property characteristics, β is the vector of characteristic prices as estimated by a hedonic regression, D_i is an indicator identifying distressed properties (comps), and γ is the regression estimate of the distress discount. Unobservable characteristics are represented by u_i and the e_i is random noise.

As previous authors have noted, the specification of X_i and the associated treatment of u_i are critical for obtaining consistent estimates of the discount. We will estimate the above equation and employ a variety of specifications for the X vector in order to approximate models from previous research. These specifications will include a baseline unconditional forced sale differential, as well as specifications that will sequentially add hedonic characteristics, time and location effects, and less common controls such as view, condition and quality.

Typically, even rich specifications of X_i have used location (and time) fixed effects to account for u_i. But appraisers have a different methodology. A property appraisal begins with a selection of a small number of recent transactions in the same development or neighborhood as the subject property. These comparable properties are chosen *precisely* because the property features contained in the u_i terms are similar to those of the subject property, and on that account each of the comparables in a given appraisal have a similar u_i.

We note that some of these comparables will be distressed sales and some will not. Our method takes advantage of the presence of the comparable transactions used in each appraisal and ultimately considers regression models that use as observations all of the comparable property sales –

and not the subject properties themselves – from all of the appraisals in our data set. Thus we can ultimately specify a regression of the form:

$$P_{ij} = X_{ij}\beta + \gamma D_{ij} + u_j + e_{ij} \tag{2}$$

where P_{ij} and X_{ij} are the transaction price and characteristics of the i^{th} comparable property from j^{th} appraisal. In this model the u_j terms become appraisal-level fixed effects. These appraisal fixed effects can account for the aforementioned unobservables that are common to all the properties used as a comparable for a given appraisal. If the regression estimate of the forced sale discount from the same set of comparables using (1) is biased because of omitted characteristics, the inclusion of the appraisal fixed effects in regression (2) will likely remove a great deal of that error.

3. The Appraisal Process

Property appraisal is a standard requirement in mortgage lending. An appraisal is supposed to provide an independent evaluation of the value of the property serving as collateral for the loan.[7] This is an estimate of the value of the property performed by a professional appraiser based on local market conditions. The appraiser, who is commissioned by the lender and typically paid a flat-fee out-of-pocket by the homeowner, generally identifies recently sold nearby properties that are comparable to the property being appraised in terms of both observable characteristics and unobservables affecting the value of the property.[8] Under the economic principle of substitution,

[7]Lenders generally require an appraisal for home purchase loans, mortgage refinancing, and even home equity loans. Appraisers are subject to state regulations. Typically, they are certified and have met both education and experience requirements. In addition, appraisers are expected to be objective and not be swayed by the participants in the transaction.

[8]When selecting comps, appraisers should consider real property rights conveyed, financing terms, conditions of sale, expenditures made immediately after purchase if applicable, market conditions, location, physical characteristics, economic characteristics, use/zoning, and non-realty components of value (Appraisal Institute, 2013). Our methodology alleviates potential bias from the omission of some of these determinants of property value that are unobservable

the sales price of a recently sold identical property should provide a good estimate of value for the subject property.[9] Of course, no two properties are identical, so the sales prices of the comparable properties may not accurately reflect the value of the subject property. To account for this, the appraiser then makes adjustments to each comp's sales price for differences in attributes between the comp and the subject property and then uses the set of adjusted comp prices to estimate the value of the subject property.[10]

The accuracy of an appraisal relies heavily on selecting suitable comps – good matches to the subject property on easily observable characteristics (e.g., beds, baths) as well as those typically not observed by outsiders (e.g., condition, granular locational features). Picking suitable matches, particularly along these latter characteristics, requires extensive property and location-specific knowledge. This knowledge is used to select comps that minimize the need for adjustments, as the correct adjustments to arrive at an adjusted price are associated with uncertainty (Conklin, Coulson, and Diop, 2021).[11] Importantly, as we use actual transaction prices of comps instead of the appraiser-adjusted prices, our methodology does not rely on appraisers' ability to make accurate adjustments, but rather on their ability to pick suitable matches.

4. Data

Our data come from a large secondary market purchaser of residential mortgage loans and include information on comparable transactions used to estimate values of subject properties in residential appraisals. A large share of financial institutions rely on the collateral valuation and mortgage underwriting platform of the data provider, resulting in broad market coverage, including appraisals

by the econometrician.

[9]This approach to appraisal, which is commonly used in home appraisals, is referred to as the sales comparison approach. Two alternative appraisal methods include the cost approach for new properties and the income approach for income producing properties.

[10]See Conklin, Coulson, and Diop (2021) for a more detailed discussion of this appraisal process.

[11]This matching process often results in tightly clustered comp sales prices within appraisals, but importantly, price itself is not a valid criterion for selecting comps

in all 50 states from 2013 to 2017. The database contains information on more than 27.3 million comps used in 7.2 million residential appraisals associated with home purchase loan applications. A sales transaction is often used as a comp in multiple appraisals, so the comps consist of 11.2 million separate sales transactions on 10.2 million unique properties.[12] See Appendix Figure A.1 for the frequencies with which each type of transaction is used in multiple appraisals.

The data record basic characteristics of the comparable homes collected by the appraiser, including information commonly used in hedonic valuation models (e.g., square footage, lot size, age, number of bedrooms, number of bathrooms, etc.) as well as other characteristics not typically captured in standard housing databases (condition, quality, location and view).[13] The data also includes a unique appraisal identifier linking comps to specific appraisals, which is essential for our matching estimation approach. The appraiser also records whether the comparable transacted as real estate owned (REO), a short sale, or an arm's length transaction.[14] The comparable transaction's sale price is also recorded.

Appendix Table A.1 describes the property characteristics used in our study. To generate our initial data set, we exclude observations with missing housing characteristics as well as those missing a geographic identifier (census tract). We include only transactions classified as arm's length, REO, or short sales.[15] We impose the following restrictions for homes to be included in

[12] The data contain a unique property identifier as well as the quarter-year when the comparable transaction occurred. We consider each comparable property/transaction quarter combination as a unique sales transaction.

[13] See HUD (2015) for a detailed description of the standardized ratings for quality, condition, location and view. The condition (quality) variable is ordinal in nature, with 1 being the best condition (quality). HUD (2015) provides explicit descriptions for what each of these ratings means. We re-code these variables from 0 to 5 with 5 being the highest quality (condition) to ease interpretation. Location and view are categorized as adverse, neutral, or beneficial relative to the subject property. Conversations with appraisers suggest that they source quality, condition, location and view information in a number of ways, including visual inspection of the property, from MLS or public records, and from communication with other industry participants (e.g., real estate agents). These variables will enter as controls into our main regression (discussed momentarily) as a series of indicator variables.

[14] For some transactions that are used as comps in multiple appraisals, the transaction type (foreclosure, short sale, arm's lenth) conflicts across appraisals. We exclude these observations from our main analysis, but we will also use two alternative methods to deal with these observations in Section 5.3.

[15] 99.2% of the comps fall into one of these three transaction categories. We exclude court ordered sales, estate sales, relocation sales, and non-arms length sales. Our main analysis also excludes comps with conflicting classifications in different appraisals.

our analysis: between 500 and 10,000 square feet of gross living area (GLA); lot size between 500 and 1,000,000 square feet; homes aged less than 150 years; less than 15 rooms total; less than nine bedrooms; sales price between \$50,000 and \$1,425,000; and sold between 2012 and 2017.[16] The final cleaned sample includes 27,302,294 comparable home transactions.

Table 1 reports descriptive statistics by transaction type.[17] As expected, arm's length transactions have the highest average price (\$276,000), followed by short sales (\$224,000) and foreclosures (\$175,000). The Table shows that arm's length sales are properties with larger square footage, especially relative to foreclosed properties. Even though distressed properties have larger lot sizes, there are no differences in number of rooms, bedrooms, and bathrooms. But they tend to be slightly newer and are less likely to have a basement. These differences in attributes could explain part of the observed price differentials between distressed and non-distressed properties; however, differences in property condition are likely to be the most important contributor. As expected, Table 1 confirms that non-distressed properties are in much better condition and of higher quality than distressed ones. Recall, that the condition and quality ratings are ordinal in nature, with 5 being the best condition or highest quality rating. Interestingly, foreclosure and short sale properties appear to be of similar condition, which likely contributes to our finding of a small difference in our foreclosure and short sale discount estimates. Finally, the summary statistics in Table 1 show no significant differences in the location and view attributes between arm's length, foreclosure, and short sale properties.

Next, we explore *within-appraisal* differences in characteristics between distressed and non-distressed properties with regressions of the following form:

$$X_{ij} = \alpha_f Foreclosure_{ij} + \alpha_s ShortSale_{ij} + U_j + e_{ij} \tag{3}$$

[16]These data cleaning procedures, which are largely meant to remove outliers, only affect 2.5% of the original sample. Our study period is dictated by the available appraisal data. Comparable transactions must occur prior to the appraisal, hence why comparable transaction year can differ from appraisal year.

[17]Appendix Table A.2 presents more extensive descriptive statistics for the whole sample.

where X_{ij} is the characteristic of the i^{th} comparable poperty from the j^{th} appraisal. $Foreclosure_{ij}$ and $ShortSale_{ij}$ indicate the type of transaction, with arm's length sale as the omitted category. e_{ij} is an error term. With the inclusion of U_j, which are appraisal fixed effects, we are asking whether foreclosure and short sale properties are different from arm's length sale properties within the same appraisal. These regressions are reported in Table 2. Although estimated differences in property attributes between distressed and non-distressed properties are statistically significant due to the large sample size, most of the magnitudes are quite small. Differences in squarefootage, age, and condition are slightly larger than the other attributes, however, they are still relatively modest.[18]

Contrasting Tables 1 and 2, reveals that while there are clear differences in observable characteristics across transaction type in the whole dataset, within appraisal those differences are markedly smaller. This emphasizes the role of the appraiser in picking similar comps within an appraisal set. Given the decrease in observable differences across transaction types within appraisal when compared to the whole dataset, it is likely that a similar decrease in unobservable differences also occurs since appraisers will likely match properties along unobservable dimensions as well. This highlights the advantage of our empirical strategy of carrying out within appraisal price comparisons across transaction types.

A related question is whether the distressed sale properties in our appraisal data are similar to the universe of distressed sale properties. To shed light on this question, we compare observable characteristics of distressed sales in our data with an alternative sample of distressed sales from RealtyTrac. Our RealtyTrac data does not track short sales, and thus our discussion here will focus on foreclosures. From the RealtyTrac data, we collect transactions classified as foreclosures from 2013 thru 2015 and calculate county-year averages of available property characteristics.[19] Similarly, we calculate averages at the county-year level for these same characteristics of foreclosed

[18] Condition and quality can take on values ranging from 0 to 5. In columns (5) and (6) they are treated as continuous variables with a higher number indicating better condition and quality, respectively. When condition and quality are used as controls in our analysis in subsequent sections, they will enter as a series of indicator variables for each possible value, with 0 as the omitted category.

[19] Although the appraisal data covers 2013 thru 2017, our RealtyTrac data end in 2015.

properties in the appraisal data over the same period. Appendix Table A.3 reports the mean values of the county-year averages across both samples. The average sale price of foreclosures is slightly higher in the appraisal data. Foreclosures in the appraisal data are also larger in terms of square footage, but smaller in terms of lot size. The biggest difference between the two samples is the age of the properties. However, appraisers often report the "effective age" of the property, which accounts for improvements to the structure, whereas age in the RealtyTrac data records the length of time since an initial structure was placed on the property. The number of bedrooms, full bathrooms, and half bathrooms are nearly identical across the two data sets. Although there are some differences across datasets, aside from age due measurement differences, they are fairly modest. Thus, at least along observable dimensions, the foreclosure sales in the appraisal data seem similar to foreclosures in an alternative sample.

5. Empirical Analysis

As discussed in Section (2), our empirical methodology relies on appraisers' superior property matching ability to produce distress discount estimates that are far less subject to omitted variable bias than estimates from traditional hedonic methods. We implement this strategy using the following model specification, where the subscripts i and j refer to comp property i used in appraisal j.

$$P_{ij} = X_{ij}\beta + \gamma_f Foreclosure_{ij} + \gamma_s ShortSale_{ij} + \lambda_i + \mathcal{T}_i + U_j + e_{ij} \tag{4}$$

As in previous studies, we regress log sales price (P_{ij}) on property characterisitics (X_{ij}), property distress condition dummies ($Forclosure_i$ and $ShortSale_i$), census tract fixed effects (λ_i), time (year-quarter) fixed effects (τ_i), and finally appraisal fixed effects (U_j); e_{ij} is the error term. This specification allows us to estimate foreclosure and short sale discounts within appraisals conditional on observables commonly used in traditional hedonic models and more – the elements of

X_{ij} are listed in Table A.1. Again, the appraisal fixed effects will significantly reduce the omitted variable problem that is present in findings from previous studies. The coefficients of interest in equation (4) are γ_f and γ_s representing the average distress discounts associated with foreclosure and short sales, respectively, relative to arm's length transactions. To demonstrate the reliability of our methodology and appease potential data concerns, we replicate findings from previous studies by progressively expanding the set of controls used in our estimation before presenting results from our matching approach.

5.1. Main Results

In this section we discuss our results and compare our foreclosure and short sale discount estimates with those in the literature. Table 3 presents estimates of the foreclosure and short sale price discounts. For the sake of concision, we only report the coefficient estimates for the foreclosure and short sale indicators in Table 3 – the coefficients on the other housing characteristics are reported in Appendix Table A.4.

The model in column (1) of Table 3 includes only the foreclosure and short sale indicators and thus it estimates the unconditional effects of these variables on sale price. The coefficient on foreclosure is precisely estimated at -0.453, meaning that on average foreclosures sell at a discount of $1 - \exp(-0.453) = 36.4\%$ relative to non-distress property sales. The short sale coefficient is precisely estimated at -0.209, corresponding to a price discount of 18.9%. This model has obviously no explanatory power because the price difference between distressed and non-distressed sale prices is unlikely to be solely caused by the distressed nature of the sale. As discussed earlier, there are a number of other differences in the two samples, and the next several columns attempt to control for these differences.

First, foreclosures and short sales likely occur in areas that differ systematically from locations where only arm's length transactions are observed. For example, neighborhoods with volatile house prices or where individuals are particularly sensitive to employment shocks are more likely to have distressed sales. Additionally, time-varying national macroeconomic trends likely impact the prevalence of foreclosures and short sales. To address these issues, column (2) includes census tract and year-quarter dummies. This model shows a strong explanatory power. Relative to column (1), the foreclosure coefficient is nearly halved, but it remains statistically and economically significant. The short sale coefficient in column (2) is similar to column (1).

As Table 1 and Table 2 show, there is a need to control for differences in observable characteristics, as small as they may be, between distressed and non-distressed sales. In column (3) we include housing characteristics in our model.[20] The foreclosure coefficient declines slightly to -0.222, which corresponds to a 19.9% price discount. However, the short sale discount remains almost unchanged at 18.4%. Basically, after finely controlling for location and national price trends, the inclusion of property characteristics only marginally improves the accuracy of our distress discount estimates. Notice that these price discounts are in line with estimates from previous studies that control for property characteristics (Campbell, Giglio, and Pathak, 2011; Chinloy, Hardin, and Wu, 2017; Forgey, Rutherford, and VanBuskirk, 1994; Shilling, Benjamin, and Sirmans, 1990).

Properties sold through foreclosure or short sale are likely of inferior condition and quality, but these characteristics are generally unobservable to the econometrician (Clauretie and Daneshvary, 2009; Lambie-Hanson, 2015). A unique feature of our data is that the condition and construction quality of the property are reported by the appraiser.[21] We also observe whether the property has a beneficial (or adverse) view, or location, relative to the subject property. After controlling for condition, quality, location and view in column (4), both the foreclosure and short sale coefficients decline, and converge in magnitude. The foreclosure discount decreases by 16.6% (from 19.9% to 16.6%) and the short discount by 14.1% (from 18.4% to 15.8%). These estimates suggest that the

[20]Appendix Table A.4 shows coefficients for these housing characteristics. Except for basement, the coefficient estimates are relatively small in this regression.

[21]Clauretie and Daneshvary (2009) estimate the foreclosure discount including dummies to account for property condition (excellent, good, fair and poor), as reported by the listing agent at the time of listing.

impact of observable quality and condition differences between foreclosed and short sale properties on the price differences across these transaction types are not as large as generally expected, at least once other observables have been controlled for.[22]

Recall that in performing the appraisal, the appraiser chooses comps that are similar to the subject property in terms of location, housing characteristics, condition, etc. The comps should also be similar along dimensions that are unobservable to researchers as well. We account for this matching by including individual appraisal fixed effects in column (5). The R-squared increases dramatically to 98%, which seems to suggest that the matching procedure by the appraiser is quite effective. Note that accounting for the appraiser's matching also results in a dramatic decline in the estimated discounts. The foreclosure and short sale discounts are 5.2% and 5.8%, respectively, or roughly one third of the estimates in column (4). Our previous estimations only control for observable differences between distressed and non-distressed properties. The addition of appraiser fixed effects allows to control for some *"unobservable"* differences by relying on appraisers' property matching expertise. As Frame (2010) notes, omitted variables represent a serious problem in hedonic distress discount estimates. These smaller estimates suggest that omitted variables likely cause an overstatement of the foreclosure and short sale discounts observed in many earlier studies. If some of the differences in unobservable attributes across properties are not a function of distress itself, namely those differences that are not induced by the current homeowners being in distress, then said differences will likely have led to an upward bias in previous estimates of distress sale discounts.

Figure 1 presents our various estimates of foreclosure and short sale price discounts as a sequence of symbols from left to right corresponding to columns 1 through 5 of Table 3. It will be seen that the values decline in absolute value in sequence. In the next section we address the role

[22]In Table A.5 we estimate in Column (1) the difference in our measure of property condition between houses that will exhibit distress in a future sale and those that will not. The differences are slight and not statistically significant. Column (2) similarly suggests that the prices of future distressed houses were not condtionally different from non-distressed counterparts. This indicates that distressed properties are indeed subject to less assiduous maintenance at some point in the housing spell that leads to distress, confirming the results in Lambie-Hanson (2015).

that deliberate comp selection might play in producing our results. In so doing, we will also be able to estimate the role that seller liquidity may play in creating the distressed discount. We find both issues to be relatively minor (which will leave stigma as the remaining source of the discount).

5.2. Addressing Sample Selection Concerns

It is important that we ascertain that our sample is devoid of selection problems that might cloud our findings. Under the reasonable assumption that the properties being appraised represent a random sample of arm's length transactions, the main concern then is whether the selection of comps introduces bias into our estimates. Appraisers do not necessarily select comps randomly, even within the subset of transactions meeting the requirement of recency and similarity to the subject property. This non-random nature of comp selection may not be problematic in our context, though, because appraisers should, and likely do, screen out many distress-sale properties that are poor matches to the subject property. Technically, identifying the causal impact of distress on price implies comparing two properties that are identical in all respects, except distress. This comparison is what we try to achieve through our matching approach. But, the possibility remains that appraisers may select comps on price, which could lead to a downward bias in our distress discount estimates, especially in thick markets where appraisers may have access to a larger pool of transactions from which to select comparables.

Note, however, that we find large distress discounts using standard hedonic procedures, which suggests that appraisers aren't selecting particularly high-value distressed comps. Nevertheless, we address this potential selection issue by examining the impact of market thickness, the level of the activity in the local housing market, on our estimates. We would expect, from the above, that the discount found in thin markets is greater in magnitude if our sample is plagued by these selection issues, because appraisers would have fewer distressed comps to choose from. For this exercise, we divide the sample into quintiles by market thickness, proxied by the number of comparable sales in the subject property's neighborhood within the twelve months prior to the appraisal. This number is reported by the appraiser, and for most appraisals, it is significantly larger than the actual number of comps used in the appraisal. Note that this number is somewhat subjective, as the appraiser reports what he deems as comparable sales. With this caveat in mind, it should still serve as a reasonable proxy for market thickness. Figure 2 shows foreclosure and short sale discounts by market thickness quintile. For all quintiles, the general pattern of estimated distressed sales discounts is as in Figure 1, thus suggesting that the appraiser's selection of comps did not introduce significant selection bias in our analysis.

Moreover, the consistency of our estimates across the market thickness quintiles suggests that market liquidity likely plays a minor role in our estimates. We have noted that sellers of distressed property (often financial institutions) may have greater urgency to sell, thus leading to a higher distressed discount, sometimes denoted as a "fire-sale" effect. In a market where properties are taking longer to transact we would expect the "fire-sale" effect to be most pronounced. Yet, we see no such differences across market thickness, suggesting the "fire-sale" effect is not impacting our discount estimate.This stands in sharp contrast to Campbell, Giglio, and Pathak (2011), where liquidity plays a major role in the discount estimates.

5.3. Robustness Checks

Using our appraisal fixed effect (matching) methodology from column (5) of Table 3, we estimate various alternative specifications and report our findings in Table 4. Columns (1) and (2) of Table 4 show that our distress discount estimates remain statistically significant when we cluster standard errors at the transaction and property level, respectively. Next, because a specific sales transaction can appear as a comp in multiple appraisals, in column (3) we use a weighted regression technique that accounts for the number of times a unique property transaction appears in the data. We use as weights the inverse of the number of times a specific transaction is used as a comp.[23,24] Again, the magnitude and significance of our estimates remain unchanged.

Consider now a variable that indicates that a comp is the lowest priced comp within its group. Given the inclusion of appraisal fixed effects this coefficient will mechanically have a negative coefficient and is, roughly speaking, the conditional price difference between the lowest-priced property in the appraisal and the average property price within the appraisal. If distressed comps are also the lowest priced comps, these effects may be confounding one another and this can lead to an *overstatement* of the distressed discount.

We provide two pieces of evidence on this point. First, distressed comps are *not* always the lowest price comps in appraisals that include a distressed comp. Conditional on a foreclosure comp being used, the foreclosure comp is the lowest price comp 48% of the time, while it is the highest price comp 20% of the time. The corresponding figures for appraisals that include short sales are 47% and 19%.[25] Second, we create dummies identifying the lowest and highest price comps within each appraisal and include them as additional controls in column (4) of Table 4. The coefficient on "lowest price" is negative, as expected and that for "highest price" is positive (again, this is a mechanical effect). The coefficient on distress is now 3.4%, which is lower than our baseline estimate of 5%, as would be expected given the fact that distressed comps are more likely than arms-length comps to be the lowest price comp within an appraisal. Taken together, these two pieces of evidence suggest that our observed distress discounts are not capturing solely a "lowest price comp" effect.

Column (5) of Table 4 controls for the intensity of foreclosure activity at the tract level using a quintile ranking system based on the share of the Census tract's housing stock that is in foreclosure as reported by RealtyTrac. It is an empirical question how foreclosure intensity will affect distress

[23]We thank an anonymous reviewer for this suggestion.

[24]Figure A.1 shows that most foreclosure and short sales are used in only one appraisal, whereas arm's length sales are more likely to appear in multiple appraisals.

[25]On average, there are four comps in an appraisal, so a distressed comp has a higher likelihood of being the lowest price comp relative to any individual arms-length comp.

discounts. More foreclosures should normally negatively affect the value of all houses in the area irrespective of their distress status. However, more foreclosure will also increase the number of foreclosure sales that are suitable as comps. As expected column (5) shows that sales prices decrease with foreclosure intensity. Although our foreclosure distress discount estimates slightly decrease with foreclosure intensity, they remain in the vicinity of our previous estimates.

We also check the robustness of our estimates along several other dimensions. First, we check the consistency of our distress discount estimates over time. The strengthening housing market recovery and improved market liquidity during our sample period might have caused a tightening of distress discounts. Therefore, we want to make sure that the magnitude of our discount estimates are not due to the improved market conditions. In Figure 3 we estimate γ_f and γ_s separately for each year of our study period using the models in Table 3. While we note small variations in the magnitudes of our estimates from the first four model specifications, by the time we get to the matching estimators our discount estimates are all very close to -0.05. Overall, the consistency of the estimates from our matching approach suggests that stigma may be the main explanation of our discount estimates than liquidity.

Campbell, Giglio, and Pathak (2011) find larger distress discounts for houses with low-priced characteristics in low-priced neighborhoods, which suggests that these larger discounts may be due to poor maintenance. Therefore, we explore whether our distress estimates vary with house prices because it might be thought that the small size of our discounts is driven by neighborhoods which are less susceptible to ill treatment of distressed property. In Figure 4, we divide our sample by quintiles in average tract price level (as measured by the American Community Survey). In contrast to Campbell, Giglio, and Pathak (2011), our foreclosure and short sale discounts estimates in low- and high-price areas are very similar in magnitude once we control for housing characteristics, property condition and quality, and time and location fixed effects using the same model specification as in column (4) of Table 3. Again, our appraisal matching approach yields distress discounts of about 5% with no apparent differences in magnitude as we move from less expensive to more expensive neighborhoods.

As an extension of our previous analysis, we explore the impact of local racial composition on distressed sale discounts given the strong correlation between neighborhood racial composition and local house prices. In Figure 5, similar results occur as we stratify our sample by census tract minority population share. Once our matching estimator is used, the discounts are nearly identical across tract minority share quintiles.

In Figure 6 we break out samples for each of the "Sand States" (Arizona, California, Florida and Nevada) and then the remainder of the states considered together. The reasons for this split are twofold. First, sand states experienced most of the excesses in mortgage lending during the housing boom of the early 2000s and have therefore faced the brunt of the subsequent foreclosure crisis (Choi et al., 2016), which could have affected the magnitude of distress discounts in those areas. Secondly, this analysis will allow us to compare our estimates with findings from previous studies relying on data from sand states (Carroll, Clauretie, and Neill, 1997; Clauretie and Daneshvary, 2009). Figure 6 shows a net decline in our discount estimates in both sand and non-sand states as we follow the same modeling sequence as in Table 3. The three foreclosure discounts estimates from the models that successively include housing characteristics, tract and year fixed effects, and condition, quality, location and view are smaller in sand states than in non-sand states, which aligns with the lower discount Clauretie and Daneshvary (2009) find. Although our matching estimation method produces as before distress discounts of roughly 5% in sand and non-sand states, there are some slight differences here as both California and Nevada have rather smaller foreclosure discounts.[26]

Figure 7 reports estimates separately for non-judicial (NJ) and judicial (J) foreclosure states.[27] Given that the foreclosure process is generally shorter in non-judicial states, lenders in those states

[26]It is of interest to note that among the lowest discounts previously estimated were those for Nevada e.g.Clauretie and Daneshvary (2009) and Carroll, Clauretie, and Neill (1997). Our estimates in that way correspond with earlier literature.

[27]State classification is based on Table 1 of Ghent and Kudlyak (2011).

may be willing to accept larger discounts when offloading repossessed properties. However, our results do not vary with this distinction.

Our final robustness check addresses the issue of comps with conflicting classifications as foreclosure, short sale, or even arm's length sales in different appraisals. As noted earlier, our baseline analysis excludes those comps. Figure 8 presents distress discount estimates based on two alternative classifications of those distressed comps. Our first measure (Alt. 1) keeps the comps' original disposition type as reported in the data, whereas the second measure (Alt. 2) reclassifies a comp transaction as a foreclosure (short sale) if it is ever recorded as a foreclosure (short sale). Figure 8 shows the resulting distress discount estimates with the corresponding regressions reported in Table A.6 and Table A.7. Again, the magnitudes of our findings remain when we use these alternative measures.

It is important to note that in each of above figures, while the distress estimates from our matching estimations, column (5), are extremely robust, with minuscule variations in the value of the estimates, those that correspond to standard hedonic methods, columns 1 through 4, exhibit much greater variations. These results, which replicate estimation methods in previous studies of distress discounts, are not robust at all, and indeed correspond to the wide variety of estimates seen in the previous literature.

6. Conclusion

The existing literature has documented astonishingly large price discounts on foreclosure and short sales, which suggest a significant degree of inefficiency in the market for such properties. Using unique nationwide home appraisal data that allow us to address the omitted variable bias that was likely present in previous estimates of forced sale discounts from standard hedonic models, we propose an innovative estimation methodology relying on appraisers' ability to match properties along both observable and unobservable attributes when performing appraisals. Our empirical approach, which relies on the use of appraisal fixed effects after controlling for a rich set of observables, produces distressed sale discounts that are substantially lower than estimates documented in the literature. Given the effectiveness of our matching technique, which allows us to account for most differences between distressed and non-distressed properties, we attribute these lower distress discounts of about 5% largely to stigma associated with distress ifself. Our findings are robust across a wide variety of subsamples and alternative estimation methods. We also show that these lower distress estimates are not affected by market liquidity or due to appraisers selecting overly favorable distressed properties as comps because we were able to reproduce using the same data discounts that are similar in magnitude to those from standard hedonic models.

There are several important implications of our findings. First and foremost, the large discounts seen in previous studies seem to imply great inefficiencies in real estate markets; our results show that this is not the case. Our findings have implications not only for real estate market efficiency, but also mortgage underwriting and housing policy. A key input in mortgage credit risk models is loss conditional on default, which depends critically on the sources and sizes of distressed sale discounts, which this study elucidates – one should note that our distress estimates are not necessarily the sole discounts that matter to lenders because property condition, which is controlled for in our estimates, may affect those discounts and be an important part of a lenders' choice regarding which property disposition method to employ. Understanding the sources and sizes of price discounts associated with forced sales, particularly foreclosure sales, also has important implications in designing policies to address negative externalities resulting from property distress. Finally, previous studies document significantly smaller discounts for short sales relative to foreclosures, which seems to imply that short sales are a dominant loss-mitigation tool from the lender's perspective. In contrast, we find that the foreclosure and short sale discounts are similar in magnitude once we account for quality, condition, and characteristics that are typically unobservable to researchers, which may help explain why short sales are not more prevalent.

Finally, our results are congruent with the general feeling in the appraisal community that foreclosure, as such, has a relatively minor impact on appraisal quality. In another paper Conklin, Coulson, and Diop (2021) document the not-uncommon use of distressed property as comps, and how the use of these comps does not unduly affect property appraisals, despite the concern of some in the residential real estate sector.

References

Appraisal Institute. 2013. Understanding the Appraisal. Working Paper, Appraisal Institute.

Campbell, J. Y., S. Giglio, and P. Pathak. 2011. Forced sales and house prices. *American Economic Review* 101:2108–31.

Carroll, T., T. Clauretie, and H. Neill. 1997. Effect of foreclosure status on residential selling price: comment. *Journal of Real Estate Research* 13:95–102.

Chinloy, P., W. Hardin, and Z. Wu. 2017. Foreclosure, reo, and market sales in residential real estate. *The Journal of Real Estate Finance and Economics* 54:188–215.

Choi, H.-S., H. G. Hong, J. D. Kubik, and J. P. Thompson. 2016. Sand states and the us housing crisis. *Available at SSRN 2373179* .

Clauretie, T. M., and N. Daneshvary. 2009. Estimating the house foreclosure discount corrected for spatial price interdependence and endogeneity of marketing time. *Real Estate Economics* 37:43–67.

Conklin, J. N., N. E. Coulson, and M. Diop. 2021. Distressed comps. *Working Paper* .

Doerner, W. M., and A. V. Leventis. 2015. Distressed sales and the FHFA house price index. *Journal of Housing Research* 24:127–46.

Donner, H. 2020. Determinants of a foreclosure discount. *Journal of Housing and the Built Environment* 35:1079–97.

Donner, H., H.-S. Song, and M. Wilhelmsson. 2016. Forced sales and their impact on real estate prices. *Journal of Housing Economics* 34:60–8.

Fannie Mae. 2020. Lender Letter (LL-2020-04). *Available at* `https://singlefamily.fanniemae.com/media/22321/display`.

FHA. 2020. Modified set of instruction, scope of work, statement of assumptions and limiting conditions, and certification for FHA desktop appraisal. *Available at* `https://www.hud.gov/sites/dfiles/SFH/documents/ModelApprDesktop_OnlyCert_03_27_20.pdf`.

Forgey, F., R. Rutherford, and M. VanBuskirk. 1994. Effect of foreclosure status on residential selling price. *Journal of Real Estate Research* 9:313–8.

Frame, W. S. 2010. Estimating the effect of mortgage foreclosures on nearby property values: A critical review of the literature. *Economic review* 95.

Ghent, A. C., and M. Kudlyak. 2011. Recourse and Residential Mortgage Default: Evidence from US States. *Review of Financial Studies* 24:3139–86. doi:10.1093/rfs/hhr055.

Harding, J. P., E. Rosenblatt, and V. W. Yao. 2012. The foreclosure discount: Myth or reality? *Journal of Urban Economics* 71:204–18.

HUD. 2015. FHA single family housing appraisal report and data delivery guide. Working Paper, Department of Housing and Urban Development.

Just, T., M. Heinrich, M. A. Maurin, and T. Schreck. 2019. Foreclosure discounts for german housing markets. *International Journal of Housing Markets and Analysis* .

Lambie-Hanson, L. 2015. When does delinquency result in neglect? mortgage distress and property maintenance. *Journal of Urban Economics* 90:1–16.

Lopez, L. A. 2021. Asymmetric information and personal affiliations in brokered housing transactions. *Real Estate Economics* 49:459–92.

Mehrotra, A., A. Nowak, and P. Smith. 2021. The effect of securitization on asset price decisions. *Working Paper* .

Mocking, R., and B. Overvest. 2017. Direct and spillover effects of forced sales on house prices: Evidence from the netherlands. *Journal of Housing Economics* 38:50–61. ISSN 1051-1377. doi:https://doi.org/10.1016/j.jhe.2017.10.002.

Renigier-Biłozor, M., M. Walacik, S. Źróbek, and M. d'Amato. 2018. Forced sale discount on property market–how to assess it? *Land use policy* 78:104–15.

Shilling, J., J. Benjamin, and C. Sirmans. 1990. Estimating net realizable value for distressed real estate. *Journal of Real Estate Research* 5:129–40.

Stroebel, J. 2016. Asymmetric information about collateral values. *The Journal of Finance* 71:1071–112.

William, H., and W. Marvin. 1996. The relationship between foreclosure status and apartment price. *Journal of Real Estate Research* 12:101–9.

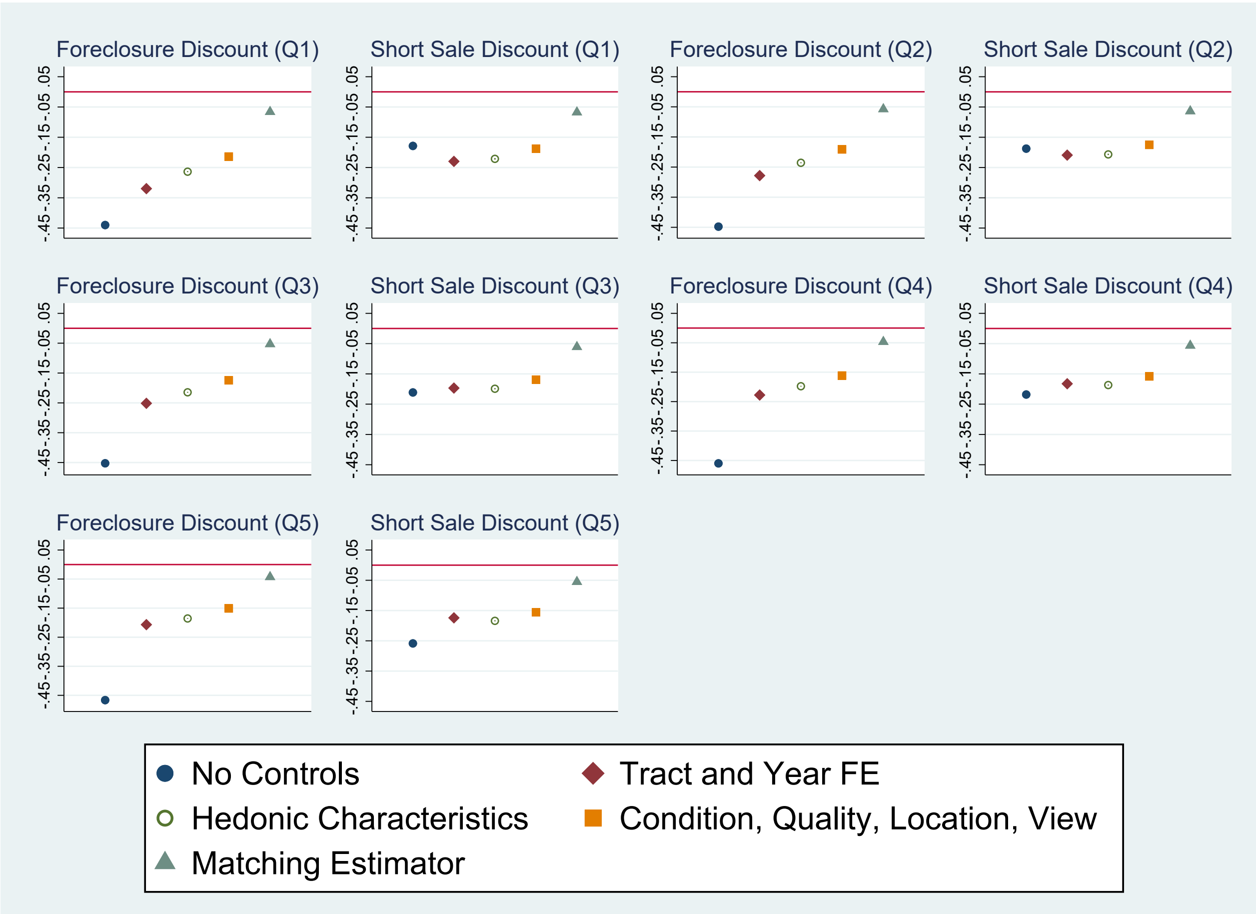
Zhou, H., Y. Yuan, C. Lako, M. Sklarz, and C. McKinney. 2015. Foreclosure discount: definition and dynamic patterns. *Real Estate Economics* 43:683–718.

7. Figures

Figure 1. Price Discount for Foreclosures and Short Sales

Note: This figure presents coefficient estimates from Table 3. Moving from left to right within each panel corresponds to moving from left to right across columns in Table 3.

Figure 2. Price Discounts for Foreclosures and Short Sales by Market Thickness

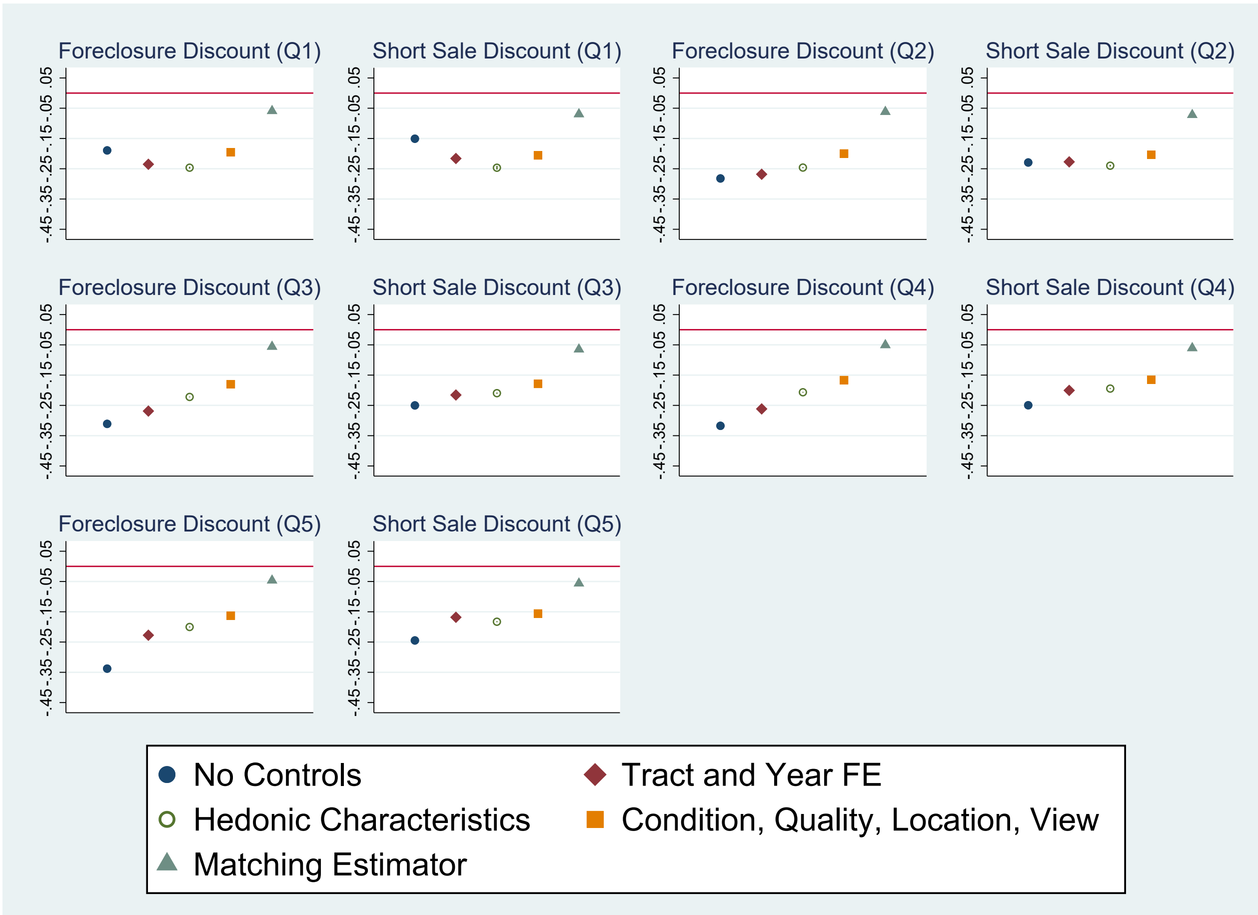


Note: This figure presents coefficient estimates of the foreclosure and short sale discounts by market thickness. The appraisal data includes the number of recent/nearby sales that are potential comps for the appraisal. We create quintiles based on this number, with Q1 and Q5 representing the quintiles with the least and most potential comps available, respectively (Q2=quintile 2; Q3=quintile 3; Q4=quintile 4;). Moving from left to right within each panel corresponds to moving from left to right across columns in Table 3.

Figure 3. Price Discounts for Foreclosures and Short Sales by Year

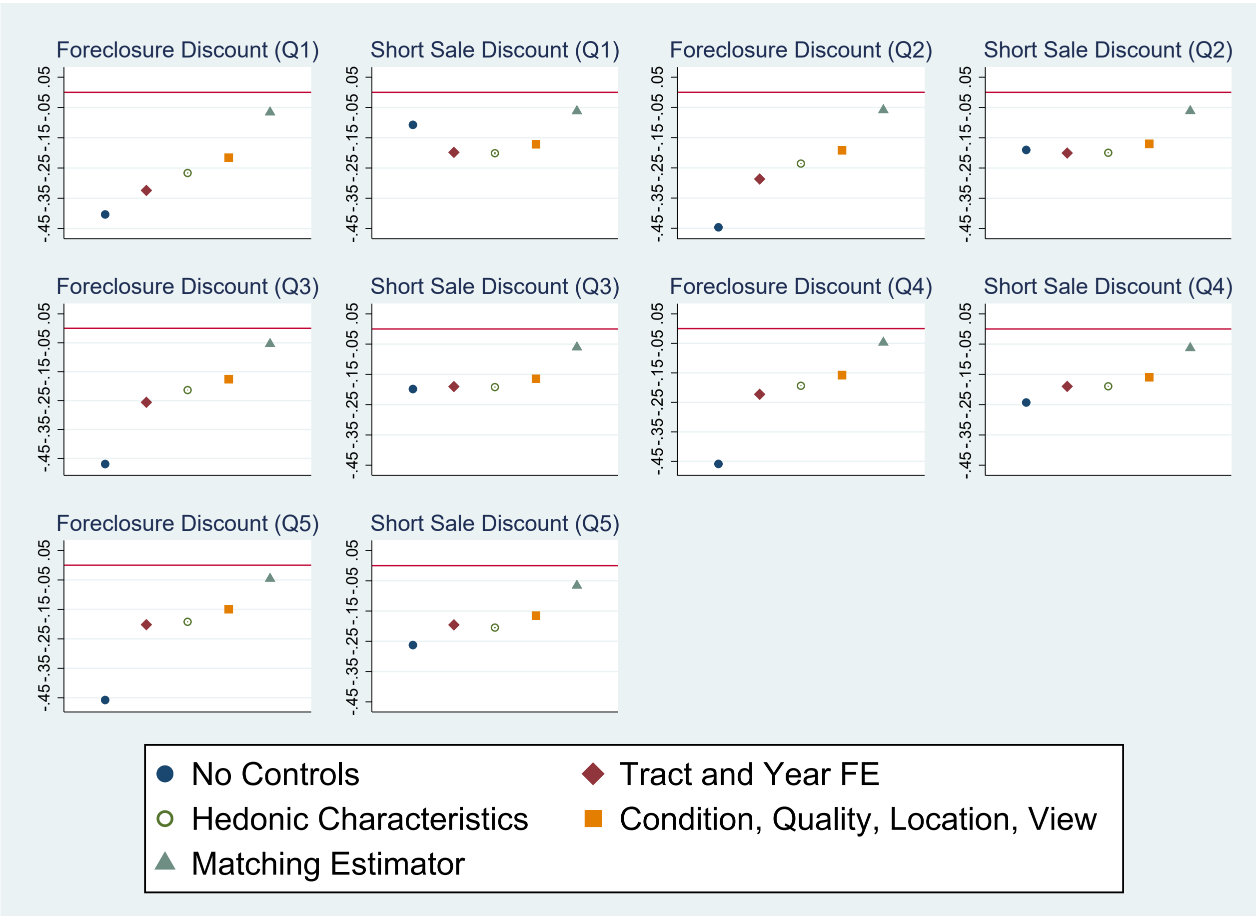
Note: This figure presents coefficient estimates of the foreclosure and short sale discounts by year of appraisal. Moving from left to right within each panel corresponds to moving from left to right across columns in Table 3.

Figure 4. Price Discounts for Foreclosures and Short Sales by Tract Price Levels



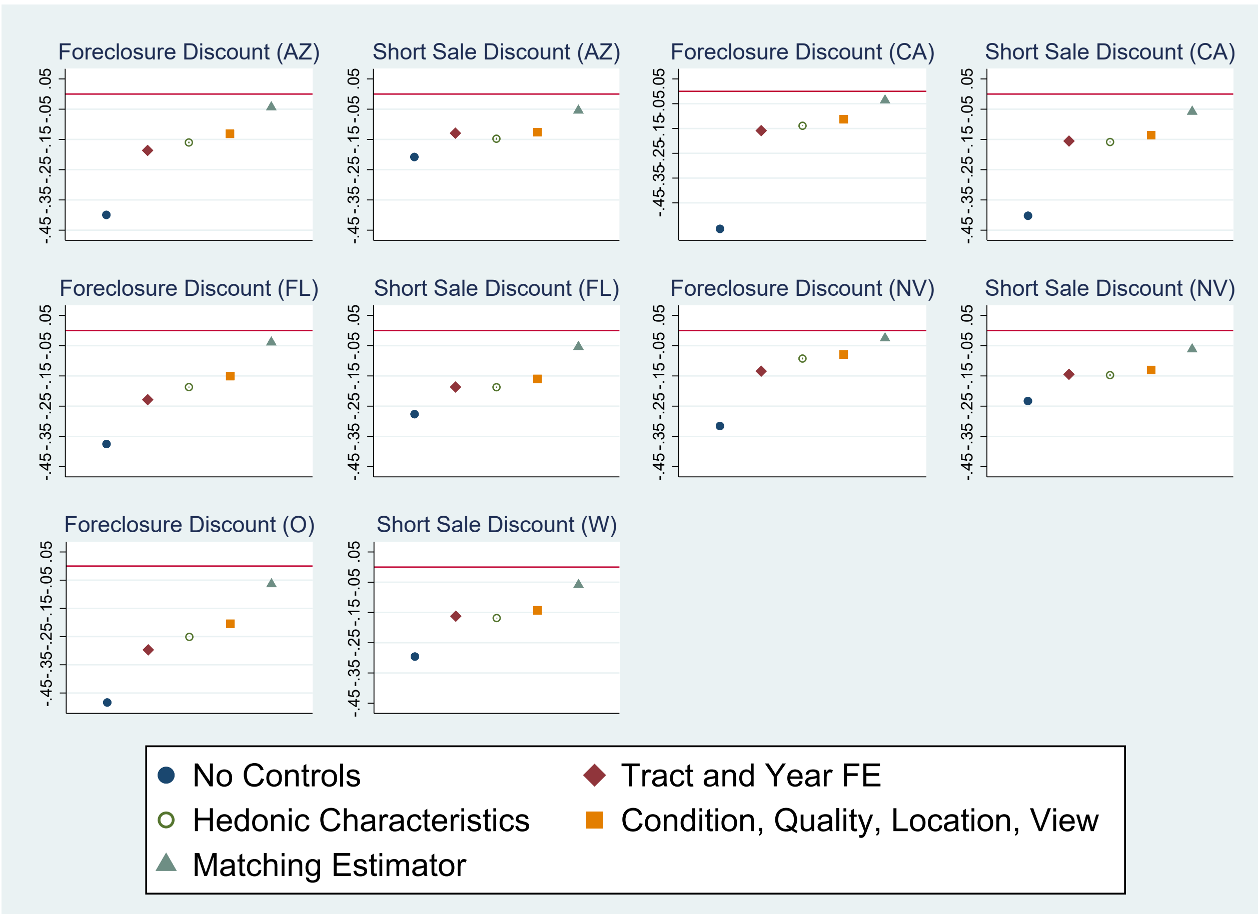
Note: This figure presents coefficient estimates of the foreclosure and short sale discounts across Census tract median house price quintiles. Each year we create Census tract median house price quintiles using ACS data, with Q1 and Q5 as the lowest and highest house price level quintile, respectively (Q2=quintile 2; Q3=quintile 3; Q4=quintile 4). Moving from left to right within each panel corresponds to moving from left to right across columns in Table 3.

Figure 5. Price Discounts for Foreclosures and Short Sales by Census Tract Minority Population Share



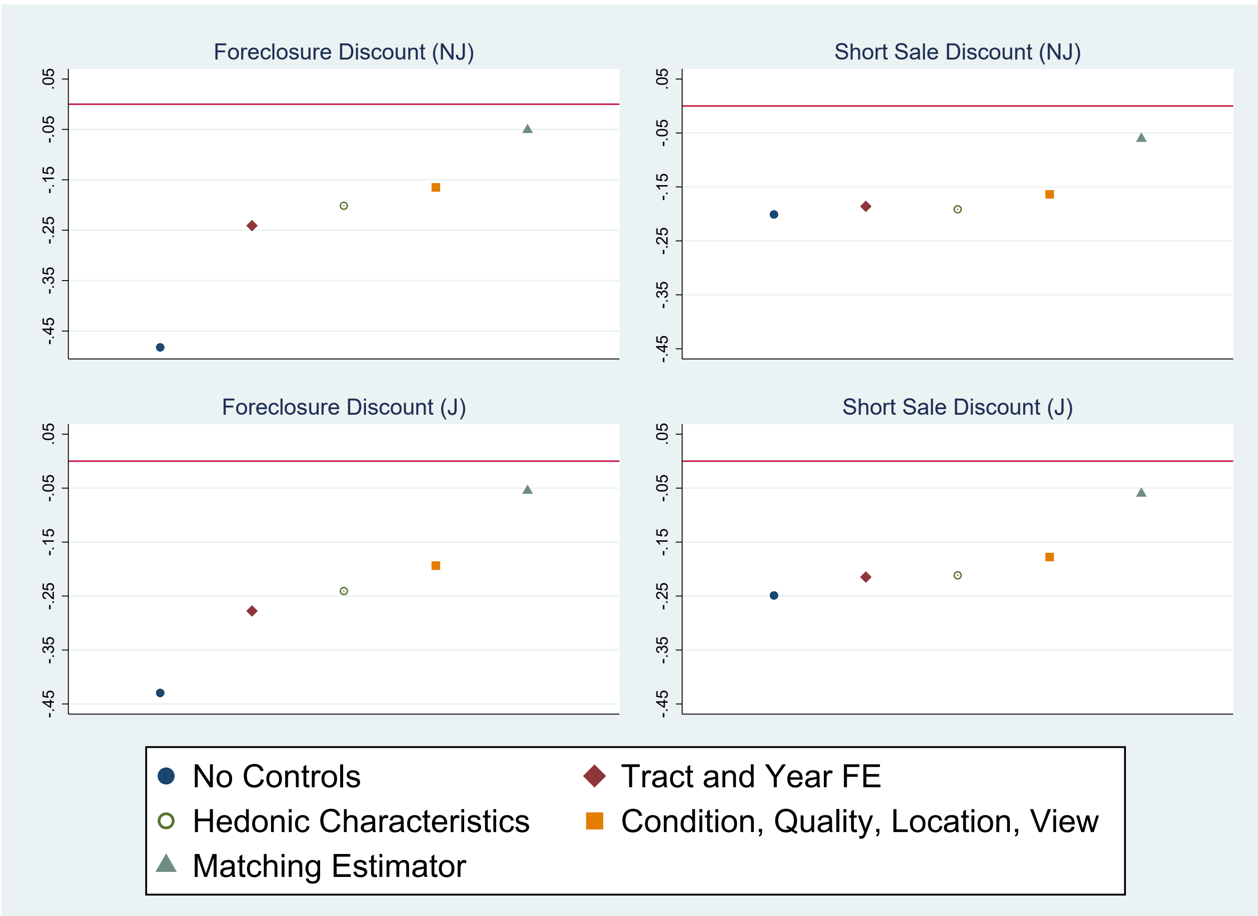
Note: This figure presents coefficient estimates of the foreclosure and short sale discounts by census tract minority population share. Minority share of the tract population is based on data from the 2010 Census, with Q1 and Q5 representing the quintiles with the lowest and highest minority population share, respectively (Q2=quintile 2; Q3=quintile 3; Q4=quintile 4;). Moving from left to right within each panel corresponds to moving from left to right across columns in Table 3.

Figure 6. Price Discounts for Foreclosures and Short Sales for Sand and Non-Sand States



Note: This figure presents coefficient estimates of the foreclosure and short sale discounts for Sand and Non-Sand States (AZ=Arizona; CA=California; FL=Florida; NV=Nevada; O=Other/Non-Sand States). Moving from left to right within each panel corresponds to moving from left to right across columns in Table 3.

Figure 7. Price Discounts for Foreclosures and Short Sales for Judicial and Non-judicial States



Note: This figure presents coefficient estimates of the foreclosure and short sale discounts for non-judicial (NJ) and judicial (J) states. State classification is based on Table 1 in Ghent and Kudlyak (2011). Moving from left to right within each panel corresponds to moving from left to right across columns in Table 3.

Figure 8. Price Discount for Alternative Measures of Foreclosures and Short Sales

Note: This figure presents coefficient estimates analogous to Table 3 using alternative measures of foreclosure and short sale. Both Alt. 1 and Alt. 2 include transactions where the recorded disposition (arm's length, foreclosure, short sale) varies across appraisals. For each observation using Alt. 1 we use the disposition type that is reported in the data. For the Alt. 2 measures, if a specific transaction is ever recorded as a foreclosure (short sale), we classify that transaction as a foreclosure (short sale) for each time that it is observed in the data.

8. Tables

Table 1. Mean Values by Transaction Type

	Transaction Type		
	Arm's Length	Foreclosure	Short Sale
Ln(Sale Price)	12.527	12.074	12.319
Foreclosure	0	1	0
Short Sale	0	0	1
Sq. ft.	1990	1834	1987
Lot size	21108	24921	22072
Age	34	32	31
Rooms	7	7	7
Bedrooms	3	3	3
Full baths	2	2	2
Half baths	0.422	0.348	0.395
Basement	0.421	0.329	0.310
Finished Basement	0.285	0.187	0.203
Condition0	0.000	0.001	0.001
Condition1	0.002	0.045	0.023
Condition2	0.178	0.468	0.390
Condition3	0.649	0.450	0.530
Condition4	0.106	0.035	0.055
Condition5	0.065	0.001	0.001
Quality0	0.000	0.000	0.000
Quality1	0.014	0.034	0.025
Quality2	0.549	0.674	0.626
Quality3	0.404	0.279	0.325
Quality4	0.032	0.013	0.023
Quality5	0.002	0.000	0.001
Neutral location	0.908	0.907	0.877
Beneficial location	0.059	0.053	0.073
Adverse location	0.033	0.039	0.050
Neutral view	0.887	0.889	0.863
Beneficial view	0.103	0.096	0.125
Adverse view	0.010	0.014	0.012
N	26,859,520	296,749	146,025

Note: Mean values by transaction type.

Table 2. Differences in Property Characteristics for Foreclosures and Short Sales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Sq. ft.	Age	Beds	Baths	Basement	Condition	Quality
Foreclosure	98.311	-2.217	0.062	0.074	0.013	-0.353	-0.025
	(0.666)	(0.024)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Short Sale	113.312	-1.854	0.067	0.081	0.010	-0.272	-0.017
	(0.932)	(0.033)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	1,986.252	33.538	3.288	2.028	0.419	3.051	2.456
	(0.057)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	27,273,234	27,273,234	27,273,234	27,273,234	27,273,234	27,273,234	27,273,234
Adjusted R-squared	0.867	0.853	0.573	0.633	0.849	0.670	0.897
Appraisal FE	Y	Y	Y	Y	Y	Y	Y

Note: This table reports estimates and standard errors, in parentheses, of a regression with a property characteristic as the dependent variable. The sample includes all comparable properties. Foreclosure and short sale indicators are mutually exclusive.

Table 3. Price Discount for Foreclosures and Short Sales

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)	(5) ln(price)
Foreclosure	-0.453	-0.261	-0.222	-0.181	-0.053
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Short Sale	-0.209	-0.201	-0.203	-0.172	-0.060
	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)
Observations	27,302,294	27,299,513	27,299,513	27,299,513	27,273,234
Adjusted R-squared	0.007	0.714	0.908	0.918	0.984
Tract FE	N	Y	Y	Y	N
Year/Qtr FE	N	Y	Y	Y	Y
Hedonic Characteristics	N	N	Y	Y	Y
CQLV	N	N	N	Y	Y
Appraisal FE	N	N	N	N	Y

Note: This table reports estimates and standard errors, in parentheses, of a regression with log sales price as the dependent variable. The sample includes all comparable properties. Foreclosure and short sale indicators are mutually exclusive.

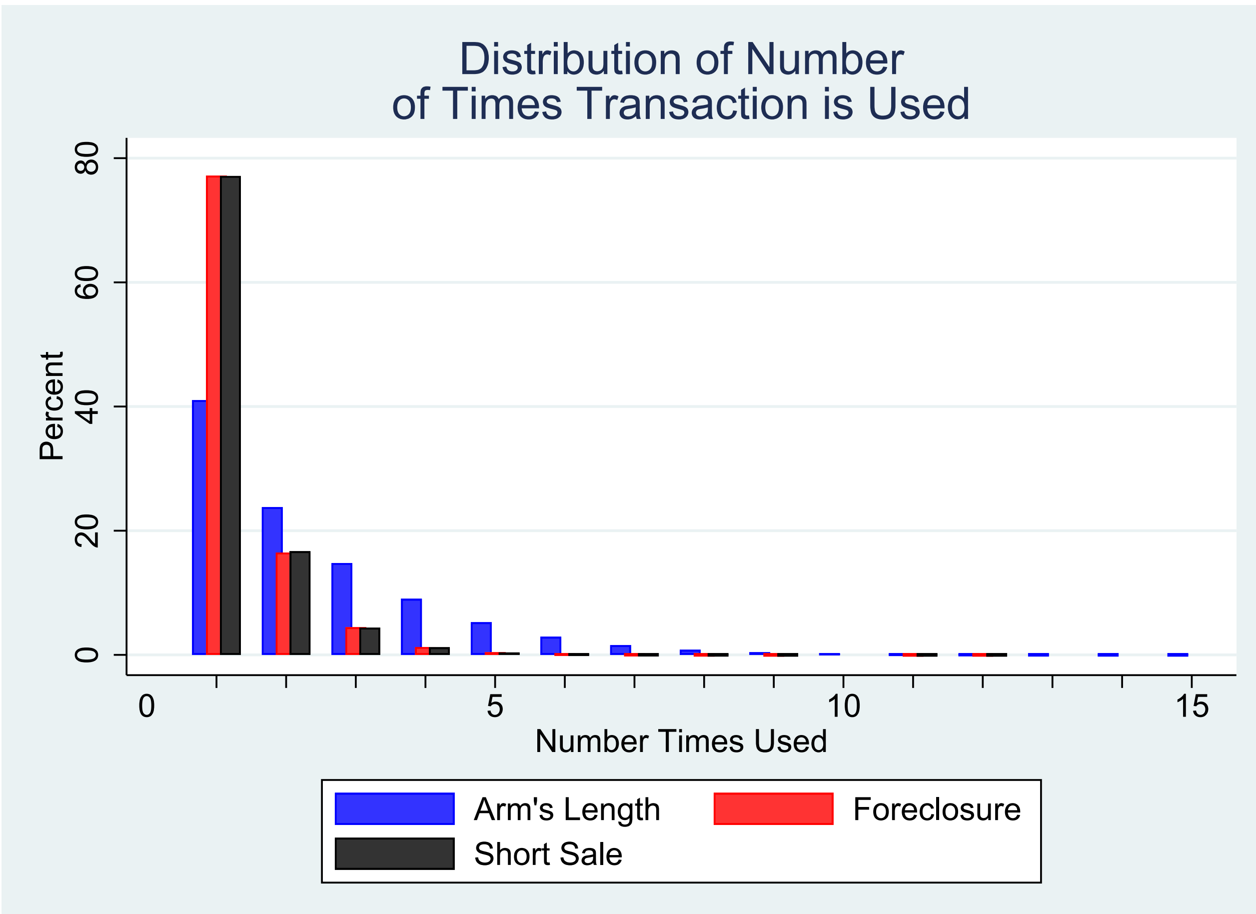
Table 4. Additional Robustness Tests

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)	(5) ln(price)
Foreclosure	-0.053	-0.053	-0.051	-0.033	-0.060
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
× Tract FC Share Q2					0.009
					(0.001)
× Tract FC Share Q3					0.014
					(0.001)
× Tract FC Share Q4					0.015
					(0.001)
× Tract FC Share Q5					0.017
					(0.001)
Short Sale	-0.060	-0.060	-0.057	-0.034	-0.059
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Lowest Price Sale				-0.066	
				(0.000)	
Highest Price Sale				0.066	
				(0.000)	
Tract FC Share Q2					-0.001
					(0.000)
Tract FC Share Q3					-0.003
					(0.000)
Tract FC Share Q4					-0.005
					(0.000)
Tract FC Share Q5					-0.008
					(0.000)
Observations	27,273,181	27,273,181	27,273,234	27,273,234	15,261,550
Adjusted R-squared	0.984	0.984	0.985	0.991	0.984
S.E. Cluster Level	Transaction	Property	N	N	N
Weighting	N	N	Y	N	N
Tract FE	N	N	N	N	N
Year/Qtr FE	Y	Y	Y	Y	Y
Hedonic Characteristics	Y	Y	Y	Y	Y
CQLV	Y	Y	Y	Y	Y
Appraisal FE	Y	Y	Y	Y	Y

Note: This table reports estimates and standard errors, in parentheses, of a regression with log sales price as the dependent variable. The sample includes all comparable properties. Columns (1) and (2) report standard errors adjusted for clustering at the transaction and property level, respectively. Column (3) uses inverse probability weighting $\frac{1}{n}$ where n is the number of times the transaction is used as a comp. Column (4) includes binary variables that indicate whether the comp was the lowest or highest price sale used in the appraisal it is associated with. Column (5) includes quintiles (1 lowest; 5 highest) for the share of the Census tract's housing stock that is in foreclosure according to the RealtyTrac data.

A.1. Online Appendix

Figure A.1 . Distribution of the Number of Times a Transaction is Used in and Appraisal



Note: This figure plots the distribution of the number of times that a specific sales transaction is used as a comp by transaction type. The figure shows that most non-arm's length transactions are used as a comp in only one appraisal.

Table A.1. Variable Names

Variable Name	Description
Dependent Variable	
Price	Sales price of the comp property.
Independent Variables of Interest	
Foreclosure	Property sold as real estate owned (REO).
Short Sale	Property sold as a short-sale.
Hedonic Characteristics	
Sq. ft.	Square footage of the property.
Sq. ft. sq.	Square footage squared.
Lot size	Lot square footage.
Lot size sq.	Lot square footage squared.
Age	Age of the subject property in years.
Age sq.	Age squared.
Rooms	Number of rooms.
Bedrooms	Number of bedrooms.
Full baths	Number of full bathrooms.
Half baths	Number of half bathrooms.
Basement	Indicates the existence of a basement.
Finished Basement	Indicates the existence of a finished basement.
Condition, Quality, Location and View	
Condition	Categorical variable defining condition of the property according to USPAP. Original variable is re-scaled to 0-5 with 5 representing the best condition. Condition enters regression models as a series of indicators.
Quality	Categorical variable defining construction quality of the property according to UPAP. Original variable is re-scaled to 0-5 with 5 representing the highest level. Quality enters regression models as a series of indicators.
Location	Location's impact on value according to USPAP, with the categories neutral, beneficial and adverse. Location enters regression models as a series of indicators.
View	Property view's impact on value according to USPAP, with the categories neutral, beneficial and adverse. View enters regression models as a series of indicators.

Note: Variable names and descriptions.

Table A.2. Descriptive Statistics

	Mean	Std. Dev.	Min	Max
Ln(Sale Price)	12.521	0.583	10.820	14.170
Foreclosure	0.011			
Short Sale	0.005			
Sq. ft.	1988	803	500	10000
Lot size	21154	52012	500	999702
Age	34	27	0	150
Rooms	7	2	1	15
Bedrooms	3	1	0	8
Full baths	2	1	1	9
Half baths	0	1	0	9
Basement	0.419			
Finished Basement	0.284			
Condition0	0.000			
Condition1	0.003			
Condition2	0.182			
Condition3	0.647			
Condition4	0.105			
Condition5	0.064			
Quality0	0.000			
Quality1	0.014			
Quality2	0.551			
Quality3	0.402			
Quality4	0.031			
Quality5	0.002			
Neutral location	0.908			
Beneficial location	0.059			
Adverse location	0.033			
Neutral view	0.887			
Beneficial view	0.103			
Adverse view	0.010			
N		27,302,294		

Note: Descriptive statistics for sample of comps. These comps consist 11,601,057 sales transactions on 10,245,574 unique properties used in 7,201,958 separate appraisals. Only the mean is reported for indicator variables.

Table A.3. Comparison of Foreclosed Property Characteristics Between Appraisal and RealtyTrac Data

	Data Source			
	Appraisal		RealtyTrac	
	N	Mean	N	Mean
Sale Price	3867	166167	1813	160061
Sq. ft.	3867	1765	2459	1638
Lot Size	3867	50484	2438	48079
Age	3867	35	2459	50
Bedrooms	3867	3	2155	3
Full Baths	3867	2	2405	2
Half Baths	3867	0.32	2405	0.27

Note: This table presents mean values of foreclosed property characteristics at the county-year level in our appraisal data and in RealtyTrac data, respectively. Some of the property characteristic fields are sparely populated in the RealtyTrac data, which results in different numbers of county-year observations across characteristics.

Table A.4. Price Discount for Foreclosures and Short Sales

	(1)	(2)	(3)	(4)	(5)
VARIABLES	ln(price)	ln(price)	ln(price)	ln(price)	ln(price)
Foreclosure	-0.453	-0.261	-0.222	-0.181	-0.053
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Short Sale	-0.209	-0.201	-0.203	-0.172	-0.060
	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)
Sq. ft.			0.001	0.001	0.000
			(0.000)	(0.000)	(0.000)
Sq. ft. sq.			-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)
Lot size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Lot size sq.			-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)
Age			-0.005	-0.003	-0.001
			(0.000)	(0.000)	(0.000)
Age sq.			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Rooms			0.004	0.004	0.003
			(0.000)	(0.000)	(0.000)
Bedrooms			-0.031	-0.024	-0.005
			(0.000)	(0.000)	(0.000)
Full baths			0.037	0.032	0.016
			(0.000)	(0.000)	(0.000)
Half baths			-0.012	-0.012	0.007
			(0.000)	(0.000)	(0.000)
Basement			0.098	0.097	0.041
			(0.000)	(0.000)	(0.000)
Finished Basement			0.109	0.103	0.039
			(0.000)	(0.000)	(0.000)
Condition1				0.100	0.102
				(0.004)	(0.002)
Condition2				0.252	0.224
				(0.004)	(0.002)
Condition3				0.316	0.303
				(0.004)	(0.002)
Condition4				0.364	0.361
				(0.004)	(0.002)
Condition5				0.386	0.356
				(0.004)	(0.002)
Quality1				-0.011	0.070
				(0.008)	(0.005)
Quality2				0.047	0.154
				(0.008)	(0.005)
Quality3				0.087	0.231
				(0.008)	(0.005)
Quality4				0.189	0.329
				(0.008)	(0.005)
Quality5				0.264	0.435
				(0.008)	(0.005)
Beneficial location				0.071	0.060
				(0.000)	(0.000)
Adverse location				-0.031	-0.037
				(0.000)	(0.000)
Beneficial view				0.079	0.041
				(0.000)	(0.000)
Adverse view				-0.034	-0.029
				(0.000)	(0.000)
Observations	27,302,294	27,299,513	27,299,513	27,299,513	27,273,234
Adjusted R-squared	0.007	0.714	0.908	0.918	0.984
Tract FE	N	Y	Y	Y	N
Year/Qtr FE	N	Y	Y	Y	Y
Hedonic Characteristics	N	N	Y	Y	Y
CQLV	N	N	N	Y	Y
Appraisal FE	N	N	N	N	Y

Note: This table reports estimates and standard errors, in parentheses, of a regression with log sales price as the dependent variable. The sample includes all comparable properties. Foreclosure and short sale indicators are mutually exclusive.

Table A.5. Differences in Property Condition and Price for Arm's Length Transactions that Sell as REO or short Sales in the Future

	(1) Condition	(2) ln(price)
Future Foreclosure	-0.011	-0.002
	(0.007)	(0.001)
Future Short Sale	-0.019	0.000
	(0.009)	(0.002)
Observations	26,679,037	26,679,037
Adjusted R-squared	0.672	0.984
Tract FE	N	Y
Year/Qtr FE	N	Y
Hedonic Characteristics	N	Y
CQLV	N	Y
Appraisal FE	Y	Y

Note: Column (1) reports estimates and standard errors, in parentheses, of a regression with property condition as the dependent variable. Column (2) reports estimates and standard errors, in parentheses, of a regression with log sales price as the dependent variable.The sample excludes REO and short sale transactions, as well as subsequent sales of these properties. Future foreclosure indicates an arm's length sale of a property that later sells as an REO. Future short sale indicates an arm's length sale of a property that later sells as a short sale.

Table A.6. Price Discount for Alternative Foreclosures and Short Sales Measures (Alt. 1)

VARIABLES	(1) ln(price)	(2) ln(price)	(3) ln(price)	(4) ln(price)	(5) ln(price)
Foreclosure	-0.412	-0.232	-0.200	-0.162	-0.046
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Short Sale	-0.172	-0.169	-0.179	-0.151	-0.053
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Observations	27,515,395	27,512,618	27,512,618	27,512,618	27,490,406
Adjusted R-squared	0.007	0.713	0.908	0.918	0.984
Tract FE	N	Y	Y	Y	N
Year/Qtr FE	N	Y	Y	Y	Y
Hedonic Characteristics	N	N	Y	Y	Y
CQLV	N	N	N	Y	Y
Appraisal FE	N	N	N	N	Y
Foreclosure Share	.01	.01	.01	.01	.01
Short Share	.01	.01	.01	.01	.01

Note: This table reports estimates and standard errors, in parentheses, of a regression with log sales price as the dependent variable. The sample includes all comparable properties including those where the disposition type for a transaction varies across appraisals. Alt. 1 use the disposition type that is reported in the data for each observation. Foreclosure and short sale indicators are mutually exclusive.

Table A.7. Price Discount for Alternative Foreclosures and Short Sales Measures (Alt. 2)

	(1)	(2)	(3)	(4)	(5)
VARIABLES	ln(price)	ln(price)	ln(price)	ln(price)	ln(price)
Foreclosure	-0.375	-0.207	-0.179	-0.146	-0.040
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Short Sale	-0.145	-0.145	-0.158	-0.134	-0.045
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Observations	27,515,395	27,512,618	27,512,618	27,512,618	27,490,406
Adjusted R-squared	0.007	0.713	0.908	0.917	0.984
Tract FE	N	Y	Y	Y	N
Year/Qtr FE	N	Y	Y	Y	Y
Hedonic Characteristics	N	N	Y	Y	Y
CQLV	N	N	N	Y	Y
Appraisal FE	N	N	N	N	Y
Foreclosure Share	.02	.02	.02	.02	.02
Short Share	.01	.01	.01	.01	.01

Note: This table reports estimates and standard errors, in parentheses, of a regression with log sales price as the dependent variable. The sample includes all comparable properties including those where the disposition type for a transaction varies across appraisals. For the Alt. 2 measures, if a specific transaction is ever recorded as a foreclosure (short sale), we classify that transaction as a foreclosure (short sale) for each time that it is observed in the data.