A Relationship Among Neighborhood Traits, Home Sales and Mortgage Fraud: The Atlanta Market Leading Into the Mortgage Crash of 2008

M. Cary Collins
Providence College

Ann Fulmer
Paladin Advisory Services

Keith D. Harvey
Boise State University

Peter J. Nigro
Bryant University

Reports of mortgage loan fraud grew exponentially in the years leading up to the 2008 financial crisis. We examine simple correlates between mortgage fraud and the economic, credit and loan traits within Atlanta zip codes. First, we find that higher home values and lower rates of owner-occupancy lead to higher levels of fraud. Second, we find that mortgage fraud is not correlated with the proportion of non-bank lenders. Third, we find that higher unemployment rates are positively correlated with higher levels of fraud. Finally, we find no correlation between the historic default rate and the prevalence of mortgage fraud.

INTRODUCTION

The financial crisis of 2008 had significant effects on the U.S and global economies. Various causes of the meltdown have been cited, including lax regulation over mortgage lending; poor risk management and weak fair value accounting rules; the absence of oversight for rating agencies, Fannie Mae and Freddie Mac; and the rise of derivative mortgage instruments such as collateralized debt obligations and credit default swaps. On the other hand, due primarily to data limitations, mortgage fraud has received little attention in academic literature, either for its role or impact on the crisis.

The U.S. Federal Bureau of Investigation (FBI) categorizes mortgage fraud as either fraud for property or fraud for profit. Fraud for the property, which the FBI estimates represent about 20 percent of all cases, involves misrepresentation solely by the borrower, usually on a single loan. In these cases, the borrower intends to occupy the home and repay the loan but misstates information the lender relies on in making the loan decision. Items such as income, debt, and property value are misrepresented to improve the likelihood of obtaining the loan.[1] Fraud for profit, which includes the vast majority of cases, may involve several parties to the transaction, including professional agents such as brokers, appraisers, builders, title examiners, escrow officers, attorney, and lenders. Common misrepresentations include inflating the appraised property value, altering the borrower credit history and debt commitments, misrepresenting property purpose as owner occupancy rather than investment, and stolen identity, among others. Fraud for profit schemes usually include several transactions where homes are quickly sold with the lenders and subsequent purchasers suffering the losses.

The consequences of mortgage fraud are felt by all participants in the mortgage market, and involve both direct and indirect costs. First and even before the mortgage fraud is discovered, deceptive behaviors harm market participants:

- Reduced affordability for new purchasers during a period of fraud for profit where housing market values are artificially inflated; and locking in homeowners who tap into artificially inflated equity values through refinancing, creating an upside-down situation with a housing market value correction.

Second, behaviors of home market participants during fraud discovery include:

- Mortgage lender losses on foreclosed homes in fraud for profit neighborhoods;
- neighboring homeowners losing property value as local foreclosure rates rise and the nature of appraisal fraud is made transparent; and
- remaining neighborhood property owners are unable to sell their homes -- even at reasonable prices -- after fraud appears in their neighborhood.

Finally, mortgage market behaviors after fraud discovery include:

- Lenders exiting housing markets or neighborhoods where fraud makes home values questionable; and
- rising lender losses associated with the repurchase of fraudulent properties previously bundled & sold in the secondary market.

Reports of mortgage loan fraud grew exponentially in the years leading up to the 2008 financial crisis, becoming one the fastest growing white-collar crimes in the United States. The FBI estimates mortgage fraud losses exceeded $1 billion in 2005, compared with only $429 million in 2004. The FBI has already identified significant fraud levels in 26 out of 50 U.S. states.[2] The 2006 Mortgage Loan Fraud Report by the Financial Crimes Enforcement Network (FinCEN) indicates the trend continued with 7,093 Suspicious Activity Reports (SARs) associated with mortgage fraud during the first quarter 2006 alone, representing a 35 percent increase over the same period in 2005.

Rapid mortgage fraud growth led the FBI to announce in 2008 an ongoing fraud probe of 26 financial firms, including Fannie Mae, Freddie Mac, Bear Stearns, Lehman Brothers, and AIG to examine the role of mortgage fraud in the subprime mortgage crisis.[3] In 2009, President Obama formed a mortgage fraud task force to investigate wrongdoing while working closely with Attorney Generals in individual U.S. states. By late 2014, the task force recovered $4.65 billion in federal funds in the wake of the mortgage crisis; including $1.85 billion from Bank of America Corporation, $614 million from JPMorgan Chase, $428 million from SunTrust Mortgage Inc. and $200 million from U.S. Bank.[4]

The mortgage fraud focus extends well beyond past fraudulent behavior. In addition to substantial financial penalties, the housing GSE's have increased their focus on preventing mortgage fraud. For example, Fannie Mae has a dedicated team that works to help both lenders (and ultimately homeowners) avoid potential fraud. The fraud team addresses conferences, provides direct training, and leads web seminars on fraud prevention. Fannie Mae also maintains a Mortgage Fraud Prevention webpage, which has been viewed more than 73,000 times since 2009. Similarly, Freddie Mac's Financial Fraud Investigation unit was expanded just after the financial crisis with the goal of assisting law enforcement in criminal investigations and making incident referrals to state licensing and regulatory boards, participating in national, regional, and industry-created mortgage-fraud task forces and providing training to the FBI, state regulators, lenders, and Freddie Mac employees.

Mortgage fraud preventive efforts are bearing fruit. For the first time in over a decade, a 2012 Financial Crimes Enforcement Network (FinCEN) report of Mortgage Fraud SAR Filings in Calendar

Year 2012, showed a decline of 25 percent in 2012 (from 92,561 to 69,277) as compared to the previous year. FinCen stated, however, that the previous three years of suspected mortgage fraud suspicious activity reports (SARs) accounted for almost one-half (approximately 46 percent) of the past decade's mortgage fraud SARs.[5]

Using the Atlanta, Georgia metropolitan area from 2005 through 2008, we examine simple correlates between mortgage fraud and the economic, credit and loan traits within the Atlanta zip codes. This is the first paper to examine mortgage fraud in a broader context given the lack of any readily available fraud data. Using the ChoicePoint® data on mortgage fraud instances, we help fill a significant literature gap and find several impressive results. First, the findings for Atlanta suggest that higher home values and a lower rate of owner-occupancy lead to higher levels of mortgage fraud. Second, we find that mortgage fraud is not correlated with the proportion of non-bank lenders in a zip code. Third, we find that zip codes with more trying economic conditions, at least as measured by the unemployment rate at the zip-code level, are positively correlated with higher levels of fraud. Finally, surprisingly we find no correlation between the historic zip code level default rate and the prevalence of mortgage fraud.

The remainder of the paper proceeds as follows. In Section II, we describe the data employed in the paper, including a description of reported mortgage fraud for limited geography, as well as a brief description of several of the variables employed. Section II also contains the descriptive statistics for the created variables. Section III provides the empirical results, while Section IV provides conclusions and generalizations.

DATA

Using the Atlanta, Georgia metropolitan area as our marketplace, we assess mortgage fraud's longer-term impact using several data sources. To control for reported mortgage fraud frequency and type by geography, we use ChoicePoint® data which document the instance of mortgage fraud for profit by zip code in the Atlanta MSA, enabling us to control for fraud likelihood when creating neighborhoods with and without high fraud instance. The ChoicePoint® database includes both public disciplinary and enforcement actions in the mortgage industry as well as non-public information on alleged fraud or other serious misconduct.[6] Since these report counts are kept confidential, we were provided the percentage of SARs reported in each of the Atlanta city zip codes at the three-digit zip-code level. As such, each percentage must be reweighted based on the total number of housing units in the zip code. The resulting fraud frequencies are then comparable across zip codes.

We control for differences in neighborhood traits by using (a) the U.S. Home Mortgage Disclosure Act (HMDA) annual data; and (b) the U.S. Census Bureau (USCB) data. We also use UFA's ForeScore(TM) index to capture default risk at the zip code level since the score is analogous to a credit score for the zip code level. We also employ U.S. Department of Housing & Urban Development (HUD) annual median family income updates to control for income differences by neighborhood. The goal is to evaluate neighborhoods controlling for similar occupancy rates, income traits, turnover rates, age of housing stock, and default likelihood *be*fore the fraud for property revelations. The HMDA data provide a mortgage 'credit flow' estimate at the census tract level over time, while also providing critical information on non-bank lending. The HMDA data also provide information on the level of high-cost loans (loans with HOEPA-reportable, or high, rate spreads over comparable U.S. Treasury securities) in a zip code area.[7]

The USCB data allow us to control for the housing stock, the age of the housing stock, turnover or average tenure in the neighborhood, as well as labor & employment patterns. These variables are necessary controls in the evaluation of a particular shock to the housing market, namely neighborhood fraud for profit.

The HUD annual median family income updates to USCB data enable us to account for changes in neighborhoods year-to-year. This set of information is most influential in geographies that are undergoing significant changes – either through blight or re-gentrification. In either case, these neighborhoods may represent the most likely areas for fraud for profit.

Table 1 provides descriptive information on the economic characteristics of residents in the Atlanta MSA by three-digit zip code cluster. There is a wide variance in values for almost every demographic characteristic by zip. For example, zip code cluster 300XX has the highest median family income ($58,727), while cluster 305XX has the highest median home value. Unemployment rates (3.3% to 9.2%) and owner occupancy rates (45.3% to 78.6%) also vary dramatically across the zip clusters. Zip code 303 stands out regarding unemployment percentage, reflecting an unemployment rate that is almost three times that of the other zip code groups while also having the lowest owner-occupied housing rate, 45.3 % versus roughly 70 % for the other geographies.

Panel A of Table 2 provides actual default by zip code cluster using UFA ForeScore data. The panel shows that the 302XX zip code cluster has the highest default probability for every term. Panel B of Table 2 shows the fraud rates by zip code cluster using the ChoicePoint fraud data. Fraud levels vary dramatically by zip code cluster and over time. Zip code 302 has the highest historical default rates, while zip code 305 has the lowest. These traits remain even across time. Also, we find in Panel 2 that zip codes 300 and 302 have higher SARs reporting rates than the other neighborhoods, but they also have two of the highest populations among the available six. This Fraud rate, however, is the rate unadjusted for the influence of population.

Table 3 employs the annual HMDA data to calculate the percentage of high costs and non-bank loans in every zip code cluster over time for these six neighborhoods. In Panel A1 we see that no zip code is immune to high-cost lending. The extremes change from one zip code to another in each year. In Panel B we find that zip code 303 has the consistently lowest use of non-bank lending when compared with the other five neighborhoods. That is, the smallest number of originations from non-banks happens in zip code 303. In each neighborhood, however, the non-bank lending presence is shown. At no time and in no zip code is the proportion of non-bank lending lower than one in five originations (see zip code 303). In some instances, the origination rate by non-banks nears one in two originations (see zip code 305 and 306).

EMPIRICAL RESULTS

In this section, we examine the differences in traits of high fraud and low fraud zip code clusters. Second, we examine the correlation between Relative Fraud and these potential explanatory factors in a simple correlation analysis. Relative Fraud represents a corrected weighting of all Suspicious Activity Reports (SARs) for mortgage lending by correcting these counts for the population in the geography provided in Table 1. As such, the incidence of relative fraud at the zip-code level is independent of the zip code's level of population.

To construct the high fraud and low fraud zips, we separate the six three-digit zip code clusters representing approximately 2.6 million people in the city of Atlanta, Georgia, into two distinct groups. Those with above-average fraud rates and those with below-average fraud rates. We labeled these groups "high-fraud" means, and "low-fraud" means, respectively. Table 4 shows there is little difference between the two groups regarding median family income or unemployment rate after each has been appropriately weighted. The first material differences between the two groups were the substantially lower owner-occupancy rate for the high fraud group and the substantially higher median home values for the high fraud group of neighborhoods. For example, the homes in the high-fraud group have median values more than $40,000 higher than the low-fraud neighborhoods.

Table 4 also finds that high-fraud and low-fraud neighborhoods do not demonstrate material differences in their rate of high-cost loan originations or the use of non-bank lenders. This finding runs contrary to the reflexive notion that mortgage fraud and mortgage default have same underpinnings. That is, mortgage default is more likely among non-bank originated loans[8], although our data do not support that correlation. Using the ForeScore zip-code default index, we find default is surprisingly lower in high-fraud neighborhoods than in low-fraud neighborhoods -- an outcome contrary to conventional wisdom. Of course, the SARs incidence rates – even after weighting and adjustment – are almost four times higher for high-fraud neighborhoods (24.6%) than their low-fraud neighborhoods (6.6%).

Unfortunately, given the small set of data for the city of Atlanta, we are unable to generalize these results to other cities or MSAs. To close our review and analysis, we undertook a comparison of relative fraud rates or SARs reporting rates with the set of independent and control variables as shown in Table 5. We find four variables significantly correlated with relative fraud. First, the unemployment rate is positively correlated with the rate of fraud reporting, while the rate of owner-occupied housing is negatively correlated with the rate of fraud reporting. While generalizations are not possible, Atlanta's fraud neighborhoods possess a higher 'mix' of housing types and are not immune to higher rates of unemployment. Not surprisingly, the relative fraud rate is highly and positively correlated with the median housing values in each neighborhood. That is, mortgage fraud happens most in areas with higher home values.

Table 5 also shows that higher cost loan origination rates (as measured by HOEPA loans) are not associated with higher rates of fraud. The higher rates of high-cost loan originations are correlated with lower rates of fraud for these six neighborhoods in the city of Atlanta.

CONCLUSIONS AND GENERALIZATIONS

These findings for the city of Atlanta lend support to the notion that higher home values and a lower rate of owner-occupancy lead to higher levels of mortgage fraud. Surprisingly, the results also suggest that fraud, even relative fraud, is not correlated with the proportion of non-bank loan originations in the neighborhood. This finding does not support the broader belief that non-bank lending led to higher mortgage default. It clarifies the differences between mortgage default and mortgage fraud. That is, mortgage fraud is an attempt at "fraud for profit." As such, neighborhoods with higher priced homes and lower owner-occupancy rates are ideal. The loan program, whether high-rate or low-rate, is immaterial to the fraud's perpetrator since no payments – or at most a few payments – were to be made no matter the loan program.

The descriptive information and correlations presented on mortgage fraud in Atlanta are the first steps in getting a handle on its determinants. More research, however, needs to be done that goes beyond correlations in a limited geographic area. Unfortunately, however, the data for a more sophisticated and broad analysis is not yet available. This paper, however, provides an essential first step in examining correlations between mortgage fraud and lender type, high-cost lending, and the economic characteristics of the borrower's geography.

ENDNOTES

1. Several research papers examine the limited area of mortgage fraud related to income overstatement on loan applications. The "liar loan" papers include Blackburn and Vermilyea (2012) and Mian and Sufi (2015).
2. Financial Crimes Report to the Public, The U.S. Federal Bureau of Investigation, (May), 2005, http://www.fbi.gov/publications/financial/fcs_report052005/fcs_report052005.htm#d1.
3. This is in addition to the thousands of "retail" fraud probes underway. See "The FBI Ramps Up: Mortgage-Fraud Probes Go Bigtime," The Wall Street Journal, September 24, 2008.
4. See, http://www.justice.gov/opa/pr/justice-department-recovers-nearly-6-billion-false-claims-act-cases-fiscal-year-2014
5. See http://www.fincen.gov/news_room/nr/html/20130820.html.
6. The ChoicePoint® data include Suspicious Activity Reports (SARs) for loans reported with Application Fraud, Tax Return/ Financial Statement Fraud, Appraisal/ Valuation Fraud, Verification of Employment Fraud, Verification of Deposit/Bank Account Fraud, Escrow/Closing Document Fraud, and Credit History Fraud.
7. The Home Ownership and Equity Protection Act of 1994 (HOEPA) addresses certain deceptive and unfair practices in home equity lending by amending the Truth in Lending Act (TILA) to establish reporting requirements for certain loans with high rates and high fees.

8. Several studies demonstrate that loans originated by mortgage brokers (non-bank) are significantly more likely to go into default than those originated by the retail arm of the bank (Ding et al. 2008; Laderman and Reid 2008).

ACKNOWLEDGEMENT

The authors wish to thank Brenna Williams for her research and editing assistance with this manuscript. Any and all errors are the responsibility of the authors.

REFERENCES

Blackburn, McKinley L. & Vermilyea, Todd, 2012. "The prevalence and impact of misstated incomes on mortgage loan applications," Journal of Housing Economics, Elsevier, vol. 21(2), pages 151-168.

Calomiris, Charles, High Loan-to-Value Mortgages: Problem or Cure?, Journal of Lending and Credit Risk Management, (September 1998), pp. 39-43.

Ding, L., J. Ratcliffe, M. Stegman, and R. Quercia (2008). "Neighborhood Patterns of High-Cost Lending: The Case of Atlanta," Journal of Affordable Housing 17(3): 193-217.

Federal Financial Institutions Examination Council, Press Release, (July 26), 2004, http://www.ffiec.gov/hmcrpr/hm072604.htm.

Financial Crimes Report to the Public, The U.S. Federal Bureau of Investigation, (May), 2005, http://www.fbi.gov/publications/financial/fcs_report052005/fcs_report052005.htm#d1.

LaCour-Little, Michael, and Gregory H. Chun, Third-Party Originators and Mortgage Prepayment Risk: An Agency Problem?, Journal of Real Estate Research, 17:1 (1999), pp. 55-70.

Laderman, Elizabeth and Carolina Reid (2009). "CRA Lending During the Subprime Meltdown," in Revisiting the CRA: Perspectives on the Future of the Community Reinvestment Act, a joint publication of the Federal Reserve Banks of Boston and San Francisco.

Mian, Atif, and Amir Sufi, "Fraudulent Income Overstatement on Mortgage Applications during the Credit Expansion of 2002 to 2005", NBER Working Paper No. 20947, Feb. 2015.

Mortgage Loan Fraud: An Industry Assessment Based Upon Suspicious Activities Report Analysis, The U.S. Financial Crimes Enforcement Network: Regulatory Policy & Programs Division, (November), 2006, http://www.fincen.gov/MortgageLoanFraud.pdf.

The Detection, Investigation, and Deterrence of Mortgage Loan Fraud Involving Third Parties: A White Paper, The Federal Financial Institutions Examination Council, (October), 2004, http://www.ffiec.gov/exam/3P_Mtg_Fraud_wp_oct04.pdf.

APPENDIX

TABLE 1
ECONOMIC CHARACTERISTICS OF THE 3-DIGIT ZIP CODES IN THE ATLANTA MSA

3-digit Zip Code Cluster	Median Family Income	Population (000s)	% Unemployment	% Owner Occupied Housing	Median Home value
300XX	**$58,727**	**1,039**	4.5	65.6	$151,300
301XX	$54,448	463	<u>3.3</u>	**78.6**	$122,853
302XX	$52,562	395	4.2	73.7	$118,651
303XX	$45,680	792	**9.2**	<u>45.3</u>	**$200,686**
305XX	$48,609	29	<u>3.3</u>	78.3	$132,200
306XX	<u>$39,747</u>	<u>43</u>	4.6	66.7	<u>$100,024</u>

In each column, the largest (smallest) value is shown in **bold** (<u>underline</u>).

This table contains information on the zip-level family income, population, unemployment rate, owner-occupied housing rate, and the median home value. These data are derived from the U.S. Census Bureau's 2000 Census, weighted to reflect the zip-code level values from the tract-level values correctly.

TABLE 2
HISTORICAL DEFAULT RATES (FORESCORETM) AND CHOICEPOINT® FRAUD RATES BY 3-DIGIT ZIP CODES IN ATLANTA

<u>Panel A</u>: Historical Default Rates (%) by 3-digit Zip Code

3-digit Zip Code Cluster	2004	2005	2006	2007
300XX	1.18	1.33	1.60	1.83
301XX	1.48	1.67	2.00	2.27
302XX	**1.55**	**1.74**	**2.08**	**2.36**
303XX	1.16	1.31	1.57	1.79
305XX	<u>0.95</u>	<u>1.07</u>	<u>1.29</u>	<u>1.47</u>
306XX	1.20	1.35	1.64	1.87

In each column, the largest (smallest) value is shown in **bold** (<u>underline</u>).
Note: These default rates are based on ForeScoreTM estimates of the default probability associated with the zip code.

Panel 1 in this table reflects the historic mortgage loss rate in each 3-digit zip code from 2004 through 2007 shown by the 3-digit Zip-code cluster.

Panel B: ChoicePoint® Fraud Rates (%) by 3-digit Zip Code
The proportion of SARs from each 3-digit zip related to entire Atlanta MSA –

3-digit Zip Code Cluster	2004	2005	2006	2007
300XX	**31.7**	**27.3**	25.0	22.2
301XX	7.5	5.0	6.4	11.1
302XX	11.0	7.7	6.0	13.9
303XX	28.0	26.5	**33.2**	**26.4**
305XX	1.4	0.8	0.0	0.0
306XX	1.0	0.5	0.0	1.4

Notes: (1) Each column represents the majority of SARs reports in the Atlanta MSA, but the column will not total 100 % since not all MSA zips are in the City of Atlanta. (2) These are 'raw,' or unweighted, SARs percentages; meaning zip codes with higher populations in owner-occupied housing will have higher SARs.

TABLE 3
AVERAGE % OF HIGH-COST LOANS AND AVERAGE % OF NON-BANK LOANS BY 3-DIGIT ZIP CODES OF THE ATLANTA MSA

Panel A: % of High-Cost Loans by 3-digit Zip Code

3-digit Zip Code Cluster	2004	2005	2006	2007
300XX	11.2	21.5	22.2	14.2
301XX	14.1	22.1	22.3	16.7
302XX	16.2	30.0	30.1	17.8
303XX	11.1	20.5	23.4	13.9
305XX	9.5	25.5	25.7	15.0
306XX	17.5	26.2	26.2	23.0

Panel B: % of Non-Bank Loans by 3-digit Zip Code

3-digit Zip Code Cluster	2004	2005	2006	2007
300XX	39.1	38.9	38.1	24.3
301XX	40.7	41.0	38.8	28.7
302XX	38.7	41.1	40.2	28.2
303XX	37.3	34.7	33.4	22.3
305XX	47.3	36.2	37.5	22.9
306XX	38.9	45.4	46.7	28.6

Panel 1 in this table reflects the average percent of high-cost HMDA-reported loans from 2004 through 2007 and shown by 3-digit zip-code. Panel 2 in this table reflects the average percent of non-bank, HMDA-reported loans from 2004 through 2007 and shown by 3-digit zip-code.

TABLE 4
COMPARING HIGH-FRAUD AND LOW-FRAUD GEOGRAPHIES BY TRAIT

Variable or Trait	High-Fraud Mean	Lower-Fraud Mean	Difference
Median Family Income	$52,990	$52,944	$46
Unemployment Rate	6.1 %	3.4 %	2.7 %
Owner Occupied Housing Rate	59.8 %	77.6%	(27.8 %)
Median Home Value	$163,077	$121,521	$41,556
High-Cost Loan Proportion	18.3 %	19.2 %	(1.1 %)
Non-Bank Loan Proportion	34.2 %	37.5 %	(3.3 %)
Historical Default Rate	1.55 %	1.79 %	(0.24 %)
SARs Fraud Reporting Rate	24.6 %	6.6 %	18.0 %

This table reflects the differences in economic/demographic traits, as well as credit & lending traits. The High Fraud column represents the 3 out of 6 zip codes with the highest levels of fraud (300, 302, 303), while the remaining three zip codes (301, 305, 306) represent lower fraud levels in the City of Atlanta.

TABLE 5
SIMPLE CORRELATES BETWEEN FRAUD RATE (ADEQUATELY WEIGHTED FOR HOUSING STOCK IN EACH ZIP CODE) AND ECONOMIC, CREDIT & LOAN TRAITS WITHIN THE ZIP CODE

Variable or Trait	Correlation
Median Family Income	0.0860
Unemployment Rate	0.3620*
Owner Occupied Housing Rate	-0.3552*
Median Home Value	0.3872*
High-Cost Loan Proportion	-0.4339**
Non-Bank Loan Proportion	-0.0990
Historical Default Rate	-0.0893

** Indicates the correlation is significant at the .05 level.
* Indicates the correlation is significant at the .10 level.

This table reflects the correlations between the economic, credit and loan traits and the rate of SARs reporting by zip code. There are four annual observations on six distinct zip codes, making for a total of 24 observations in the dataset.