

Attribute (Mis)Reporting and Appraisal Bias ^{*}

Michael D. Eriksen^{†1}, Chun Kuang^{‡2}, and Wenyu Zhu^{§3}

¹Department of Finance and Real Estate, University of Cincinnati

²School of Banking and Finance, University of International Business and Economics

³School of Finance, Renmin University of China

Abstract

We assemble a property-level panel of appraiser reported attributes associated with 4.6 million loan applications from 2013 to 2017 to test whether attributes were consistently reported. Appraisers have an incentive to misreport property attributes to justify higher appraised values to ensure associated mortgage loans are approved. We find evidence of discrepancies in reported attributes consistent with an intention to inflate valuations, even when restricting to those recently reported by the same appraiser. Purposeful misreporting is identified in cases where appraisers who report attributes of the same comparable transaction as inferior when the associated subject property's contract price exceeds alternative value estimates. We estimate that highly leveraged borrowers whose appraisals had inconsistently reported comparable transaction attributes were 13.3-14.3% more likely to become seriously delinquent in their loan payments.

Keywords: appraisal, real estate market, mortgage, collateral valuation
JEL Classification: R31, G21, G28, K1

^{*}We are grateful to Bob Barclay and Eric Rosenblatt for their generous assistance in obtaining and working with the data. We thank Tony Yezer and seminar participants at the Federal Reserve Bank of Cleveland, the University of Louisville, and Purdue University for their comments. All errors are our own.

[†]Corresponding Author. 2906 Woodside Drive, Cincinnati, OH 45221, USA (mike.eriksen@uc.edu)

[‡]10 Huixin E Street, Beijing, China (chunkuang@uibe.edu.cn)

[§]59 Zhongguancun Avenue, Beijing, China (wenyuzhu@ruc.edu.cn)

1. Introduction

The 2008 financial crisis illustrated the importance of unbiased valuation of financial assets, especially collateral associated with securitized mortgage loans.¹ Previous research has shown that asymmetric information among buyers, lenders, and regulators resulted in multiple errors in valuation and subsequently higher than expected loan defaults catalyzing the crisis (Ben-David, 2011; Griffin and Maturana, 2016). In response, subsequent regulations have attempted to reform the valuation production process by implementing stricter rules for information acquisition and disclosure in mortgage markets. It remains unclear whether such new regulations improve information quality and reduce the likelihood of future crises (Bongaerts et al., 2012). In this paper, we identify a new channel through which appraisal inflation occurs and provide evidence that current financial reforms have not eliminated bias in the disclosed collateral values associated with mortgage loans.

Similar to credit rating agencies, appraisers of residential property are crucial in mitigating information asymmetries in the mortgage market by providing an independent assessment of value. An independent assessment of value is important because prior research has shown that more informed lenders make superior decisions, and that independence enhances the liquidity of originated loans on the secondary mortgage market (Stroebel, 2016; Eriksen et al., 2019). It is well established that appraised values before the 2008 crisis were routinely not independent and often severely biased upwards as compared to alternative Automated Valuation Model

¹See Griffin et al. (2020) for a review of the contributing factors of the 2008 financial crisis.

(AVM) estimates (Kruger and Maturana, 2021). Moreover, it remains unclear how bias was introduced, and whether such channels remain open after reforms were implemented following the 2008 financial crisis. In particular, the Home Valuation Code of Conduct (HVCC) in 2009 and further appraisal-specific reforms within the Dodd-Frank Act of 2010 were implemented with the specific intent to limit previously identified financial conflicts of interest between lenders and property appraisers, but arguably did little to address how appraised values were actually manipulated. Better understanding of the motivations, mechanisms, and harm of appraisal bias is important in both designing further reforms and assessing the degree to which the economy remains exposed to future crises.

We test for the use of a specific mechanism that appraisers might employ to manipulate valuations, namely misreporting property attributes used to anchor their estimates. This manipulation can occur through an appraiser reporting either a subject property (i.e., the property being appraised) as more desirable than reality, or decreasing the relative appeal of a selected comparable transaction (or “comp” for short) to justify a larger than otherwise price adjustment. Virtually all appraisers of residential properties use the sales-comparable approach to valuation which requires their explicit statement of attributes of both the subject and selected comps to make necessary price adjustments. Common differences in property attributes used by appraisers to justify subject properties transacting for more or less than recent comparable transactions include subjective (e.g., property condition) and objective (e.g., bedrooms, living area) elements. Finding some appraisers not only misreport attributes,

but doing so in a strategic fashion to increase valuations of properties that may not appraise otherwise would be the strongest evidence of purposeful manipulation in the literature to date.

Such analysis is possible because residential appraisers have been required to report property and other attributes of both subject and selected comparable transactions to financial institutions originating conforming mortgage loans since 2011. We assemble a property-level panel with a unique identifier assigned by a secondary market purchaser.² While we find at least one discrepancy occurred in the 7 most commonly reported attributes (i.e., above-ground gross living area (GLA), bedrooms, bathrooms, lot size, property condition, quality of design, and view) for 94.55% of comparable transaction compared to when originally reported as a subject, it is important to recognize that a significant share of the reporting discrepancies are economically small and can be explained by either rounding discretion, or differences in interpretation of or access to data on comparable transactions by different appraisers. While appraisers of the property when originally a subject are required to visit the property to conduct their own measurements and assessment of attributes, appraisers reporting attributes only as a comp rely on external sources of data. This is especially true since the sales-comparison method of valuation requires appraisers to report and anchor their estimates based on the attributes as of the specific date when the comparable property itself transacted in the past (i.e., a comparable *transaction*).³ Known sources of data on comp

²We describe in detail the sample construction process in Section 3 and the Appendix.

³Empirically, each comparable transaction is identified by the combination of the unique property ID and

attributes include other appraisers, property tax assessors, and real estate brokers as reported in the local multiple listing service (MLS) when the comparable property was originally listed for sale.

Given the differential access to information about comparable transactions, the majority of our analysis focuses on two subsets of the sample when the same appraisers report attributes of the same properties more than once. The first restricted sample includes 644,525 comparable transactions of 494,225 unique properties when the original appraiser reporting attributes as a subject subsequently reused the resulting sale as a comp and reported a new set of attributes for the same property when appraising a different subject in the future. We first examine attribute reporting consistency through the subject-comp comparisons of the same transaction reported by the same appraiser. Although occasional and sometimes economically large discrepancies in reported attributes are observed, there is limited evidence that appraisers strategically over-report the quality of subject properties to inflate valuations. More specifically, at least some of the observed subject-comp reporting discrepancies by the same appraiser could be explained by appraisers using external data for attributes of all comparable transactions regardless if they appraised the original transaction as a subject.

Our strongest evidence of purposeful misreporting is found when we limit the sample to the 1.25 million properties where an appraiser used the same comparable transaction at least twice to value different properties. This restriction controls for potential discrepancies due to

transaction quarter. See the appendix for more information about how the data are assembled.

measurement error, differences in interpretation and purpose, as appraisers rely exclusively on external data to acquire property attributes to base valuations. While appraisers are most likely to report attributes consistently for repeated comparable transactions, occasional and often economically large differences persisted when discrepancies did occur.

Most importantly, we find evidence that appraisers are more likely to under-report attributes of the same comparable transaction selectively when the contract price of the subject property exceeded the alternative AVM estimate, suggesting the appraised value would not have met contract price without such manipulation.⁴ This result is robust to restricting the sample comparable transactions reported multiple times by the same appraiser in the same quarter, and does not appear sensitive to the number of potential comparable transactions in the immediate area or to the local housing market conditions. We also show that more than 80% of appraisers made at least 1 significant under-reporting error that inflated the appraised value of subject properties, and more than 30% made such an error for at least 20% of their appraisals.⁵

We explore the economic consequences of attribute misreporting in Section 6. This analysis focuses on the performance of 2,992,969 loans originated between 2013 and 2017 where at least one comparable transaction was used multiple times by the same appraiser

⁴The AVMs used in the analyses are part of the dataset provided by our data provider, and are created using their own valuation models and internal data.

⁵A significant reporting error is defined as at least a 100 square foot discrepancy in gross living area or lot size, or a different value reported for property condition, design quality, bedrooms, full bathrooms, or view.

to value other properties. Controlling for lender, appraiser, and market-year-quarter fixed effects as well as borrowers' credit worthiness and loan term, we estimate that borrowers of highly-leveraged loans associated with an under-reported comp were 13.33-to-14.29% more likely to become seriously delinquent in their loan payments by the end of 2018. Taken together, our findings suggest that there are real economic consequences of appraisal bias due to attribute misreporting in mortgage appraisals, especially given the negative consequences of foreclosures on borrowers and their neighbors (Mian and Sufi (2017)). For those who misreport property attributes in order to manipulate appraisals, not only have they failed their obligation to provide independent and accurate valuation of the collateral, but they have further increased the systemic risks in the market by inflating the collateral values for the most financially constrained borrowers purchasing over-valued properties.⁶

We make two main contributions to the literature. First, we document a specific channel through which appraisal bias occurs, and provide strong empirical evidence of intentional manipulation by at least some appraisers to inflate property valuations. Second, we show a real economic consequence of appraisal bias in that borrowers of highly-leveraged loans associated with attribute misreporting were more likely to become delinquent in their loan payments. Such evidence of purposeful manipulation remains pervasive even after reforms were

⁶Even if original mortgage lenders are able to capture some of the higher expected default risks due to valuation bias through differential loan pricing, understanding the linkage between appraisal bias and loan performance is still crucial to policymakers and mortgage practitioners. We further discuss these consequences in more detail in the conclusion.

enacted after the 2008 financial crisis suggest that further oversight and policy innovation is necessary to curb future crises. The paper concludes with policy recommendations, including calling for the creation of a nationwide database of reported property attributes to promote accountability.

2. Background and Recent Research

The purpose of an appraisal is to provide an independent assessment of the stated value of the collateral that a borrower pledges to the mortgage lender. The Real Estate Appraisal Reform Act of 1988 regulates appraisals for real estate loans originated by federally regulated financial institutions. The appraisal provides crucial information for mortgage lenders and regulators to assess a borrower's equity stake in the property and the resulting default risk of underlying loans (Ding and Nakamura, 2016).

A growing body of literature has illustrated the presence of appraisal bias using a variety of methods. The most common approach has been to compare the appraised values of properties with an alternative value calculated by an Automated Valuation Model, where the percentage difference between the two estimates is used as a measure of appraisal bias. Kruger and Maturana (2021) find that 60% of mortgages have appraisals higher than the AVM valuations and Griffin and Maturana (2016) further document that 44.9% of properties have appraisals that are 5% higher than estimates by AVMs. Demiroglu and James (2016) also find that 19.7% of refinance loans have overstated appraisals compared to the AVM

valuations, much higher than that of purchase loans.

An important limitation when comparing appraised values and AVM estimates is that current property-specific data available to calibrate AVMs are often limited. This is especially true for AVM estimates using publicly available records, such as local property tax assessors databases, as information about the property may only be available based on that reported from the previous sale. Major and minor renovations of the property, especially those commonly done before a sale, would create unobserved heterogeneity and result in AVM estimates themselves being biased. In contrast, residential appraisers are required to visit the property being appraised and are aware of the current state of the property due to recent renovations and other changes otherwise undocumented in public records.⁷

Three key questions remain unresolved in the literature. The first is whether appraisal bias persists after reforms were implemented following the 2008 financial crisis. The HVCC in 2009 and subsequent reforms further codified in the Dodd-Frank Act of 2010 were implemented with the intent to limit known conflicts of interest between lenders and appraisers. Several studies have found the HVCC and other reforms to be effective in reducing the share of appraisals exactly confirming contract price, as well as other outcomes correlated to appraisal bias, but the level of residual bias remains unclear (Agarwal et al., 2015; Calem et al., 2015;

⁷Two recent examples addressing these data limitations of public records include Agarwal et al. (2015) in comparing appraised values with future transaction prices, and Eriksen et al. (2019) showing that appraisers still target contract prices even after adjusting for their own observed differences in property attributes during reconciliation.

Ding and Nakamura, 2016).⁸

The second question to be answered is whether the presence of biased appraisals is a result of a purposeful act, or an unintentional consequence of current appraisal practices. Title XI of the Financial Institutions Reform, Recovery, and Enforcement Act of 1989 empowers the Appraiser Qualifications Board to establish the minimum licensure requirements for an appraiser in each state and requires appraisers to follow the Universal Standards of Professional Appraisal Practice (USPAP) to justify their valuation estimate. Virtually all appraisers of residential properties use the sales-comparison approach to valuation, which requires them to define attributes of the subject property, select comparable transactions in the last 12 months of similar properties, adjust for observable differences between the subject property and comparable transactions, and then weight adjusted values of comps to reach a final opinion of property value (Vandell, 1991). While financial conflicts of interest have been identified as a motivating force for some appraisers to bias their estimates, that evidence is largely circumstantial. The exact mechanisms through which ill-intentioned appraisers introduce bias remains largely unidentified (Eriksen et al., 2019). This is important because a significant share of biased values identified in the previous literature could be the result of well-intentioned appraisers following flawed appraisal practices.⁹

⁸For example, Ding and Nakamura (2016) provide empirical evidence that the HVCC led to a 3% increased probability of appraisal values being less than contract price, suggesting they were at least partially effective.

⁹ For example, appraisers of properties associated with home purchase loans are provided the contract price upfront and may subconsciously select evidence to confirm that value (Eriksen et al., 2020). More

The last unresolved question is whether appraisal bias itself results in decreased loan performance and subsequent economic losses for mortgage investors. The main constraint in the previous literature is that loans associated with biased appraised values before the 2008 financial crisis were mostly likely deficient along multiple other dimensions. [Griffin \(2021\)](#) provides a review of the literature and argues that overstated incomes and the presence of subordinate mortgage debt were prevalent in the period immediately prior to the 2008 financial crisis. The effect of these unobserved loan and borrower attributes associated with the cohort have made it difficult to separately identify the effect of appraisal bias on loan performance outcomes.

3. Data

The empirical analysis in this paper is based on property attributes reported by appraisers to financial institutions to support their valuation estimate of collateral associated with conforming mortgage loans as required by the Federal Housing Finance Agency since 2011. Reported attributes include both objective (i.e., gross living area, lot size, bedrooms, bathrooms) and subjective measures (i.e., condition of property, quality of design, and view) of

recently, [Agarwal et al. \(2020\)](#) suggest that decreasing returns to effort given current incentives may also result in similar unintentional biases.

the properties.¹⁰ These attributes could be reported for the same property more than once – initially when it was the subject of an appraisal associated with loan approval process and subsequently when used as a comparable transaction to value other properties. We are able to construct a property-level panel of reported attributes because the large secondary market purchaser of mortgage loans from whom we obtain our data assigned a unique identification number to each property and reporting appraiser. Properties are included in the sample if either reported as a subject of a pending loan application, or used as a comparable transaction to support the value estimate of another property between 2013 and 2017.¹¹

Three main restrictions are applied when constructing the sample. First, we restrict the sample to properties with at least two sets of reported attributes. Attributes could be reported as a subject or a comp by the same or a different appraiser. Second, only the initial sale of properties that transacted multiple times in the sample period is retained.¹² For similar reasons, properties transacting twice within the same quarter are removed from the sample.¹³ Last, we only retain reported attributes of comparable transactions if made within one year of the associated sale to guard against possible post-transaction property

¹⁰Condition of property and quality of design are reported on a 6-point scale, while view is reported on a 3-point scale.

¹¹Properties exist in our data even if the associated mortgage loan was not originated, or sold to the secondary market institution providing the data, as attributes were captured during the loan approval process.

¹²Although comparison of repeated transactions of the same property are interesting, it is impossible for us to separately identify potential modifications made to the property between transactions.

¹³Approximately, 5.9% of properties (431,106) were excluded as a result of these restrictions.

improvements being reflected in tax assessor records. The resulting sample includes 4,594,672 unique properties. Consistent with GSE-published guidelines requiring appraisers to use comparable transactions within the last 12 months without extenuating circumstances, Figure 1 illustrates that 99.5% of all selected comps are within 4 quarters of the original transaction.

Panel A of Table 1 illustrates summary attributes of properties that were originally the subject of a pending home purchase appraisal, and used at least once as a comparable transaction in a future appraisal. This restriction results in 23.5 million comparable transactions based on 4.6 million unique properties (subject transactions) from 2013Q1 to 2017Q4. On average, each subject property has been used as a comparable transaction in 5 subsequent appraisals. Column (2) in Panel A of Table 1 shows that the median property has 1,559 square feet of GLA, 3 bedrooms, 2 full bathrooms, and is situated on a 9,309-square-foot lot. In addition to reporting objective attributes of the property, appraisers are also required to indicate subjective attributes, including the condition of the property, quality of the design, and the view of the property based on uniform definitions set by USPAP.¹⁴ Column (1) of Table 1 reports the median price adjustment made by appraisers per unit of reported difference between subject and comparable transaction for each attribute .

Given attributes of the same transaction as a subject and as a comp may be reported

¹⁴Appraisers report on a 6-point scale the condition and quality of the property based on specific attributes and features. View is reported on a 3-point scale: (1) property does not have a view, (2) property has some view, and (3) property has a view. For ease of exposition, all subjective attributes were recoded such that a greater value indicates a superior feature.

by different appraisers, some discrepancies are expected due to differences in interpretation, methods of measurement, and access to information. A comparable transaction is considered consistently reported when the attributes reported are the same as when itself was the subject of an appraisal. Column (1) of Panel A of Table 1 illustrates the percent of comps with consistently reported attributes for each of the 7 most commonly reported features.¹⁵ GLA was the least consistently reported attribute with only 19.78% of comparable transactions having it consistently reported compared to when the subject of an appraisal with an average discrepancy of 109.98 square feet. This discrepancy alone accounts for a difference in value of \$25,515, given the average price adjustment per square foot difference in GLA across the sample was \$232 as displayed in column (1).

Property condition (64.87%), design quality (58.10%), and lot size (60.48%) had similar discrepancy rates, whereas view (88.67%), bedrooms (91.44%), and bathrooms (95.48%) were the most consistently reported. Only 5.45% of all subject-comp pairs had all 7 attributes reported consistently. Multiplying the average reporting discrepancy in column (4) by the average price adjustment reported in column (1) for each of the 7 attributes results in a variation in value of \$67,933 in potential adjustments, amounting to approximately 20% of the average appraised value of \$335,782.

¹⁵Reporting consistencies are tabulated at the comp level, rather than at the property level. In other words, the sample used to calculate the share of reporting (in)consistencies includes only comps without the original transaction when themselves were the subject of an appraisal.

While the above differences are sizable and could lead to large differences in adjusted prices for the same property, it is important to recognize they are not necessarily reflective of purposeful misreporting. The original appraiser of the property is required by USPAP to visit the property to either make or verify the measurements and assessments of it. In contrast, an appraiser only using it as a comparable transaction would have not access to the property, and therefore must rely on measurements, assessments, and interpretation of attributes available in other data sources. Potential sources of property information used by appraisers for comps include the original appraiser when a subject, property tax assessors, and real estate brokers when the comparable property was originally listed for sale as recorded in the local MLS.¹⁶ In the next section, we focus on subject-comp pairs of attributes reported by the *same appraiser* to help answer these questions.

4. Subject-Comparable Analyses

Appraisers can potentially manipulate the appraised value of a subject property by misreporting attributes of either the subject property as superior (i.e., over-reporting the subject), or by misreporting the attributes of the comparable transaction as inferior (i.e., under-reporting

¹⁶Despite the clear financial incentive for real estate brokers to inflate the attractiveness of a property, using the MLS is especially appealing for appraisers because the listings often include additional features and pictures of the property to allow assessment of subjective qualities at the time of transaction, which is often not available from using property tax records alone.

a comp). In this section, we focus on the magnitude of the subject-comp gaps in attributes reported for the same transaction.

The most obvious explanation for why discrepancies exist in reported attributes is that different appraisers have different access to information about the property. Therefore, we start by focusing on attributes reported for comparable transactions by the same appraiser who originally appraised the same property as a subject, and then generalize to the larger unrestricted sample later in the section.

Results in Table 2 indicate that while significant reporting discrepancies occurred for attributes reported for the same property as a subject and a comp, appraisers did not uniformly report the subject as superior. Focusing on Panel B for properties with discrepancies reported by the same appraiser, 61.1% of subject-comp pairs had the subject GLA reported as higher, compared to 38.9% of observations had the comp reported as larger. The grey bars in the top panel of Figure 2 illustrate the distribution of differences in GLA when reported as a subject minus as a comparable transaction by the same appraiser. While at least some of the discrepancies could be attributable to small differences in rounding, the top panel of Figure 2 shows that, for 11.9% of comps, the same appraiser reported the subject as having at least 100 additional square feet of GLA, and more than 200 square feet for 6.5% of comps. For GLA alone, these discrepancies would result in an economically large average price adjustment of \$23,200, and \$46,400, or approximately 7.1% and 14.2% of the average appraised value in the sample, respectively.

Total price adjustments for some properties can be even larger if multiple reporting discrepancies occurred for the same property. Conditional upon a discrepancy, appraisers were also more likely to report the same property when a subject as superior for condition (58.6%), design quality (58.2%), view (57.2%), and lot size (54.6%). The bottom panels of Figure 2 illustrate the distribution of errors for condition, design quality, and view. While the majority of the reporting errors for condition and design quality were within 1 point of the 6 point scale, these seemingly slight errors still resulted in an additional sizable price adjustment of \$12,500 and \$11,500, respectively. Finally, the same appraiser reported 4.1% of subjects as having a view (i.e., view = 3) when they reported the same property had no view as a comp (i.e., view = 1), resulting in an additional price adjustment of \$15,000.

However, appraisers do not appear to exclusively report subjects a superior as expected if purposeful misreporting to inflate subject's appraised value was the only explanation for discrepancies. In addition to bedrooms (41.1%) and bathrooms (41.3%) being less likely to be misreported as superior for the same property as a subject, some of the negative errors for the other reported attributes were themselves economically large. For example, 7.9% of subject properties were reported as having 100 fewer square feet of GLA than when reported in the future as a comp by the same appraiser.

Potential reasons why such a high percentage of comps are reported as superior than subjects are explored further in Section 4.3. The remainder of this section is dedicated to measuring the average effect of over-reporting attributes of subjects on appraised values, and

whether such over-reporting occurs strategically.

4.1. Average Subject-Comp Gaps in Reported Attributes

Table 3 reports the average subject-comp gaps in reported attributes. The averages were estimated using all properties used at least once as a subject and a comparable transaction by the same appraiser:

$$A_{ijk} = \beta * Subject_{ijk} + \mu_{ij} + \epsilon_{ijk}, \quad (1)$$

where A_{ijk} is the value of one of the attributes reported for property i by appraiser j in appraisal k , $Subject_{ijk}$ is an indicator for the property being the subject of appraisal k , μ_{ij} is a vector of unique transaction-appraiser pair fixed effects, and ϵ_{ijk} is a mean zero idiosyncratic error term. The estimated coefficient on $Subject_{ijk}$ therefore captures the average difference in attributes reported for the same transaction when the subject of an appraisal, compared to when reported as a comp subsequently. Potential renovations to the comparable property between appraisals should remain irrelevant as the identifying requirement of the sales-comparison method of valuation is to report attributes of the comparable property at the time of its previous transaction.

Panel B of Table 3 reports the estimates for GLA, condition, quality, and view with two-way clustered standard errors based on unique transactions and appraisers reported in parentheses.¹⁷ The sample of Panel B includes the 1,138,750 sets of reported attributes for the

¹⁷We focus on gross living area and the 3 subjective attributes for the remainder of the paper given these

494,225 subject properties used by the same appraiser as a comp at least once. The estimates in column (1) of Panel B indicate the same appraiser reporting the same property as having an additional 5.1 square feet of living area when the property is a subject compared when reported as a comp subsequently. The dependent variables of the 3 subjective attributes are log-transformed to allow for interpretation as percent differences. Accordingly, an appraiser is estimated to report the same property when a subject to be superior in condition (0.5%), quality (0.4%), and view (0.2%). While statistically significant, the estimates reported in Panel B of Table 3 are not economically large. Given the median price adjustment per unit difference used by appraiser reported in Table 1, these 4 combined estimates correspond to subject properties being overvalued on average by \$1,608.

Panel A of Table 3 presents estimates when relaxing the prior same-appraiser restriction, and thus expands the sample to 28,096,842 sets of reported attributes for 4,595,092 subject properties used at least once as a comp. Panels A and B of Table 1 and Figure 2 show that the distributions of discrepancies for the two samples are almost identical with appraisers reporting when the subject as slightly more superior than as a comp. The average estimated over-statement of GLA for subjects for the larger sample was 11.6 square feet, with condition (1.2%), quality (0.8%), and view (0.9%) - more than double of the more restricted estimates. While the combined average price adjustments increases to \$3,682 and approximately 1% of average appraised values, at least half is explained by appraisers' differential access to subject attributes have the greatest reporting discrepancies.

attributes.

4.2. Potential Explanations for Observed Differences

The possibility that differences in reported attributes might reflect actual changes in features of the subject property when reported again as a comparable transaction is, in our opinion, the gravest threat to our identification strategy. However, appraisers are required by USPAP to report and use attributes as of the date of the comparable sale (t_0), and are instructed to disregard any property changes of the comparable property between its own transaction date (t_0) and the date of current appraisal it is being used as a comp (t). Put differently, even if the appraisers are in a position to report the update-to-date attributes of the comparable property, they should not, per the sales-comparable method of valuation. This is because doing so would create a disconnection between the transaction value of the comp and the attributes supporting it. Also, assuming that most subsequent changes to property attributes made immediately after a transaction are value-enhancing, this would result in an appraiser appearing to report the same property as superior as a comp - the opposite of what we find.

We guard against this possibility by restricting the sample to comparable transactions reported within 4 quarters of the original transaction as described in Section 3.¹⁸ There

¹⁸With the restriction of no more than 1 year between the subject and comparable transactions, 154,712 and 7,073 observations (or 0.55% and 0.62% of the samples) are deleted from the unrestricted and the restricted same-appraiser samples, respectively, confirming that only a very small portion of comparable transactions used were of sales more than 1 year old.

would presumably be some lag before actual changes in property features were reflected in county property tax assessors records, and this 4 quarter restriction most likely captures the majority of potential changes were they to occur. We also show that the discrepancy rate for the same-appraiser sample does not increase the longer the gap between attributes reported as a comp and original transaction.¹⁹ Our strongest evidence is presented in Panel C of Table 3, where Equation (1) is re-estimated using the 277,317 instances when the same appraiser that appraised the property used it as a comp in the *same quarter*. Although reduced in magnitude, the subject-comp gaps in reported attributes remain significant and positive for GLA (from 5.1 to 4.2 square feet) and the 3 subjective attributes. Under the assumption that physical attributes are unlikely to instantly change during the same quarter when the property transacted, these results suggest that the observed subject-comp gaps in reported attributes do not appear to be due to actual property changes.

4.3. Strategic Over-Reporting of Subject Attributes

While the above analyses documented robust subject-comp gaps in reported attributes on average, it remains unclear whether these observed gaps are due to strategic misreporting. In particular, we next investigate whether appraisers inflate attributes of subject properties that

¹⁹Figure A4 in the Appendix plots the percent of comps with an inconsistently reported GLA by the number of quarters between its own transaction (t_0) and the reuse as a comp (t) by the same appraiser. A larger percent of discrepancies would occur as the number of quarters increase if explained by actual changes in property attributes and records, which we do not find.

would otherwise not have been appraised above the contract price. As a matter of practice, appraisers are always informed upfront of the contract price negotiated by the buyer and seller (Conklin et al., 2020). An appraised value below contract price jeopardizes a loan origination because either the seller would have to accept a lower price than originally agreed, or the buyer would have to contribute a larger amount of equity to purchase the property (Ben-David, 2011). A failed origination would therefore have a direct financial consequence for the loan officer and real estate broker(s) associated with the transaction, who may themselves apply pressure on appraisers to justify an appraised value at least equal to the contract price (Eriksen et al., 2019). As the unbiased appraised value is unobserved, we proxy for the risk an appraised value might be below contract price by using the AVM estimate of the subject property by the financial institution providing our data. Of the 494,225 subject properties with attributes also reported as a comp by the same appraiser, more than half (52.3%) had a contract price above the AVM estimate, and 17.5% had contract prices more than 10% above.

If the observed positive subject-comp gaps in reported attributes are indeed due to strategic over-reporting of the subject properties, the share of comps with a positive subject-comp gap should increase as the associated subject properties' contract prices increase relative to the AVM estimates. We examine this hypothesis in Figure 3, where the x-axis is the % the contract price of subject is above AVM in 5 percentage point (pp) bins, and the y-axis is the percent of comps where GLA is reported larger as a subject than as a comp by same appraiser. Of subjects with a contract price below AVM and thus most likely to have an

appraised value above contract without manipulation, 21.5% had a GLA reported larger as a subject than as a comp. This percent decreased to 20.3% if the contract price was between 0 and 5% above AVM, 18.7% if between 5% and 10%, and was the lowest at 15.5% if the contract price was more than 10% above AVM estimate. Contrary to the strategic over-reporting of subject hypothesis, these estimates imply that appraisers were actually less likely to over-report GLA for properties at greatest risk for an appraised value below contract price. Similar patterns are repeated for condition, quality, and view as presented in the other panels of Figure 3. While over-reporting of subject attributes might occur for some appraisers, the majority of the observed positive subject-comp gaps in reported attributes do not seem to be strategic, indicating that it might be an artifact of something else.

A potential explanation for the above pattern is that appraisers were themselves aware of AVM estimates and that subject properties with a contract price above that alternative value were the most likely to be scrutinized. This additional scrutiny could be by either borrowers and related parties upset for having an appraised value below contract price, or by regulators attempting to detect their bias. For example, appraisers might only misreport attributes of properties otherwise close to having an appraised value equal to contract price without manipulation and resort to other ways to introduce bias (e.g., cherry-picking comps, using unequal weights during reconciliation) as relatively more attractive given their subjective nature when larger adjustments are required. In the next section, we provide evidence of a previously unexplored channel used by at least some appraisers to inflate values of subjects

that is similarly difficult to detect.

5. Repeated-Comparable Analyses

The results presented in the previous section indicate systematic misreporting by at least some appraisers consistent with an intention to inflate property valuations. However, it remains unclear whether such misreporting is intentional, as the observed discrepancies could be a result of differences in appraisers' practices in data collection for subject properties versus comparable transactions. After all, appraisers may create and report assessment of the subject property attributes gathered during their own inspection, but always rely on property tax records for all comp attributes. To eliminate this possibility, we next focus exclusively on properties selected multiple times by the same appraiser as a comparable transaction to value different properties (i.e., a repeated-comp sample).

Restricting the sample to comparable transactions used at least twice by the same appraiser is advantageous as it mechanically eliminates the potential for differences in appraisers' assessment approaches to explain misreporting. The exercise also allows us to better understand appraisers' motivations by testing how reporting consistency varies with attributes of a unrelated subject property. Misreporting attributes of a comp as inferior (i.e., under-reporting) would result in a larger price adjustment and higher appraised value of the subject property than otherwise.

Column (1) of Table 4 presents the average attributes of the 1,249,149 comparable

transactions used at least twice by the same appraiser. The average comparable transaction in the sub-sample was used 2.3 times and reported to be slightly larger with a GLA of 1,977.5 square feet, but otherwise similar to the other samples in terms of reported attributes. As expected, appraisers were the most likely to consistently report property attributes of repeated comparable transactions with 81.1% having no discrepancies for all 7 attributes. Appraisers were the most likely to report the number of full bathrooms consistently (99.0%), and least likely to consistently report property condition (91.8%), design quality (94.3%), lot size (96.2%), and GLA (96.6%).

When a discrepancy in reported attributes exists for the same comparable transaction, it is unclear whether the higher- or lower-value of the pair represents reality. We define comps with the lower-indicated-values by the same appraiser as being under-reported. Given the paired relationship of repeated-comps, approximately half of the discrepancies reported for each attribute in column (3) of Table 4 are therefore defined as being under-reported. Appraisers were most likely to under-report property condition (4.1%) and design quality (2.9%) and least likely to under-report bedrooms (0.7%) and full bathrooms (0.5%). For GLA, 1.7% of all repeated-comps had an under-reporting discrepancy.

Columns (4) and (5) of Table 4 report the median and average under-reporting errors when a discrepancy occurs. For the 1.7% of repeated-comps when GLA was under-reported, the median under-reported comp was indicated as having 85 fewer square feet than what the appraiser reported for the same comparable transaction on a different appraisal. This

translates into an additional \$19,720 (6.1%) for the appraised value of the subject property, compared with when the appraiser reported GLA consistently. The average under-reporting errors were more than double the median, and can be explained by appraisers selectively including unfinished ancillary living areas (e.g., basements, attics, and garages) in GLA for comps in some appraisals, while excluding it from others.

The above patterns would not reflect purposeful under-reporting of comps if the external data source used by appraisers to obtain attributes of comparable transactions reflect actual updates to property after the original transaction. Figure 4 explores this possibility and suggests that it explains a relatively small share of under-reporting at best. That figure illustrates the number of quarters between when the maximum and minimum GLA was reported for comps with discrepancies. If under-reporting was indeed due to actual property improvements and records used by appraisers being updated accordingly, the distribution is likely to be skewed to the right. Instead, the distribution is almost symmetrical with the greatest share of under-reporting (29.9%) occurring when appraisers reported different attributes of the same comparable transaction on different appraisals during the same quarter. 89.8% of discrepancies occurring within plus or minus 2 quarters. Such short time period between the majority of discrepancies makes it unlikely that record-updating alone is sufficient to explain the observed under-reporting.

5.1. Strategic Under-Reporting of Comps

As curious as it is that reporting discrepancies persist in the repeated-comp sample, it remains unclear whether appraisers were more likely to under-report comps strategically to inflate the appraised value of subject property that would otherwise have been appraised below contract price without manipulation. Similar to Section 4.3, we proxy for the risk an appraised value might be below contract price by using the AVM estimate of the unrelated subject property. Unlike Section 4.3, the subject for each repeated comp are different properties the appraiser was hired to provide an opinion of value of collateral associated with a pending loan application in the future. Our hypothesis is that appraisers are more likely to under-report the attributes of a comp when the contract price of the subject property exceeds the AVM estimate.

We investigate such strategic under-reporting of comp attributes using the following empirical model:

$$A_{ijkt} = \theta * I(Contract > AVM)_k + \sigma_{ijt} + \epsilon_{ijkt}, \quad (2)$$

where A_{ijkt} is the value of one of the attributes reported for transaction i by appraiser j in appraisal k in quarter t , $I(Contract > AVM)_k$ is an indicator for the associated subject property in appraisal k having a contract price higher than the AVM estimate. The reference group in this model, therefore, are comparable transactions whose associated subject property has a contract prices less than or equal to the corresponding AVM estimate.

Table 5 report the results. Panel A is estimated with transaction-appraiser fixed effects (σ_{ij}) and Panel B transaction-appraiser-quarter fixed effects (σ_{ijt}). Estimates in Panel A indicate that an appraiser on average reported the same comp as having 3.9 fewer square feet of GLA, and an inferior condition (-0.9%), quality (-0.6%), and view (-0.2%) if the subject property's contract price exceeded the AVM estimate. Each of these estimates is statistically significant based on clustered standard errors robust to potential sources of non-independence within appraisal, appraiser, comparable transaction, and quarter. These estimated differences account for an additional \$1,622 in appraised value of all subject properties associated with repeated-comps with a contract price above AVM. These estimated differences are in addition to under-reporting that happens independent of contract price and can be substantially higher for some individual properties when under-reporting occurs.

The inclusion of transaction-appraiser-quarter fixed effects in Panel B of Table 5 controls for potential changes in property records accessed by appraisers to explain the observed reporting discrepancies, and are, therefore, our strongest evidence of purposeful misreporting of comps by appraisers in order to inflate subject appraised values. The estimation sample in Panel B is however significantly smaller, as the same appraiser must have multiple sets of reported attributes for the same comp in the same quarter to be included. While the comp reporting discrepancies presented in Panel B are approximately 8.3-25.0% smaller than those in Panel A, they nonetheless indicate that the same appraiser in the same quarter on average reports the same comp to be significantly inferior when there is an incentive to inflate the

subject valuation.

5.2. The Extensive and Intensive Margins of Comp Under-Reporting

If the under-reporting of comp attributes are indeed strategic, we expect both the likelihood of comp under-reporting and the magnitude of reporting discrepancy to be increasing with the difference between the contract price and the AVM estimate. Panel A of Figure 5 illustrates the share of all repeated-comps with GLA under-reported as the percent differences between the subject's contract price and the AVM estimate vary in 5pp bins. Of the repeated-comp sample, 47.1% of the subjects had a contract price above the AVM estimate, and 15.0% had a contract price more than 10% above. A monotonic pattern is evident across the panels. While only 1.2% of subject properties with a contract price more than 10% below the AVM estimate have an under-reported comp, this share is 1.9% for subject properties with a contract price exceeding the AVM estimate by more than 10%. This 0.7pp difference represents a 58.3% increase in likelihood of a comp being under-reported.

The pattern is repeated in Panels B-D of Figure 5 for property condition, design quality, and view. For each of these attributes, the shares of comp under-reporting were relatively constant when the subject's contract price is below the AVM estimate, and increased significantly as the contract price exceeded AVM. Comparing subject properties with a contract price either below or above 10% of the contract price, the likelihood of under-reporting comps increased for property condition by 2.1pp (72.7%), design quality by 1.5pp (71.1%) and view by 0.2pp

(27.2%).

Figure 6 summarizes the median and average discrepancies by which GLA is under-reported conditional upon a GLA discrepancy for each of the 5pp bins of the subject's contract price relative to AVM estimate. The bars indicates that, in addition to the likelihood increasing, the magnitude of errors also increase with the contract price relative to the AVM. A median error of 98 square feet when the contract price exceeded the AVM estimate by no more than 5% increases to 122 square feet when contract price was more than 10% above the AVM estimate. These errors correspond with an appraiser justifying an additional \$22,736 (6.9%) and \$28,304 (8.6%) in price adjustments based on the same comp. Average under-reporting discrepancies were more than double the median errors, and suggest that appraisers were more likely to selectively exclude some large ancillary living areas from GLA as subject contract price increases relative to the AVM. These exclusions are, however, difficult to detect without a detailed breakdown of how GLA was determined for each comparable transaction.

6. The Consequences of Attribute Misreporting

Recent studies have shown that properties with an appraised value greater than alternative AVM estimates were more likely to default during the Great Recession. However, [Griffin \(2021\)](#) and others have shown that in addition to appraiser bias this cohort of loans were also poorly documented with often missing and overstated borrower incomes, lower credit

scores, and a common presence of unknown simultaneously originated second mortgage loans. The joint occurrence of the above factors makes it difficult to identify the true effect specific to appraiser introduced valuation biases (Ding and Nakamura, 2016; Kruger and Maturana, 2021). Significant reforms have subsequently been enacted to improve reporting and income documentation, but we have shown in previous sections that purposeful appraisal bias persists at least through property attribute misreporting. In this section, we illustrate that, controlling for differences in borrowers' credit worthiness, lenders' underwriting standards, and local market conditions, attribute misreporting is still directly associated with borrowers being more likely to become seriously delinquent in their loan payments.

We test the linkage between comp under-reporting and the subsequent loan performance of the subject property. Inspired by our previous findings, we focus particularly on the under-reporting of GLA in this section, and test whether there is a systemic and significant connection between the misreporting of comp attributes and the subsequent loan performance. The analysis is carried out using the repeated-comparable sample, where the sample is further restricted to the 2,992,969 loans originated and eventually purchased by the secondary market financial institution providing our data. Table 6 presents summary statistics on the loan-level attributes of our estimation samples for the delinquency analyses, where Panel A includes both purchase money and refinance loans and Panel B with purchase money loans only, which is about 44.01% of the full sample. Consistent with the general conservativeness of underwriting standards during the analysis period as noted by Rappaport and Willen

(2014) and others, only 0.69% of the full sample and 0.70% of purchase money loans in the sample became seriously delinquent over this period. Table 6 also summarizes proxies of the borrowers' credit worthiness, such as FICO score and Debt-to-Income (DTI) ratio, and information about the loan, including loan term (e.g., 30-year fixed rate) and purpose of the purchase (i.e., owner-occupied or investment). The median borrower has a FICO score of 740, with an LTV of 79% and a DTI of 34.34, buying an owner-occupied unit with a 30-year fixed term loan.

The dependent variable for our loan performance analyses is a delinquency indicator, which takes the value of 1 if the borrower of the loan associated with the subject property appraisal k became 60-days delinquent in their loan payments by the end of 2018.²⁰ Our variable of interest, $I(\text{Underreported})_{ijk}$, is an indicator for the GLA of comparable transaction i in appraisal k being reported as smaller by appraiser j , compared with the same comp i 's GLA reported by the same appraiser j in other appraisals g (where $g \neq k$). The construction of this measure ensures that reporting discrepancies come only from within each transaction-apraiser pair. In addition, such a measure of comp under-reporting could only be constructed first at the comp level i before being aggregated to the loan (i.e., appraisal) level k . Since loan performance measures does not vary within appraisal across comps, we collapse the comp level under-reporting measure to the loan level.²¹ $I(\text{Underreported})_k$ takes the value

²⁰60 days is the minimum period when a borrower is considered seriously delinquent and foreclosure proceedings initiate.

²¹Conducting the analysis at the loan level also makes sure that all appraisals receive equal weighting.

of 1 when at least one of comps in appraisal k is under-reported.²²

As the incentive to and the likelihood of default vary directly with the equity of a borrower, which is most often measured by loan-to-value (LTV) at origination, we investigate whether more financially constrained borrowers (i.e., with an LTV greater than 90%) with under-reported comp GLAs were more likely to become delinquent.²³ Empirically, we interact borrowers' LTV status indicator with a loan-level indicator for comp under-reporting, and estimate the following specification:

$$\begin{aligned}
 \text{Delinquent}_{jklst} = & \alpha * I(\text{Underreported})_k + \beta * I(\text{LTV} > 90)_k & (3) \\
 & + \delta * I(\text{Underreported})_k * I(\text{LTV} > 90)_k \\
 & + \gamma * \text{Controls}_k + \lambda_j + \theta_l + \sigma_{st} + \epsilon_{jklst},
 \end{aligned}$$

where $I(\text{Underreported})_k$ indicates that at least one of comps' GLA in appraisal k is under-reported, $I(\text{LTV} > 90)_k$ is an indicator for having an LTV greater than 90%, and the vector of Controls_k contains both proxies of the borrowers' credit worthiness and information about

Alternatively, if the loan performance regression were carried out at the comparable transaction level, the level at which the original under-reporting measure is constructed, appraisals with more comparable transactions would receive higher weights than those with fewer.

²²Results using the loan-level average of $\text{Underreported}_{ijk}$ for all comps in appraisal k , $\text{Underreported}_k^{ave}$, instead are reported in Table A1 and are highly similar.

²³Borrowers with an LTV over 90% are most likely liquidity and/or credit constrained as they have to pay a higher effective interest rate associated with private mortgage insurance (Calem et al., 2015).

the loan, as summarized in Table 6. We also include an array of high-dimensional fixed effects to account for various heterogeneities: (1) λ_j are appraiser fixed effects, stripping away differences in average loan performance across appraisers; (2) θ_l are lender fixed effects, intended to capture any differences in underwriting standards across mortgage lenders,²⁴ and (3) σ_{st} represents a vector of county-year-quarter fixed effects to control for any variations in local housing market conditions that might change the likelihood of delinquency. Our main coefficient of interest δ is defined as the differences in the probability of delinquency, for borrowers with an $LTV > 90$, when at least one of the comps' GLA was under-reported. To allow for correlations in the residuals along different dimensions, robust standard errors are multi-way clustered at the market, appraiser, lender, and time levels simultaneously.

Results are presented in Table 7, with columns (1)-(2) for the full sample and columns (3)-(4) for purchase loans. The omitted reference category are borrowers with an LTV less than or equal to 90% and with no comps' GLA being under-reported, which have an average probability of 60-day delinquency of 0.62% in the full sample and 0.55% for purchase money loans. Borrowers with an LTV greater than 90% are more likely to default, with a significant fraction of the variations being explained by credit worthiness. Suggestive of a possible exchange in positive soft borrower information between lenders and appraisers, borrowers of properties with a misreported GLA of comparable transaction are found to be 0.05-0.08

²⁴In particular, a misreporting appraiser may also be more likely to work with a mortgage lender with less strict underwriting standard making it difficult to identify separate effects.

percentage points less likely to default if they have an initial LTV less than or equal to 90%. Given that an average 0.62% (0.55%) of borrowers with an initial LTV less than or equal to 90% with no under-reported GLA of comps became seriously delinquent over this period in the full sample (purchase money loan sample), the estimates imply that, for borrowers with a greater than 90% LTV, the under-reporting of comps' GLA is associated with a 13.33%-14.29% increase in the probability of becoming seriously delinquent in their loan payments.²⁵ Alternatively, the combined effects of under-reporting for at least one comps' GLA with a greater than 90% LTV in the magnitude of 0.0012 percentage points correspond to 19.35% (21.82%) of the full sample (purchase loan sample) mean of delinquency probability. Taken together, our results demonstrate the real consequences of misreporting of property attributes, which is especially profound in the highly-leveraged segment of the market.

7. The Prevalence of Attribute Misreporting

Previous sections have demonstrated that purposeful misreporting of property attributes exist and the under-reporting of comparable transactions has real consequences on subsequent mortgage loan performance. In this section, we show that attribute misreporting is not just

²⁵More specifically, the combined effect of at least one of the comps' GLA being under-reporting with a greater than 90% LTV is 0.0012 (=0.0017-0.0005) and 0.0012 (=0.0019-0.0007) for the full sample and purchase money sample correspondingly, which divided by (0.0062 + 0.0028) and (0.0055 + 0.0029) are 13.33% and 14.29%, respectively.

limited to a few bad actors or occurs only when the scope for biased comparable selection is limited.

7.1. Prevalence Across Appraisers

We first aim to determine whether a large fraction of misreporting is by a few misbehaving appraisers being exceptionally active, or that such behavior is more pervasive across the industry. Figure 7 illustrates the distribution of appraisers by the share of properties with a reporting inconsistency. The solid bar represents the share of appraisers who reported the same property's GLA at least 100 square feet larger as a subject than as a comp. The grey bar represents the share of appraisers who made a similar understatement of GLA for at least one comp in the repeated-comp sample. Only appraisers with at least 10 appraisals are included to limit the influence of less active or experienced appraisers on our findings. We also adopt a more conservative definition of misreporting and focus only on significant reporting errors of continuous GLA and lot size in our tabulations of misreporting appraisers.

The bars left of 0 in Figure 7 indicate the percent of appraisers without a single-instance of misreporting. This share is 11.5% for the subject-comp reporting comparisons and 19.4% for the repeated-comps. Almost half of appraisers (49.4%) have at least one instance of comp under-reporting for less than 10% of appraisals they had worked on, while 23.4% of appraisers under-reported between 10 and 20% of their appraisals using repeated-comps. 7.5% of appraisers under-reported a comp attribute for more than 20% of their appraisals.

The share of misreporting appraisers increases as we examine the subject-comp reporting consistencies. Approximately one-quarter of appraisers each misreported 1-to-10% and 10-to-20% of appraisals they had worked on. 39% of appraisers made significant reporting discrepancies for more than 20% of their appraisals.

7.2. Prevalence Across Housing Markets

Finally, we turn to investigate the prevalence of misreporting across market conditions. For example, the relative attractiveness of under-reporting may depend on the availability of comparable transactions or market-wide house price growth. Under the assumption that attribute misreporting is easier to detect than biased comparable selection, it may only be a backup means of appraisal manipulation when appraisers' ability to "cherry-pick" suitable comparable transactions is limited.²⁶

Ideally, the pool of potential comparable transactions each appraiser considered but did not select, along with their adjusted sale price relative to the subject, would be identified. Such information is unfortunately not available. We test instead whether appraisers are more likely to engage in misreporting when the availability of comparable transactions, i.e., "similar" property sales occurred within the last 4 quarters in the ZIP code where the property is located, varies.²⁷

²⁶We thank an anonymous reviewer for this suggestion.

²⁷This measure of comp availability was provided by our data provider and uncertainty exists as to how "similar" was defined for their purposes. We obtain similar results using the number of transactions at the

The median property in our database had 19 potential comparable transactions of which the appraisers selected between 3-7 actual property sales to use. Figure 8 illustrates the likelihood of comp under-reporting as the availability of comps changes. The x-axis of Panel A in Figure 8 indicates the quartiles of comp availability ranked from lowest to highest, and the y-axis is defined as the percent of appraisals with at least one significantly under-reported comparable transaction. While it is natural that a slightly larger fraction of appraisals are subject to significant under-reporting of comps when there is fewer to select from (19.7%), this figure remains as high as 17.7% even when there is an abundance of comps available for selection. In summary, appraisers appear to under-report comp attributes regardless of the availability of comps to select from.

As a final exercise, we stratify the sample into four quartiles based on the annual growth of county-level house prices.²⁸ The median county in our sample enjoyed a 3.00% appreciation in HPI in the previous year, with the 25th and 75th percentile of 1.14% and 5.28%, respectively. The x-axis of Panel B in Figure 8 indicates the quartiles of recent house price growth ranked from lowest to highest, and the y-axis is defined as the same above. Similar to what we find with the availability of comparable transactions, there is little difference across the 4 quartiles of HPI growth (from 17.6% to 15.9%) suggesting misreporting appears to persist across county-level estimated in our sample, but prefer the micro-sample analysis given significant heterogeneity in population and other attributes across counties.

²⁸County-level house price index (HPI) annual changes are produced by the Federal Housing Finance Agency (FHFA).

across market conditions.

Taken together, our findings suggest that attribute-misreporting is not limited to a few misbehaving appraisers nor is it a tactic that the appraisers had to resort to when the possibility of cherry-picking is low.

8. Conclusions

In this paper, we document a specific channel through which appraisal bias occurs, and provide strong evidence that some appraisers continue to manipulate appraised values even after reforms enacted following the 2008 financial crisis to curb such behavior. The specific channel identified is that appraisers inflate the valuation of the subject property by under-reporting property attributes of comparable transactions, especially when the alternative appraisal would have been below the contract price suggested by an AVM estimate. This misreporting has real economic consequences as borrowers of the loans for properties associated with attribute manipulation were found to be more likely to become seriously delinquent in their loan payments, especially those with the highest leverage.

Several important caveats of the research are important to disclose. First, while the dataset used to arrive at the above findings is compiled from 21 million loan applications from 2013-2017, it does not represent the complete universe of loan applications over this period. Most notable absentees from the data are jumbo and other non-conforming loan applications. Also, as most housing markets were relatively stable during this period, it is

unclear how the results would generalize to periods with greater levels of price volatility. Last, while a relatively high percentage of appraisers have engaged in at least some type of misreporting, at least some of this misreporting is unavoidable as appraisers were faced with the difficult task of doing more with increasingly limited compensation and time.

Our findings have important implications for the economy and financial institutions that are directly and indirectly involved in mortgage investment and the MBS securitization process, especially issuers who often provide guarantee to principal and interest payment of the underlying mortgage. Even if original mortgage lenders were able to capture some of the higher expected default risks through differential loan pricing, it is unclear this would be priced correctly by regulators and issuers unaware of such biases. This will lead to lower capital reserves than otherwise required, and greater exposure for taxpayers. More importantly, a non-biased valuation may have prevented some inexperienced borrowers from paying a purchase price above market value, and the eventual economic costs of foreclosure upon themselves and surrounding property owners.

An important implication of our research is that a national database of property transactions and associated reported attributes of subject properties should be created. Our results show that appraisers were relatively consistent and non-strategic in reporting attributes of subject properties, but did occasionally engage in under-reporting attributes of comparable transactions. Currently, appraisers must rely on publicly available data records for attributes of comparable transactions since they are not granted interior access to the property unless

they were themselves or able to obtain attributes from the original appraiser of the property. Providing appraisers access and then holding them accountable to use consistent attributes of comparable transactions reported by appraisers with access to the interior of the property could lower potential appraisal biases and enhance the liquidity of financial products supported by such valuations. This database would also be helpful to the further automation of valuation processes, who often use data indiscriminately regardless of source. The potential for this database and the degree to which appraisal biases are priced into loans remain open questions of research.

References

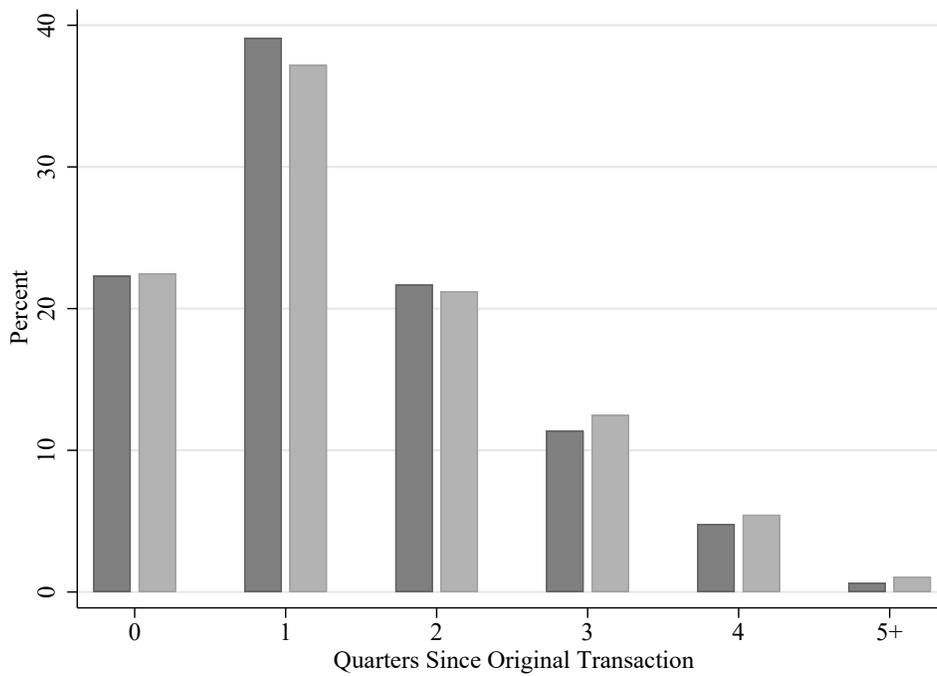
- Agarwal, S., Ambrose, B. W., and Yao, V. W. (2020). Can regulation de-bias appraisers? *Journal of Financial Intermediation*, 44, 100827.
- Agarwal, S., Ben-David, I., and Yao, V. (2015). Collateral valuation and borrower financial constraints: Evidence from the residential real estate market. *Management Science*, 61(9), 2220–2240.
- Ben-David, I. (2011). Financial constraints and inflated home prices during the real estate boom. *American Economic Journal: Applied Economics*, 3(3), 55–87.
- Bongaerts, D., Cremers, K. M., and Goetzmann, W. N. (2012). Tiebreaker: Certification and multiple credit ratings. *Journal of Finance*, 67(1), 113–152.
- Calem, P. S., Lambie-Hanson, L., and Nakamura, L. I. (2015). Information losses in home purchase appraisals. *FRB of Philadelphia Working Paper No. 15-11*.
- Conklin, J., Coulson, N. E., Diop, M., and Le, T. (2020). Competition and appraisal inflation. *Journal of Real Estate Finance and Economics*, 61, 1–38.
- Demiroglu, C. and James, C. (2016). Indicators of collateral misreporting. *Management Science*, 64(4), 1747–1760.
- Ding, L. and Nakamura, L. (2016). The impact of the home valuation code of conduct on appraisal and mortgage outcomes. *Real Estate Economics*, 44(3), 658–690.

- Eriksen, M. D., Fout, H. B., Palim, M., and Rosenblatt, E. (2019). The influence of contract prices and relationships on appraisal bias. *Journal of Urban Economics*, 111, 132–143.
- Eriksen, M. D., Fout, H. B., Palim, M., and Rosenblatt, E. (2020). Contract price confirmation bias: Evidence from repeat appraisals. *Journal of Real Estate Finance and Economics*, 60, 77–98.
- Griffin, J. M. (2021). Ten years of evidence: Was fraud a force in the financial crisis? *Journal of Economic Literature*, 59(4), 1293–1321.
- Griffin, J. M., Kruger, S., and Maturana, G. (2020). What drove the 2003–2006 house price boom and subsequent collapse? disentangling competing explanations. *Journal of Financial Economics*.
- Griffin, J. M. and Maturana, G. (2016). Who facilitated misreporting in securitized loans? *Review of Financial Studies*, 29(2), 384–419.
- Kruger, S. and Maturana, G. (2021). Collateral misreporting in the residential mortgage-backed security market. *Management Science*, 67(5), 2729–2750.
- Mian, A. and Sufi, A. (2017). Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005. *Review of Financial Studies*, 30, 1832–1864.
- Rappaport, J. and Willen, P. S. (2014). Tight credit conditions continue to constrain the housing recovery. *Macro Bulletin*, (pp. 1–2).

Stroebel, J. (2016). Asymmetric information about collateral values. *Journal of Finance*, 71(3), 1071–1112.

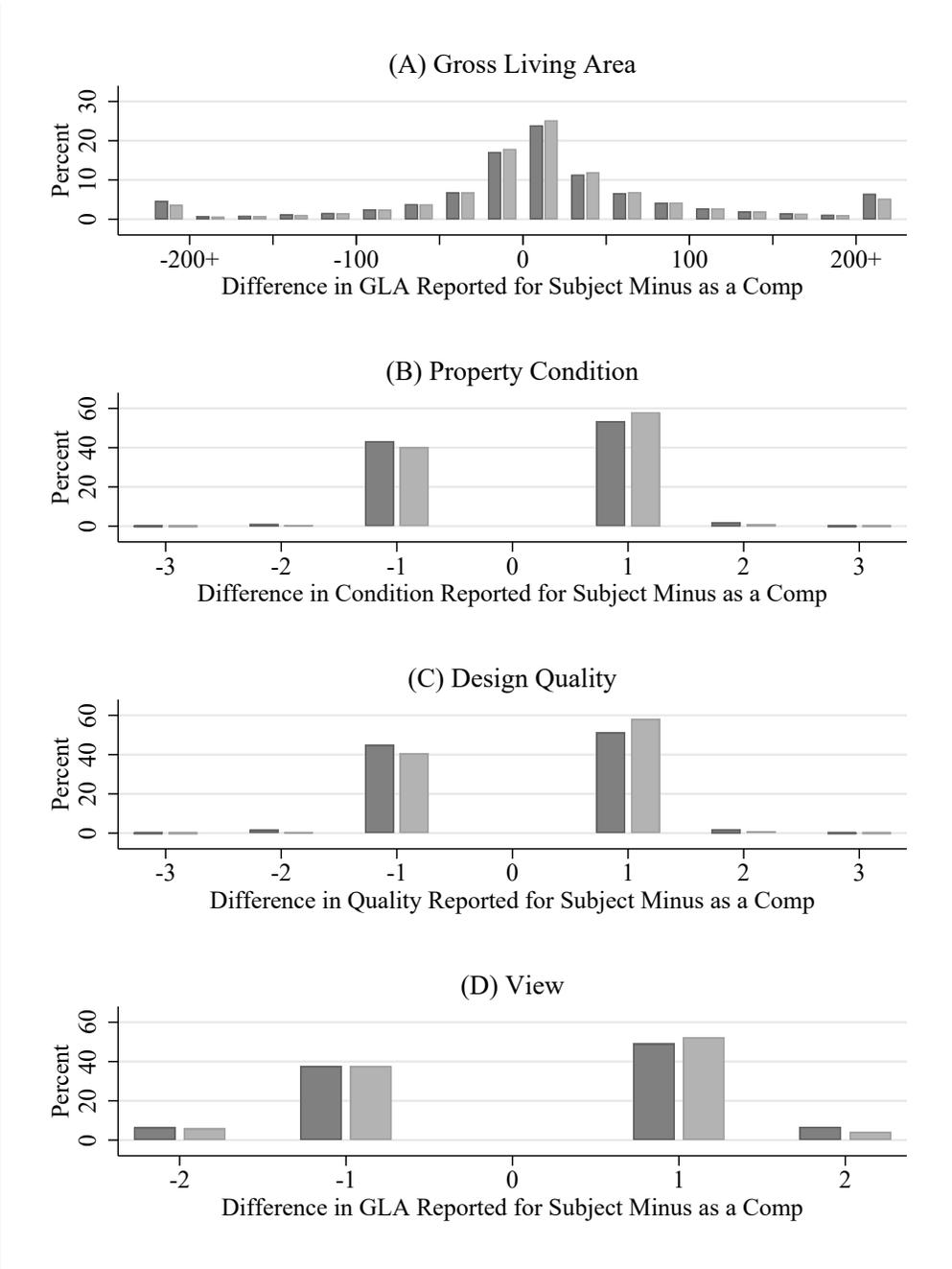
Vandell, K. D. (1991). Optimal comparable selection and weighting in real property valuation. *Real Estate Economics*, 19(2), 213–239.

Figure 1: Distribution of comparable transactions with respect to number of quarters since original transaction.



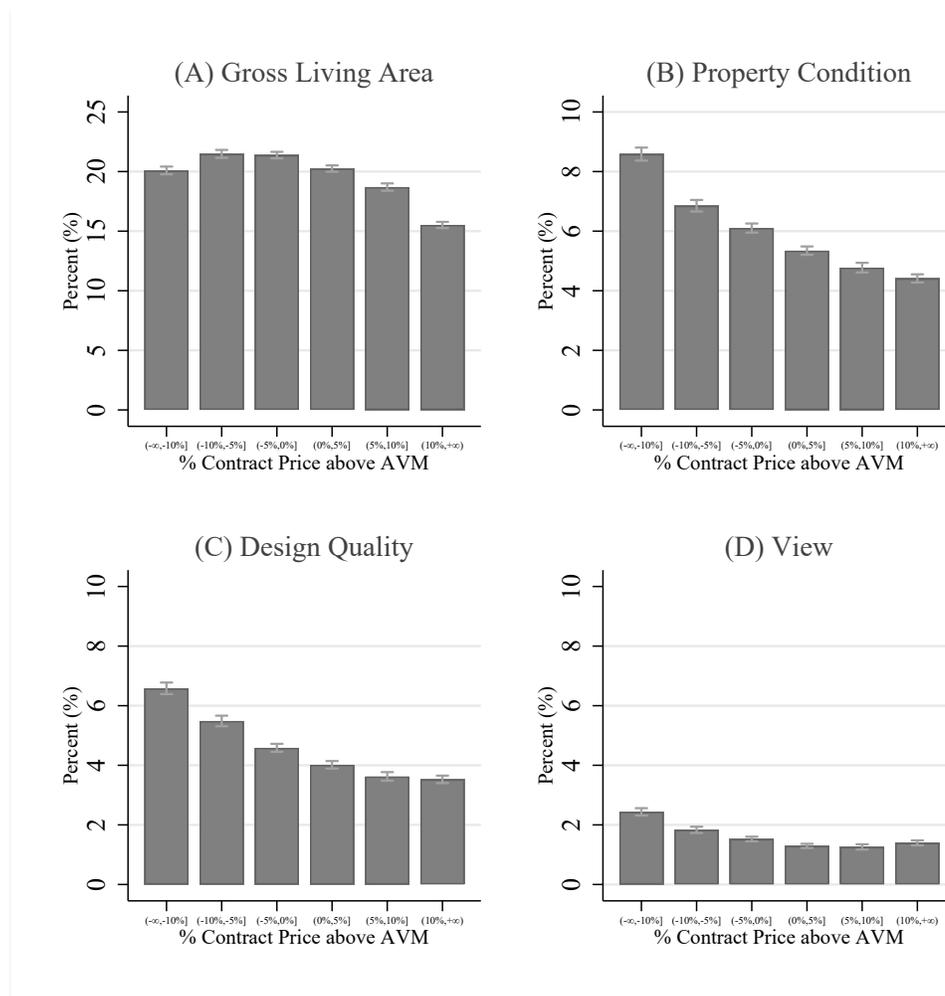
Notes: All estimation samples in the paper are restricted to comparable transactions used within 4 quarters from the original transaction. This deletes 154,732 and 7,074 observations from the two samples, respectively. Dark grey bars represent comparable transactions with attributes also reported as a subject ($n = 28,258,168$) and light grey bars represent attributes reported as a subject by the same appraiser ($n = 1,149,417$).

Figure 2: Differences in reported attributes for the same property as a subject and as a comparable transaction with a discrepancy.



Notes: Dark grey bars represent all subject-comparable transaction pairs and light grey bars represents subject-comparable transaction pairs with attributes reported by the same appraiser

Figure 3: Percent of comparable transactions with a superior stated property attribute as a subject by the same appraiser with respect to the percent contract price of subject above AVM estimate.



Notes: Bars indicate 95% confidence intervals.

Figure 4: Percent of comparable transactions with under-reported GLA by the number of quarters between when the maximum and minimum GLA were reported by the same appraiser.

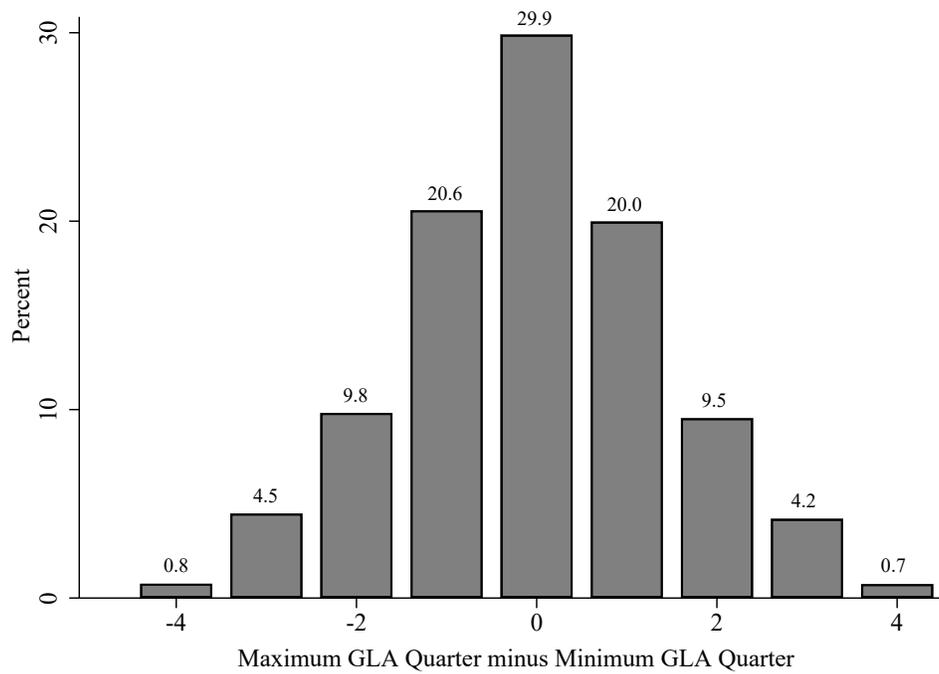
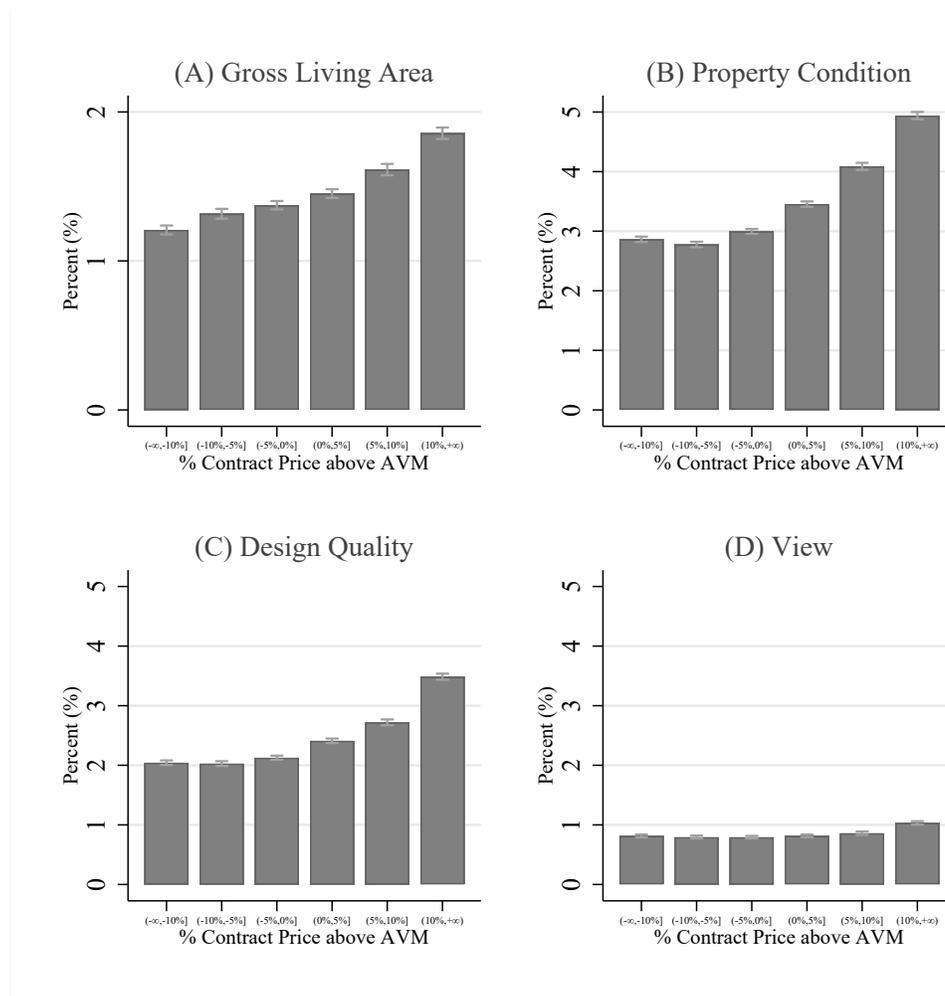
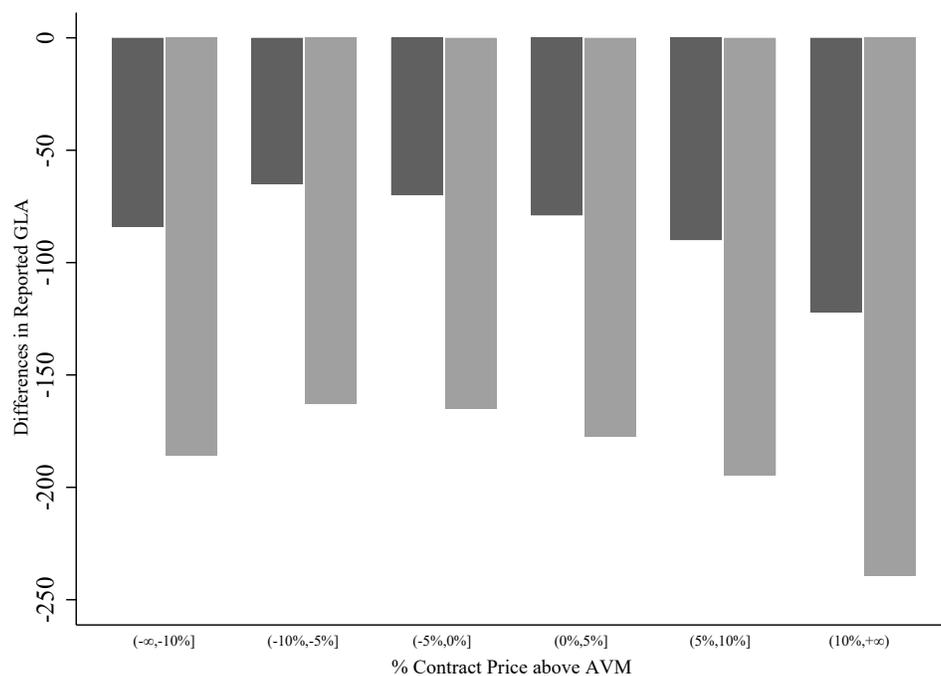


Figure 5: The percent of repeated-comps with an under-reported attribute by the same appraiser with respect to the percent subject's contract price above AVM estimate.



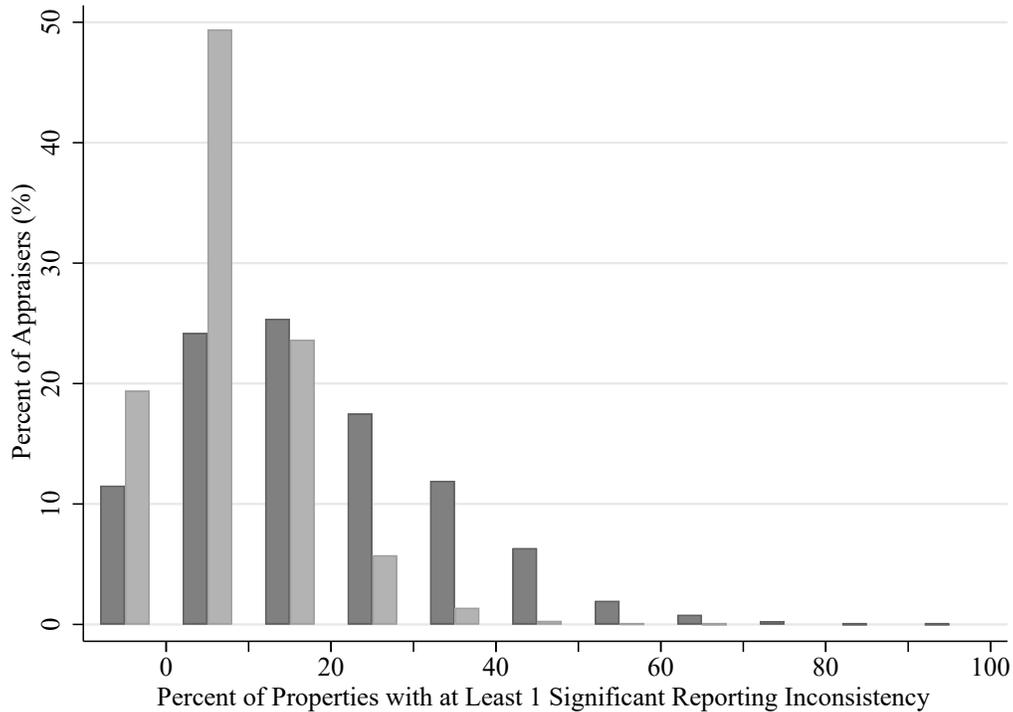
Notes: Bars indicates 95% confidence intervals.

Figure 6: Magnitude of under-reporting of Gross Living Area (GLA) of a comparable transaction by the same appraiser with respect to the percent subject's contract price above AVM estimate.



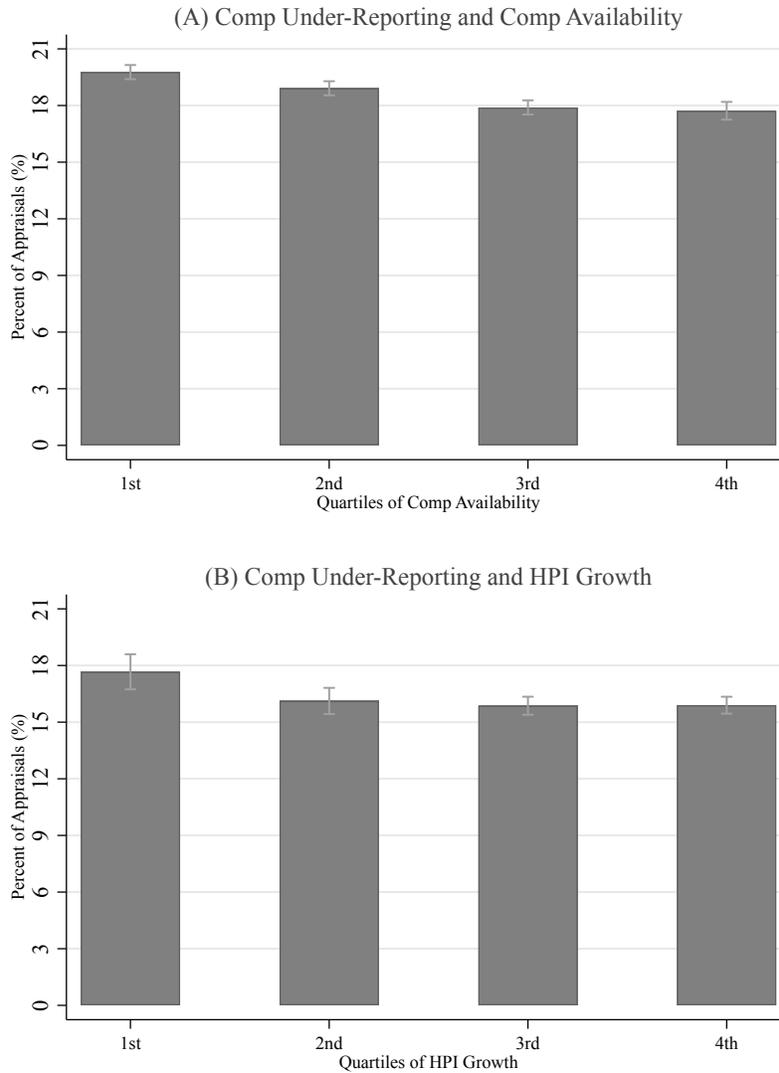
Notes: Dark grey bars represent median differences and light grey bars represents averages.

Figure 7: Distribution of appraisers by the % of appraised properties with at least one significant reporting inconsistency. Dark grey bars represents distribution of appraisers over-reporting attributes of subject properties and light grey bars represents distribution of appraisers under-reporting attributes of a comparable transaction



Notes: Both samples are restricted to appraisers with at least 10 properties of each potential type of reporting error. A significant reporting error is defined as at least a 100 square foot discrepancy in gross living area or lot size, or a different value reported for property condition, design quality, bedrooms, full bathrooms, or view. The bars left of 0 indicate the percent of appraisers with no reporting inconsistencies.

Figure 8: The prevalence of under-reporting attributes of a comparable transaction by the same appraiser across different market conditions.



Notes: The graph plots percentage of appraisals with an significantly under-reported comp across markets with different levels of: (A) potential comps; and (B) county-level house price index (HPI) growth as measured by the Federal Housing Finance Agency. An appraisal is defined as being under-reported when at least 1 of all 7 attributes of one of its comparable transactions was reported as being inferior compared to as reported by the same appraiser for a different subject appraisal. For GLA and lot size, the reporting discrepancies need to exceed 100 sqft to be counted as significant under-reporting. Bars represent 95% confidence intervals.

Table 1: Summary Statistics of Comparable Transactions also Reported as a Subject

	Panel A: Properties Used as a Subject and Comparable at Least Once			Panel B: Properties Used as a Subject and Comparable at Least Once by Same Appraiser			
	Median Price Adjustments per Unit Difference (\$)	Median Reported Attributes	Consistently Reported (%)	Average Range of Reported Discrepancies	Median Reported Attributes	Consistently Reported (%)	Average Range of Reported Discrepancies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gross Living Area (square feet)	232	1,559	19.78	109.98	1,780	66.61	26.35
Property Condition (scale: 1-6)	12,500	4	64.87	0.85	4	89.86	0.12
Quality of Design (scale: 1-6)	11,500	3	58.10	0.90	3	92.17	0.09
View (scale: 1-3)	7,500	1	88.67	0.30	1	97.15	0.04
Bedrooms (counts)	5,000	3	91.44	0.18	3	96.58	0.04
Full Bathrooms (counts)	4,000	2	95.48	0.11	2	98.32	0.02
Lot Size (square feet)	40	9,309	60.48	446.93	10,890	84.14	116.94
All Attributes			5.45			53.82	
Reported Consistently (%)							
Average Contract Price (\$)	-		332,805.16			323,832.72	
Average Appraised Value (\$)	-		335,782.24			326,639.97	
Average AVM (\$)	-		336,197.64			323,215.04	
Unique Properties	-		4,594,672			494,225	
Comparable Transactions	-		23,499,763			644,525	
Average Times	-		5.11			1.30	
Subject Used as Comparable	-						
Unique Appraisers	-		57,344			34,973	

Notes: GLA and Lot Size are in square feet. Property Condition and Quality of Design are on a six-point scale and View is on a three-point scale, where a higher value always indicates a property with a superior attribute. Consistently Reported (%) are tabulated at the property level.

Table 2: Comparison of Reporting Discrepancies of Comparable Transactions also Reported as a Subject

	Panel A: Properties Used as a Subject and Comparable at Least Once			
	Subject > Comparable		Subject < Comparable	
	% of	Average	% of	Average
	Discrepancies	Discrepancies	Discrepancies	Discrepancies
	(1)	(2)	(3)	(4)
Gross Living Area (square feet)	59.91	85.95	40.09	-91.78
Property Condition (scale: 1-6)	55.41	1.04	44.59	-1.03
Quality of Design (scale: 1-6)	53.15	1.04	46.85	-1.04
View (scale: 1-3)	55.56	1.14	44.44	-1.17
Bedrooms (counts)	39.63	1.07	60.37	-1.06
Full Bathrooms (counts)	41.67	1.03	58.33	-1.03
Lot Size (square feet)	53.45	609.39	46.55	-583.13

	Panel B: Properties Used as a Subject and Comparable by Same Appraiser at Least Once			
	Subject > Comparable		Subject < Comparable	
	% of	Average	% of	Average
	Discrepancies	Discrepancies	Discrepancies	Discrepancies
	(1)	(2)	(3)	(4)
Gross Living Area (square feet)	61.14	73.85	38.86	-77.33
Property Condition (scale: 1-6)	58.64	1.02	41.36	-1.01
Quality of Design (scale: 1-6)	58.19	1.01	41.81	-1.01
View (scale: 1-3)	57.17	1.08	42.83	-1.14
Bedrooms (counts)	41.06	1.06	58.94	-1.05
Full Bathrooms (counts)	41.27	1.04	58.73	-1.02
Lot Size (square feet)	54.58	696.75	45.42	-695.35

Notes: GLA and Lot Size are in square feet. Property Condition and Quality of Design are on a six-point scale and View is on a three-point scale, where a higher value always indicates a property with a superior attribute.

Table 3: Estimated Difference in Reported Property Attributes: Subject vs. Comparable

	GLA (1)	Condition (2)	Quality (3)	View (4)
Panel A. Subject transaction used at least once as a comparable transaction				
Subject of Appraisal	11.6078*** (0.0639)	0.0117*** (0.0001)	0.0084*** (0.0001)	0.0091*** (0.0001)
Property Fixed Effects	✓	✓	✓	✓
# of Observations	28,094,435	28,094,435	28,094,435	28,094,435
# of Unique Properties	4,594,672	4,594,672	4,594,672	4,594,672
# of Unique Appraisers	57,344	57,344	57,344	57,344
Panel B. Subject transaction used at least once as a comparable transaction by the <i>same appraiser</i>				
Subject of Appraisal	5.1423*** (0.1560)	0.0052*** (0.0001)	0.0038*** (0.0001)	0.0024*** (0.0002)
Property-Appraiser Fixed Effects	✓	✓	✓	✓
# of Observations	1,138,750	1,138,750	1,138,750	1,138,750
# of Unique Properties	494,225	494,225	494,225	494,225
# of Unique Appraisers	34,973	34,973	34,973	34,973
Panel C. Subject transaction used at least once as a comparable transaction by the <i>same appraiser</i> in the <i>same quarter</i> as its original transaction				
Subject of Appraisal	4.1989*** (0.2281)	0.0031*** (0.0002)	0.0021*** (0.0002)	0.0018*** (0.0003)
Property-Appraiser Fixed Effects	✓	✓	✓	✓
# of Observations	277,317	277,317	277,317	277,317
# of Unique Properties	130,746	130,746	130,746	130,746
# of Unique Appraisers	25,152	25,152	25,152	25,152

Notes: Dependent variables are the reported property attributes indicated in the column titles. GLA is gross living area and specified in square feet. Condition, quality, and view are log-transformed. A positive coefficient indicates that appraisers reported the same property as superior when the subject of an appraisal. Robust standard errors are clustered at the transaction level in Panel A and at the transaction and appraiser levels in Panels B and C.

Table 4: Summary Statistics of Comparable Transactions Used at Least Twice by Same Appraiser

	Average	% Consistently	%	Under-Reporting	
	Attributes	Reported	Under-Reported	Discrepancies	
	(1)	(2)	(3)	Median	Average
Gross Living Area (square feet)	1,977.51	96.56	1.72	85	189.52
Property Condition (scale: 1-6)	4.12	91.80	4.10	1	1.03
Quality of Design (scale: 1-6)	3.48	94.25	2.88	1	1.02
View (scale: 1-3)	1.12	98.01	1.00	1	1.12
Bedrooms (counts)	3.27	98.67	0.67	1	1.08
Full Bathrooms (counts)	2.03	99.04	0.48	1	1.02
Lot Size (square feet)	14,326.95	96.19	1.91	200	958.83
All Attributes		81.12		-	
Reported Consistently (%)				-	
Average Subject Contract Price (\$)		317,199.05		-	
Average Subject Appraised Value (\$)		320,632.03		-	
Average Subject AVM (\$)		320,598.26		-	
Unique Properties		1,249,147		-	
Comparable Transactions		2,907,582		-	
Average Times		2.33		-	
Subject Used as Comparable				-	
Unique Appraisers		39,254		-	

Notes: GLA and Lot Size are in square feet. Property Condition and Quality of Design are on a six-point scale and View is on a three-point scale, where a higher value always indicates a property with a superior attribute.

Table 5: Difference in Reported Attributes for the Same Transaction used as Comparable Transaction Repeatedly by Same Appraiser when There are Incentives to Inflate Subject Appraisal Value

	GLA (1)	Condition (2)	Quality (3)	View (4)
Panel A. Within-Transaction-Appraiser Estimation				
I (Contract Price >AVM)	-3.8602*** (0.2642)	-0.0087*** (0.0007)	-0.0064*** (0.0006)	-0.0024*** (0.0004)
Transaction-Appraiser Fixed Effects	✓	✓	✓	✓
# of Observations	2,907,582	2,907,582	2,907,582	2,907,582
# of Unique Properties	1,230,545	1,230,545	1,230,545	1,230,545
Panel B. Within-Transaction-Appraiser-Year-Quarter Estimation				
I (Contract Price >AVM)	-2.9806*** (0.3028)	-0.0071*** (0.0008)	-0.0048*** (0.0006)	-0.0022*** (0.0004)
Transaction-Appraiser-Year-Quarter Fixed Effects	✓	✓	✓	✓
# of Observations	1,724,994	1,724,994	1,724,994	1,724,994
# of Unique Properties	778,687	778,687	778,687	778,687

Notes: Dependent variables are the reported property attributes indicated in the column titles. *GLA* is gross living area and specified in square feet. *Condition*, *Quality*, and *View* are log-transformed. A coefficient value less than 0 indicates that the same appraiser reported the same comparable transaction as inferior. Contract price is the accepted offer price of the property and AVM is the estimate generated by an automated valuation model. Robust standard errors are clustered at the appraisal, transaction, appraiser, and year-quarter levels.

Table 6: Summary Statistics of Loan Attributes

	Full Sample			
	Mean (1)	25 th percentile (2)	Median (3)	75 th percentile (4)
I(60-Day Delinquency)	0.01	0	0	0
Borrower FICO Score	742.98	717	740	783
Loan-to-Value (LTV)	74.30	66	79	85
Debt-to-Income (DTI)	33.92	26.68	34.34	40.87
I(30-Year Term)	0.69	0	1	1
I(Owner-Occupied)	0.78	1	1	1
I(Renter-Occupied)	0.04	0	0	0
I(Vacant)	0.18	0	0	0
I(Purchase Money Loan)	0.44	0	0	1
# of Observations	2,992,969			

	Purchase Money Loans			
	Mean	25 th percentile	Median	75 th percentile
I(60-Day Delinquency)	0.01	0	0	0
Borrower FICO Score	758.49	730	770	794
Loan-to-Value (LTV)	81.88	79	80	95
Debt-to-Income (DTI)	33.68	27.33	34.80	41.02
I(30-Year Term)	0.84	1	1	1
I(Owner-Occupied)	0.57	0	1	1
I(Renter-Occupied)	0.03	0	0	0
I(Vacant)	0.40	0	0	1
# of Observations	1,317,226			

Notes: The sample includes only mortgage loan applications with non-missing information on loan performance and borrower credit worthiness. I(60-Day Delinquency) is an indicator for 60-day delinquency. Borrower FICO Score refers to the FICO score of the associated borrower (i.e., loan applicant). Loan-to-Value (LTV) and Debt-to-Income (DTI) ratios are also reported in the mortgage application. I(30-Year Term) is an indicator for a 30-year term and amortization length. I(Owner-Occupied), I(Renter-Occupied), and I(Vacant) are indicators for owner-occupied, renter-occupied, and vacant occupancy status at the time of the appraisal, respectively.

Table 7: The Effect of Under-Reporting Attributes of a Comparable Transaction on Loan Delinquency

	60-Day Delinquency			
	Full Sample		Purchase Money Loans	
	(1)	(2)	(3)	(4)
I(Underreported)	-0.0006** (0.0003)	-0.0005* (0.0003)	-0.0008* (0.0004)	-0.0007* (0.0004)
I(LTV >90)	0.0043*** (0.0006)	0.0028*** (0.0005)	0.0049*** (0.0007)	0.0029*** (0.0005)
I(Underreported) × I(LTV >90)	0.0017** (0.0006)	0.0017** (0.0006)	0.0019** (0.0008)	0.0019** (0.0008)
I(Renter-Occupied)		-0.0002 (0.0003)		0.0026*** (0.0005)
I(Vacant)		0.0007*** (0.0002)		0.0010*** (0.0002)
I($FICO^{med}$)		-0.0082*** (0.0011)		-0.0056*** (0.0010)
I($FICO^{high}$)		-0.0129*** (0.0014)		-0.0119*** (0.0013)
I(DTI^{med})		0.0088*** (0.0011)		0.0097*** (0.0014)
I(DTI^{high})		0.0127*** (0.0016)		0.0134*** (0.0019)
I(30-year Term)		0.0031*** (0.0003)		0.0027*** (0.0004)
I(Purchase Money)	-0.0020*** (0.0004)	-0.0052*** (0.0005)	—	—
County-Year-Quarter Fixed Effects	✓	✓	✓	✓
Appraiser Fixed Effects	✓	✓	✓	✓
Lender Fixed Effects	✓	✓	✓	✓
# of Observations	2,992,969	2,992,969	1,317,226	1,317,226

Notes: The dependent variable is an indicator if the borrower of the loan of the subject property becomes 60-day delinquent in their payments. All regressions are at the loan-level and focus on misreporting in GLA. $I(Underreported)$ is a loan-level indicator for having at least one comp's GLA being under-reported. $I(Renter-Occupied)$ indicates partially owner-occupied, and $I(Vacant)$ indicates investment purchase. $I(FICO^{med})$ and $I(FICO^{high})$ are indicators for having a FICO score in [700, 760) and ≤ 760 , respectively. $I(DTI^{med})$ and $I(DTI^{high})$ are indicators for having a debt-to-income ratio in [30%, 40%) and $\leq 40\%$, respectively. $I(30\text{-year Term})$ is an indicator for 30-year term and amortization schedule. $I(Purchase Money)$ indicates loan's purpose was to purchase a new home. Robust standard errors are clustered at the market (county), time (year-quarter), and appraiser levels.

Appendix

A1.1. Sample Construction

Identifying Unique Transactions. The empirical analyses in this paper are based on property attributes reported by appraisers to support his/her valuation estimates as required by the Federal Housing Finance Agency for conforming loans since 2011. These attributes could be reported for the same property more than once – initially when property i was the subject of an appraisal k associated with loan approval process at t_0 and subsequently when the resulting transaction of property i at t_0 was used as a comparable transaction to value other properties j at t in other appraisals g (where $j \neq i$, $g \neq k$, and $t \geq t_0$).

Unique transactions are identified by the combination of property IDs and their corresponding transaction quarters in our sample.²⁹ We exclude properties transacting more than once in the same quarter from the sample.³⁰ To ensure that the correct transaction is identified, we keep only the first transaction of each unique property if

²⁹Only the quarter of the appraisal and resulting property transaction were identified in the data for confidentiality reasons, and the two quarters may differ. For example, a property transacted in Q4 of 2014 may have been originally appraised in Q3 of 2014. The quarter of transaction of each subject property is identified when this resulting transaction was later used as a comp in other appraisals, as the quarter of own transaction, t_0 , is reported for all comparable transactions. We exclude from our sample the 1,735,773 (out of 7,265,791) observations whose transaction quarters were not identified.

³⁰431,106 observations were excluded as a result of this restriction.

multiple transactions with the same property ID are observed in the sample.³¹

Creating Empirical Samples. The main filter was to limit to transactions with at least two sets of reported attributes between 2013 and 2017. These reported attributes could be for a pending transaction of property itself (i.e., a subject) or as a comparable transaction to support the appraisers opinion of value for a different property in the future. For the subject-comparable analyses, a subject property i is included when the original transaction was observed at least once as a comp in the sample. To limit the influence of differential access to information as well as interpretation and method, we employ a restricted same-appraiser sample for the majority of our subject-comparable analyses. Only when the same appraiser in the original transaction of property i subsequently used the original transaction as a comp in other appraisals to value other properties. In Section 4.2, we further restrict to comps used by the same appraiser in the same quarter as the original transaction to eliminate the possibility of actual property feature changes. Attributes reported for the same transaction, first as a subject then as a comp, either by the same or a different appraiser, form the basis of our subject-comparable analyses.

To address the possibility that the observed discrepancies in reported attributes when a subject versus as a comparable transaction might be unavoidable due to flows in the existing practice, we create a sample of repeated comparable transactions. This repeated-comp sample consists comp observations when the same appraiser used the

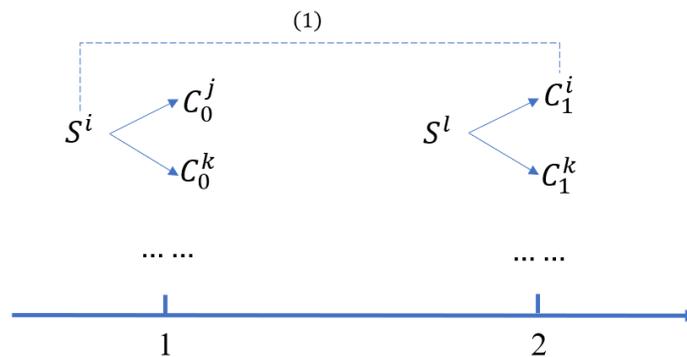
³¹Another 391,009 observations were excluded as a result of this restriction.

same comparable transaction more than once in different appraisals.

A1.2. Sample Description

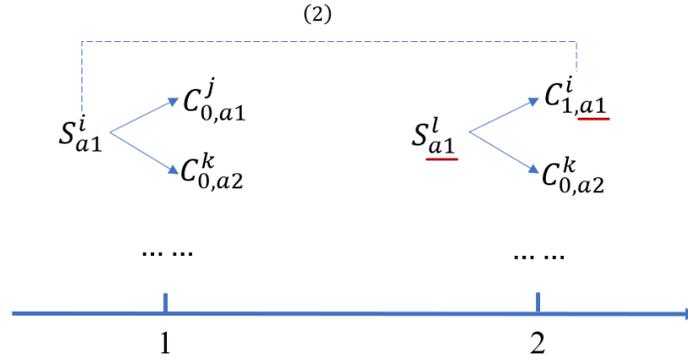
We illustrate our appraisal data structure below. The horizontal axis indicates time periods, corresponding to quarters in our data. Subject properties are denoted by S , with the superscripts (e.g., i, j , etc.) indicating unique property IDs. According to the sales-comparison approach of valuation, each subject property's valuation is supported by multiple comparable transactions, indicated by C in the figure. The subscript of each comp indicates its original transaction quarter. Appraisers are required not only to report attributes for the subject of the appraisal, but also for the comparable transactions they have selected to justify their valuation of the subject property. For the subject-comparable analyses in Section 4, we first identify subject-comp pairs (i.e., connected pair (1) in Figure A1) of attributes reported for the same property, both as the subject of an appraisal (e.g., S^i) and as a comp to evaluate other properties (e.g., C^i).

Figure A1: Subject-comp comparison



Appraiser IDs are added to Figure A2, and are indicated by the subscripts $a1$, $a2$, and so on. Subscripts of S indicate the appraiser working on the appraisal, reporting attributes for both the subject property and the corresponding comps. For example, appraiser $a1$ reported attributes for both S^i , C^j , and C^k . The original appraiser when the comparable property was itself the subject of an appraisal is indicated by the subscripts of C_{a1} . To eliminate the influence of differences in information access, methods of interpretation and assessment, we also restrict subject-comp analyses to subject-comp pairs (i.e., connected pair (2) in Figure A2) reported by the same appraiser (i.e., appraiser $a1$).

Figure A2: Subject-comp comparison with same appraiser



In the most restricted subject-comp comparison in Section 4.2, we keep only subject-comp pairs where the comparable transaction was used and reported by the same appraiser who originally appraised it as a subject in the same quarter. For example, we could compare attributes reported by appraiser $a1$ for property i as a subject (i.e., S^i) with those as a comp, if it was also used by appraiser $a1$ to evaluate a different property in the same quarter as its original transaction (i.e., quarter 1).

Finally, to eliminate the possibility that the observed subject-comp gaps of reported attributes are due to differences in data collection methods for subjects and comps, we focus exclusively on properties selected multiple times by the same appraiser as a comparable transaction to value different properties (i.e., a repeated-comp sample). Connect pair (3) in Figure A3 is an example of our repeated-comp analyses. We compare attributes reported for the same transaction as comps (i.e., $C_{0,a2}^k$) by the same appraiser (i.e., appraiser $a1$) in different appraisals (i.e., i versus l).

Figure A3: Repeated-comp comparison

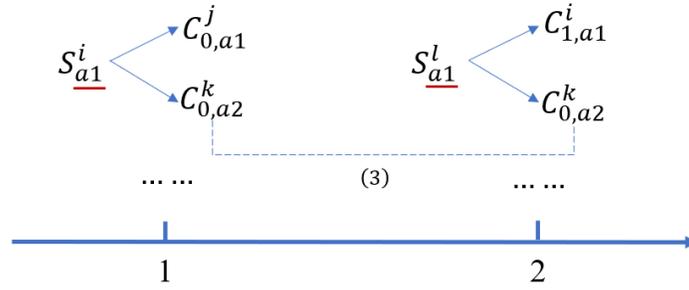
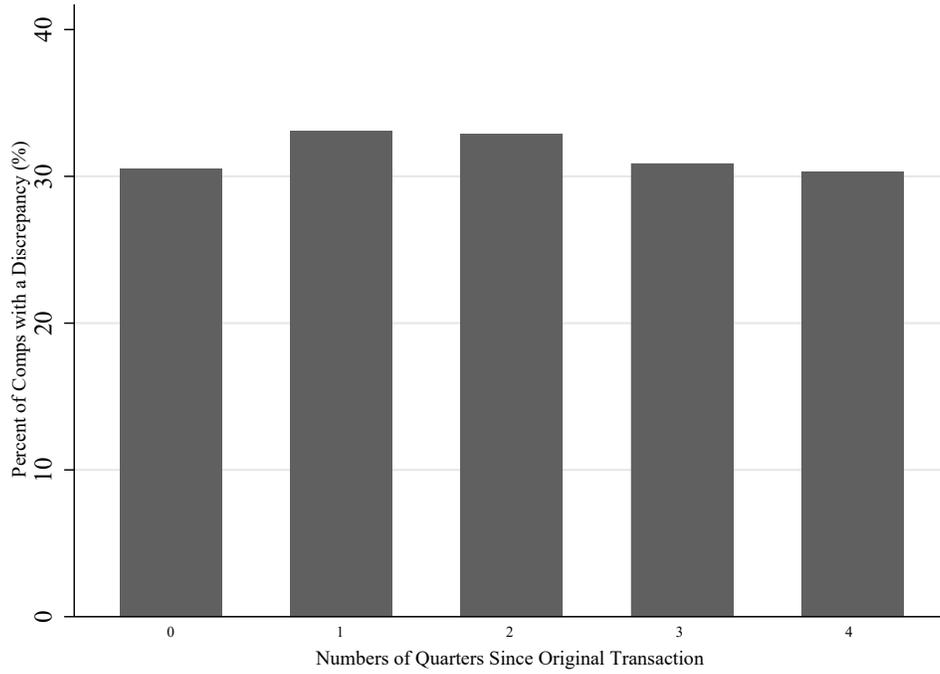


Figure A4: The percent of comparable transactions with different GLA reported than as a subject by the same appraiser for each quarter since original transaction



Notes: A larger percent of discrepancies would occur as the number of quarters increases if explained by actual changes in property attributes and records.

Table A1: The Effect of Misreporting of Comparable Transaction Attribute(s) on Loan Delinquency

	60-Day Delinquency			
	Full Sample		Purchase Money Loans	
	(1)	(2)	(3)	(4)
$Underreported^{ave}$	-0.0006*	-0.0005	-0.0007	-0.0006
	(0.0003)	(0.0003)	(0.0005)	(0.0005)
I(LTV >90)	0.0043***	0.0029***	0.0049***	0.0030***
	(0.0006)	(0.0005)	(0.0007)	(0.0005)
$Underreported^{ave} \times I(\text{LTV} >90)$	0.0021**	0.0021**	0.0020*	0.0020*
	(0.0008)	(0.0008)	(0.0010)	(0.0010)
I(Renter-Occupied)		-0.0002		0.0026***
		(0.0003)		(0.0005)
I(Vacant)		0.0007***		0.0010***
		(0.0002)		(0.0002)
I($FICCO^{med}$)		-0.0082***		-0.0056***
		(0.0011)		(0.0010)
I($FICCO^{high}$)		-0.0129***		-0.0119***
		(0.0014)		(0.0013)
I(DTI^{med})		0.0088***		0.0097***
		(0.0011)		(0.0014)
I(DTI^{high})		0.0127***		0.0134***
		(0.0016)		(0.0019)
I(30y Fixed Term)		0.0031***		0.0027***
		(0.0003)		(0.0004)
I(Purchase Money)	-0.0020***	-0.0052***	—	—
	(0.0004)	(0.0005)		
County-Year-Quarter Fixed Effects	✓	✓	✓	✓
Appraiser Fixed Effects	✓	✓	✓	✓
Lender Fixed Effects	✓	✓	✓	✓
# of Observations	2,992,969	2,992,969	1,317,226	1,317,226

Notes: The dependent variable is an indicator if the borrower of the loan of the subject property becomes 60-day delinquent in their payments. All regressions are at the loan-level and focus on misreporting in GLA. $Underreported^{ave}$ is the loan-level average of indicators for having a comp's GLA being under-reported. I(Renter-Occupied) indicates partially owner-occupied, and I(Vacant) indicates investment purchase. I($FICCO^{med}$) and I($FICCO^{high}$) are indicators for having a FICCO score in [700, 760) and ≤ 760 , respectively. I(DTI^{med}) and I(DTI^{high}) are indicators for having a debt-to-income ratio in [30%, 40%) and $\leq 40\%$, respectively. I(30y Fixed Term) is an indicator for 30-year fixed-rate mortgage. I(Purchase Money) indicates purchase money loans. Robust standard errors are clustered at the market (county), time (year-quarter), and appraiser levels.