



WORKING PAPERS

RESEARCH DEPARTMENT

WORKING PAPER NO. 15-11
INFORMATION LOSSES IN HOME PURCHASE APPRAISALS

Paul S. Calem
Federal Reserve Bank of Philadelphia

Lauren Lambie-Hanson
Federal Reserve Bank of Philadelphia

Leonard I. Nakamura
Federal Reserve Bank of Philadelphia

March 2015

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Information Losses in Home Purchase Appraisals

Paul S. Calem, Lauren Lambie-Hanson, and Leonard I. Nakamura¹

Federal Reserve Bank of Philadelphia

February 2015

Abstract

Home appraisals are produced for millions of residential mortgage transactions each year, but appraisals are rarely below the transaction price. We exploit a unique data set to show that the mortgage application process creates an incentive to substitute the transaction price for the true appraised value when the latter is lower. We relate the frequency of information loss (appraisals set equal to transaction price) to market conditions and other factors that plausibly determine the degree of distortion. Information loss in appraisals may increase the procyclicality of housing booms and busts.

Keywords: Information, mortgage, regulation, appraisal

JEL Codes: D81, G14, G21, G28, L85

¹ The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. We thank Jan Brueckner, Kris Gerardi, Tim Lambie-Hanson, Doug McManus, Albert Saiz, and seminar participants at the Federal Reserve Bank of Philadelphia, the Federal Reserve System Regional conference, and the American Real Estate and Urban Economics Association for their helpful comments. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers/.

Information Losses in Home Purchase Appraisals

Paul S. Calem, Lauren Lambie-Hanson, and Leonard I. Nakamura

1. Introduction

Home appraisals are a standard feature of the U.S. mortgage underwriting process. Yet, since the 1990s, it has been well known that the vast majority of appraisals — typically about nine out of 10 — are at or above the transaction price. Thus, it appears that appraisals are either biased or provide little informational value. Appraisals are supposed to provide the answer to an important informational question millions of times each year. However, as we show, the answers are clearly deficient as a consequence of well-intentioned but poorly designed regulation.

We construct a stylized model to argue that the standard mortgage application review process — under which the loan-to-value ratio is calculated with a home value that is the lesser of the appraised value and the transaction price — can cause an appraisal to have a distortionary impact. As a result, the lending opportunity might be lost when the originally requested loan terms are rejected. This situation creates an incentive for the appraiser to substitute the transaction price for the actual appraised value if the latter is below the transaction price (a “negative appraisal”). We call this substitution, motivated by the potential cost of a lost lending opportunity that accompanies a negative appraisal, *information loss*.

In detail, our model argues that, when the underlying true appraisal is above the transaction price, the appraiser reports the appraisal without bias. However, if the underlying appraisal is below the transaction price, the appraiser substitutes the transaction price or biases the appraisal substantially upward.

We demonstrate that our model can replicate the distribution of the ratios of appraised values to transaction prices observed in the data. Additional support for the model is obtained from an empirical analysis in which we show that the frequency of reported negative appraisals can be predicted by variables that influence the decision of whether to substitute a sales price for an appraised value. Moreover, the factors that contribute to greater information loss in appraisals appear, on balance, to increase the procyclicality of housing booms and busts. We develop this supporting evidence using a unique database that contains nationwide information on single-family home sale transactions and associated appraisals for the period from 2007 through mid-

2012. During this period, more than 20 million home purchase appraisals were conducted in the U.S.²

To our knowledge, our study, together with a companion paper by Ding and Nakamura (2014), is the first to rely on a national sample of presale, premortgage transactions data that includes both reported appraised values and accepted offer prices. Prior empirical studies on appraisals have relied primarily on the Fannie Mae appraisal database. That database is constructed from appraisals after the mortgage has been completed; it does not include appraisals that result in failed mortgages. Thus, the appraisals contained in the Fannie Mae database may show a bias due to selection. Cho and Megbolugbe (1996) pioneered in this area, providing some of the earliest empirical evidence that appraisers rarely report values below the offer prices. Agarwal, Ben-David, and Yao (forthcoming) use the Fannie Mae database to explore the bias of appraisals for refinances, in which there is no accepted offer price to anchor on. Ding and Nakamura (2014), who use the same database as we do in our study, focus on the impact of the 2009 Home Valuation Code of Conduct (HVCC), a regulatory change that sought to reduce appraisal bias.

The role of an appraisal is to provide an independent estimate of the underlying home value that constitutes the collateral for the mortgage loan. The appraisal is especially needed to identify instances in which the accepted offer price may be too high due to fraud or too low due to a less-than-arm's-length relationship, such as a sale to a relative. The true underlying value of the home as collateral is difficult to know because of the uncertainty about the value of the land at the home's location or because of idiosyncratic aspects of the property.³ Recent transactions on nearby properties constitute valuable information about the underlying value and, hence, are a primary input into the appraisal process.

An independent appraisal estimate typically should not equal the accepted offer price exactly because each may be affected by idiosyncratic factors. The accepted offer price may be affected by the parties' respective preferences, knowledge, and bargaining ability, while the

² According to Home Mortgage Disclosure Act (HMDA) data, 20.3 million applications for first-lien purchase mortgages were made between 2007 and 2011. This figure excludes applications that were withdrawn or closed because of incompleteness.

³ The dependence on recent neighboring transactions creates a dynamic information externality, as argued in Lang and Nakamura (1993). When the flow of transactions falters, the precision of an appraisal falls and the loan becomes riskier. The empirical importance of this information externality has been explored in several papers, notably Blackburn and Vermilyea (2007) but also Calem (1996), Ling and Wachter (1998), Avery et al. (1999), and Ding (2014).

appraisal may overlook nonstandard features of the home that affect its market value. We build upon the theoretical approach to appraisals of Quan and Quigley (1991) and Lang and Nakamura (1993), assuming that appraisers use all available information in a Kalman filter, updating to arrive at an optimal (in a mean-squared-loss sense) appraised value and a confidence interval around it.⁴ Under this approach, appraisers take into consideration the accepted offer price together with the information in recent comparable transactions to estimate the underlying true resale value of the property, which is unknown. Our theoretical approach assumes that appraisers determine the optimal appraised value in accordance with this literature but may choose to report a different value, often the transaction price itself.

Ideally, the lender would consider the appraised value relative to the contract price and, in the event the two values diverge, come to an optimal decision whether to alter the amount or terms of the loan in question and by how much. Because of regulatory requirements, however, the value of a property is taken to be the lesser of the transaction price or the appraised value in the mortgage application review process. As a consequence, an appraisal that is below the transaction price can result in the denial of the originally requested mortgage terms. The ultimate consequence could be an unsuccessful transaction, and the cost of the missed lending opportunity could provide an incentive not to report an appraisal that falls short of the transaction price. Basing the approved loan amount on the higher of the two valuations, however, can lead to an increased default rate.

We demonstrate that the model can explain the empirical frequency with which reported appraised values equal accepted offer prices. Empirically, less than 10 percent of reported appraisals are below the accepted offer price. About 50 percent of reported appraisals equal the accepted offer price or fall within 1 percent above it, while roughly 40 percent are 1 percent or more above the accepted offer price. We show that our model can replicate this observed distribution.

According to our framework, the decision whether to report the actual appraised value or substitute the contract price reflects a tradeoff between mitigating default costs by relying on the actual appraisal and the potential for a mortgage application and property transaction failing

⁴ Lang and Nakamura (1993) draw the explicit conclusion, which has been supported by considerable evidence, that the number of recent transactions increases the precision of appraisals. In our empirical analysis, we apply this theory as a basis for the expectation that the variance of appraisals is negatively dependent on the number of recent transactions.

because of a negative appraisal. Thus, factors that reduce the credit risk of the mortgage (e.g., greater expected house price appreciation) should reduce the value of appraisal information with respect to default costs, and thereby strengthen the incentive for information loss.

Our empirical analysis yields evidence consistent with this implication of the model. In particular, we find that rising house prices and decreased foreclosure rates, which we interpret as indicators of lower expected default costs, are associated with increased information loss. We also find that having a high transaction price relative to the median price within the same ZIP code, which we view as increasing the perceived risk of default, appears to weaken the incentive for information loss.

Our empirical analysis also examines relationships between institutional or regulatory factors and information loss. Appraisal management companies (AMCs), by acting as intermediaries between the lender and appraiser, may increase the objectivity of the appraiser by distancing the appraiser from the lender and its incentive to complete the mortgage origination.⁵ Similarly, the HVCC was, as discussed in Ding and Nakamura (2014), an effort to insulate appraisers from bias. Consistent with these expectations, we find a reduction in information loss for appraisals conducted by AMCs, and information loss was less common after the HVCC took effect.

In Section 2, we provide basic information about the institutional framework for appraisals. In Section 3, we present the simple theoretical framework that formalizes the incentive for information loss in appraisals. In Section 4, we discuss the data and report some basic empirical findings on the distribution of reported appraisals relative to accepted offers. We demonstrate that the model can be used to simulate the main features of the empirical distribution. In Sections 5 and 6, we analyze the determinants of information loss, and in Section 7, we offer conclusions and a suggestion for reducing information losses in home purchase appraisals.

2. Institutional Aspects of Appraisals

In the U.S., appraisals must be performed to provide a valuation for collateral — for the purposes of calculating the loan-to-value ratio — when mortgages are to be guaranteed by a

⁵ By definition, appraisal management companies (AMCs) rely on a network of appraisers. This breaks the reliance of any individual appraiser on any particular lender for repeat business, in as much as individual appraisers work for various AMCs or AMCs serve multiple lenders.

government-sponsored enterprise (GSE) (Freddie Mac or Fannie Mae) or the federal government (Federal Housing Administration [FHA] or Department of Veterans Affairs [VA]), or when the mortgages are originated by a federally insured commercial bank or savings and loan institution. The collateral value in these cases is required to be equated to the lesser of the transaction price and the appraised value.⁶

The requirement to value the collateral at the minimum of the appraised value and the sale price has significant implications because the loan-to-value ratio is a crucial indicator of the credit risk of the mortgage, and it determines the interest rate and terms the lender is willing to offer. A lowered home valuation due to an appraisal at or below the transaction price can result in cancellation of the transaction if the home seller is unwilling to lower the sale price, the buyer is unable to provide a larger down payment, or the borrower is unwilling to pay the mortgage insurance premium and/or higher interest rate associated with a low down payment loan.

The New York State Attorney General's office performed an investigation in response to the mortgage crisis and indications that reported appraisals had been biased upward. The outcome was an agreement by the GSEs (Fannie Mae and Freddie Mac) to implement the HVCC in May 2009; an action intended to curtail practices that generate appraisal bias. One of the most significant implications of the HVCC is that it compelled lenders to hire AMCs, rather than working directly with appraisers. Ding and Nakamura (2014) use a difference-in-differences methodology to show that, in the wake of the HVCC, mortgages qualifying for GSE backing showed less bias relative to jumbo loans that were not subject to the HVCC.

Appraisal management companies. AMCs are intermediaries standing between lenders and appraisers, specializing in appraisal quality control and strengthening appraiser independence. As such, AMCs are expected to reduce information loss in appraisals. AMCs proliferated in the wake of the mortgage crisis to reduce the possibility that lenders or realtors might attempt to influence appraisal reports. In particular, many lenders have turned to AMCs to help ensure compliance with the HVCC, with the appraiser independence rules in the Dodd-

⁶ These requirements are enconced in regulations governing the real estate lending activities of federally regulated banking institutions and in the underwriting standards for loans purchased by Fannie Mae and Freddie Mac or insured by the FHA. For example, the table that gives the method for calculating the loan-to-value ratio in the Fannie Mae Selling Guide (2014), pp. 171–172, reads: “Divide the loan amount by the property value. (Property value is the lower of the sales price or the current appraised value.)”

Frank Wall Street Reform and Consumer Protection Act, and with Interagency Guidelines on appraisal conduct.⁷

3. Model

In this section, we address from a theoretical perspective how the requirement to use the lesser of the appraised value and the accepted offer price as the collateral value can lead to a suboptimal outcome when the appraised value is lower. Therefore, the appraiser, serving in the interest of the lender, may substitute the contract price for the actual appraised value in order to mitigate the distortion.

Consider a property under contract such that the buyer and seller have agreed upon a price v_o , the accepted offer price. The buyer applies for a mortgage loan of amount L_o , with pricing and terms determined by the implied loan-to-value ratio $\lambda_o = L_o/v_o$. The lender proceeds to evaluate the mortgage application and commissions an appraisal of the property.

The appraised value of the property, a , would be the appraiser's best estimate of the market value of the property, which is not observable. Likewise, on average, offer prices for comparable properties should equate to their underlying market value.

Ideally, the lender would take into consideration the appraised value a , contract price v_o , and other relevant information to weight the information value of each measure and decide on a final valuation, v , for the purpose of evaluating the loan application. However, the minimum value rule imposes the constraint that $v = \min(v_o, a)$, potentially introducing two types of distortion into the loan approval process. One type is informational: The resulting valuation of the property will be downward biased. The second is transactional: The lender may have to reject the loan application when $a < v_o$, even if the consequent risk benefit is too small to justify the cost of disrupting the transaction. We posit that these two distortions combine to create a strong incentive for the appraiser to act in the service of the lender (and the optimal market outcome) by reporting an appraised value equal to the contract price, in the event that the true appraised value is lower.

Informational impact. One form of distortion is informational. The minimum value rule implies that if both the appraisal and the contract price are unbiased estimates of the underlying

⁷ See, for example, National Association of Realtors (2013).

home value, then a valuation based on the lesser of the two will be biased downward and the bias may be appreciable.

For example, suppose that the appraised value (a) and the accepted offer (v_o) are distributed lognormally relative to the true market value. That is, $\ln a$ and $\ln v_o$ are distributed bivariate normally, with both means \bar{v} equal to the underlying value, with variances σ_a^2 and σ_o^2 , and correlation coefficient ρ . Then the expected value of

$$(1) \min(\ln a, \ln v_o) = \bar{v} - \sqrt{\sigma_a^2 + \sigma_o^2 - 2\rho\sigma_a\sigma_o} \phi(0),$$

where ϕ is the pdf of a standard normal distribution, so that $\phi(0) = 1/\sqrt{2\pi} \approx 0.4$.⁸ Thus, the effect of this rule is to bias low the expected value by 0.4 times the standard deviation of $(\ln a - \ln v_o)$.

Imposing such downward bias on the lender's valuation of the collateral clearly is suboptimal, and when $a < v_o$, the lender is compelled to reject the requested loan amount L_o and offer in its stead an amount L that reflects the downward biased valuation:

$$(2) L = \lambda_0 \times v = \lambda_0 \times \min(v_o, a)$$

Thus, when $a < v_o$, the appraiser has an incentive to mitigate this distortionary impact of the rule by incorporating upward bias (relative to the true appraised value) into the reported appraisal, which we shall denote as \tilde{a} (implying $\tilde{a} > a$).

Note from equation (1) that a greater variance of actual appraised values relative to offer prices implies larger bias from applying the minimum value rule, and thus a stronger impetus to upwardly adjust the reported appraised value when $a < v_o$. Consistent with this notion, we find in our following empirical analysis that, in county-quarter-aggregated data, our proxy for less precise appraisals is positively related to the degree of information loss.

Transactional impact. As noted, when $a < v_o$ the lender is compelled to (1) reject the requested loan amount L_o and offer in its stead the amount $L < L_o$, (2) offer an alternative, higher-priced product characterized by a higher loan-to-value ratio $\lambda > \lambda_0$, or (3) simply reject the application outright. A smaller loan amount increases the likelihood that the property sale will be canceled and the mortgage will not be completed because the seller may be unwilling to reduce the transaction price and the borrower may be unable or unwilling to provide a larger down payment. Likewise, the borrower may be unwilling to bear the higher cost (insurance or risk premium) associated with a higher loan-to-value ratio product.

⁸ Nadarajah and Kotz (2008)

Thus, an appraisal below the accepted offer price increases the likelihood that the mortgage will not be made, in which case a cost is borne by the buyer, seller, and lender. The buyer and seller must bear the cost of renegotiating their contract or resuming their respective searches, and at least part of this cost may be borne by the lender as a reputational cost. The lender will also bear the cost of foregone fee income that would offset the application processing expenses it had already incurred.

We shall denote the expected cost to the lender from loss of the loan as $f(v_o - a)$. If $v_o \leq a$, then $f(v_o - a) = 0$, as then the appraisal does not affect the loan-to-value calculation, and the original loan amount is retained. If $v_o > a$, then $f'(v_o - a) > 0$, since the likelihood of the transaction failing increases as the offered loan amount L declines further.

Again, the appraiser (acting in the lender's interest) can mitigate this cost by incorporating upward bias into the reported appraised value, and reporting $\tilde{a} > a$ in lieu of a , thus reducing the expected cost from potential loss of the transaction to $f(v_o - \tilde{a})$. There will be a tradeoff, however, in deviating from the true appraised value. From equation (2), replacing a with the reported appraisal $\tilde{a} > a$ will result in a higher offered loan amount, lower down payment, and higher risk of default. We denote the increase in potential credit loss as $g(\tilde{a} - a)$, where $g(0) = 0$ and $g'(\tilde{a} - a) > 0$ if $\tilde{a} > a$.

Note that we opt not to complicate the model by introducing an agency problem. We assume that the appraiser fully internalizes the costs faced by the lender and seeks to minimize the total cost. Thus, if $a \geq v_o$, then $\tilde{a} = a$, and if $a < v_o$, then the appraiser seeks to minimize:

$$(3) f(v_o - \tilde{a}) + g(\tilde{a} - a)$$

A deviation from the true appraised value impacts the probability of default and the loss-given-default of the associated mortgage, both of which increase with the rise in loan-to-value ratio, with the impact on expected loss being multiplicative. Thus, it seems plausible to assume that the cost of the misreported appraisal rises more than proportionally to the gap between the reported and true appraisal. In contrast, the primary impact of an appraisal below the transaction price is on the probability that the loan will not be made. This distinction helps to motivate a highly simplified, linear-quadratic version of the model, as follows:

$$g(\tilde{a} - a) = d(\tilde{a} - a)^2$$

$$f(v_o - \tilde{a}) = b(v_o - \tilde{a}) \text{ if } v_o > a \text{ and } f(v_o - \tilde{a}) = 0 \text{ otherwise,}$$

where b and d are strictly positive constants.

Proposition: With these costs, the appraiser determines the reported appraisal as follows:

- i. If $a \geq v_o$, then $\tilde{a} = a$.
- ii. If $a < v_o$ and $a > v_o - b/2d$, then $\tilde{a} = v_o$.
- iii. If $a < v_o - b/2d$, then $\tilde{a} = a + b/2d$.

The first statement is obvious; the other two are straightforward to derive and are proved in the Appendix.

We interpret this model as having three main implications. First, when the true appraisal is greater than the transaction price, then the reported appraisal is equal to the true underlying appraisal. There is no incentive for deviation. Second, when the true underlying appraisal is within a distance of the transaction price such that the expected cost of a canceled transaction exceeds the risk benefit from a reduced loan-to-value ratio, then the reported appraisal is identical to the transaction price. The size of this range depends positively on the perceived cost of losing the transaction and negatively on marginal credit losses. Several of the empirical results described in the following section are consistent with this implication of the model. Third, if the true appraisal is sufficiently below the transaction price, then the reported appraisal will be between the true appraisal and the transaction price, and it will exceed the true appraisal.

In summary, the distribution of appraisals should include an unaltered portion (the appraisals greater than the accepted offer price), which we can test. There should also be a substantial proportion of appraisals precisely at the accepted offer price. The proportion of appraisals precisely at the transaction price should be larger when the cost of the loan application failing is higher, and smaller when the cost of inaccuracy is higher. Finally, the few reported appraisals that fall below the transaction price will be upwardly biased from the tail of the distribution and thus will be highly asymmetrically distributed relative to the reported appraisals above the transaction price.

Note that if the reported appraisal \tilde{a} deviates above a (such that a cannot be precisely inferred from \tilde{a}), then this inaccuracy will reduce the information value of the appraisal. According to our model, negative appraisals are somewhat informative as long as they are interpreted as being biased upward by $b/2d$. However, substantial information is lost when appraisals are set equal to the offer price. These appraisals are biased upward by an unknown amount up to $b/2d$. Assuming, as seems likely, that lenders cannot infer the true appraisal in these cases, information loss occurs.

4. The Empirical Distribution of Appraised Values Relative to Transactions Prices

We explore the model's conclusions using a data set of approximately 800,000 appraisals completed from 2007 through early 2012 on single-family homes across the U.S. The data vendor, a real estate mortgage technology company called FNC, Inc., provides information on the date of each appraisal, the ZIP code of the property, the offer price in the sale contract rounded to the nearest \$50,000, the ratio of the (precise) contract price to the reported appraised value, and a code signifying the lender requesting the appraisal. This lender code distinguishes between appraisals coordinated by an AMC and those contracted directly by the lender.

Using these data, we examine the distribution of reported ratios of appraised values relative to contract prices for elements of consistency with our stylized model. Specifically, we calculate the natural log of the ratio of the reported appraised value to the contract price. We then compare the distribution of these values with those we would expect to observe if reported appraised values never deviate from true appraised values ($\tilde{a} = a$) and the log appraisal-price ratio were normally distributed. Table 1 presents this comparison.

After winsorizing at 1 percent and 99 percent, this distribution has a mean of 0.02 and a standard deviation of 0.07. We also present two lognormal distributions in Table 1. The first is a theoretical lognormal distribution, assuming a mean of zero (indicating that the reported appraisal is unbiased relative to the contract price, consistent with both being unbiased in relation to the underlying value) and the empirical standard deviation of 0.07. The second is a theoretical lognormal whose mean and standard deviation agree with the empirical distribution.

The most striking aspect of the observed distribution relative to the lognormal distributions is that there is a large mass point at exactly zero; approximately one-third of appraisals are identical to the offer price. Also striking is the degree of asymmetry. The right-hand portion of the distribution, where the ratio exceeds zero (reported appraisal exceeds transaction price), has a shape that roughly resembles a normal distribution (although somewhat thicker tailed, as discussed later). Only a small portion of the distribution falls on the left-hand side.

In comparison to the lognormal distributions, the tail of the empirical distribution on the right-hand side, which according to our model should match the true distribution, is somewhat too thick. One plausible interpretation is that the empirical distribution is a mixture of appraisals

with different standard deviations; indeed, a mixture distribution as such would generate thicker tails. Lang and Nakamura (1993) imply that different homes should have appraisals with differing precision, potentially justifying the view that the empirical distribution is best represented as a mixture distribution.⁹

Accordingly, Table 1 also displays corresponding values from a theoretical mixture distribution (labeled “Mix”), with a mean of zero and with half the distribution having a standard deviation of 0.02 and half with 0.10. This fits the right-hand side of the observed distribution reasonably well, though there are still 5 percentage points too many observations falling just above zero (but less than 0.01). Almost all these excess observations, relative to the mixture, however, are quite small, between 0 and 0.005.¹⁰

Finally, the “Left Side” version of the distribution assumes that our model is exactly correct, with $b/2d = 0.08$ and that the underlying distribution has the mixture normal we have just described. That is, it adds 0.08 to the part of the left-side tail of the mixture distribution that falls below -0.08 . This version is also displayed in the right-side panel of Figure 1. Although 6 percentage points too many observations fall exactly at zero, the mass point, this distribution generally fits the data well. Moreover, this excess 6 percentage point share assigned to the mass point at zero by the model is comparable with the excess 5-percentage-point share in the empirical distribution that lies just above the offer price, which, as noted, may represent appraisals subject to a small amount of added noise. The close fit between the modeled and observed distribution in this case supports the theoretical model presented in Section 3 and suggests that, on average in this period, $b/2d$ is equal to 0.08. Thus, this replication exercise suggests that for the most negative values (where offer prices most exceed appraisals), the appraisal-price ratio is biased upward by 0.08.

Implied true appraisal valuation. Under our hypothesis, we can partially reverse-engineer reported appraisals to obtain the true underlying appraisal valuation. On average, if the reported appraisal is less than the accepted offer price, then we should lower it by 8 percent of the accepted offer price to obtain the true appraisal. If an appraisal is reported as being equal to (or just slightly above) the accepted offer price, then we can interpret the true appraisal as being

⁹ Alternatively, the mixture can arise because of differences in the *relative* precision of the appraisal compared with the transaction price as estimates of the property value under differing circumstances.

¹⁰ Thus, the average difference is about \$625 on a median accepted offer price of \$250,000. Some appraisers (in the vicinity of one-sixth of our sample) may choose to produce an appraisal very slightly above the accepted offer price when our model would specify that the appraisal should be exactly at the accepted offer price.

between that value and 8 percent below the accepted offer price. Under our model, this is the information loss to the lender: We lose the fine-grained detail of the underlying true appraisal when the appraisals are bunched at zero or just above it.

5. Panel Regression Analysis of Factors Influencing the Degree of Information Loss

We test the model’s implications on the factors affecting the degree of information loss using two empirical approaches distinguished by their level of aggregation. In this section, we apply a panel regression analysis of the degree of information loss by county and quarter. In Section 6, we conduct a logit analysis of individual appraisal outcomes.

One advantage to aggregating by county and quarter is that we can directly test the relationship between information loss and underlying appraisal variance as measured by the right-hand side of the observed distribution of appraised values relative to contract prices (where $a > v_o$) within each county and quarter. As explained previously, for this portion of the distribution, $\tilde{a} = a$. In addition, the panel regression approach allows us to include county and time fixed effects to control for unobservable factors. The individual outcomes analysis, however, can more fully exploit the variation in the circumstances of individual appraisals, particularly in regard to neighborhood characteristics.

Information loss measure. The dependent variable for our panel regression analysis is a summary measure of information loss by county and quarter in which the appraisal was conducted. The information loss to which we refer is the proportion of the underlying appraisals in a county-quarter whose value is set at the accepted offer price or very slightly above. On the assumption that half the underlying unbiased appraisals should be below the accepted offer price and half above, information loss is 0.5 less the proportion of all appraisals that are below the accepted offer price:

$$Information\ loss_{it} = 0.5 - \frac{number\ of\ negative\ appraisals_{it}}{number\ of\ total\ appraisals_{it}}$$

If the measured information loss is less than zero, it is set equal to zero (this restriction has no qualitative impact).

Our theory implies that this summary measure of information loss is determined by the cost of a lost lending opportunity (represented previously by f), relative to the expected increase in default costs from basing the approved loan amount on a reported appraisal that deviates

above the true appraised value (represented by g). In this context, we test several factors in relation to information loss.

Expectations of house prices rising or falling. Brueckner et al. (2012, forthcoming) argue that because house price inflation is positively serially correlated, a rising house price reduces the expected default cost of a mortgage. Using the previous year's house price appreciation as a proxy for this year's expected appreciation, they find that as expected appreciation rises, the supply of subprime and alternative, riskier mortgage products expands, consistent with the hypothesized relation to expected default cost.

We likewise expect rising house prices to reduce expected default costs and, in turn, strengthen the incentive for information loss. Here, we aggregate ZIP code-level Zillow house price appreciation rates to the county level, weighting by the ZIP code share of the sample population of appraisals, to measure expected mean house price inflation. We specify this, as in Brueckner et al. (forthcoming), as the four-quarter house price inflation rate, lagged four quarters.¹¹

Foreclosure rates. A high rate of foreclosures in a neighborhood is likely to increase the perceived riskiness of mortgage lending for homes in that neighborhood, reducing the incentive for information loss. We calculate the proportion of mortgage loans in the foreclosure process by quarter using McDash mortgage data from Black Knight Financial Services at the ZIP code-level, again aggregated to the county level using the proportion of appraisals in each ZIP code.¹²

Relative price. If a home has a high price relative to its neighbors, there is likely to be more risk that the buyer has overpaid for the house or that fraud is occurring, thus increasing expected default costs and reducing the incentive for information loss. Conversely, when a home has a relatively low price, there is less risk that the buyer has overpaid for the house, and it is possible that it is a less-than-arm's-length transaction. We measure relative price as the percent difference of an appraised home's accepted offer price (as reported in our data, rounded to the nearest \$50,000) from the average single-family home value in that ZIP code (as measured by Zillow). We use the mean for all appraisals in the county-quarter after winsorizing at the first and 99th percentiles.

¹¹ For example, for an appraisal conducted in May 2007, we factor in the house price change between May 2005 and May 2006.

¹² The ZIP code foreclosure rate is the percentage of all loans 90 days or more past due, in foreclosure, or bank owned.

HVCC and AMCs. We control for the impact of the HVCC by including the proportion of appraisals that are not subject to the HVCC, interacted with a dummy variable for dates in or after the third quarter of 2009, when the HVCC took effect. Appraisals not subject to the HVCC are those associated with a loan amount above the GSE (conforming) limit. We can only identify these approximately because observations in which the contract price (reported in our data as rounded to the nearest \$50,000) is more than 1.25 times the local conforming loan limit (on the assumption that a standard mortgage loan has an 80 percent loan-to-value ratio).

As noted, AMCs are intermediaries specializing in appraisal quality control and ensuring appraiser independence. As such, AMCs are expected to reduce information loss in appraisals. We use the proportion of AMC appraisals in a given county-quarter as our measure of the influence of AMCs on information loss.

Underlying appraisal precision. A final right-hand side variable is the underlying variance of the distribution of the appraised values to accepted offer prices (after applying a log transformation to the ratio). Our underlying theory suggests that there is no incentive for appraisers to misreport the underlying appraisal if it is above the accepted offer price. Under this theory, the true underlying appraisal variance can be recovered by measuring the observed variance of the distribution using the appraisals that are greater than the *accepted* offer price. This observed variance is equal to the sum of the mean squared and the variance of the log-transformed ratio, a procedure that assumes that the mean of the appraisals equals the mean of the accepted offer prices.¹³

We have noted that a greater variance of actual appraised values relative to contract prices implies larger bias from applying the minimum value rule and, consequently, stronger impetus to adjust the reported appraised value upwardly when it is less than the accepted offer price. Therefore, we expect to see a positive association between our measures of information loss and the underlying variance of appraised values relative to contract prices.

Model specifications. County fixed effects are included in some specifications, along with quarter fixed effects or a time trend. In specifications where we control for the HVCC, we also include separately the proportion of jumbo loans as an additional variable and, where time dummies are not present, a dummy variable for dates in or after the quarter when the HVCC took

¹³ What is observed is $E(X - 0)^2 = E(X^2)$. Since it is well known that $\text{Var}(X) = E(X^2) - (E(X))^2$, then $E(X^2) = \text{Var}(X) + (E(X))^2$. See, for example, Rice (2007), p. 133.

effect. The regressions are limited to county-quarters with 10 or more total appraisals, and we weight the regressions with frequency weights using the total number of appraisals. Means and standard deviations of the dependent and independent variables are provided in Table 2.

Regression results. The regression estimates are reported in Table 3. The first column in Table 3 shows the coefficient results from our regressions in which we include only the year-over-year house price inflation rate lagged four quarters and the foreclosure frequency. As expected, we find an inverse association between default costs and information loss.

A higher expected house price inflation rate, which should lower default costs, is associated with more information loss. An increase in the house price inflation rate by one standard deviation increases information loss by 3.0 percentage points. Since the standard deviation of information loss is 6.7 percentage points, this is an economically significant amount.

A higher area foreclosure rate, which is expected to increase default costs, is associated with reduced information loss. An increase in the foreclosure rate by one standard deviation reduces information loss by 1.6 percentage points. Note that including foreclosure rates and house price inflation together result in an R-squared of 40 percent, so we are accounting for a large proportion of the movements in information loss with these two variables alone. Moreover, both factors are likely to increase information loss during housing booms, when home price inflation is high and foreclosures are low.

The second column in Table 3 provides the model estimates when we add in the underlying variance of appraisals and the proportion of appraisals conducted by AMCs. The coefficients on house price inflation and foreclosures show little change. The coefficient on the underlying variance of appraisals is positive, as expected. Thus, as underlying appraisals become less precise, appraisers tend to react by increasing information loss. This makes sense conceptually, in that the appraisals are less reliable, but it is worrisome in that it is precisely when information is scarce that the appraisal is most important. An increase in the underlying variance by one standard deviation results in an increase in information loss by 0.8 percentage point.

A higher proportion of AMC appraisals appears to have the desired effect of less information loss occurring. A one-standard-deviation increase in AMC appraisals results in a 0.6-percentage-point decrease in information loss.

The third column in Table 3 presents results after incorporating the control for HVCC impact along with a time trend. These results suggest that once appraisals became subject to the HVCC, information loss decreased for loans under the GSE loan limit relative to jumbo loans. Thus the HVCC appears to have had the desired effect, consistent with Ding and Nakamura (2014). A one-standard-deviation increase in the difference-in-differences variable results in a 0.7-percentage-point decrease in information loss for the affected, non-jumbo loans relative to the jumbo loans.

Also in this regression, there is a notable increase in the strength of the AMC effect, as a one-standard-deviation increase in AMCs now accounts for a one-percentage-point decrease in information loss. Other coefficients are largely unaffected. However, it is notable that the time trend shows a very significant erosion of these gains over time, with information loss trending back toward a higher level. Overall, in this regression, our variables account for some 45 percent of the squared errors, as measured in R-squared. Thus, we are able to account for a large proportion of the movement in information loss.

The fourth column in Table 3 shows the impact of adding state and time dummy variables into the simplest (column one) regression with expected house price inflation and foreclosures. The coefficient of the house price inflation rate drops by about one-third but remains highly significant, while the coefficient on foreclosures remains roughly the same and the R-squared rises to 57 percent.

In columns five and six in Table 3, we first add a set of state dummies (column five) and then state and quarter dummies (column six) to the HVCC difference-in-differences specification (column three). In the latter case, we drop the time trend and the HVCC period dummy, which are vitiated by the full set of time dummies. In our most complete regression, we are able to account for 60 percent of the variation in information loss across county-quarters, as measured by R-squared. Adding state and time dummies does not change the signs of any of our coefficients of interest, so our qualitative conclusions remain intact.

6. Individual Appraisal Analysis

To more thoroughly investigate the role of transaction and neighborhood characteristics on appraisal outcomes, we now turn to an analysis at the appraisal level, focusing on the probability that the appraised value equals the offer price (the event of information loss),

conditional on it being equal to or falling short of the offer price. We relate this outcome to factors viewed as determinants of information loss, and we examine how these relationships may have changed over time. These factors are viewed as indicative of costs of inaccuracy (likelihood and cost of default) or the cost of a lost lending opportunity.

As discussed previously, appraisals exceeding the offer price are assumed to have $\tilde{a} = a$. In other words, the appraisal represents the appraiser’s unbiased estimate of the value of the home, so there is no information loss experienced with these appraisals. The remaining appraisals may be subject to an appraiser’s efforts to improve the chance that the mortgage application is successful. We note from the distribution of appraisal-to-price ratios shown in Table 1 that there is a pile-up of values with appraisals just above the offer price. Because of this, we treat appraised values between 100 percent and 101 percent of the offer price as being identical.¹⁴

With this rounding, about one-half of all appraisals in our data set matched the offer price agreed upon by the buyer and seller, which indicates that information loss was prevalent. As shown in Table 4, while negative appraisals approximately doubled from 5 percent in 2007 to 10 percent to 13 percent in 2009–2012, appraisals equal to the offer price hovered around 50 percent.

We specify a set of logit models to estimate the probability that a particular appraisal, i , “matches” the offer price:

$$prob(v_{oi} \leq \tilde{a}_i < 1.01 v_{oi} \mid \tilde{a} < 1.01 v_{oi}) = \frac{\exp(\beta' x_i)}{1 + \exp(\beta' x_i)}.$$

x_i represents a vector of explanatory variables summarized in Table 4 that are likely to influence the cost of a default or the cost of a missed lending opportunity. We estimate the model five times, once for each appraisal year (2007–2011).

Factors transferred from the panel regressions. Three indicators of default costs included in the county-quarter panel regressions are also included: rate of house price appreciation, foreclosure rates, and relative home value. For the individual appraisal analysis, we measure these at the more granular ZIP code level. We also include dummy variables denoting appraisals coordinated by AMCs and jumbo mortgages.

¹⁴ We present results in the Appendix that show this distinction is not driving our results. Our results are similar when we require that appraisals strictly match the offer price to be considered equivalent.

In particular, we include a dummy variable indicating that the appraised property is located in a ZIP code with a foreclosure rate between 3 percent and 10 percent and a dummy variable indicating a foreclosure rate in excess of 10 percent. ZIP code foreclosure rates of 3 percent and 10 percent in the McDash data from Black Knight correspond to the 50th and 90th percentile values, respectively, for the sample as a whole. We include dummy variables signifying that the price exceeded the median single-family home value in the ZIP code by at least 50 percent or fell short by at least 33 percent, but any additional variation in the relative price may be partly captured by a contract price variable.

Neighborhood and property-specific factors. The individual appraisal analysis is conducive to a more granular analysis of neighborhood or property-specific factors that might affect the appraiser's reporting incentives. In particular, we control for low frequency of home sales in the neighborhood, which may imply greater uncertainty around the appraisal as an indicator of the value of the property, as there will be fewer "comparables" available to the appraiser. As such, there could be larger bias from applying the minimum value rule in neighborhoods with slow sales activity, so appraisers in thin markets may have stronger impetus to counter the bias when the appraised value is below the contract price. Sales activity is measured as the percentage of homes in a given ZIP code sold in the year leading up to the appraisal.

We also include the natural log of the rounded contract price, though we are agnostic about how it impacts information loss. In general, there is more heterogeneity of characteristics among higher-priced homes, which could imply more difficulty for the appraiser in identifying comparable properties and greater uncertainty around the appraisal as an indicator of value of the property. As such, there could be larger bias from applying the minimum value rule to higher-priced homes.

We additionally incorporate several neighborhood- (ZIP code-) level variables derived from Home Mortgage Disclosure Act (HMDA) data. These variables include the share of home purchase loan applications that are for mortgages insured by the government (FHA or VA), the share of applications that are associated with an application for a "piggyback" second lien, and the share of home purchase loan applications that will involve the use of private mortgage insurance (PMI) if originated. The expected relationship to information loss for these variables is ambiguous a priori. The higher preappraisal loan-to-value ratios characterizing these

applications, or the presence of a third party bearing some of the default risk, may affect the incentive to report an appraised value that exceeds the true appraised value.

We also include the percentage of home purchase loan applications in the ZIP code that were filed with local (in-market) depository institutions, defined as institutions with a branch in the county where the sought-after property is located. We also interact this variable with an indicator variable for ZIP codes experiencing substantial house price depreciation (annualized rate of price decline in excess of 10 percent). In-market institutions are expected to have ongoing relationships with appraisers and real estate agents, which, in general, should increase the cost of a lost lending opportunity resulting from a low appraisal, due to reputational impacts and potential disruption of such relationships. Thus, we expect the share of applications going to in-market lenders to be positively related to the likelihood of information loss. However, because of their local presence, these lenders may be more responsive to a declining price market, or in other words, becoming more concerned with default risk than with lending opportunities, thus weakening or reversing this expected relationship.

The final HMDA-derived variable included in the analysis is the percentage of home purchase loan applications in the ZIP code that are for properties located in low- or moderate-income (LMI) census tracts. Loans originated in LMI neighborhoods may have elevated default risk; thus, we expect a lower likelihood of information loss in ZIP codes in which LMI borrowers are more predominant.

Finally, we include a set of state dummies. In each of our logit models, standard errors are clustered by ZIP code.

Estimation results. Table 5 displays the estimated odds ratios and z-statistics from the logit models. The results are consistent with the implications of our theoretical model and our hypothesized relationships for the explanatory variables. Specifically, house price inflation in a given ZIP code, measured as the four-quarter lagged year-over-year rate of change in Zillow median home values, is positively associated with information loss, particularly in 2009 and 2010, well into the housing market downturn. Figure 2 displays fitted probabilities of a reported appraisal matching the offer price, based on different values of the variables, calculated using the median values for the continuous predictors and modal categories for the other controls. The top-left panel of the figure displays the fitted probabilities under three scenarios: house prices rising by 5 percent annually, remaining stable, or falling by 10 percent. The relationship between prices

and appraisal outcomes was weaker in 2007, 2008, and 2011. The weaker results for these years may reflect the limitations of using lagged four-quarter change to represent contemporaneous house price expectations in these years because of the onset of downturn conditions in 2007 and 2008 and the incipient housing market recovery in 2011.

Having greater foreclosure activity in a neighborhood reduces information loss. Appraisals in areas with foreclosure rates of 3 percent to 10 percent of mortgages have 69 percent to 84 percent as much information loss, depending on the year. In areas with foreclosure rates higher than 10 percent, appraised values have 46 percent to 77 percent as much information loss. As shown in the top-right panel of Figure 2, the relationship was strong and fairly consistent in magnitude over time, even though the effect, when represented in odds, varied over the sample period. This also provides evidence of a monotonic relationship between higher area foreclosure rates and less information loss in the appraisals. Interestingly, information loss is prevalent even in areas with high foreclosure rates; our fitted estimate of information loss does not fall below 83 percent of nonpositive appraisals for any year.¹⁵

As expected, appraisals conducted through an AMC were less likely to match the transaction price. When the HVCC took effect in 2009, AMC appraised values were about 80 percent as likely to be identical to the contract price, as appraised values were submitted by appraisers who were hired directly by the lender. However, this gap has narrowed over time, as shown in the lower-left panel of Figure 2.

We find only a very small negative relationship between the frequency of sales in a ZIP code (measured as the percentage of all homes sold in the previous year) and greater incidence of the appraisal matching the contract price. This relationship is largest and most significant in 2010 and 2011. As we note previously, fewer sales mean greater uncertainty around appraisals as indicators of property value, which could strengthen the incentive to report the contract price in place of the true appraised value. A weak neighborhood housing market, however, could also

¹⁵The McDash mortgage data include both prime and nonprime loans, including those securitized and those held in portfolio. However, the data set is not perfectly representative in its composition. In particular, subprime securitized mortgages are less likely to be included, which we expect to result in the McDash measure of foreclosure rates being an underestimate of foreclosure rates in the population. See Table A-2 for a description of a robustness test we conduct using an alternative foreclosure measure from Zillow.

increase concern about credit risk, which would mitigate that incentive. These competing effects may explain the relatively small effect of having few transactions.¹⁶

Having a high relative price (the contract price exceeding the single-family median value in a ZIP code by 50 percent or more) consistently means reduced information loss. Having a higher contract price was also associated with reduced information loss. Finally, considering the variables derived from the HDMA data, we observe the expected relationships to information loss for the share of in-market lenders and its interaction with the indicator for declining house prices. A larger share of loan applications associated with in-market lenders implies increased information loss, except where prices were declining (and in 2007, when the market was about to decline). We note, however, that these effects are not economically large. We also observe the expected inverse relationship between the share of applications for properties in LMI areas and information loss. Finally, we note that appraisals carried out in ZIP codes with a large share of loan applications for FHA and VA programs are less likely to suffer from information loss. In contrast, other types of high-loan-to-value lending in a ZIP code are, in most years, positively correlated with information loss. This finding could indicate that FHA and VA appraisals are higher quality or subject to greater scrutiny. This topic, however, deserves further exploration, ideally using mortgage-level (rather than ZIP code-level) indicators of FHA and VA status.

7. Summary and Conclusion

We have demonstrated that the current mortgage practice of setting the property valuation to the lesser of the transaction price and the appraised value provides incentives for substantial information loss. Although this information loss was somewhat reduced by the implementation of the HVCC and by the advent of AMCs intermediating between lenders and appraisers, information loss continues to be prevalent.

Moreover, information loss is greater during boom times in the housing market, when prices are rising, and smaller during weak markets. These effects were exacerbated by the HVCC. A likely consequence was that the home price boom was extended by these practices and that the home price bust was similarly worsened. Thus, appraisals can be added to the list of

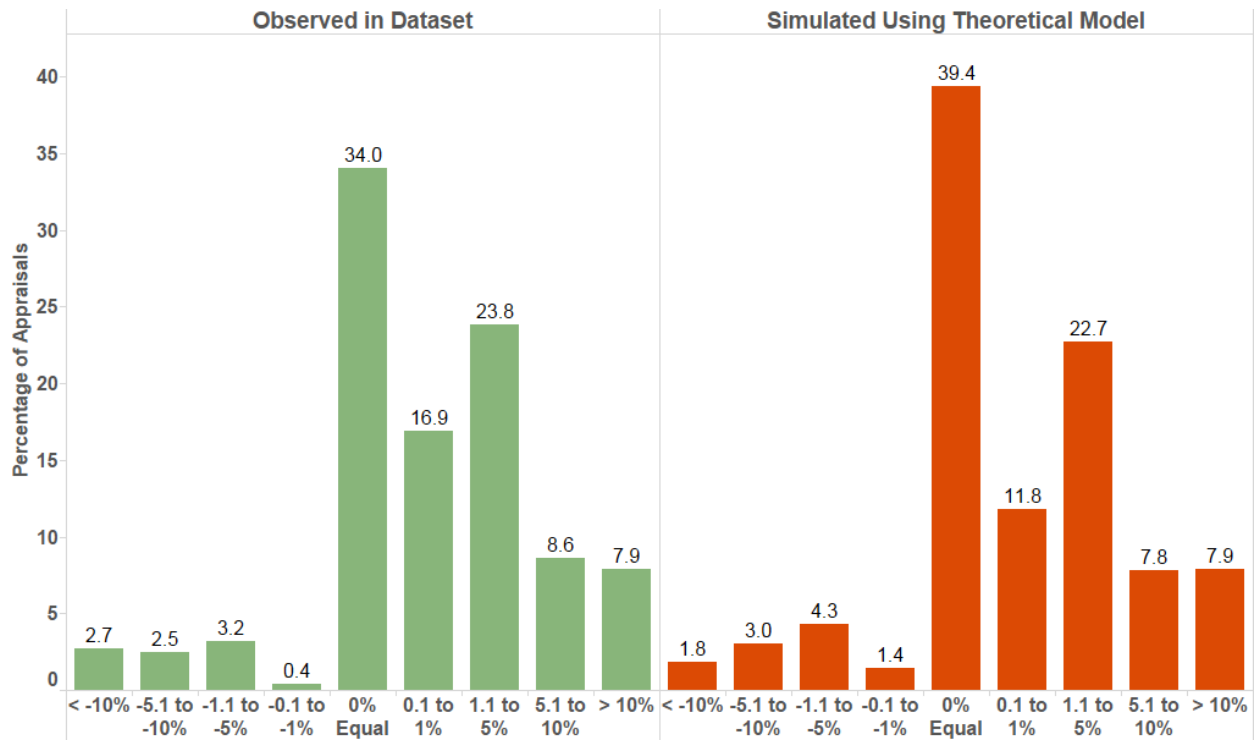
¹⁶ Another factor to consider is that the measure of sales captures all properties — it is not restricted to single-family home sales. As a result, this measure is only a rough proxy for relevant market activity and the availability of comparable sales for the appraisals in our data set. At this time, the share of *single-family* homes sold is not available.

practices that tend to exaggerate the natural home price cycle and, therefore, tend to lead to what are perceived ex post as bubbles and to economic crises.

We have not set up our framework to determine the optimal contract. We believe, however, that we have created a strong argument that the current arrangements are far from optimal. Thus, it might be valuable for securitizers and regulators to reevaluate the method for property valuation and perhaps to engage in experimentation. For example, suppose the property valuations were to be set to equal to transaction price, with the appraisal reported as an additional characteristic of the property. This would likely reduce somewhat the tendency of appraisals to be reported as exactly the transaction price. Over time, this might lead to more accurate, and less biased, appraisals.

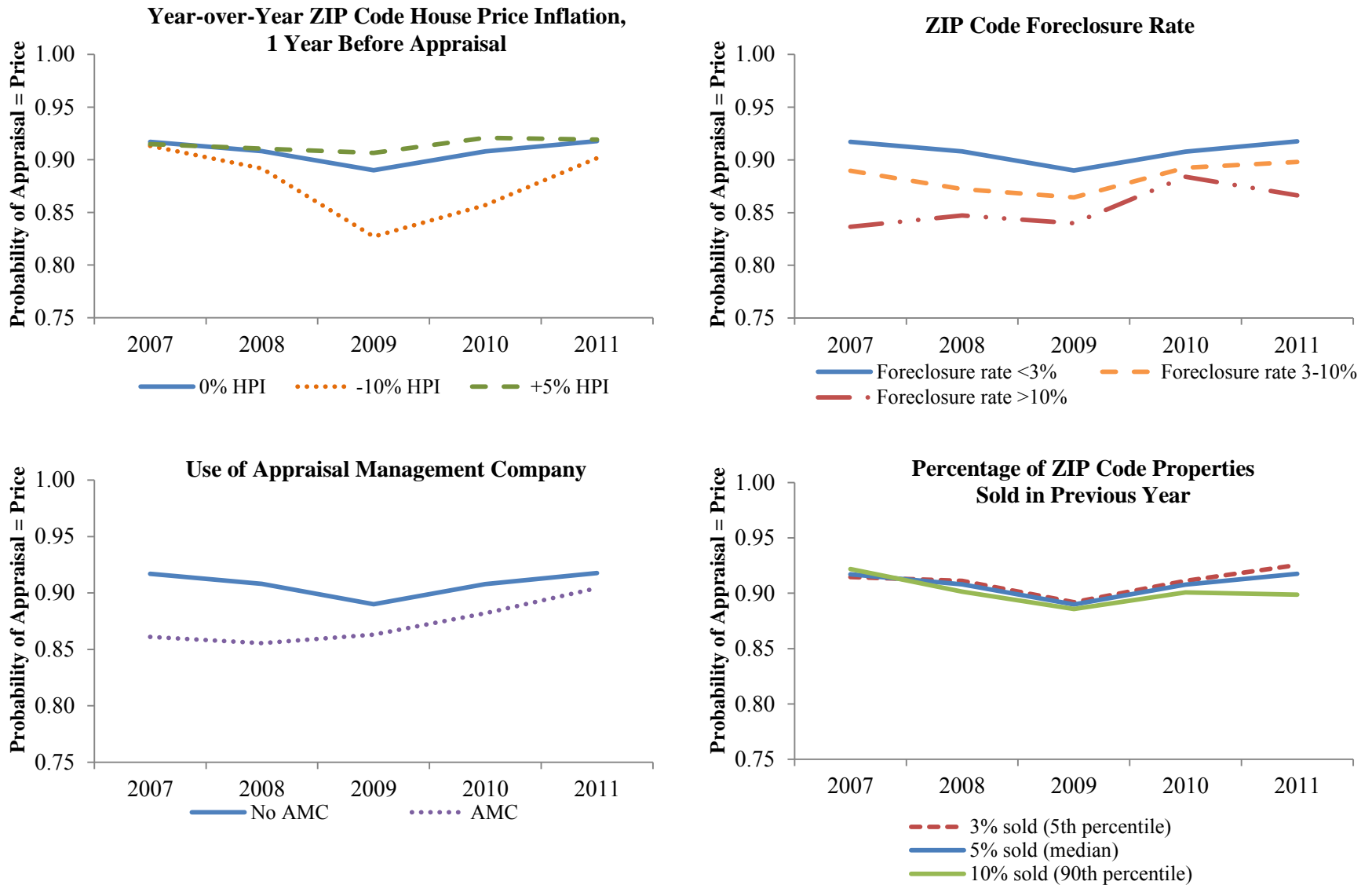
Figures

Figure 1: Comparison of the Distributions of Appraised Values Relative to Contract Prices as Observed in the Data Set and Simulated Using a Stylized Model



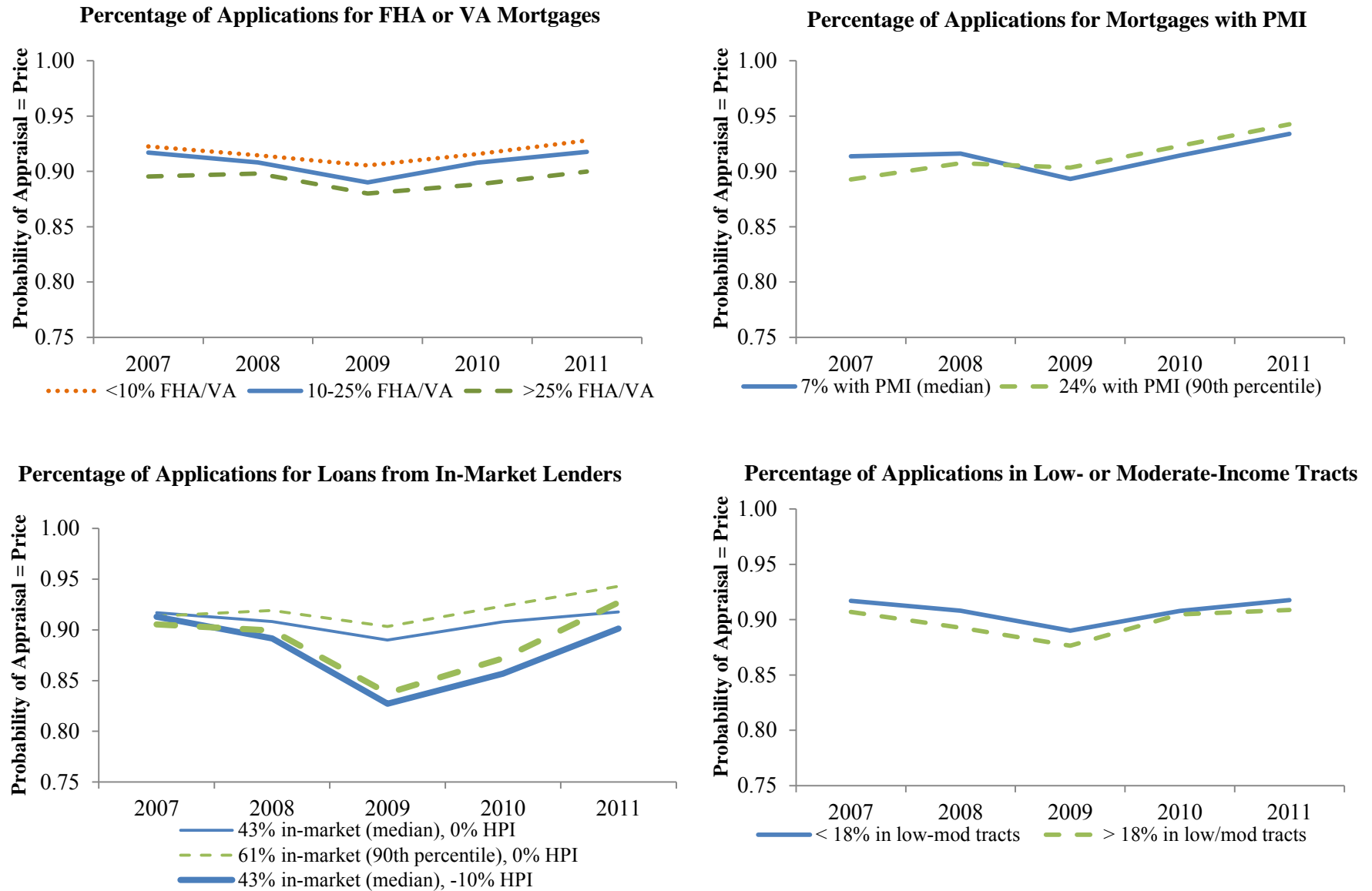
Source: Authors' calculations based on data from FNC

Figure 2: Investigating How Appraisal Characteristics Influence the Prevalence of Information Loss Outcomes Conditional on Appraisal ≤ 1.01 Transaction Price



Note: Information loss is defined as an appraisal value matching the contract price (or exceeding it by no more than 1%), conditional on the appraisal falling below 101% of the transaction price. Fitted values are calculated using median values of covariates, unless indicated otherwise. HPI = house price inflation, AMC = appraisal management company. Source: Authors' calculations based on data from FNC, Zillow, Federal Financial Institutions Examination Council (Home Mortgage Disclosure Act), and Black Knight Financial Services

Figure 3: Investigating How ZIP Code Characteristics Influence the Prevalence of Information Loss Outcomes Conditional on Appraisal ≤ 1.01 Transaction Price



Note: Information loss is defined as an appraisal value matching the contract price (or exceeding it by no more than 1%), conditional on the appraisal falling below 101% of the transaction price. Fitted values are calculated using median values of covariates, unless indicated otherwise. FHA = Federal Housing Administration, VA = Department of Veterans Affairs, PMI = private mortgage insurance, HPI = house price inflation. Source: Authors' calculations based on data from FNC, Zillow, Federal Financial Institutions Examination Council (Home Mortgage Disclosure Act), and Black Knight Financial Services

Tables

Table 1: Distribution of the Natural Log of Appraisals to Price Ratio

	Percent of Values Falling Within Each Band								
	of ln (Appraisal/Offer Price)								
	< -0.1	< -0.05 and ≥ -0.1	< -0.01 and ≥ -0.05	< 0 and ≥ -0.01	Exactly Equal to 0	> 0 and ≤ 0.01	> 0.01 and ≤ 0.05	> 0.05 and ≤ 0.10	> 0.10
<u>Year</u>									
2007	1.2	1.1	1.8	0.4	32.7	19.5	25.7	9.0	8.5
2008	2.1	1.7	2.0	0.3	31.1	16.5	24.6	10.3	11.3
2009	4.3	3.5	3.9	0.4	33.2	15.9	22.7	8.5	7.6
2010	2.9	2.9	3.8	0.5	34.6	17.0	23.4	8.1	6.8
2011	2.5	2.4	3.4	0.5	37.0	16.0	23.4	8.0	6.8
2012	2.6	2.9	4.0	0.6	36.4	16.4	23.3	7.4	6.4
Total	2.7	2.5	3.2	0.4	34.0	16.9	23.8	8.6	7.9
<u>Theoretical</u>									
Normal (0, 0.07)	7.7	16.1	20.6	5.7	0	5.7	20.6	16.1	7.7
Normal (0.02, 0.07)	4.3	11.5	17.5	5.3	0	5.6	22.3	20.8	12.7
Mix	7.9	7.8	22.7	11.6	0	11.8	22.7	7.8	7.9
Left Side	1.8	3.0	4.3	1.4	39.4				

Source: Authors' calculations based on data from FNC

Table 2: County-Quarter Data Summary Statistics, Weighted by Number of Appraisals

	Mean	Standard Deviation	Number of County- Quarter Observations	Number of Appraisals Included
Information loss, percent	40.91	6.65	6,645	573,028
House Price Inflation rate, previous year, percent	-5.52	10.20	6,645	573,028
Foreclosure rate	15.60	15.16	6,645	573,028
Underlying variance of appraisals	0.0054	0.0034	6,642	572,993
AMC proportion	0.1284	0.1845	6,645	573,028
Proportion over GSE limit	0.1048	0.1285	6,645	573,028
Post-HVCC dummy * Proportion over GSE limit	0.0527	0.0807	6,645	573,028

Note: HPI = house price inflation, AMC = appraisal mortgage company, HVCC = Home Valuation Code of Conduct, GSE = government-sponsored enterprise. Source: Authors' calculations based on data from FNC, Zillow, and Black Knight Financial Services

Table 3: Estimating Information Loss by County and Quarter as a Function of Risk, Return, and Regulation

	(1)	(2)	(3)	(4)	(5)	(6)
House Price Inflation Rate,	0.293***	.292***	.292***	.142***	.255***	.156***
Previous year	(.0009)	(.0009)	(.0010)	(.0012)	(.0012)	(.0012)
Foreclosure Rate	-.108***	-.122***	-.119***	-.133***	-.054***	-.105***
	(.0008)	(.0008)	(.0008)	(.0009)	(.0008)	(.0008)
Underlying variance of appraisals		237.5***	215.2***		205.9***	123.8***
		(2.78)	(2.79)		(3.155)	(2.691)
Price relative to ZIP code		-.0225***	-.0236***		-.0438***	-.0420***
		(.0004)	(.0004)		(.0005)	(.0005)
AMC proportion		-3.423***	-5.672***		-4.597***	-4.597***
		(.0941)	(.0607)		(.0947)	(.0947)
Proportion over GSE limit			-.1656***		3.527***	3.111***
			(.0454)		(.0475)	(.0496)
HVCC dummy			-3.514***		-3.425**	
			(.0291)		(.0291)	
HVCC dummy* Proportion over GSE limit			8.843***		8.885***	8.062***
			(.1075)		(.1149)	(.1149)
Time trend			.2300***			
			(.0026)			
Constant	44.21***	44,30***	.6321		10.09***	48.92***
	(.0103)	(.0220)	(.4999)		(.4732)	(.0969)
State dummy variables				✓	✓	✓
Time dummy variables				✓		✓
N	573,028	572,993	572,993	572,993	572,993	572,993
R-square	.3956	.4260	.4481	.5747	.5390	.6002

Note: Observations are weighted by number of appraisals in given county-quarter and robust standard errors are in parentheses. Information loss is defined to be 0.5 minus the proportion of appraisals below the accepted offer price. HPI = home price inflation, AMC = appraisal management company, HVCC = Home Valuation Code of Conduct, GSE = government-sponsored enterprise. *** coefficients are significant at the 0.001% level. Source: Authors' calculations based on data from FNC, Zillow, and Black Knight Financial Services

Table 4: Appraisal-Level Data Summary Statistics
Percentages of observations each year are reported, unless otherwise noted.

	2007	2008	2009	2010	2011	Total
Transaction-Specific Characteristics						
<i>Outcomes</i>						
Negative appraisal ($A < P$)	5	7	13	11	10	10
Appraisal approx. equal to price ($1.01P > A \geq P$)	52	46	49	50	49	49
Positive appraisal ($A \geq 1.01P$)	43	47	38	39	40	41
<i>Controls</i>						
% Appraisal management company	6	6	5	9	28	11
% Appraisal requested directly by lender ⁺	94	94	95	91	72	89
% Jumbo	18	9	7	8	8	9
% Conforming ⁺	82	91	93	92	92	91
Median ln contract price	12.6	12.4	12.4	12.4	12.4	12.4
High relative price ($P > 50\%$ above ZIP SF median)	23	22	22	24	27	23
Contract price similar to ZIP SF median ⁺	71	69	70	69	66	69
Low relative price ($P > 33\%$ below ZIP SF median)	6	9	8	7	7	7
ZIP Code Characteristics						
Median % change in SF price index 24–12 months before appraisal	7	-4	-12	-10	-2	-5
< 3% of mortgages in foreclosure ⁺	83	51	35	32	29	43
3–10% of mortgages in foreclosure	16	37	47	55	58	45
> 10% of mortgages in foreclosure	1	12	18	12	13	12
% of homes sold in 12 months prior to appraisal (median)	7	6	5	5	5	5
< 10% of mortgage applications FHA/VA	72	16	4	5	7	18
10–25% of mortgage applications FHA/VA ⁺	3	59	83	81	75	64
> 25% of mortgage applications FHA/VA	1	12	18	13	13	12
% of mortgage applications requiring PMI (median)	3.1	2.9	1.9	1.6	1.9	2.1
% of mortgage applications with piggyback (median)	2.4	0.7	0.0	0.0	0.0	0.0
% of mortgage applications with in-market lender (median)	32	50	46	43	41	43
% of applications in low- or moderate-income tracts (median)	31	27	25	23	21	25
Observations	119,322	104,267	180,909	169,612	145,520	719,630

ZIP foreclosure is loans 90 days or more past due, in foreclosure, or bank owned. ⁺ indicates categories treated as base cases in Tables 2-4. Due to rounding, percentages may not sum to 100. FHA/VA lending cutoffs were chosen to ensure that each model is estimated using a minimum of 1,800 loans per group per year. A = appraisal value, P = transaction price, AMC = appraisal management company, SF = single family, FHA = Federal Housing Authority, VA = Department of Veterans Affairs, PMI = private mortgage insurance. Source: Authors' calculations based on data from FNC, Zillow, and Black Knight Financial Services

Table 5: Estimating the Prevalence of Information Loss, Defined as the Probability of an Appraisal Equaling the Offer Price (or Being Within 1% Above It), Conditional on Not Exceeding the Offer Price (or Being Within 1% Above It)

	2007	2008	2009	2010	2011
Transaction-Specific Characteristics					
Appraisal management co. dummy	0.561*** (-10.23)	0.599*** (-9.28)	0.779*** (-6.11)	0.759*** (-8.77)	0.849*** (-7.15)
Jumbo mortgage dummy	1.014 (0.27)	0.99 (-0.18)	1.154*** (3.66)	0.959 (-1.00)	1.068 (1.55)
In contract price	0.756*** (-12.88)	0.732*** (-13.09)	0.760*** (-21.89)	0.775*** (-18.27)	0.826*** (-16.65)
High relative price	0.883** (-3.17)	0.737*** (-8.42)	0.838** (-7.81)	0.805*** (-8.91)	0.797*** (-8.90)
Low relative price	1.059 (0.61)	1.095 (1.27)	0.98 (-0.49)	0.887* (-2.56)	0.878** (-2.89)
ZIP Code Characteristics					
House price inflation (percentage)	0.994* (-2.45)	1.006~ (1.88)	1.037*** (18.67)	1.033*** (20.06)	1.004* (2.11)
3-10% foreclosure rate	0.729*** (-6.60)	0.691*** (-8.69)	0.789** (-8.61)	0.842*** (-6.15)	0.791*** (-7.25)
10% + foreclosure rate	0.463*** (-4.66)	0.561*** (-9.25)	0.648*** (-10.41)	0.773*** (-4.92)	0.582*** (-8.26)
% of homes sold	1.014~ (1.96)	0.984* (-2.43)	0.991~ (-1.91)	0.983** (-3.13)	0.953*** (-6.92)
< 10% FHA/VA	1.076~ (1.84)	1.082 (1.54)	1.184** (2.97)	1.102 (1.59)	1.155* (2.54)
> 25% FHA/VA	0.774** (-3.00)	0.892** (-2.91)	0.907** (-2.98)	0.807*** (-5.76)	0.807*** (-6.17)
In % with PMI	0.823*** (-6.50)	0.917*** (-3.37)	1.095*** (6.68)	1.100*** (6.85)	1.130*** (8.31)
In % with piggyback mortgage	1.250*** (8.73)	1.084*** (3.32)	1.071** (3.12)	1.080** (2.87)	1.155*** (4.87)
% with in-market lender	1.002 (0.96)	1.003~ (1.85)	1.005*** (5.13)	1.006*** (5.98)	1.008*** (7.11)
% in-market lender interacted prices falling by 10% +	0.997~ (-1.77)	0.997*** (-3.35)	0.996*** (-8.14)	0.996*** (-5.14)	0.996*** (-4.90)
% of applications in LMI tracts	0.881*** (-3.44)	0.841*** (-4.95)	0.878*** (-5.33)	0.965 (-1.36)	0.894*** (-4.01)
State dummy variables	✓	✓	✓	✓	✓
N	68,055	55,702	111,939	104,322	86,914
Log likelihood	-19,325.2	-20,404.7	-54,279.3	-46,087.6	-3,740.0

Note: Odds ratios are displayed, along with z-statistics in parentheses. Standard errors are clustered by ZIP code. ~, **, and *** represent statistical significance at the 10, 5, 1, and 0.1% levels, respectively. AMC = appraisal management company, HPI = house price inflation, FHA = Federal Housing Administration, VA = Department of Veterans Affairs, PMI = private mortgage insurance, LMI = low- and moderate-income. Source: Authors' calculations based on data from FNC, Zillow, the Federal Financial Institutions Examination Council (Home Mortgage Disclosure Act), and Black Knight Financial Services

References

- Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao. "Collateral Valuation and Institutional Pressures: Evidence from the Residential Real-Estate Market," *Management Science* (forthcoming).
- Avery, Robert B., Patricia E. Beeson, and Mark S. Sniderman, 1999. "Neighborhood Information and Home Mortgage Lending," *Journal of Urban Economics* 45, 287–310.
- Blackburn, McKinlay, and Todd Vermilyea, 2007. "The Role of Information Externalities and Scale Economies in Home Mortgage Lending Decisions," *Journal of Urban Economics* 61, 71–85.
- Brueckner, Jan K., Paul S. Calem, and Leonard I. Nakamura. "House-Price Expectations, Alternative Mortgage Products and Default," *Journal of Money, Credit and Banking* (forthcoming).
- Brueckner, Jan K., Paul S. Calem, and Leonard I. Nakamura, 2012. "Subprime Mortgages and the Housing Bubble," *Journal of Urban Economics* 71, 230–243.
- Calem, Paul S., 1996. "Mortgage Credit Availability in Low- and Moderate-Income Minority Neighborhoods: Are Information Externalities Critical?" *Journal of Real Estate Finance & Economics* 13, 71–89.
- Cho, Man, and Isaac F. Megbolugbe, 1996. "An Empirical Analysis of Property Appraisal and Mortgage Redlining," *Journal of Real Estate Finance & Economics* 13, 45–55.
- Ding, Lei, 2014. "Information Externalities and Residential Mortgage Lending in the Hardest Hit Housing Market: The Case of Detroit," *Cityscape* 16, 233–252.
- Ding, Lei, and Leonard I. Nakamura, 2014. "The Impact of the Home Valuation Code of Conduct on Appraisal and Mortgage Outcomes," Federal Reserve Bank of Philadelphia Working Paper 14-23.
- Fannie Mae, 2014. *Selling Guide: Fannie Mae Single Family*. Washington, D.C.: Fannie Mae, April 15.
- Lang, William W., and Leonard I. Nakamura, 1993. "A Model of Redlining," *Journal of Urban Economics* 33, 223–234.
- Ling, David, and Susan Wachter, 1998. "Information Externalities and Home Mortgage Underwriting," *Journal of Urban Economics* 44, 317–332.
- Nadarajah, S., and S. Kotz, 2008. "Exact Distribution of the Max/Min of Two Gaussian Random Variables," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 16, 210–212.

National Association of Realtors, 2013. NAR Issue Brief: Appraisal Management Company Q&A, July 2, 2013, Washington, D.C.

Quan, Daniel C., and John M. Quigley, 1991. "Price Formation and the Appraisal Function in Real Estate Markets," *Journal of Real Estate Finance & Economics* 4, 127–146.

Rice, John A., 2007. *Mathematical Statistics and Data Analysis*, 3rd ed., Thomson Brooks/Cole: Belmont, CA.

Appendix

Proof of the proposition.

The goal is to minimize the total cost (C):

$$C = d(\tilde{a} - a)^2 + \max(b(v_o - a), 0)$$

If $a \geq v_o$, then C is minimized with $\tilde{a} = a$, where $C = 0$, establishing (i).

Now note that in regions where $v_o > a$, C is strictly positive, with:

$$C = d(\tilde{a} - a)^2 + b(v_o - \tilde{a})$$

$$\frac{dC}{d\tilde{a}} = 2d(\tilde{a} - a) - b = 0$$

implies $\tilde{a} = a + b/2d$, is a local minimum as:

$$\frac{d^2C}{d\tilde{a}^2} = 2d > 0$$

If $a < v_o$, then if the appraiser reports (ii), $\tilde{a} = a + b/2d$, total cost is:

$$C = \frac{b^2}{4d} + b(v_o - a - b/2d) = b(v_o - a) - b^2/4d$$

On the other hand, if the appraiser reports (iii), $\tilde{a} = v_o$, then

$$C = d(v_o - a)^2$$

The minimum cost of these two is then (ii) when

$$d(a - v_o)^2 > b(v_o - a) - b^2/4d$$

$$(a - v_o)^2 - \frac{b}{d(v_o - a)} + \frac{b^2}{4d^2} > 0$$

$$(v_o - a - b/2d)^2 > 0$$

$$v_o - a > b/2d$$

And conversely, (iii) is the minimum cost of the two when this does not hold.

Table A-1: Estimating the Prevalence of Information Loss, Defined as the Probability of an Appraisal Exactly Equaling the Offer Price, Conditional on Not Exceeding the Offer Price

	2007	2008	2009	2010	2011
Transaction-Specific Characteristics					
Appraisal management co. dummy	0.570 ^{***} (-9.39)	0.623 ^{***} (-7.88)	0.796 ^{***} (-5.25)	0.765 ^{***} (-8.10)	0.870 ^{***} (-5.83)
Jumbo mortgage dummy	1.110 [~] (1.87)	1.129 [*] (1.97)	1.296 ^{***} (6.26)	1.108 [*] (2.33)	1.173 ^{***} (3.61)
ln contract price	0.710 ^{***} (-13.69)	0.712 ^{***} (-15.07)	0.737 ^{***} (-24.78)	0.750 ^{***} (-21.28)	0.798 ^{***} (-19.50)
High relative price	0.841 ^{***} (-4.27)	0.675 ^{***} (-10.01)	0.780 ^{***} (-10.63)	0.734 ^{***} (-12.68)	0.716 ^{***} (-12.66)
Low relative price	1.217 [*] (2.15)	1.198 [*] (2.48)	1.092 [*] (2.16)	1.00 0.00	0.983 (-0.37)
ZIP Code Characteristics					
House price inflation (percentage)	0.999 (-0.36)	1.004 (1.16)	1.029 ^{***} (14.36)	1.027 ^{***} (16.19)	1.004 [~] (1.84)
3-10% foreclosure rate	0.821 ^{***} (-3.99)	0.762 ^{***} (-6.00)	0.831 ^{***} (-6.30)	0.895 ^{***} (-3.68)	0.862 ^{***} (-4.47)
10% + foreclosure rate	0.754 [*] (-2.11)	0.695 ^{***} (-5.43)	0.724 ^{***} (-7.29)	0.925 (-1.41)	0.727 ^{***} (-4.70)
% of homes sold	0.988 (-1.60)	0.959 ^{***} (-5.85)	0.968 ^{***} (-6.21)	0.953 ^{***} (-7.34)	0.934 ^{***} (-9.45)
< 10% FHA/VA	1.201 ^{***} (4.35)	1.112 [~] (1.93)	1.254 ^{***} (3.87)	1.141 [*] (2.07)	1.207 ^{**} (3.23)
> 25% FHA/VA	0.752 ^{**} (-3.11)	0.828 ^{***} (-4.61)	0.867 ^{**} (-4.09)	0.770 ^{***} (-7.03)	0.772 ^{***} (-7.33)
ln % with PMI	0.752 ^{***} (-9.10)	0.897 ^{***} (-3.98)	1.105 ^{***} (7.12)	1.097 ^{***} (6.38)	1.120 ^{***} (7.36)
ln % with piggyback mortgage	1.328 ^{***} (10.81)	1.044 [~] (1.67)	1.067 ^{**} (2.79)	1.089 ^{**} (3.02)	1.160 ^{***} (4.80)
% with in-market lender	1.003 [~] (1.87)	1.007 ^{***} (4.27)	1.007 ^{***} (5.98)	1.006 ^{***} (5.61)	1.009 ^{***} (8.45)
% in-market lender interacted prices falling by 10% +	0.996 ^{**} (-2.81)	0.998 [*] (-2.16)	0.997 ^{***} (-6.00)	0.996 ^{***} (-4.76)	0.997 ^{***} (-4.36)
% of applications in LMI tracts	0.976 (-0.63)	0.873 ^{***} (-3.61)	0.926 ^{**} (-3.12)	0.987 (-0.46)	0.943 [*] (-2.06)
State dummy variables	✓	✓	✓	✓	✓
N	42,613	37,381	82,442	73,678	60,388
Log likelihood	-16,027.3	-16,853.8	-45,852.9	-38,577.6	-31,105.2

Odds ratios are displayed, along with z-statistics in parentheses. ~, *, **, and *** represent statistical significance at the 10, 5, 1, and 0.1% levels, respectively. AMC = appraisal management company, HPI = house price inflation, FHA = Federal Housing Administration, VA = Department of Veterans Affairs, PMI = private mortgage insurance, LMI = low- and moderate-income. Source: Authors' calculations based on data from FNC, Zillow, the Federal Financial Institutions Examination Council (Home Mortgage Disclosure Act), and Black Knight Financial Services

Table A-2: Examining Different Controls for Area Foreclosure Rates in Estimating the Prevalence of Information Loss, Defined as the Probability of an Appraisal equaling the Offer Price (or Being Within 1% Above It), Conditional on Not Exceeding the Offer Price (or Being Within 1% Above It)

	2007	2008	2009	2010	2011
McDash measure, using full sample (main model results)					
3-10% foreclosure rate (stock)	0.729 ^{***} (-6.60)	0.691 ^{***} (-8.69)	0.789 ^{***} (-8.61)	0.842 ^{***} (-6.15)	0.791 ^{***} (-7.25)
10% + foreclosure rate (stock)	0.463 ^{***} (-4.66)	0.561 ^{***} (-9.25)	0.648 ^{***} (-10.41)	0.773 ^{***} (-4.92)	0.582 ^{***} (-8.26)
N	68,055	55,702	111,939	104,322	86,914
McDash measure, restricting sample to observations with Zillow foreclosure data					
3-10% foreclosure rate (stock)	0.739 ^{***} (-4.72)	0.749 ^{***} (-5.41)	0.814 ^{***} (-6.31)	0.873 ^{***} (-4.17)	0.813 ^{***} (-5.25)
10% + foreclosure rate (stock)	0.558 ^{**} (-2.76)	0.627 ^{***} (-6.00)	0.713 ^{***} (-6.83)	0.797 ^{**} (-3.14)	0.674 ^{***} (-3.31)
N	41,558	40,269	81,078	76,120	59,948
Zillow measure (foreclosures completed per 10,000 homes in ZIP)					
9-37 foreclosures completed (flow)	0.760 ^{***} (-4.13)	0.863 [*] (-2.57)	0.839 ^{***} (-5.12)	0.859 ^{***} (-4.35)	0.777 ^{***} (-6.56)
> 37 foreclosures completed (flow)	0.626 ^{**} (-2.68)	0.724 ^{***} (-3.99)	0.811 ^{***} (-3.87)	0.871 [*] (-2.06)	0.655 ^{***} (-5.80)
N	41,558	40,269	81,078	76,120	59,948

Odds ratios are displayed, along with z-statistics in parentheses. ~, *, **, and *** represent statistical significance at the 10, 5, 1, and 0.1% levels, respectively. ZIP code foreclosure rates of 3 and 10% in the McDash data from Black Knight correspond to the 50th and 90th percentile values, respectively, for the sample as a whole. Correspondingly, 9 and 37 foreclosures completed per 10,000 homes represent the 50th and 90th percentile values in the Zillow data. Source: Authors' calculations based on data from FNC, Zillow, the Federal Financial Institutions Examination Council (Home Mortgage Disclosure Act), and Black Knight Financial Services

The McDash mortgage data include both prime and nonprime loans, including those securitized and those held in portfolio. However, the data set is not perfectly representative in its composition. In particular, subprime securitized mortgages are less likely to be included, which means the McDash measure of foreclosure rates are likely an underestimate of foreclosure rates in the population, particularly in areas with high rates of subprime lending. As a robustness check, we display the foreclosure odds ratios in Table A-2 using the McDash data and the full sample used in the analysis, followed by a comparison of the McDash foreclosure measure and Zillow's rate of foreclosures (calculated as the number of *completed* foreclosures per 10,000 homes), estimated using the subset of the sample for which both McDash and Zillow data are available.