The Subprime Virus*

Sumit Agarwal, National University of Singapore†

Brent W. Ambrose, Pennsylvania State University‡

Yildiray Yildirim, Syracuse University§

Original version: June 18, 2012

Revised version: May 6, 2013

*We thank Zahi Ben-David, John Clapp, Lynn Fisher, John Leahy, Amit Seru, Shane Sherlund, and the participants at the 2010 AREUEA Mid-year meeting and the 2010 AREUEA International Conference for their helpful comments and suggestions. We also thank Thomas Emmerling, Rob McMenamin and Caitlin Kearns for their research assistance. The views expressed in this research are those of the authors and not necessarily those of the Federal Reserve System or the Federal Reserve Bank of Chicago.

†Departments of Finance and Real Estate, National University of Singapore, Singapore, ushakri@yahoo.com

‡Institute for Real Estate Studies, The Pennsylvania State University, University Park, PA 16802-3306, bwa10@psu.edu

§Whitman School of Management, Syracuse University, 721 University Ave Suite 500, Syracuse, NY 13244, yildiray@syr.edu.

The Subprime Virus

Abstract

We examine the correlation between the increase in mortgage default risk on prime mortgages and the introduction of subprime mortgages in a local area. We motivate our analysis with a model of a default contagion effect that spreads the impact of a mortgage foreclosure from one property to surrounding properties. Through numerical analysis, we demonstrate the impact of the origination of subprime mortgages to the risk of a portfolio of prime mortgages. Finally, we offer empirical support for our model by examining the spatial variation in MSA prime mortgage default rates correlated with the level of subprime mortgage activity.

Key words: Subprime, Default, Portfolio Risk, Contagion

JEL Classification: G2; R2

1 Introduction

One of the catalysts that is often blamed for the most recent boom and subsequent bust in the U.S. housing market is the rapid expansion of alternative or subprime mortgages. As a result, numerous studies have examined the role that subprime mortgages played in the current financial crisis.[1] Given the risk characteristics associated with subprime borrowers, it is not surprising that these loans have experienced significantly higher default rates than prime mortgages. For example, Schloemer et al (2006) document that 12.5 percent of all subprime mortgages originated between 1998 and 2004 ended in foreclosure.

In addition to research focusing on the causes and consequences of the housing crisis, attention is now turning to the externalities or spillovers that accompanied the growth in subprime mortgage origination activity. A variety of channels exist whereby subprime origination activity could impose negative externalities on the prime market. For example, initial subprime defaults have the potential to destabilize local housing markets leading to a cascade effect of falling property values that increases the default risk for prime mortgages. Consistent with this theory, Agarwal et. al (2012) show that the distribution of subprime mortgages across geographic areas is not uniform and that areas with higher concentrations of subprime mortgages experienced greater house price volatility.[2] Furthermore, recent evidence shows that the presence of higher risk, alterna-

[1]For example, see Ben-David (2011), Demyanyk and Van Hemert (2011), Keys et al (2010), Mian and Sufi (2009), Ashcraft and Schuermann (2008) and Mayer and Pence (2008) among others. Furthermore, Brueckner, Calem, and Nakamura (2012) offer theoretical and empirical support for the connection between subprime lending and house price increases during the previous decade.

[2]Consistent with this finding, Guerrieri, Hartley, and Hurst (2010) document significant differences in house prices *within* cities.

tive mortgages that subsequently default can have destabilizing effects on surrounding properties. For example, Campbell, Giglio, and Pathak (2011) document that houses sold in foreclosure sell at an average 28 percent discount. As a result of this discount, Immergluck and Smith (2006) note that foreclosures on conventional loans within one-eighth mile depress house prices between 0.9 and 1.1 percent while Lin, Rosenblatt and Yao (2007) document that a foreclosure within a 0.9km radius resulted in an 8.7 percent value discount on neighboring properties.[3] In addition, Mian, Sufi, and Trebbi (2010), using data from 2008 and 2009, estimate that a one standard deviation increase in foreclosures per homeowner results in an 8 percent to 12 percent relative decline in house price growth.[4] Furthermore, consistent with the evidence that foreclosure impact nearby property values, Towe and Lawley (2013) document that a foreclosure sale increases the probability of default for neighboring properties by as much as 28 percent.[5]

Yet, it remains unclear exactly how the growth in high-risk subprime mortgages may have affected the risk of prime mortgages.[6] Recent research on default correlations in fixed income securities suggests that defaults by subprime borrowers may increase the risk of default by prime borrowers.[7] For example, Ascheberg et. al (2011) develop a

[3] In addition to empirical evidence that foreclosures depress neighborhood property values, Ellen, Lacoe, and Sharygin (2012) document that foreclosures lead to increases in neighborhood crime rates.

[4] Schuetz, Been, and Ellen (2008) and Lin, Rosenblatt, and Yao (2009) find similar spillover effects of foreclosures on prices. Lee (2008) and Frame (2010) provide critical reviews of this literature and note that the dispersion in foreclosure effect estimates may be due to differences in data and empirical methods employed by the various studies.

[5] In addition, Goodstein et. al (2011) report that the probability of mortgage default can increase up to 24 percent with a one standard deviation increase in the foreclosure rate in the borrower's zip-code.

[6] In addition to the house price volatility channel discussed above, subprime originations could also alter the risk on prime mortgages through relaxed underwriting standards as lenders compete to retain market share in the face of new competition (Keys, et al., 2010) or through reduced social costs associated with default as foreclosures become more prevalent.

[7] See Duffie (1998) and Zhou (2001) for a discussion of default correlations that result from linkages between individual firms via industry specific and general macro economic conditions.

dynamic simulation model for the evolution of aggregate home prices so that they can analyze the impact of subprime mortgage defaults on prime defaults, and the relative impact of various government policies. Their analysis theoretically demonstrates how subprime mortgage originations can increase the default risk for prime mortgages. Thus, the rise of high-risk mortgages raises an interesting research question: What is the correlation between the presence of subprime mortgages in a geographic area and the risk profile of 'prime' mortgages in the same area? The answer has direct implications for the effectiveness of financial regulations. For example, current bank capital regulations require that financial institutions hold capital based on the riskiness of the assets in their portfolio. However, what happens to the portfolio risk of a 'safe' or 'conservative' institution that currently holds adequate capital when a competitor enters the market and originates a portfolio of high-risk mortgages?

We address these questions by first simulating the effect of the introduction of a new high-risk mortgage loan to a closed market. We utilize Merton's (1974) framework to create a simple model of a bank portfolio of prime (low-risk) mortgages. We then demonstrate how the spillover effect of the origination of new high-risk mortgages increases the riskiness of existing prime, lower-risk mortgages. Our numerical analysis reveals that increasing the subprime mortgage market share from 0 percent to 50 percent increases the default risk on a prime mortgage between 1.8 and 2.3 times (depending upon assumptions regarding house price volatility.)

We recognize that our theoretical model and empirical analysis of the effect of subprime origination on a prime mortgage implicitly assumes the existence of lender segmentation along product type in mortgage origination. In other words, we implicitly assume that certain lenders specialized in the origination of subprime mortgages while other lenders concentrated on the prime market. While this assumption may appear overly strong, we note that empirical evidence supports our assumption of segmentation in mortgage origination. For example, Mayer and Pence (2008) note that the majority of subprime mortgages were originated by specialized lenders that did not compete in the prime mortgage market. In addition, Agarwal et al. (2011) provide a detailed discussion of segmentation in the mortgage industry, again noting the clear distinction between lenders who originated and held prime mortgages and those who originated subprime mortgages.

We empirically investigate the correlation between prime mortgage default risk and subprime origination activity using data from LPS Applied Analytics on mortgages originated between 2003 and 2008. Although the appropriate level of analysis is the lender portfolio level, unfortunately we are unable to obtain micro data at the lender level, and thus we conduct the analysis of default and foreclosure rates based on the zip-code level concentration of subprime origination activity. By conditioning our analysis on the level of subprime activity in 2003 (prior to the subsequent growth in subprime originations that began in 2004), we are able to isolate the analysis to geographic areas where prime lenders dominated the market before subprime lender entry to that market. We identify 8,620 zip-codes that had less than 7.5 percent subprime mortgage originations in 2003. We then track the quarterly default rate of these zip-codes through 2008. Confirming the theoretical model's predictions, the empirical results indicate that the increase in prime mortgage default rates is highly correlated with areas that experienced significant increases in subprime mortgage origination activity, even after controlling for differences in area riskiness. The estimated elasticities indicate that a one point increase in the subprime origination rate increases the prime mortgage portfolio default rate by 0.3 to 0.5 percent and a one point increase in the subprime default rate increases the prime mortgage portfolio default rate by 8.6 to 9.2 percent.

2 Theoretical Model

In this section, we use a simple model to show the impact of the introduction of subprime mortgages on prime mortgage default probabilities to motivate our empirical analysis. The motivation behind the model is to demonstrate in a simple context how subprime originations in a geographic area can alter the riskiness of prime mortgages originated independently of the subprime market. The channel that causes the increase in risk arises from the negative externalities associated with foreclosures. Thus, to the extent that riskier subprime mortgages have higher incidents of foreclosure, our model shows how this risk will have consequences for the risk associated with prime mortgages.

Consider a geographic area with N houses where the average house price level $\bar{H}_t$ moves according to the following stochastic process:

$$\frac{d\bar{H}_t}{\bar{H}_t} = \mu_{\bar{H}}\mathrm{d}t + \sigma_{\bar{H}}\mathrm{d}W_t^{\bar{H}} - L_{\bar{H}}\mathrm{d}U_t \tag{1}$$

where $\mu_{\bar{H}}$ is the drift of the average house price process, $\sigma_{\bar{H}}$ is the corresponding average house price volatility, $L_{\bar{H}}$ represents the amount that the average house price level decreases based on the aggregate number of mortgage defaults (U_t), and $W_t^{\bar{H}}$ is a standard Brownian motion.

Individual house prices (H_t^i) move according to the following stochastic process

$$\mathrm{d}H_t^i = \kappa(\bar{H}_t - H_t^i)\ \mathrm{d}t + \sigma_{H^i} H_t^i \mathrm{d}W_t^{H^i} - L_{H^i} H_t^i \mathrm{d}U_t^i, \qquad (2)$$

where κ is the speed of reversion, $W_t^{H^i}$ $(i = 1, .., N)$ are independent Brownian motions that mean revert around the average house price level $(\bar{H}_t)$, σ_{H^i} is the volatility associated with the i^{th} house, and L_{H^i} represents the individual foreclosure discount associated with a mortgage default. The process U_t^i counts the number of defaults associated with house i.[8] Thus, the cumulative default counting process U_t for the market is defined by $U_t = \sum_{i=1}^N U_t^i$. Based on the empirical evidence about the foreclosure discount associated with mortgage defaults, we assume that house prices will decline by $L_{H^i} = 20\%$ if the borrower defaults.[9] Our default structure explicitly captures the observed externalities associated with the recent foreclosure crisis. In our model, as the number of mortgage defaults increases, the cumulative default process (U_t) increases, which in turn causes a decrease in the average house price process $(\bar{H}_t)$ producing a feedback effect in the mean reverting level of the individual house price processes (H_t^i). In the simulation below, we

[8]For simplicity in the simulations, we allow one default per house.

[9]For example, Mian, Sufi and Trebbi (2010) and Campbell, Giglio and Pathak (2011) provide evidence suggesting that the foreclosure discount ranges between 8 percent and 28 percent.

assume that households finance their houses with interest-only mortgages having loan balances of P^i due at maturity (T).

In structural models, default is determined by the underlying process describing the house value. If the house value is less than the face value of the debt at maturity, the borrower defaults and the debt holders receive the total value of the house. Otherwise, the borrower does not default, and the debt is repaid in full. This is also called the Merton (1974) model and captures the essence that negative equity is a necessary condition for borrower default. Note that in order to highlight the role of neighboring property defaults in determining optimal default, we explicitly ignore the role of prepayment in our model.

To parameterize the model, we assume that a conservative bank originates 5-year, interest-only mortgages to prime borrowers with 80 percent loan-to-value (LTV) ratios, and we normalize the house prices to 100 at time zero. In equation (2), we set the mean reversion speed to $\kappa = 6.1$ and the loss amount due to default to $L_{H^i} = 20\%$. In equation (1), we set the average house price process parameters to $\mu_{\bar{H}} = 3\%$, $\sigma_{\bar{H}} = 7\%$, and $L_{\bar{H}} = 1\%$. Next, we assume that a new lender enters the market and originates 5-year, interest-only high-risk mortgages characterized as having high-LTV ratios ($LTV = 99\%$) to some number i households (where $i < N$). For ease of exposition, we assume that the 'prime' and 'subprime' labels reflect the mortgage risk as captured by the low and high loan-to-value ratios.

Since the goal of our simulation is to investigate how originations by subprime mortgage borrowers affect the risk associated with prime mortgages, we first specify the percentage of subprime borrowers within the N households. We then assume that all subprime mortgage borrowers (out of the N households) have the same loan maturity date, which occurs prior to T. Thus, we can study the impact of changes in the percentage of subprime borrowers on both housing prices and the likelihood of prime borrowers defaulting at time T.

We approximate the continuous dynamics of $\bar{H}$ and H^i using a simple Euler discretization. After simulating all price paths for $\bar{H}$ and H^i, we then focus upon one prime borrower and check whether the borrower defaults at T (i.e.$H^i_T < P^i_T$). We record whether the borrower defaults and rerun the simulated house price paths for another borrower. We report the percentage of defaults that occur out of 2,500 trials.

Although simplistic, our characterization of the market as having a conservative bank originating prime (low-risk) loans and a subprime lender originating high-risk (subprime) mortgages broadly reflects the lender segmentation that existed between prime (GSE) and subprime (non-GSE) loans. Mayer and Pence (2008) and Agarwal et al (2011) provide empirical justification for this characterization by noting that most subprime mortgages were originated by specialized subprime lenders.

To consider the impact of a subprime mortgage lender entering the market, we populate the area with an increasing percentage of subprime mortgages and examine the impact of these originations on the default risk for a prime mortgage. Table 1 shows the impact of increasing subprime market shares and assumptions regarding asset volatility on a default risk for a prime mortgage. Table 1 provides several empirically testable hypotheses. First, consistent with traditional Merton (1974) models, we see that the prime mortgage probability of default increases as the house price volatility increases. For example, in the base case with no subprime activity, increasing the house price volatility by a factor of three (from 10% to 30%) increases the prime mortgage default probability by 1.5 times (from 1.4% to 2.15%). Second, table 1 shows that the prime loan's default risk increases with an increase subprime mortgage market share. For example, in the low volatility environment ($\sigma_{H^i} = 10\%$), the prime mortgage default probability increases by 3.5 times (from 1.4% to 4.9%) as subprime market share increases from 0% to 75%. Furthermore, we see that the impact of subprime origination activity is muted during periods with higher house price volatility.

It is important to recall that the increase in the prime portfolio risk is beyond the prime lender's control. Essentially, the prime portfolio value is reduced through an externality outside the control of the prime lender. As a result, we provide an economic rational for the existence of financial regulations in the market. In the above economy, the actions of the subprime lender imposed a negative externality on the prime lender. Furthermore, to the extent that the subprime mortgages defaulted and these defaults further reduced surrounding property values, then the actions of the subprime lender and borrowers harmed the prime borrowers.

3 Empirical Analysis

3.1 Data

To test the hypothesis that the increase in prime mortgage default risk is correlated with subprime origination activity, our empirical strategy is to classify markets based on their respective subprime market shares. In order to determine market concentration, we collect data from Lender Processing Service (LPS) Applied Analytics.[10] We then determine the share of subprime mortgages originated in each zip-code by quarter as well as the default rate of prime mortgages in each zip-code by quarter.[11] LPS Applied Analytics advertises that it collects data from nine of the ten largest mortgage servicers, although the breadth and depth of its coverage have varied over time. Currently the data base delivers approximately 45 million active loans with over 80 loan level attributes.[12] The LPS data have grown over the years by adding more servicers and requiring servicers to report more variables. When a servicer begins reporting to LPS Applied Analytics, it must report all active mortgages in its portfolio. This information includes data on mortgages that were originated prior to joining LPS Applied Analytics, but it does not include mortgages that were terminated before joining. For example, a servicer that joined LPS Applied Analytics in January 2005 currently uploads active mortgages that originated in 2003, but not the 2003 mortgages that were either prepaid or foreclosed before January 2005 (that is, before the beginning of the servicer's LPS reporting agree-

[10] Mian and Sufi (2009) provide a detailed description of the LPS data.

[11] Subprime classification is reported by the servicers contributing to LPS explicitly.

[12] LPS indicates that the database covers over 65 percent of the total residential mortgage market.

ment). Thus, we restrict the LPS data to first-lien mortgages where LPS reports data within 120-days of origination. The 120-day cutoff controls for back filling of data as servicers enter the sample. We then calculate the subprime percentage of loans originated in each zip-code in each quarter from 2003 to 2008. We also calculate the percentage of prime loans that are in default (90-days or more delinquent) for each quarter between 2003 and 2008.

Table 2 provides a comparison by year of the prime and subprime mortgages contained in the LPS database. At the peak of the subprime lending boom, we see that approximately 9 percent of mortgages tracked by LPS were subprime.[13] Consistent with the definition of subprime, we see that the average loan amount for subprime mortgages was less than the average prime loan amount and the average subprime borrower's credit score (FICO) was less than the average prime borrower's credit score. Furthermore, consistent with subprime mortgages being considered higher risk, we note that subprime mortgages had higher loan-to-value ratios and were more likely to be adjustable-rate mortgages.

3.2 Subprime Concentration

In order to test our hypothesis, we classify zip-codes based on their average exposure to subprime mortgages in 2003. First, we select all zip-codes that had at least 10 mortgages

[13]One concern with using LPS data is that it is known to undercount the population of subprime mortgages. Agarwal et al. (2012) provide a detailed discussion of the underreporting of subprime mortgages in the LPS data. However, to the extent that LPS under counting of subprime mortgages creates a bias, we believe this bias works against our finding an effect. Since our analysis focuses on prime zip-codes, the potential underreporting of subprime mortgages in LPS will result in misclassifying true subprime zip-codes as prime zip-codes, thus biasing our results against finding an effect.

originated in 2003 producing a sample of 10,000 zip codes. Second, we divide the sample into 8,620 zip codes that had subprime mortgage exposure in 2003 less than 7.5 percent of their total 2003 mortgage origination activity (the "qualified" mortgage zip-code sample) and 1,380 zip-codes with subprime activity greater than 7.5 percent (the "non-qualified" zip-code sample.)[14] Finally, we matched each zip-code with the 2000 decennial census resulting in 8,501 qualified zip-codes and 1,370 non-qualified zip-codes. The majority of our analysis is conducted on the qualified zip-code sample. In essence, this sample corresponds to the portfolio of 'prime' mortgages originated by the 'conservative' bank modeled in the theory section.

Table 3 provides a comparison of the demographic characteristics of the non-qualified zip-codes and the qualified zip-codes. Given our classification screen, the non-qualified zip-codes represent the areas that were targeted by subprime lenders prior to 2004. Table 3 shows that the areas with significant subprime exposure in 2003 are different from our qualified, 'prime' areas.[15] For example, the non-qualified areas have substantially lower median household incomes ($37,730 versus $51,071 for the qualified sample), were more rural (78 percent urbanized versus 81 percent urbanized for the qualified sample), had a higher percentage of vacant property (9 percent versus 8 percent), and had older homes (average median year built was 1965 versus 1972 for the qualified sample.) In addition,

[14] Although admittedly arbitrary, our choice of 7.5 percent cutoff criteria was motivated by sample size and the nature of the subprime market. Using a 5 percent cutoff criteria resulted in very few zip-codes in the "qualified" mortgage sample while a cutoff criteria of 10 percent resulted in the majority of zip-codes being classified as "qualified" in 2003. Thus, our choice of 7.5 percent represented a balancing the need for sample size with testing power.

[15] The differences in mean values are statistically significant at the 1 percent level.

we note that the qualified sample has a lower average minority presence (24 percent) than the non-qualified sample (33 percent.)

Next, we classify the "prime" zip-code sample into two segments based on the growth in subprime lending in that area. Once a zip-code's subprime mortgage origination activity exceeds 7.5 percent of any particular quarter's total origination activity, we reclassify that zip-code as a 'non-prime' area. For example, in the first-quarter of 2004, 300 (or 3.5 percent) of the 8,620 'prime' zip-codes experienced subprime origination activity that exceeded 7.5 percent of the total origination activity in that quarter. As is well documented, subprime mortgage origination activity exploded in the U.S. between 2004 and 2007. Thus, by the first quarter of 2007 (the peak of the subprime market), fully 81 percent of the 'prime' zip-codes are now classified as non-prime. Figure 5 shows this explosive growth in subprime origination activity by zip-codes. We note that the majority of the expansion in subprime origination occurred between the third quarter of 2004 and the second quarter of 2005.

To gain a greater feel for the overall spatial growth in subprime origination activity between 2004 and 2008, Figures 1, 2, 3, and 4 show the geographical changes in subprime activity by zip code for Atlanta, Chicago, Philadelphia, and Washington, DC, respectively. For example, the maps for Atlanta (Figure 1) reveal that the high-priced areas of Buckhead and the northern suburbs surrounding Roswell avoided significant subprime activity during the housing bubble period, but the remainder of the Atlanta metropolitan area saw a significant increase in subprime activity. Figures 2 and 4 reveal a similar patter of subprime growth in Chicago and Washington, D.C., respectively. For example, in Chicago only the high-price areas in the north-west suburbs and the area along north Lake Michigan remained subprime free. In contrast, Figure 3 shows that large sections of Philadelphia appear to have escaped the subprime virus.

We focus on the 90+ day 'prime' mortgage delinquency rate experienced by each zip-code as the measure of risk. The 90+ day delinquency rate is the typical measure of mortgage default. As a baseline, we note that the quarterly prime mortgage default rate for these areas averaged 1.57 percent in 2003. In contrast, the average 2003 quarterly prime mortgage default rate in the non-qualified zip-codes was 3.15 percent, or almost twice as high as the default rate in the prime zip-codes. Next, we track the 'prime' mortgage default rate (90+ days delinquency) and the percent of subprime mortgages originated for the 8,501 qualified zip-codes for each quarter starting with the first quarter of 2004 through the fourth quarter of 2008.

Figure 6 shows the quarterly prime mortgage default rates for the 'prime,' 'non-prime,' and non-qualifying zip-codes. Consistent with the theoretical predictions from our model, we see that the default rates in the areas that experienced subprime activity are uniformly higher than the zip-codes without subprime exposure. For example, the default rate for the non-prime zip-codes in the first quarter of 2004 is 98 basis points higher than the average default rate in the prime zip-codes (2.49 percent versus 1.51 percent, respectively).[16] Figure 6 also shows the effects of the housing and financial crisis as the default rates for both prime and non-prime areas increase rapidly in 2007 and 2008. However, we note that the default rates in the non-prime zip-codes increase

[16]Standard t-statistics confirm that the default rates are significantly different from each other.

at a faster rate than the prime zip-code, converging toward the default rates experienced by the zip-codes that failed the initial 2003 subprime screen. Figure 7 confirms this by showing the difference in the quarterly default rates and indicates that the default rate differential was steadily increasing over time such that by the fourth quarter of 2008, the non-prime zip-codes had an average default rate that was 252 basis points higher than the prime zip-codes. Quarterly t-tests confirm that the difference in the default rates is statistically significant.

While the simple univariate comparison of default rates appears to confirm our hypothesis that subprime origination activity alters the risk profile of prime mortgages, it does not control for the endogenous relation that subprime activity increased in areas with substantial house price appreciation and increased volatility. Furthermore, it is possible that systematic differences in risk characteristics may exist between the zip-codes that experienced subprime activity and the 'prime' only zip-codes. Thus, to control for these effects we estimate the following regression of mortgage default rates:

$$\begin{aligned}\delta_{i,t} \;=\; & \alpha + \beta_1 \sum_{k=1}^{t-1} Sub_{i,t-1} + \beta_2 \Delta U_{i,t} + \beta_3 \Delta HPI_{i,t} + \beta_4 \sigma_{i,t}^{HPI} + \beta_5 Sub\delta_{i,t} \\ & + \beta_6 R_{i,t} + \beta_7 \frac{HPI_{i,t}}{\overline{HPI_i}} + \beta_8 X_i + \theta T + \lambda L_i + \epsilon_{i,t} \end{aligned} \tag{3}$$

where $\delta_{i,t}$ is the period t prime mortgage default rate for zip-code i, $\sum_{k=1}^{t-1} Sub_{i,t-1}$ represents the lagged cumulative percentage of subprime mortgages originated in zip-code i (at time $t-1$ beginning with the first quarter of 2004), $\Delta U_{i,t}$ is the quarterly change in the MSA-level unemployment rate at time t that corresponds to zip-code i's location, $\Delta HPI_{i,t}$ is the quarterly change in the MSA-level repeat sales index for zip-code i's respective MSA, $\sigma_{i,t}^{HPI}$ is the standard deviation in the MSA-level repeat sales index for zip-code i's respective MSA, $Sub\delta_{i,t}$ is the subprime default rate for zip-code i at time t, $R_{i,t}$ is the mortgage refinance rate for zip-code i at time t, and $HPI_{i,t}/\overline{HPI_i}$ is the average percentage increase (or decrease) in zip-code i's respective MSA level house price index at time t, X_i is a matrix of demographic characteristics, and T and L_i represent time and location (CBSA) fixed-effects.

We use the FHFA (formerly OFHEO) MSA level repeat sales index to capture changes in house prices. For individual zip-code's that do not map onto a MSA covered by the FHFA index, we use the corresponding state-level MSA HPI index. We obtain the unemployment rate ($U_{i,t}$) from the percent metropolitan area unemployment rates reported by the Bureau of Labor Statistics (BLS) and match to the zip code level mortgages data. For those zip codes that are not part of a metropolitan area, we use the state unemployment rate. The BLS derives their measures of unemployment from various data provided by state employment security agencies, including unemployment insurance claims. Data is benchmarked annually to the CPS estimates to maintain consistency among local areas. The demographic characteristics in X_i include the percentage minority representation in the zip-code, the median household income, the percent of the zip-code that is in an urban area, the percentage of the housing stock that is vacant, and the median home age. These variables are obtained from the 2000 Census ZCTA aggregates, which are static geographical regions that closely match to the year 2000 zip-code areas.

Table 4 reports the demographic characteristics of the prime and non-prime zip-codes (as of the fourth quarter of 2008). Clearly, we see that differences do exist between the prime and non-prime areas.[17] For example, households in the prime areas have higher incomes than non-prime areas ($65,135 versus $47,752, respectively). We find that the non-prime areas have a higher minority concentration than prime areas (25 percent versus 19 percent, respectively). This is not surprising given the evidence that subprime mortgages are over represented in minority communities. We also see that a higher percentage of the prime-only zip-codes are urbanized than the non-prime zip-codes (88 percent versus 80 percent) and the prime-only zip codes have a higher property vacancy rate than non-prime zip codes (9 percent versus 7 percent, respectively.) However, in the other risk measure (mean property age), the two groups are not different.

Columns 1 and 3 in Table 5 report the estimated coefficients for equation (3).[18] As expected, the negative and significant (at the 1 percent level) coefficient for ΔHPI indicates that areas experiencing positive house price growth have lower prime mortgage default rates. Furthermore, consistent with our theoretical model we find that areas with higher house price volatility (HPI Standard Deviation) have higher default rates. In addition, the positive and significant coefficient for ΔU indicates that areas with increasing unemployment rates (a proxy for increasing local economic risk or uncertainty) have higher prime mortgage default rates. The coefficients for percent minority, and percent vacant are positive and significant. These coefficients are consistent with previous empirical research showing that the presence of vacant properties increases risk.

[17]With the exception of population and median year built, the differences in mean values are statistically significant at the 1 percent level.

[18]We report robust clustered standard errors in parentheses.

In addition, we find a positive and significant coefficient for percent urban indicating that urban areas tend to have higher default rates. In column 1 we include the mean current FICO score and in column 3 we present the results using the mean FICO score at origination. Both measures of average credit quality are negative and statistically significant indicating that zip-codes with borrowers having higher credit quality scores (higher FICO scores) are positively correlated with lower default rates. Finally, we note that our model has a high degree of explanatory power with adjusted R^2's of 81% and 79%, respectively.

Turning to the variables of interest for our analysis, the positive and significant coefficients on the subprime mortgage origination activity variable ($\sum_{k=1}^{t-1} Sub_{i,t-1}$) confirms the predictions from our theoretical model that a positive correlation exists between an increase in subprime mortgage originations and the risk of prime mortgages. The estimated coefficient indicates that every one point increase in the subprime origination rate increases the prime mortgage portfolio default rate by 0.3% and 0.5%, respectively. In addition, the estimated coefficients for subprime mortgage default rate are positive and significant, confirming the hypothesis that subprime mortgages may have a spillover effect to prime mortgage performance. The estimated coefficients imply that a one point increase in the subprime default rate increases the prime mortgage portfolio default rate by 8.6 percent and 9.2 percent, respectively.

3.3 Robustness Checks

As noted earlier, one concern with our finding is the possibility that the observed relation between area default rates and subprime origination activity could be endogenous. Although our empirical method attempted to control for differences in area risk through the inclusion of a variety of demographic risk factors, it is possible that our results may still reflect unobserved risk factors. Thus, to control for this possibility, in this section we report two robustness checks.

Our first robustness check begins with the observation that the 2003 (baseline) default rates for zip-codes that we subsequently identify as non-prime may be higher than the 2003 (baseline) default rates for the always prime zip-codes. In other words, it is possible that zip-codes that attract subprime origination activity have some unobserved characteristic that results in higher default rates for all mortgages, and thus, the presence of subprime activity is a spurious correlation. To control for the possible differences in the 2003 baseline default rates, we recast equation (3) as follows:

$$\begin{aligned} \delta_{i,t} - \overline{\delta_{i,03Q4}} &= \alpha + \beta_1 \sum_{k=1}^{t-1} Sub_{i,t-1} + \beta_2 \Delta U_{i,t} + \beta_3 \Delta HPI_{i,t} + \beta_4 \sigma_{i,t}^{HPI} + \beta_5 Sub\delta_{i,t} \\ &\quad + \beta_6 R_{i,t} + \beta_7 \frac{HPI_{i,t}}{\overline{HPI_i}} + \beta_8 X_i + \theta T + \lambda L_i + \epsilon_{i,t} \end{aligned} \tag{4}$$

where $\overline{\delta_{i,03Q4}}$ represents the default rate in the fourth-quarter of 2003 for zip-code i. Thus, equation (4) estimates the impact of the growth in subprime origination activity ($\sum_{k=1}^{t-1} Sub_{i,t-1}$) on the increase (or decrease) in zip-code i's default rate relative to the default rate prior to the subprime boom period (2004 to 2007).

Columns 2 and 4 of Table 5 report the estimated coefficients from equation (4). Consistent with the results discussed above, the positive and significant coefficients for the subprime mortgage origination activity variable ($\sum_{k=1}^{t-1} Sub_{i,t-1}$) confirms that as subprime origination activity in a zip-code increased, the zip-code's default rate increased. The estimated coefficient implies that for every one percent increase in subprime market share, the default rate increases 0.1 to 0.3 basis points above the 2003 baseline default rate. For example, the zip-code 60614 (Chicago) saw a cumulative increase in the subprime origination market share from the fourth-quarter of 2003 to the fourth quarter of 2004 of 183 basis points. Thus, the estimated coefficient implies that the 2004Q4 prime mortgage default rate in zip-code 60614 increased between 0.186 and 0.549 basis points over the baseline 2003Q4 default rate.

Our second robustness check accounts for the potential endogeneity between subprime market share and prime default rates. Again, we are concerned with the potential that subprime activity is reflecting unobserved area risk characteristics that impact prime mortgage default rates. Thus, to control for the potential endogenous relation between subprime origination activity and prime mortgage default rates, we estimate the following two-stage least squares (2SLS) model:

$$\begin{aligned} Sub_{i,t} &= \alpha_0 + \alpha_1 Sub_{i,t-1} + \alpha_2 \Delta U_{i,t} + \alpha_3 \Delta HPI_{i,t} + \alpha_4 \sigma_{i,t}^{HPI} \\ &\quad + \alpha_5 R_{i,t} + \alpha_6 \frac{HPI_{i,t}}{\overline{HPI_i}} + \alpha_7 X_i + \epsilon_{i,t} \end{aligned} \tag{5}$$

$$\delta_{i,t} = \alpha + \beta_1 \sum_{k=1}^{t-1} \widehat{Sub_{i,t-1}} + \beta_2 \Delta U_{i,t-1} + \beta_3 \Delta HPI_{i,t-1} + \beta_4 \sigma_{i,t-1}^{HPI} + \beta_5 Sub\delta_{i,t-1}$$
$$+\beta_6 R_{i,t-1} + \beta_7 \frac{HPI_{i,t-1}}{\overline{HPI_i}} + \beta_8 X_i + \theta T + \lambda L_i + \xi_{i,t} \quad (6)$$

where again, $\delta_{i,t}$ is the period t prime mortgage default rate for zip-code i, $Sub_{i,t}$ represents the percentage of subprime mortgages originated in zip-code i at time t, and the other variables are defined above. We assume that $Sub_{i,t-1}$ serves as the instrument for the endogenous variable $Sub_{i,t}$.

Table 6 reports the estimated coefficients from the 2SLS estimation. Column (1) reports the results using mean current credit scores while column (2) reports the results using mean FICO score at origination. In the first stage, we find positive coefficients for the change in house prices ($HPI_{i,t}/\overline{HPI_i}$) and ($\Delta HPI_{i,t}$) suggesting that prime areas in 2003 that experienced significant house price increases had higher subprime origination activity. However, we note that the negative coefficient on house price index volatility ($\sigma_{i,t}^{HPI}$) implies that areas with higher house price risk had lower subprime origination activity.[19] In terms of area demographic characteristics, we see that higher minority concentrations and more urban areas are positively correlated with subprime origination activity while higher income and more vacant property are associated with lower subprime activity. Finally, we note that areas experiencing higher growth in unemployment ($\Delta U_{i,t}$) and higher average credit scores have lower subprime activity.

[19] We also estimated the models using a zip-code level house price index and found qualitatively the same results.

The second stage model shows the effects of the predicted cumulative subprime origination activity ($\sum_{k=1}^{t-1} \widehat{Sub_{i,t-1}}$) on the prime mortgage default rate. Again, we find a positive and significant effect indicating that subprime origination activity is highly correlated with prime mortgage default rates. The estimated coefficients imply that a one point increase in the cumulative predicted subprime origination rate results in a 30 to 60 basis point increase in the prime default rate. In addition, we also confirm that higher subprime default rates ($Sub\delta_{i,t-1}$) are correlated with greater prime default rates. The estimated coefficients suggest that a one point increase in the subprime default rate leads to between a 8.8 percent and 9.3 percent increase in the prime mortgage portfolio default rate. The negative coefficients for the change in house prices ($\frac{HPI_{i,t-1}}{HPI_i}$) suggest that prime areas in 2003 that experienced significant house price appreciation had lower prime mortgage default rates. In addition, the estimated coefficients confirm the previous findings that areas that experienced greater refinancing activity and positive house price growth had lower prime mortgage default rates.

4 Conclusions

This paper focuses on the simple question: Was the introduction of subprime mortgages correlated with a change in the risk profile of prime mortgages in the same area? To answer this question, we present a simple theoretical model based on Merton's (1974) framework that demonstrates the potential spillover effects associated with the introduction of risky assets into a market. Consistent with the empirical research documenting foreclosure discounts in the single-family home market (e.g. Campbell, Giglio and Pathak; 2011), we introduce a default transmission mechanism in our model that leads to lower asset values if a mortgage defaults.[20] Through numerical analysis, we demonstrate the impact of the origination of subprime mortgages on the risk of a prime mortgage. Consistent with similar models of default correlation, the numerical analysis shows a positive shift in the prime mortgage default probability as subprime mortgages market share increases.

Finally, we offer empirical support for our model by examining the spatial variation in MSA prime mortgage default rates correlated with the level of subprime mortgage activity. We focus our analysis on the 8,620 zip-codes that had subprime mortgage exposure in 2003 less than 7.5 percent of their total 2003 mortgage origination activity. We then track these zip-codes from 2004 through 2008 and classify them into 'prime' and 'non-prime' areas when the level of subprime mortgage origination activity exceeds 7.5 percent. We then focus on the 90+ day 'prime' mortgage delinquency rate experienced by each zip-code in the prime and non-prime groups. Consistent with the theoretical predictions from our model, the default rates in the areas that experienced subprime activity are uniformly higher than in the zip-codes without subprime exposure. The estimated elasticities indicate that a one point increase in the subprime origination rate increases the prime mortgage default rate by 30 to 50 basis points while a one point increase in the subprime default rate increases the prime mortgage default rate by 8.6 percent to 9.2 percent.

[20] Our transmission mechanism is similar to the way income shocks affect land prices as documented in Guerrieri, Hartley, and Hurst (2010).

The results from our study provide an economic rational for the existence of financial regulations. We demonstrate how the actions of a subprime lender impose negative externalities on prime lenders through increased property volatilities that increased default risk of a prime mortgage portfolio. This increase in the prime portfolio risk is beyond the prime lender's control as they are unable to prevent the subprime lender from entering their geographic market. Furthermore, to the extent that future subprime mortgage origination activity was not anticipated, then the effect of the introduction of subprime mortgages on the risk of prime mortgages was not priced at origination.

5 References

Agarwal, Sumit, Brent W. Ambrose, Souphala Chomsisengphet, and Anthony B. Sanders, 2012. "Thy Neighbor's Mortgage: Does Living in a Subprime Neighborhood Impact Your Probability of Default?" *Real Estate Economics* 40(1), 1-22.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff, 2011, "The Role of Securitization in Mortgage Renegotiation," *Journal of Financial Economics,* 102, 554-578.

Agarwal, Sumit, Yan Chang, and Abdullah Yavas, 2012, "Adverse Selection in Mortgage Securitization," *Journal of Financial Economics,* 105, 640-660.

Ascheberg, Marius, Robert A. Jarrow, Holger Kraft, and Yildiray Yildirim, 2011. "Government Policies, Residential Mortgage Defaults, and the Boom and Bust Cycle of Housing Prices," Working paper.

Ashcraft, Adam B. and Til Schuermann, 2008. "Understanding the Securitization of Subprime Mortgage Credit," Federal Reserve Bank of New York Staff Reports, no. 318.

Ben-David, Itzhak, 2011. "Financial Constraints and Inflated Home Prices during the Real-Estate Boom," *American Economic Journal: Applied Economics,* 3(3), 55-78.

Brueckner, Jan K., Paul S. Calem and Leonard I. Nakamura, 2012. "Subprime Mortgages and the Housing Bubble," *Journal of Urban Economics,* 71(2), 230-243.

Campbell, John, Stefano Giglio, and Parag Pathak, 2011. "Forced Sales and House Prices," *American Economic Review,* 101(5): 2108-2131.

Demyanyk, Y. and O. Van Hemert, 2011. "Understanding the Subprime Mortgage Crisis," *Review of Financial Studies,* 24:6, 1848-1880.

Duffie, G., 1998. "The Relation Between Treasury Yields and Corporate Yield Spreads," *Journal of Finance,* 55:6, 2225-2243.

Ellen, Ingrid Gould, Johanna Lacoe, and Claudia Ayanna Sharygin, 2012. "Do Foreclosures Cause Crime?" *Journal of Urban Economics,* Forthcoming.

Frame, W. Scott, 2010. "Estimating the Effect of Mortgage Foreclosures on Nearby Property Values: A Critical Review of the Literature," *Federal Reserve Bank of Atlanta Economic Review* 95:3, 1-9.

Goodstein, Ryan, Paul Hanouna, Carlos D. Ramirez, and Christof W. Stahel, 2011. "Are Foreclosures Contagious?" GMU Working Paper in Economics No. 11-12.

Guerrieri, Veronica, Daniel Hartley, and Eric Hurst, 2010. "Endogenous Gentrification and Housing Price Dynamics," University of Chicago working paper.

Immergluck, Dan and Geoff Smith, 2006. "The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values," *Housing Policy Debate,* 17(1), 57-79.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans," *Quarterly Journal of Economics* 125(1), 307-362.

Lee, Kai-yan, 2008. "Foreclosure's Price Depressing Spillover Effects on Local Properties: A Literature Review," Federal Reserve Bank of Boston Community Affairs Discussion Paper Number 2008-1.

Lin, Zhengou, Eric Rosenblatt, and Vincent Yao, 2009. "Spillover Effects of Foreclosures on Neighborhood Property Values ," *Journal of Real Estate Finance and Economics,* 38, 387-407.

Mayer, C.J., and K. Pence, 2008. “Subprime Mortgages: What, Where, and to Whom?” NBER Working Paper No. 14083.

Merton, R.C., 1974. “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates,” *Journal of Finance*, 2: 449-470.

Mian, Atif, and Amir Sufi, 2009. “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis,” *Quarterly Journal of Economics*, November 2009, 124(4), 1449-1496.

Mian, Atif, Amir Sufi, and Francesco Trebbi, 2010. “Foreclosure, House Prices, and the Real Economy,” University of California-Berkeley working paper.

Schloemer, E., W. Li, K. Ernst, and K. Keest, 2006. *Losing Ground: Foreclosures in the Subprime Market and Their Cost to Homeowners.* Center for Responsible Lending.

Schuetz, Jenny, Vicki Been, and Ingrid Gould Ellen, 2008. “Neighborhood Effects of Concentrated Mortgage Foreclosures,” *Journal of Housing Economics,* 17, 206-319.

Towe, Charles, and Chad Lawley, 2012. “The Contagion Effect of Neighboring Foreclosures,” *American Economic Journal: Economic Policy,* 5(2): 313-335.

Zhou, Chunsheng, 2001. “An Analysis of Default Correlations and Multiple Defaults,” *Review of Financial Studies*, 14:2, 555-576.

Table 1: **Impact of asset volatility and subprime origination activity on prime mortgage default probability**

Subprime Mortgage Market Share	House Price Volatility $\sigma_{H_i} = 0.1$	$\sigma_{H_i} = 0.2$	$\sigma_{H_i} = 0.3$
0	1.40%	1.55%	2.15%
25	1.50%	1.95%	3.85%
50	2.55%	3.50%	4.60%
75	4.90%	5.20%	5.80%

Note: Each column represents the default probability for a prime mortgage loan for different subprime mortgage market shares.

Table 2: **Mean Prime and Subprime Characteristics by Year from LPS Applied Analytics**

	Number of Loans		Loan Amount		FICO Score		LTV Ratio		% ARM	
	Subprime	Prime	Subprime	Prime	Subprime	Prime	Subprime	Prime	Subprime	Prime
2003	95,863	5,708,546	148,088	174,422	638	721	76.0	70.5	55%	14%
			(93,977)	(139,498)	(71)	(59)	(15)	(19)		
2004	186,142	4,383,158	178,842	202,286	617	713	78.2	73.3	62%	29%
			(112,225)	(201,083)	(61)	(62)	(13)	(18)		
2005	581,775	5,617,842	194,855	231,244	613	717	79.3	72.8	62%	18%
			(127,656)	(205,365)	(56)	(59)	(12)	(17)		
2006	471,371	5,019,945	201,467	243,465	610	712	78.9	73.7	51%	13%
			(146,945)	(214,689)	(54)	(62)	(13)	(17)		
2007	180,363	4,620,254	200,853	241,540	602	712	78.9	75.0	11%	4%
			(148,518)	(225,525)	(52)	(65)	(14)	(18)		
2008	6,394	3,529,959	181,970	219,108	605	717	76.6	76.9	2%	2%
			(134,934)	(162,885)	(50)	(66)	(15)	(19)		

Note: Standard deviations reported in parentheses.

Table 3: **Descriptive Statistics of the Qualified and Non-Qualified Samples**

	Mean	**Standard Deviation**	**25th Percentile**	**Median**	**75th Percentile**
Panel A: Qualified Zip codes (8,501 zip codes)					
Population	23,098	15,142	11,412	20,218	31,215
% Minority	24%	22%	8%	16%	33%
Median Household Income	$51,071	$18,264	$38,049	$47,258	$60,226
Number of Housing Units	9,363	5,861	4,703	8,393	12,835
% Urban	81%	28%	72%	96%	100%
% Vacant	8%	9%	3%	5%	8%
Median Year Built	1972	13	1963	1974	1982
Panel B: Non-Qualified Zip codes (1,370 zip codes)					
Population	21,193	14,403	10,886	18,055	28,396
% Minority	33%	31%	7%	20%	55%
Median Household Income	$37,730	$11,315	$30,243	$35,728	$42,734
Number of Housing Units	8,548	5,336	4,449	7,602	11,664
% Urban	78%	29%	66%	92%	100%
% Vacant	9%	7%	5%	7%	10%
Median Year Built	1965	13	1955	1966	1975

Note: Zip-codes are classified based on their average exposure to subprime mortgages in 2003 using the following screens: First, we select all zip-codes that had at least 10 mortgages originated in 2003 producing a sample of 10,000 zip codes. Second, we divide the sample into 8,620 zip codes that had subprime mortgage exposure in 2003 less than 7.5 percent of their total 2003 mortgage origination activity (the "qualified" mortgage zip-code sample) and 1,380 zip-codes with subprime activity greater than 7.5 percent (the "non-qualified" zip-code sample.) Finally, we matched each zip-code with the 2000 decennial census resulting in 8,501 qualified zip-codes and 1,370 non-qualified zip-codes.

Table 4: **Demographic Information for the Qualified Zip Codes**

	Mean	**Standard Deviation**	**25th Percentile**	**Median**	**75th Percentile**
Panel A: Prime-Only Zip codes (1,623 zip codes)					
Population	23,191	14,563	12,433	20,480	31,045
% Minority	19%	15%	8%	14%	24%
Median Household Income	$ 65,135	$ 23,699	$ 47,547	$ 61,475	$ 77,851
Number of Housing Units	10,136	6,326	5,662	9,099	13,496
% Urban	88%	22%	87%	99%	100%
% Vacant	9%	13%	3%	4%	8%
Median Year Built	1972	16	1961	1975	1984
Panel B: Prime Zip codes That Became Non-Prime Zip codes (6,878 zip codes)					
Population	23,076	15,276	11,256	20,193	31,231
% Minority	25%	24%	8%	17%	36%
Median Household Income	$ 47,752	$ 14,904	$ 36,955	$ 45,268	$ 55,838
Number of Housing Units	9,181	5,732	4,479	8,211	12,678
% Urban	80%	29%	68%	94%	100%
% Vacant	7%	7%	4%	5%	8%
Median Year Built	1972	13	1963	1974	1981

Note: Zip-codes are classified based on their average exposure to subprime mortgages in 2003 using the following screens: First, we select all zip-codes that had at least 10 mortgages originated in 2003 producing a sample of 10,000 zip codes. Second, we divide the sample into 8,620 zip codes that had subprime mortgage exposure in 2003 less than 7.5 percent of their total 2003 mortgage origination activity (the "qualified" mortgage zip-code sample) and 1,380 zip-codes with subprime activity greater than 7.5 percent (the "non-qualified" zip-code sample.) Finally, we matched each zip-code with the 2000 decennial census resulting in 8,501 qualified zip-codes and 1,370 non-qualified zip-codes. Panel A covers the zip-codes that never had more than 7.5 percent subprime origination activity between 2004 and 2008. Panel B covers the zip-codes that were prime-only in 2003 but subsequently saw more than 7.5 percent subprime origination activity by 2008.

Table 5: **Estimated Regression Coefficients**

	(1) 90+ Day Prime Default Rate	(2) Change in 90+ Day Prime Default Rate from 2003	(3) 90+ Day Prime Default Rate	(4) Change in 90+ Day Prime Default Rate from 2003
Sum of past predicted Subprime Origination Rates ($\sum_{k=1}^{t-1} Sub_{i,t-1}$)	0.003***	0.001***	0.005***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
1 Quarter change in Unemployment ($\Delta U_{i,t}$)	0.044***	0.040***	0.062***	0.039***
	(0.005)	(0.007)	(0.007)	(0.006)
HPI Annualized rate ($\Delta HPI_{i,t}$)	-0.102***	-0.068***	-0.103***	-0.071***
	(0.001)	(0.001)	(0.001)	(0.001)
HPI Standard Deviation ($\sigma_{i,t}^{HPI}$)	0.038***	0.037***	0.033***	0.037***
	(0.001)	(0.001)	(0.001)	(0.001)
Current HPI over Average Origination HPI ($\frac{HPI_{i,t}}{\overline{HPI_i}}$)	0.003***	-0.008***	0.003***	-0.007***
	(0.000)	(0.000)	(0.000)	(0.000)
Refinance Rate ($R_{i,t}$)	-1.914***	-2.050***	-2.486***	-2.003***
	(0.034)	(0.033)	(0.036)	(0.037)
Subprime Default Rate ($Sub\delta_{i,t}$)	0.086***	0.088***	0.092***	0.089***
	(0.001)	(0.001)	(0.001)	(0.001)
Demographic Controls				
% Minority	0.002***	-0.002***	0.004***	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Median income (in $1,000s)	0.012***	0.004***	0.012***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
% Urban	0.002***	0.001***	0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
% Vacant	0.025***	-0.003***	0.030***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Median Home Age	0.009***	0.003***	0.011***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Mean current FICO score	-0.054***	0.004***		
	(0.000)	(0.000)		
Mean FICO score at origination			-0.061***	0.012***
			(0.000)	(0.000)
Date Fixed Effects	Yes	Yes	Yes	Yes
CBSA Fixed Effects	Yes	Yes	Yes	Yes
Clustered S.E.	Yes	Yes	Yes	Yes
Constant	39.380***	-0.918***	44.752***	-8.149***
	(0.172)	(0.180)	(0.232)	(0.238)
Observations	123,145	123,145	123,145	123,145
R-squared	0.812	0.635	0.793	0.639
Adjusted R-squared	0.811	0.633	0.792	0.637

Standard errors in parentheses

Notes: Columns 1 and 3 report the estimated coefficients for the following regression of mortgage default rates:

$$\begin{aligned}\delta_{i,t} &= \alpha + \beta_1 \sum_{k=1}^{t-1} Sub_{i,t-1} + \beta_2 \Delta U_{i,t} + \beta_3 \Delta HPI_{i,t} + \beta_4 \sigma_{i,t}^{HPI} + \beta_5 Sub\delta_{i,t} \\ &\quad + \beta_6 R_{i,t} + \beta_7 \frac{HPI_{i,t}}{\overline{HPI_i}} + \beta_8 X_i + \theta T + \lambda L_i + \epsilon_{i,t}\end{aligned}$$

where $\delta_{i,t}$ is the period t prime mortgage default rate for zip-code i.

Columns 2 and 4 report the estimated coefficients for the following model:

$$\begin{aligned}\delta_{i,t} - \overline{\delta_{i,03Q4}} &= \alpha + \beta_1 \sum_{k=1}^{t-1} Sub_{i,t-1} + \beta_2 \Delta U_{i,t} + \beta_3 \Delta HPI_{i,t} + \beta_4 \sigma_{i,t}^{HPI} + \beta_5 Sub\delta_{i,t} \\ &\quad + \beta_6 R_{i,t} + \beta_7 \frac{HPI_{i,t}}{\overline{HPI_i}} + \beta_8 X_i + \theta T + \lambda L_i + \epsilon_{i,t}\end{aligned}$$

where $\overline{\delta_{i,03Q4}}$ represents the default rate in the fourth-quarter of 2003 for zip-code i. $\sum_{k=1}^{t-1} Sub_{i,t-1}$ represents the lagged cumulative percentage of subprime mortgages originated in zip-code i (at time $t-1$ beginning with the first quarter of 2004), $\Delta U_{i,t}$ is the quarterly change in the MSA-level unemployment rate at time t that corresponds to zip-code i's location, $\Delta HPI_{i,t}$ is the quarterly change in the MSA-level repeat sales index for zip-code i's respective MSA, $\sigma_{i,t}^{HPI}$ is the standard deviation in the MSA-level repeat sales index for zip-code i's respective MSA, $Sub\delta_{i,t}$ is the subprime default rate for zip-code i at time t, $R_{i,t}$ is the mortgage refinance rate for zip-code i at time t, and $HPI_{i,t}/\overline{HPI_i}$ is the average percentage increase (or decrease) in zip-code i's respective MSA level house price index at time t, X_i is a matrix of demographic characteristics, and T and L_i represent time and location (CBSA) fixed-effects. The dependent variables are the prime-mortgage 90+ day default rate (columns 1 and 3) and the change in default rates from the average default rate in 2003 (columns 2 and 4.) Robust clustered standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Two-stage Least Squares Regression

	(1)		(3)	
	Stage 1 Subprime Origination Rate	Stage 2 90+ Day Prime Default Rate	Stage 1 Subprime Origination Rate	Stage 2 90+ Day Prime Default Rate
Sum of past predicted Subprime Origination Rates ($\sum_{k=1}^{t-1} \widehat{Sub_{i,t-1}}$)		0.003***		0.006***
		(0.000)		(0.000)
Lagged 1 Quarter change in Unemployment ($\Delta U_{i,t-1}$)		0.121***		0.125***
		(0.007)		(0.008)
Lagged HPI over Average Origination HPI ($\frac{HPI_{i,t-1}}{\overline{HPI_i}}$)		-0.003***		-0.004***
		(0.000)		(0.000)
Lagged HPI Standard Deviation ($\sigma_{i,t-1}^{HPI}$)		0.039***		0.034***
		(0.001)		(0.001)
Subprime Default Rate ($Sub\delta_{i,t}$)		0.088***		0.093***
		(0.001)		(0.001)
Lagged Refinance Rate ($R_{i,t-1}$)		-1.449***		-2.044***
		(0.036)		(0.037)
Lagged HPI Annualized rate ($\Delta HPI_{i,t-1}$)		-0.087***		-0.089***
		(0.001)		(0.001)

continued on next page ...

Continued from previous page ...

	(1)		(2)	
	Stage 1 90+ Day Origination Rate	Stage 2 Subprime Prime Default Rate	Stage 1 90+ Day Origination Rate	Stage 2 Prime Default Rate
Demographic Controls				
% Minority	0.008***	0.002***	0.005***	0.004***
	(0.001)	(0.000)	(0.001)	(0.000)
Median Income (in $1,000) -0.005***	-0.005***	0.012***	0.001	0.012***
	(0.001)	(0.000)	(0.001)	(0.000)
% Urban	0.001	0.002***	0.002***	0.001***
	(0.001)	(0.000)	(0.001)	(0.000)
% Vacant	-0.004***	0.025***	0.001	0.030***
	(0.002)	(0.001)	(0.002)	(0.001)
Median Home Age	-0.003***	0.009***	-0.002***	0.011***
	(0.001)	(0.000)	(0.001)	(0.000)
Lagged Subprime Origination Rate ($Sub_{i,t-1}$)	0.706***		0.680***	
	(0.002)		(0.002)	
1 Quarter change in Unemployment ($\Delta U_{i,t}$)	-0.310***		-0.322***	
	(0.019)		(0.018)	
Current HPI over Average Origination HPI ($\frac{HPI_{i,t}}{\overline{HPI_i}}$)	0.026***		0.023***	
	(0.001)		(0.001)	
HPI Standard Deviation ($\sigma_{i,t}^{HPI}$)	-0.032***		-0.035***	
	(0.001)		(0.001)	
Refinance Rate ($R_{i,t}$)	1.211***		1.166***	
	(0.076)		(0.076)	
HPI Annualized rate ($\Delta HPI_{i,t}$)	0.039***		0.054***	
	(0.003)		(0.003)	
Mean current FICO score	-0.019***	-0.054***		
	(0.001)	(0.000)		
Mean FICO score at origination			-0.038***	-0.060***
			(0.001)	(0.000)
Date Fixed Effects	No	Yes	No	Yes
CBSA Fixed Effects	No	Yes	No	Yes
Clustered S.E.	Yes	Yes	Yes	Yes
Constant	11.598***	39.166***	25.081***	43.320***
	(0.358)	(0.182)	(0.478)	(0.246)
Observations	129,401	123,153	129,401	123,153
R-squared	0.653	0.809	0.657	0.790
Adjusted R-squared	0.653	0.808	0.657	0.789

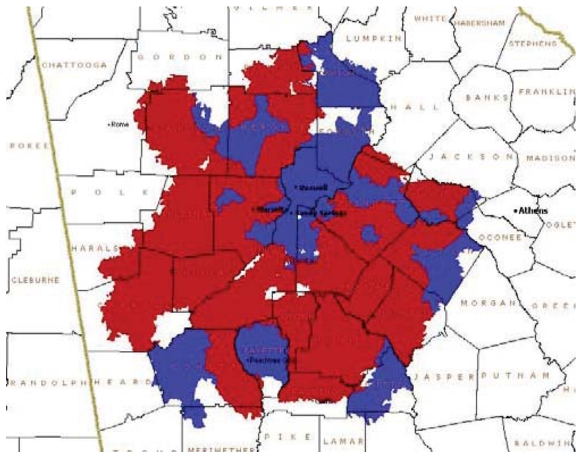
Clustered standard errors in parentheses

Notes: This table presents the estimated coefficients from the following two-stage least squares (2SLS) model:

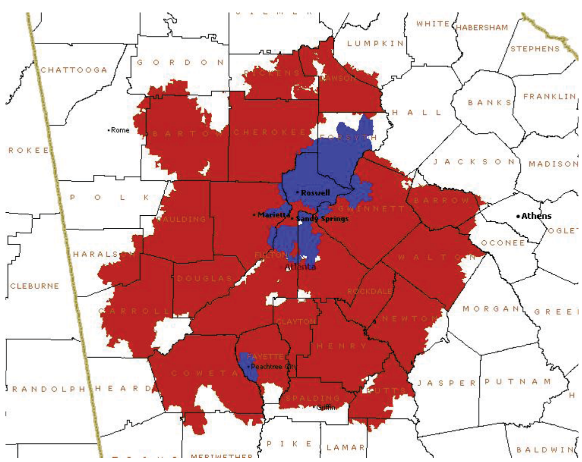
$$\begin{aligned} Sub_{i,t} &= \alpha_0 + \alpha_1 Sub_{i,t-1} + \alpha_2 \Delta U_{i,t} + \alpha_3 \Delta HPI_{i,t} + \alpha_4 \sigma_{i,t}^{HPI} \\ &\quad + \alpha_5 R_{i,t} + \alpha_6 \frac{HPI_{i,t}}{\overline{HPI_i}} + \alpha_7 X_i + \epsilon_{i,t} \end{aligned}$$

$$\begin{aligned} \delta_{i,t} &= \alpha + \beta_1 \sum_{k=1}^{t-1} \widehat{Sub_{i,t-1}} + \beta_2 \Delta U_{i,t-1} + \beta_3 \Delta HPI_{i,t-1} + \beta_4 \sigma_{i,t-1}^{HPI} + \beta_5 Sub\delta_{i,t-1} \\ &\quad + \beta_6 R_{i,t-1} + \beta_7 \frac{HPI_{i,t-1}}{\overline{HPI_i}} + \beta_8 X_i + \theta T + \lambda L_i + \xi_{i,t} \end{aligned}$$

where $\delta_{i,t}$ is the period t prime mortgage default rate for zip-code i, $Sub_{i,t}$ represents the percentage of subprime mortgages originated in zip-code i at time t, $\sum_{k=1}^{t-1} Sub_{i,t-1}$ represents the lagged cumulative percentage of subprime mortgages originated in zip-code i (at time $t-1$ beginning with the first quarter of 2004), $\Delta U_{i,t}$ is the quarterly change in the MSA-level unemployment rate at time t that corresponds to zip-code i's location, $\Delta HPI_{i,t}$ is the quarterly change in the MSA-level repeat sales index for zip-code i's respective MSA, $\sigma_{i,t}^{HPI}$ is the standard deviation in the MSA-level repeat sales index for zip-code i's respective MSA, $Sub\delta_{i,t}$ is the subprime default rate for zip-code i at time t, $R_{i,t}$ is the mortgage refinance rate for zip-code i at time t, and $HPI_{i,t}/\overline{HPI_i}$ is the percentage increase (or decrease) in zip-code i's respective MSA level house price index at time t, X_i is a matrix of demographic characteristics, and T and L_i represent time and location (CBSA) fixed-effects. *** p<0.01, ** p<0.05, * p<0.1



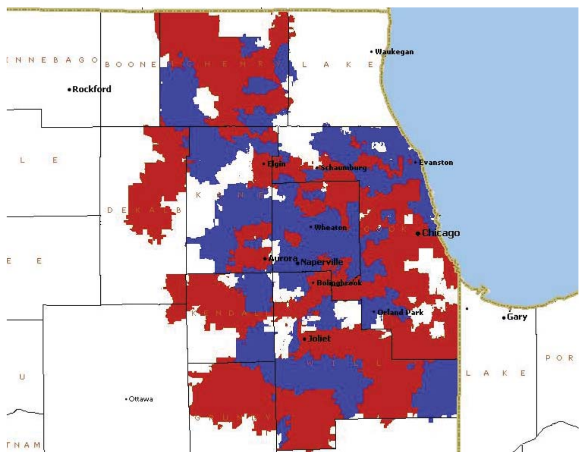
Atlanta 2004:Q4



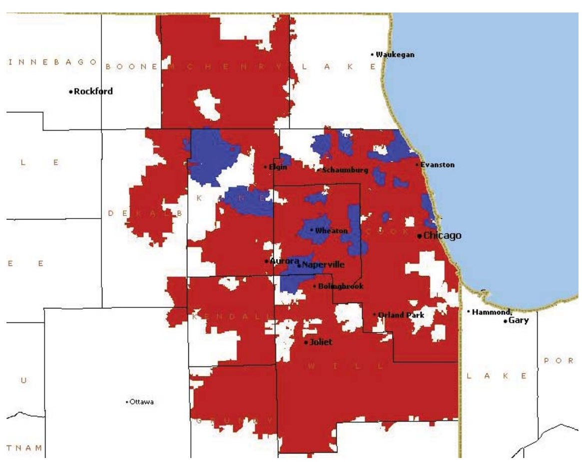
Atlanta 2008:Q4

Red Shading are Subprime, Blue Shading are Prime

Figure 1: **Change in Atlanta subprime and prime zip-codes between 2004 and 2008**



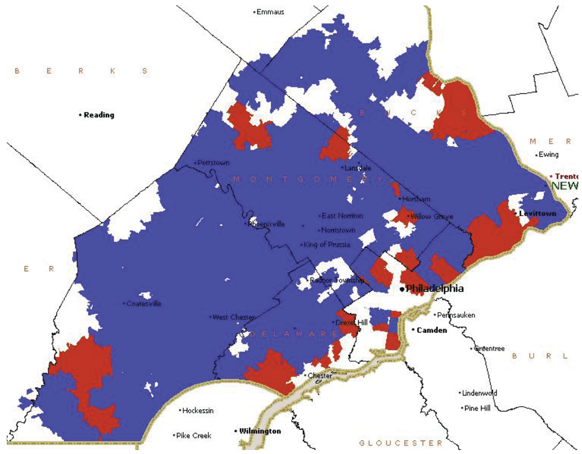
Chicago 2004:Q4



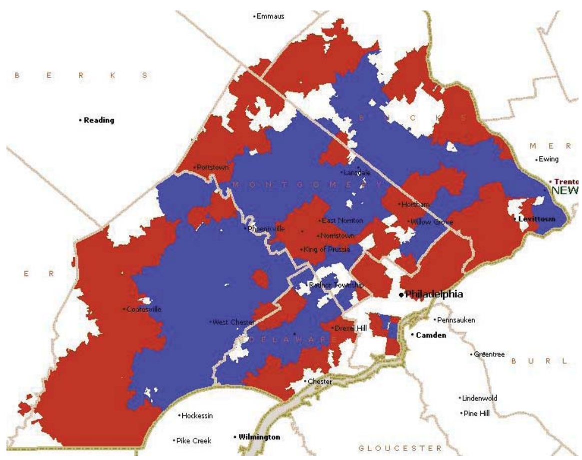
Chicago 20008:Q4

Red Shading are Subprime, Blue Shading are Prime

Figure 2: **Change in Chicago subprime and prime zip-codes between 2004 and 2008**



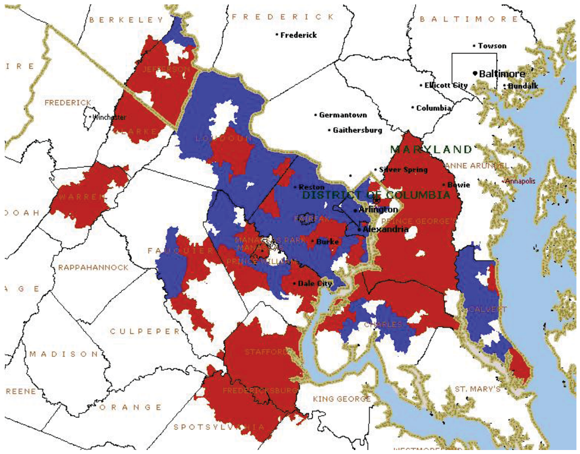
Philadelphia 2004:Q4



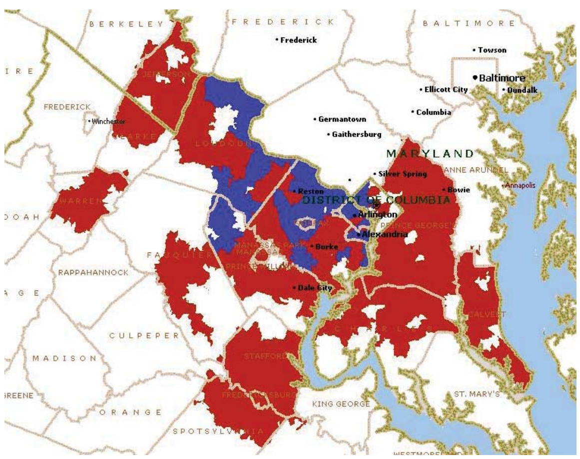
Philadelphia 2008:Q4

Red Shading are Subprime, Blue Shading are Prime

Figure 3: **Change in Philadelphia subprime and prime zip-codes between 2004 and 2008**



Washington, DC 2004:Q4



Washington, DC 2008:Q4

Red Shading are Subprime, Blue Shading are Prime

Figure 4: **Change in Washington, DC subprime and prime zip-codes between 2004 and 2008**

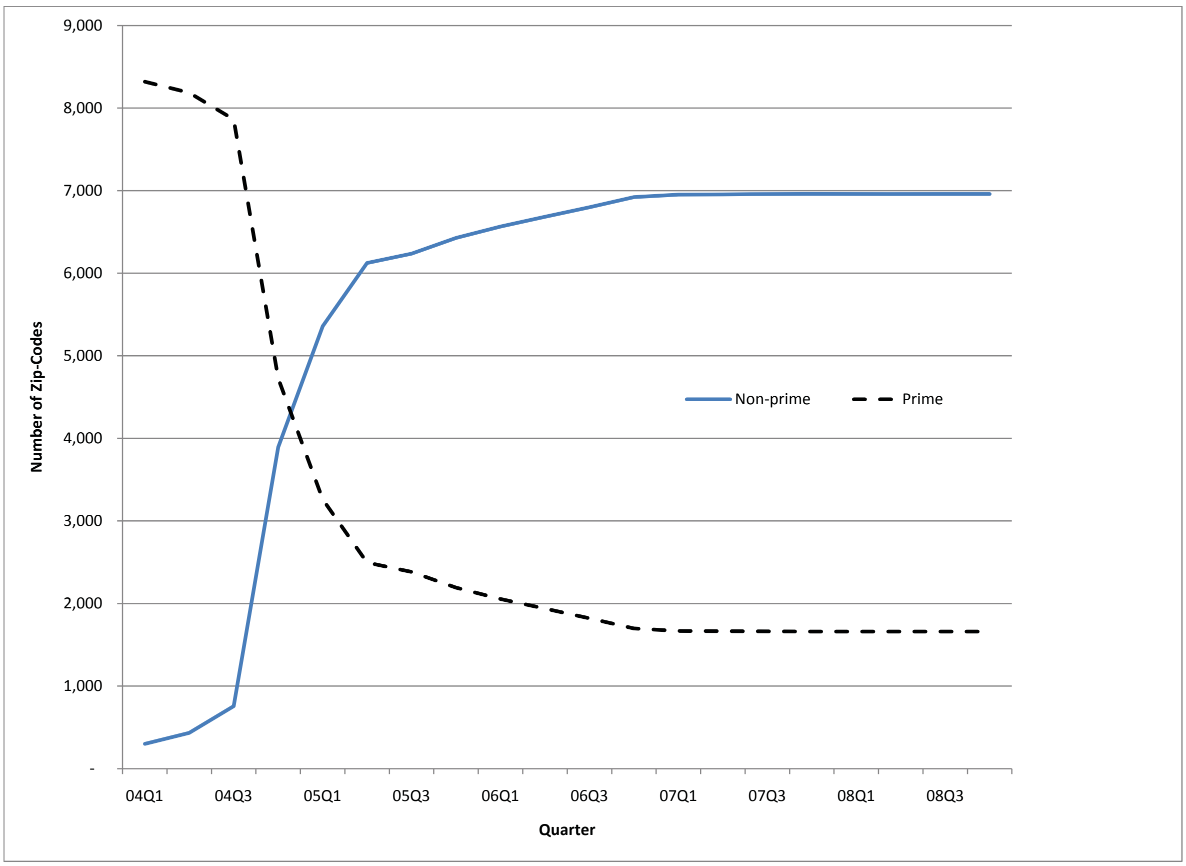


Figure 5: **Number of qualified sample zip-codes classified as prime and non-prime**

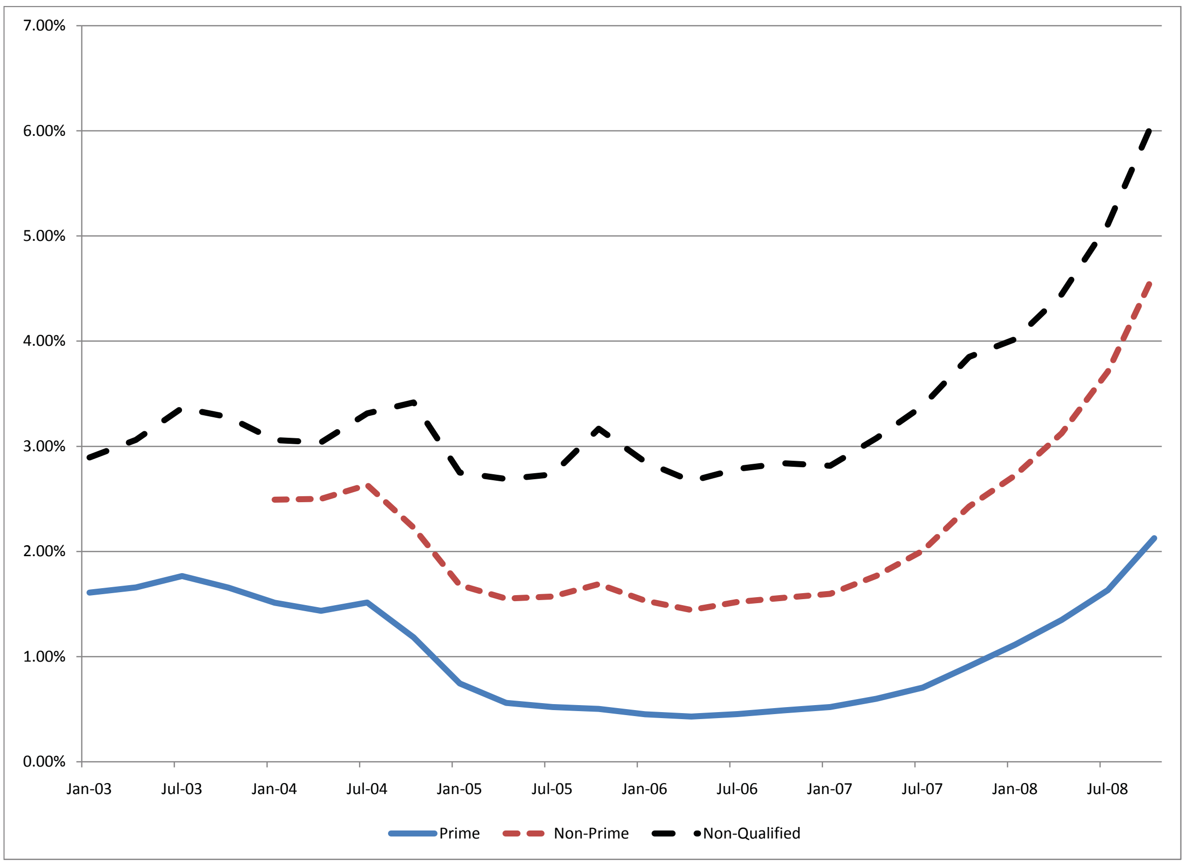


Figure 6: **90-Day default rate for prime, non-prime, and non-qualified zip-codes**

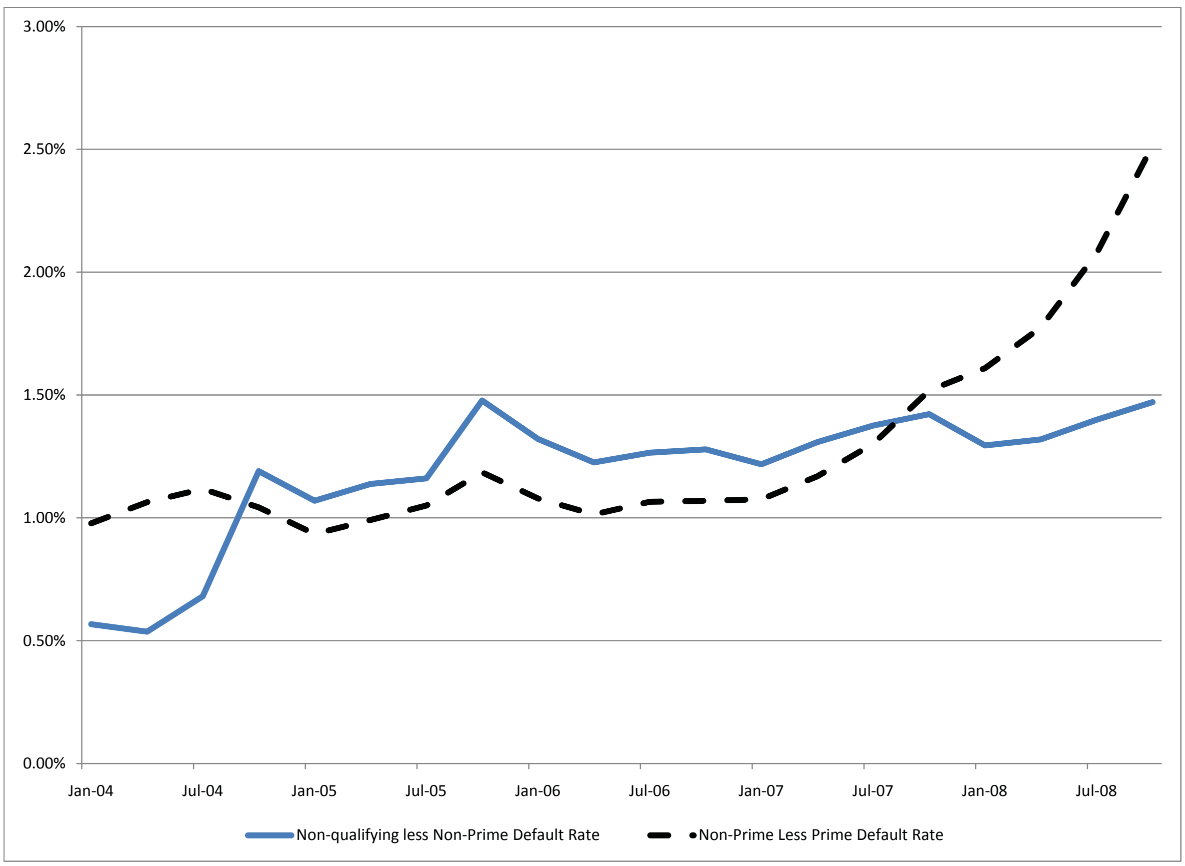


Figure 7: **Difference between the non-qualifying and non-prime zip-code default rates and the prime and non-prime zip-code default rates**