The Subprime Virus: Theory and Evidence

Sumit Agarwal, Brent W. Ambrose

and Yildiray Yildirim*

Current version: February 3, 2011

*Agarwal, ushakri@yahoo.com, Economic Research, Federal Reserve Bank of Chicago, 230 South LaSalle Street Chicago, IL 60604; Ambrose, bwa10@psu.edu, Institute for Real Estate Studies, The Pennsylvania State University, University Park, PA 16802-3306; Yildirim, yildiray@syr.edu, Whitman School of Management, Syracuse University, 721 University Ave Suite 500, Syracuse, NY 13244. We thank John Clapp, Lynn Fisher, 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 Rob McMenamin for his 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.

The Subprime Virus: Theory and Evidence

Abstract

We develop a theoretical model demonstrating the potential spillover effects associated with the introduction of risky assets. Specifically, we examine the potential increase in mortgage default risk on prime mortgages that results from the introduction of subprime mortgages in a local area. The transmission mechanism is that the higher incidence of default by subprime borrowers leads to greater asset price volatility, which in turn impacts the default probability of prime mortgages in the same geographic area. 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

JEL Classification: G2; R2

1 Introduction

The implosion of the US housing market has created a record number of mortgage defaults. Many borrowers with negative equity now find it optimal to default on their mortgages rather than continue to make payments. Recent research now demonstrates that mortgage defaults and foreclosures impose significant negative externalities on mortgaged properties and surrounding property owners. For example, Campbell, Giglio, and Pathak (2009) 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. Furthermore, 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.[1]

One of the catalysts that is often blamed for the most recent boom and subsequent bust in the housing market is the rapid expansion of alternative or subprime mortgages. Numerous studies have examined the role that subprime mortgages played in the current financial crisis.[2] 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. Furthermore, Agarwal et. al (2010) show that the distribution of subprime mortgages across a geographic area is not uniform and that areas with higher concentrations of subprime mortgages experienced greater house price volatility.

[1]In addition, Schuetz, Been, and Ellen (2008), Lin, Rosenblatt, and Yao (2009), and Leonard and Murdoch (2009) find similar spillover effects of foreclosures on prices. Furthermore, Baxter and Lauria (2000) note that foreclosures have a negative impact on communities, while Moreno (1995) documents the direct cost of foreclosures on cities and neighborhoods. 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.

[2]For example, see Ben-David (2011), Demyanyk and Van Hemert (2010), Keys et al (2010), Mian and Sufi (2009), Ashcraft and Schuermann (2008) and Mayer and Pence (2008) among others.

As recent evidence shows, the presence of higher risk, alterative mortgages that subsequently default can have destabilizing effects on surrounding properties. This result is consistent with recent research in finance on default correlations in fixed income securities that results from linkages between individual firms via industry specific and general macro economic conditions (see Zhou, 2001). Thus, the rise of high-risk mortgages raises an interesting research question: To what extent does the presence of subprime mortgages in a geographic area alter the risk profile of 'prime' mortgages in the same area? In other words, what is the impact of the introduction of higher risk mortgage products on a portfolio of prime mortgages? The answer to these questions is directly related to 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 (through default correlation) of the origination of new high-risk mortgages increases the riskiness of the prime portfolio. Our numerical analysis reveals that adding one subprime mortgage to a market comprising three prime mortgages will increase the probability of default for the prime mortgage portfolio by 1.4 times.

Next, we empirically test the model's predictions by examining the default and foreclosure rates at the zip-code level based on their level of subprime origination concentration. Using data on mortgages originated between 2003 and 2008 from LPS Applied Analytics, 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 prime mortgage default rates increased substantially in 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 1.8 percent and a one point increase in the subprime default rate increases the prime mortgage portfolio default rate by 7.5 percent.

2 Theoretical Model

In this section, using a simple model we show the impact of the introduction of subprime mortgages on the prime mortgage default probabilities. Consider a geographic area with N houses. We assume that house prices move according to the following standard geometric Brownian motion processes:

$$\frac{dH_i(t)}{H_i(t)} = \mu_i dt + \sigma_i dW_i(t) \quad i = 1, .., N \text{ and } t = 1, .., T \tag{1}$$

where μ_i is the instantaneous drift coefficient, σ_i is the instantaneous volatility, and we assume that the individual house price processes are correlated with one another, $dW_i dW_j = \rho_{ij} dt$. Now, we set up a simple mortgage model that captures the default risk similar to a structural model. For ease of exposition, we assume a simple interest-only mortgage with a loan balance (P_i) due at maturity.

In structural models, the default time is determined by an 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.[3]

[3]The first-passage-time approach (Black and Cox (1976) extends the original Merton model by allowing the default to occur not only at the debt's maturity, but also prior to this date.

As in the Merton (1974) model, we assume that each household has a constant barrier. We begin by assuming that each household begins as a prime borrower with a 75 percent loan-to-value (LTV) mortgage ($\frac{P_i}{H_i(0)} = 0.75$) from a conservative bank. Next, we assume that some number i households (where $i < N$) originate highrisk subprime mortgages denoted as having high-LTV ratios ($\frac{P_i}{H_i(0)} = 0.95$) from a new lender.[4] We assume that borrowers default when the house price process crosses the constant barrier, P_i, at the maturity.[5] Default can be defined as:

$$D_i = \begin{cases} 1 & \text{if } H_i(T) < P_i \text{ (default)}, \\ 0 & \text{otherwise (no default).} \end{cases} \quad \text{or similarly } D_i = \begin{cases} 1 & \text{if } y_i < 1 \text{ (default)}, \\ 0 & \text{otherwise (no default).} \end{cases}$$

where $y_i = \frac{H_i(T)}{P_i}$. The face value of the loan (e.g. P_i) and y_i change depending on whether the loans are subprime or prime. For prime loans, the barrier is a downward parallel shift. Our aim is to calculate the probability of joint default at T, and study the relationship of subprime loans and prime loans in changing economic conditions.

Based on the Merton model, if only one house exists in the area, $i = 1$, the default probability is given by

$$P(D_i(T) = 1) = P(y_i \leq 1) = \Phi(d_2^i), \text{where } d_2^i = -\frac{\text{Ln}y_i + \left(\mu_i - \frac{\sigma_i^2}{2}\right)T}{\sigma_i\sqrt{T}}, \tag{2}$$

and Φ is the cumulative normal distribution function. One can also drive the joint probability of $N = 2$ based on Zhou (2001). However, for our purpose, $N > 2$, and thus, the probability of joint default at time T is given by

$$P(y_1 \leq 1, y_2 \leq 1, ...y_N \leq 1) = \Phi_N(d_2^1, d_2^2, d_2^3, ..., d_2^N, \rho_{12}, \rho_{13}, ...), \tag{3}$$

where Φ_N is the multinomial cumulative distribution function.

[4]For ease of modeling, we assume that the 'prime' and 'subprime' labels reflect the mortgage risk as captured by the low or high loan-to-value ratio.

[5]Our set up more suitable for the interest only mortgages, and when the price process hits the principal amount then default occurs at the maturity.

We can now conduct a numerical analysis of the above equation. The lending market consists of two mortgage types: 10-year, interest only prime loans with a 75 percent LTV ratio and 10-year, interest only subprime loans with a 95 percent LTV ratio. We also set the drift and the correlation of individual house price process $\mu_i = 0.05$ and $\rho_{ij} = 0.4$ respectively. Figure 1 shows the results for the numerical analysis demonstrating the impact on the prime portfolio of a lender entering the market and originating high-risk mortgages.

First, consider the default probabilities for portfolios consisting of one prime or subprime mortgage, labeled $(n_p = 1; n_s = 0)$ and $(n_p = 0; n_s = 1)$ respectively.[6] As the subprime mortgage has a higher initial loan-to-value ratio, we see that it has a correspondingly higher initial default probability. Furthermore, given that we simplified the analysis by restricting the mortgages to interest-only loans, we see that both the prime and subprime mortgages have increasing probabilities of default as time passes, reflecting the volatility (σ_i) in the underlying state variable.

Next, we consider a portfolio consisting of three prime mortgages in a market with no subprime mortgages, labeled $(n_p = 3; n_s = 0)$. The default probability for this portfolio is substantially lower than the default probability for the single prime mortgage portfolio. For example, at time 0 the probability of default for the $(n_p = 3; n_s = 0)$ portfolio is 0.2 percent whereas the probability of default for one prime mortgage is 4.1 percent. This substantial reduction in the default probability represents the effects of diversification that mortgage originators or investors can achieve.

We now consider the impact of a subprime mortgage lender entering the market. The lines labeled $(n_p = 3; n_s = 1)$, $(n_p = 3; n_s = 2)$, and $(n_p = 3; n_s = 3)$ show the impact on the prime portfolio's default probability when a new bank enters the market originating one, two, and three high-risk loans, respectively. It is important to note that the lines in Figure 1 labeled $(n_p = 3; n_s = 1)$, $(n_p = 3; n_s = 2)$, and $(n_p = 3; n_s = 3)$ show the default probabilities on a portfolio of three prime mortgages only, and not the default

[6]For ease of analysis, we model the distinction between prime and subprime mortgages based on LTV and assume that all prime mortgages have LTVs of 75 percent while subprime mortgages have LTVs of 95 percent.

probabilities of a portfolio consisting of prime and subprime mortgages. For example, the line labeled ($n_p = 3; n_s = 3$) represents the default probability for a portfolio of three prime mortgages when three subprime mortgages are included in the market. It does not represent the default probability of a portfolio of three prime mortgages and three subprime mortgages combined. In essence, our analysis captures the impact on a conservative bank that originated only prime mortgages when a second bank entered the market and originated subprime mortgages. The analysis clearly reveals a positive shift in the prime portfolio's default probability as the number of subprime mortgages in the market increase. For example, at period 5 the probability of default for the prime portfolio when three subprime mortgages are in the market ($n_p = 3; n_s = 3$) is 16.1 percent versus 8.5 percent for the prime portfolio with no subprime mortgages present ($n_p = 3; n_s = 0$). By period 10, the probability of default for the ($n_p = 3; n_s = 3$) portfolio has increased to 35.9 percent while the ($n_p = 3; n_s = 0$) portfolio default probability has increased to 18.9 percent. On a relative basis, adding one subprime mortgage to the market increases the prime portfolio's default probability by 1.4 times. Furthermore, adding two subprime mortgages increases the default probability by 1.7 times on average and adding three subprime mortgages increases the default probability on average by 1.9 times.

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 prime mortgage default risk increased as a result of subprime origination activity, our empirical strategy is to classify markets based on subprime origination activity. In order to determine market concentration, we collect data from Lender Processing Service (LPS) Applied Analytics. 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.[7]

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.[8] 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 agreement). 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.

[7]Subprime classification is reported by the servicers contributing to LPS explicitly.
[8]LPS indicates that its database covers over 65 percent of the total residential mortgage market.

Table 1 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. 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 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. 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 2 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 2 shows that the areas with significant subprime exposure in 2003 are different from our qualified, 'prime' areas.[9] For example, the non-qualified areas have substantially lower

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

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, 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 6 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 2, 3, 4, and 5 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 2) 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 3 and 5 reveal a similar patter of subprime growth. 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 4 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 7 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 is 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).[10] Figure 7 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 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 8 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

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

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:

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

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 monthly 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 3 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.[11] 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.

Table 4 reports the estimated coefficients. 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. We also 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 while area median income is negative and significant. Again, these coefficients are consistent with previous empirical research showing that higher income areas are less risky while the presence of vacant properties increases risk. In addition, we find a negative and significant coefficient for percent urban indicating that urban areas tend to have lower default rates.

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

Finally, the positive and significant coefficient on the subprime mortgage origination activity variable confirms the predictions from our theoretical model that an increase in subprime mortgages originations has a positive impact on 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 1.8 percent. In addition, the estimated coefficient for subprime mortgage default rate is positive and significant, confirming the hypothesis that subprime mortgages may have a spillover effect to prime mortgage performance. The estimated coefficient implies that a one point increase in the subprime default rate increases the prime mortgage portfolio default rate by 7.5 percent.

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 the presence of subprime activity is a spurious correlation. Thus, to control for the possible differences in the 2003 baseline default rates, we recast equation 8 as follows:

$$\begin{aligned} \delta_{i,t} - \overline{\delta_{i,03Q4}} &= \alpha + \beta_1 \sum_{k=1}^{t-1} Sub_{i,t-1} + \beta_2 \Delta U_{i,t} + \beta_3 \Delta HPI_{i,t} + \beta_4 \sigma_{i,t}^{HPI} + \beta_5 Sub\delta_{i,t} \\ &\quad + \beta_6 R_{i,t} + \beta_7 \frac{HPI_{i,t}}{\overline{HPI_i}} + \beta_8 X_i + \theta T + \lambda L_i + \epsilon_{i,t} \qquad (5) \end{aligned}$$

where $\overline{\delta_{i,03Q4}}$ represents the default rate in the fourth-quarter of 2003 for zip-code i. Thus, equation 5 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).

The second column of Table 4 reports the estimated coefficients from equation 5. Consistent with the results discussed above, the positive and significant coefficient for the subprime mortgage origination activity variable 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 2 basis points above the 2003 baseline default rate. For example, the zip-code 60614 (Chicago) saw a cumulative increase in the subprime origination market shares 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 36.6 basis points over the baseline 2003Q4 default rate as a result of the increase in subprime origination activity.

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:

$$\begin{aligned} Sub_{i,t} &= \alpha + \beta_1 \Delta U_{i,t} + \beta_2 \Delta HPI_{i,t} + \beta_3 \sigma_{i,t}^{HPI} + \beta_4 Sub\delta_{i,t} \\ &\quad + \beta_5 R_{i,t} + \beta_6 \frac{HPI_{i,t}}{\overline{HPI_i}} + \beta_7 X_i + \epsilon_{i,t} \end{aligned} \tag{6}$$

$$\delta_{i,t} = \gamma + \gamma_1 \sum_{k=1}^{t-1} Sub_{i,t-1} + \gamma_2 Sub\delta_{i,t} + u_{i,t} \tag{7}$$

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.

Table 5 reports the estimated coefficients from the 2SLS estimation. In the first stage, we find a negative coefficient for the change in house prices suggesting that prime areas in 2003 that experienced significant house price increases had lower subprime origination activity. However, we note that the positive coefficient on house price index volatility implies that areas with higher house price risk attracted greater subprime origination activity. In terms of area demographic characteristics, we see that higher minority concentrations are positively correlated with subprime origination while higher income, more urban areas, and more vacant property are associated with lower subprime activity. Finally, we note that areas experiencing higher unemployment have greater subprime activity.

The second stage model shows the effects of the cumulative subprime origination activity on the default rate. As before, we find a positive and significant effect indicating that subprime origination activity is highly correlated with prime mortgage default rates. The estimated coefficient implies that a one point increase in the subprime origination rate results in a 1.9 percent increase in the prime default rate. In addition, we also confirm that higher subprime default rates lead to greater prime default rates with a one point increase in the subprime default rate leading to a 5.9 percent increase in the prime mortgage portfolio default rate.

4 Conclusions

This paper focuses on the simple question: Did the introduction of subprime mortgages alter 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; 2009), the transmission mechanism in our model that leads to the default spillover is that the higher incidence of subprime borrower defaults result in greater asset price volatility and uncertainty regarding all homes, including houses collateralizing prime mortgages in the same geographic area.[12] Through numerical analysis, we demonstrate the impact of the origination of subprime mortgages on the risk of a portfolio of prime mortgages. Consistent with similar models of default correlation, the numerical analysis shows a positive shift in the prime mortgage portfolio default probability as the number of subprime mortgages increase.

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 1.8 percent while a one point increase in the subprime default rate increases the prime mortgage default rate by 7.5 percent.

The results from this 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

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

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, 2010. "Thy Neighbor's Mortgage: Does Living in a Subprime Neighborhood Impact Your Probability of Default?" The Pennsylvania State University 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,* forthcoming.

Campbell, John, Stefano Giglio, and Parag Pathak, 2009. "Forced Sales and House Prices," NBER Working Paper Number 14866.

Carroll, Thomas, Terrence Clauritie, and Helen Neill, 1997. "Effect of Foreclosure Status on Residential Selling Price: Comment," *Journal of Real Estate Research*, 13(1), 95-102.

Demyanyk, Y. and O. Van Hemert, 2010. "Understanding the Subprime Mortgage Crisis," *Review of Financial Studies,* 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.

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

Harding, John, Eric Rosenblatt, and Vincent Yao, 2009. "The Contagion Effect of Foreclosed Properties," *Journal of Urban Economics,* 66, 164-178.

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.

Pennington-Cross, Anthony, 2006. "The Value of Foreclosed Property," *Journal of Real Estate Research,* 28(2), 193-214.

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.

Shilling, James, John Benjamin, and C.F. Sirmans, 1990. "Estimating Net Realizable Value for Distressed Real Estate," *Journal of Real Estate Research,* 5(1), 129-140.

Sumell, Albert, 2009. "The Determinants of Foreclosed Property Values: Evidence from Inner-City Cleveland," *Journal of Housing Research,* 18(1), 45-61.

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

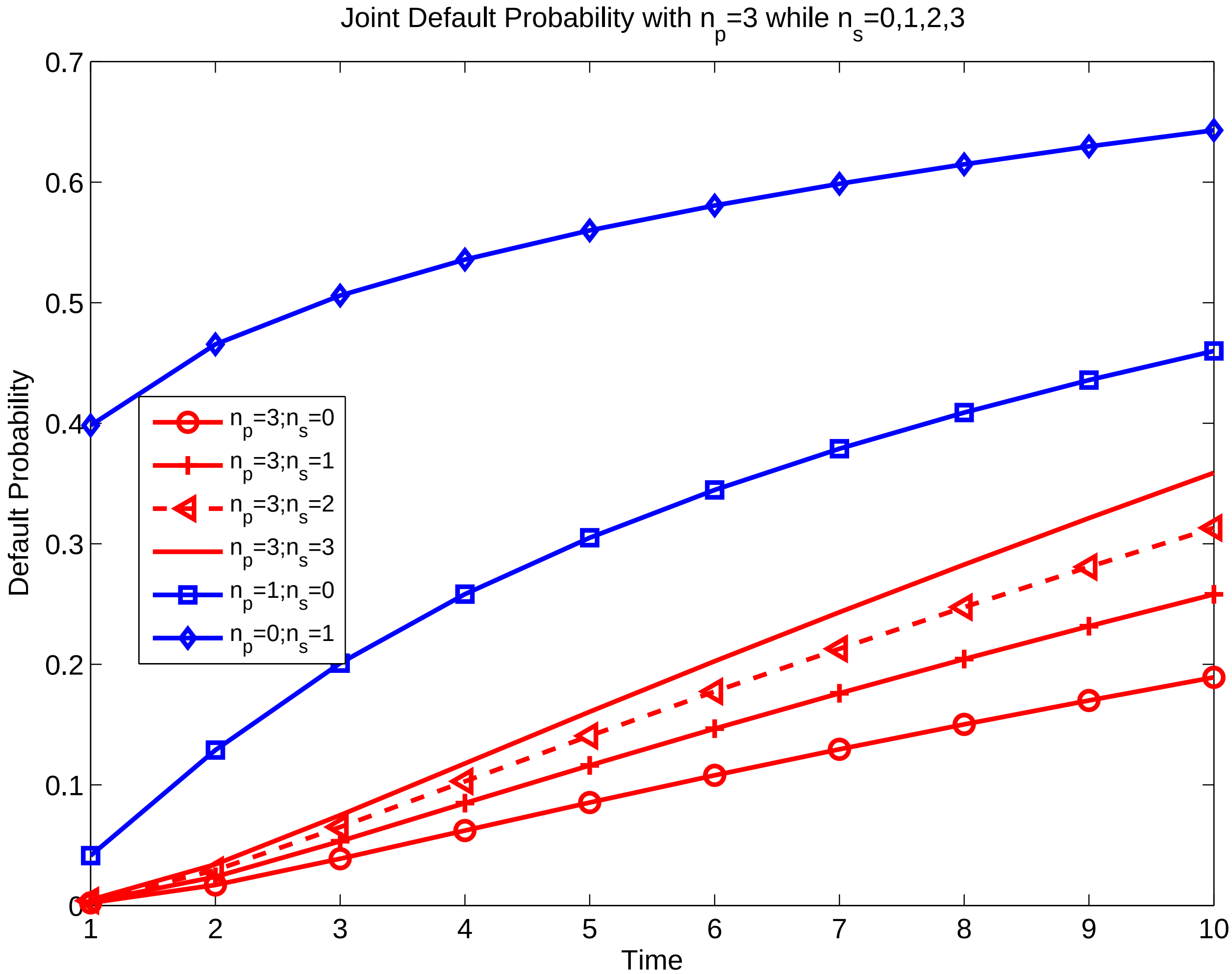
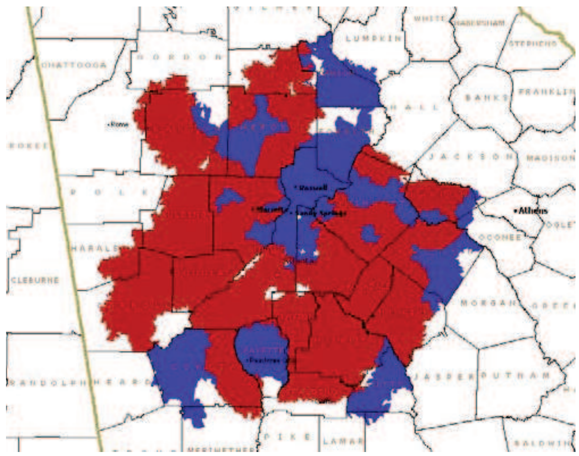
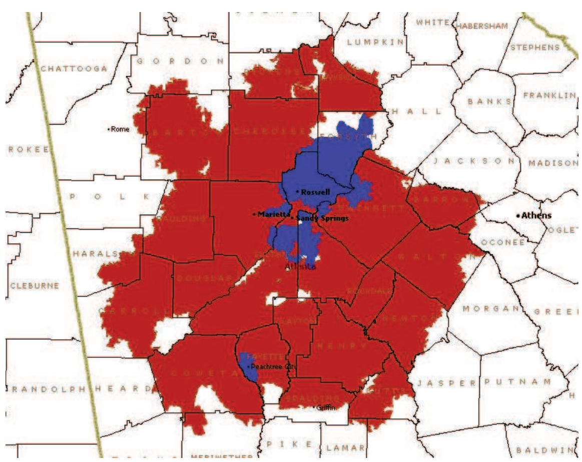


Figure 1: **Numerical Results**



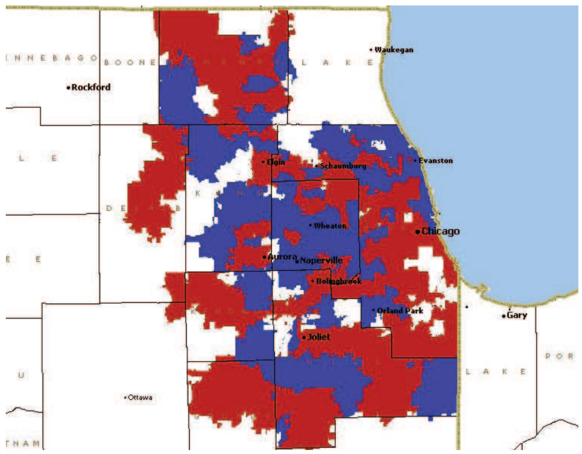
Atlanta 2004:Q4



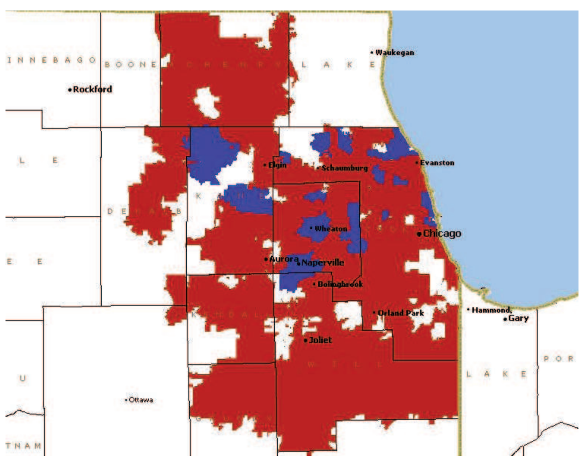
Atlanta 2008:Q4

Red Shading are Subprime, Blue Shading are Prime

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



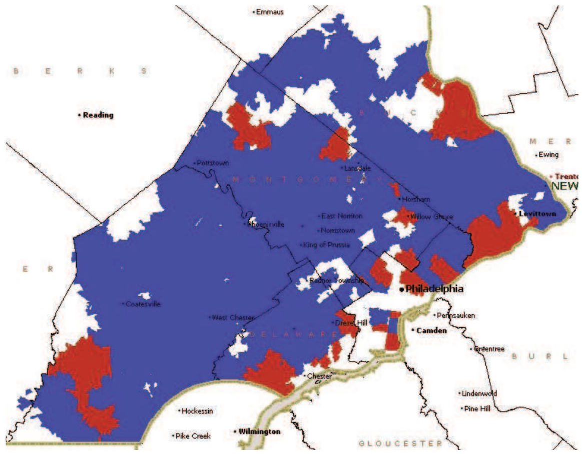
Chicago 2004:Q4



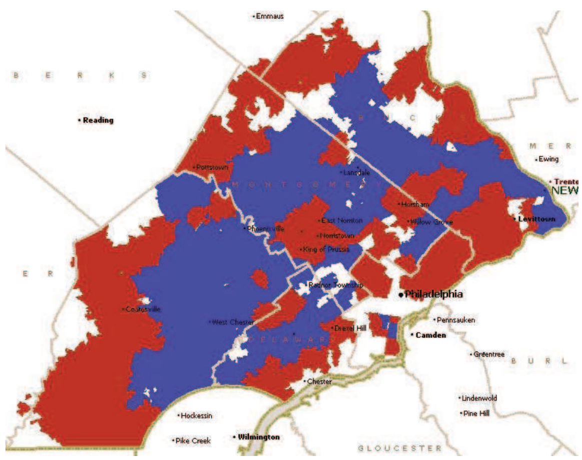
Chicago 20008:Q4

Red Shading are Subprime, Blue Shading are Prime

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



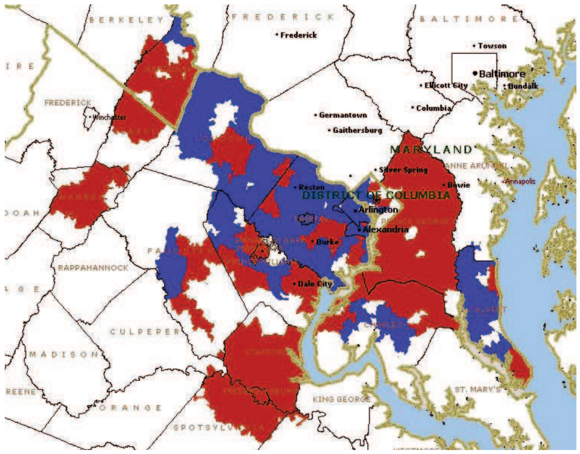
Philadelphia 2004:Q4



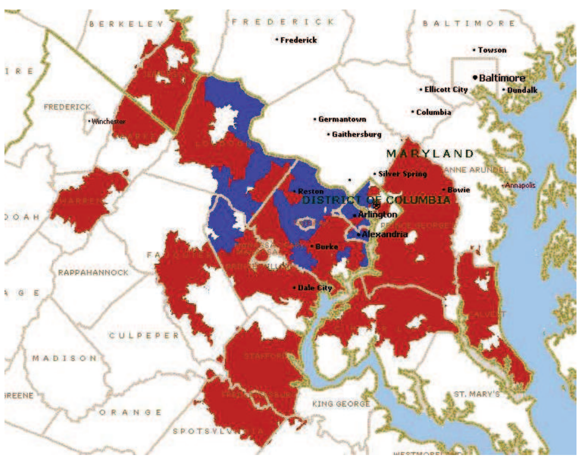
Philadelphia 2008:Q4

Red Shading are Subprime, Blue Shading are Prime

Figure 4: **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 5: **Change in Washington, DC subprime and prime zip-codes between 2004 and 2008**

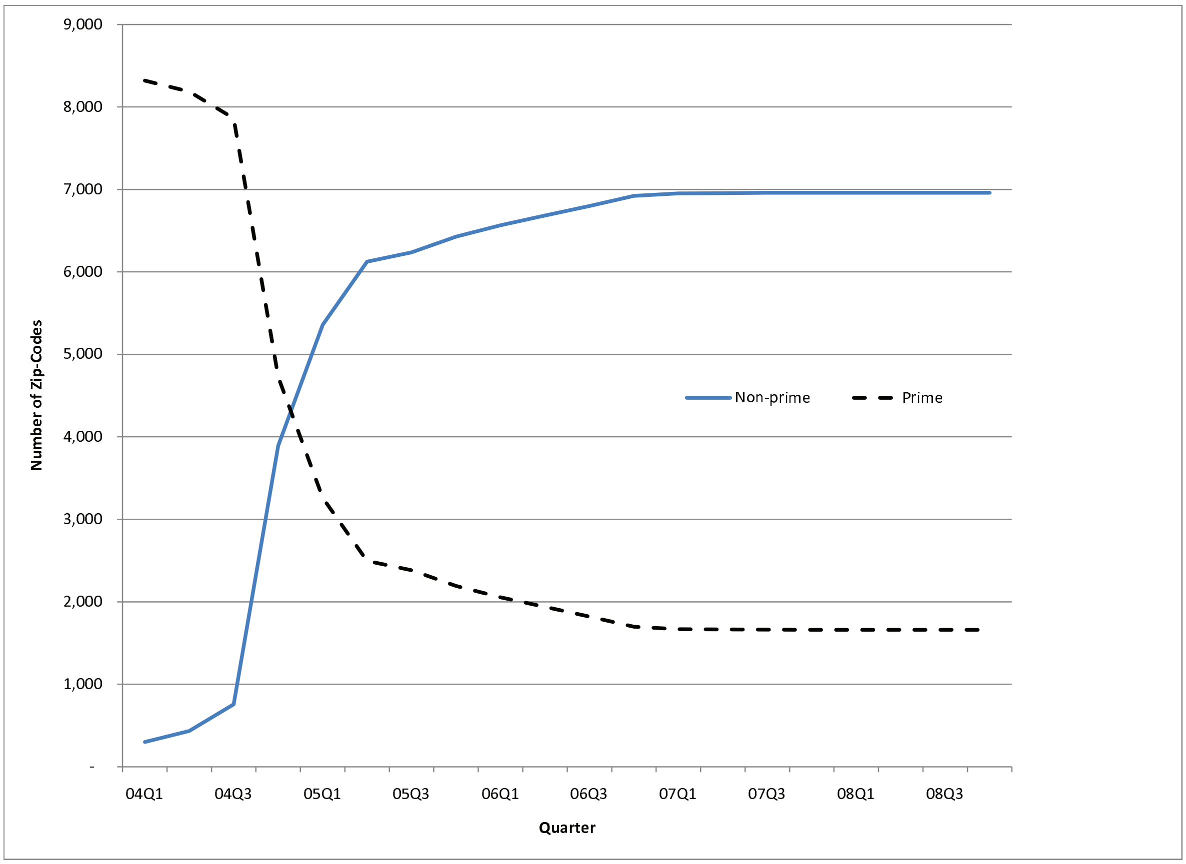


Figure 6: **Number of qualified sample zip-codes classified as prime and non-prime**

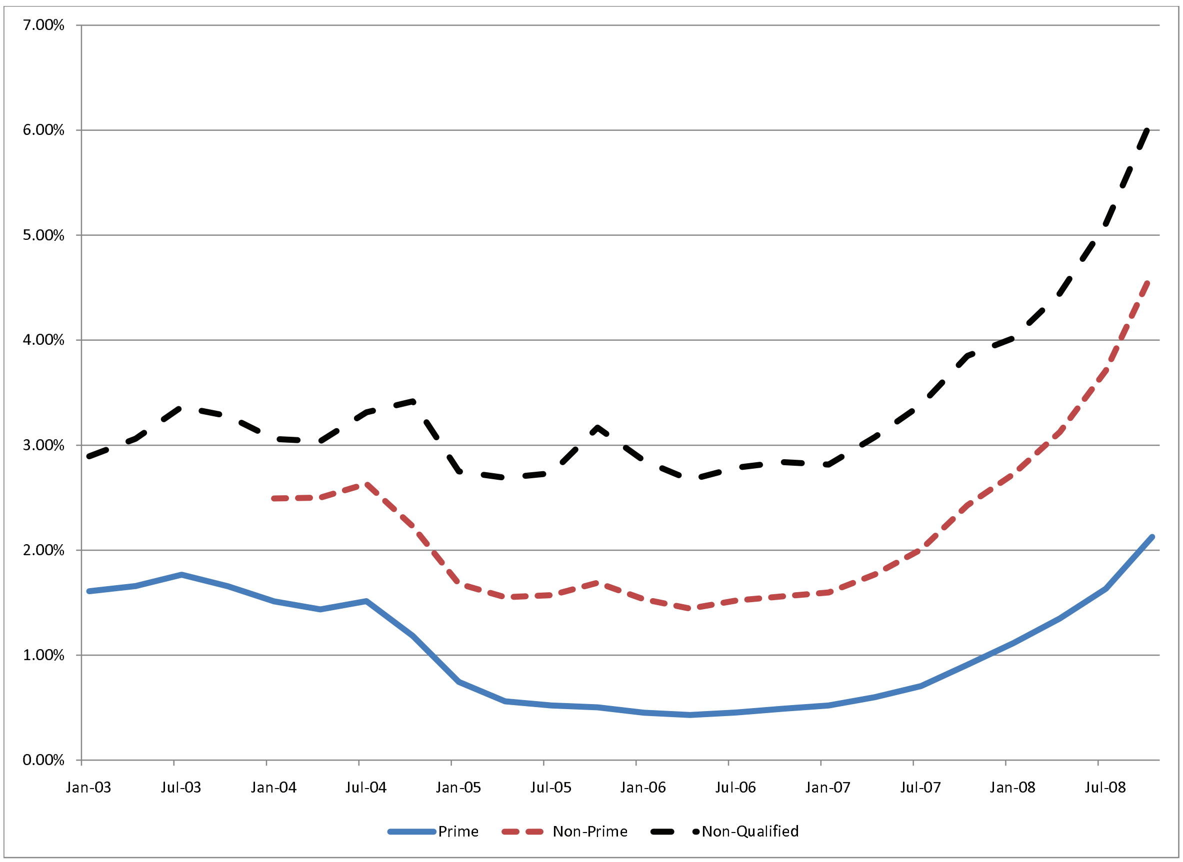


Figure 7: **90-Day default rate for prime, non-prime, and non-qualified zip-codes**

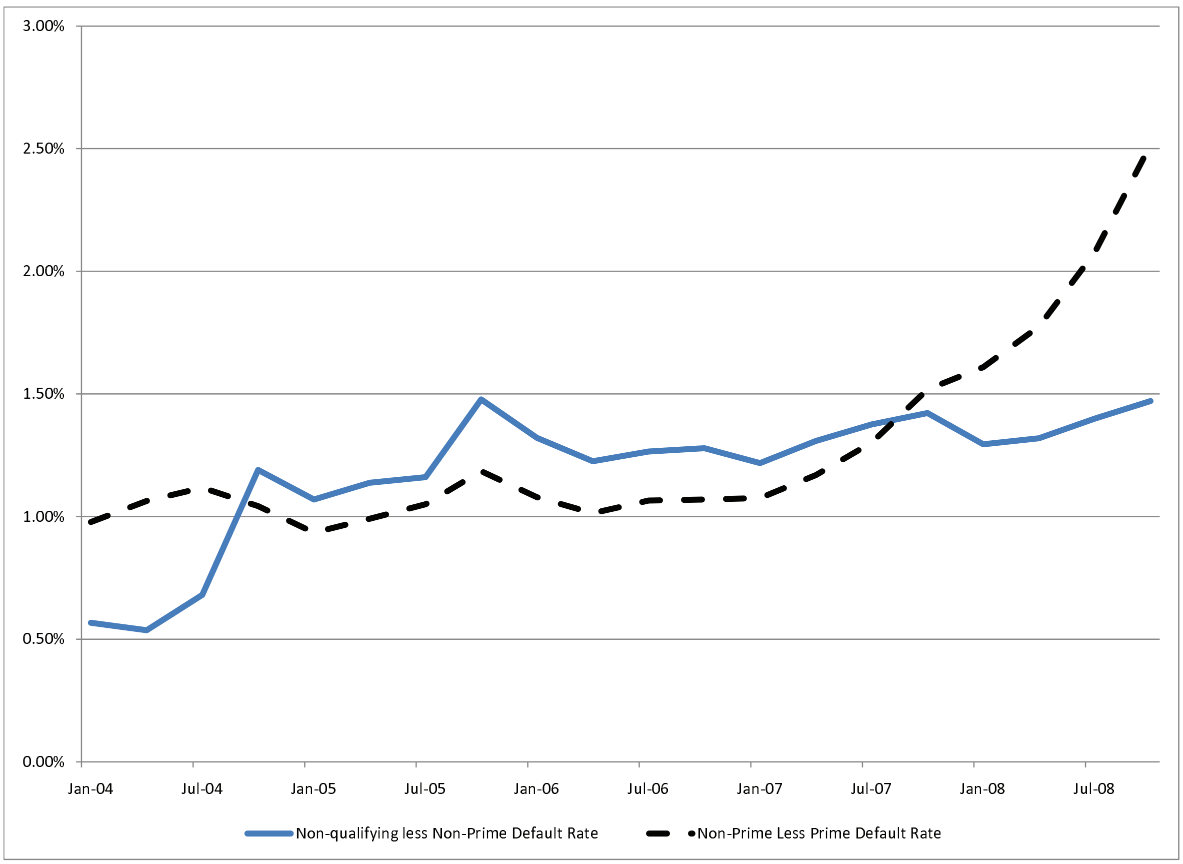


Figure 8: **Difference between the non-qualifying and non-prime zip-code default rates and the prime and non-prime zip-code default rates**

Table 1: 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 2: **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 3: **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 4: **Estimated Regression Coefficients**

VARIABLES	90+ Day Default Rate	Change in Default Rate from 2003
Subprime: Sum of past Subprime Rates	0.018***	0.002***
1 Quarter Lag	(0.000)	(0.000)
1 Quarter change in Unemployment	0.052***	0.047***
	(0.000)	(0.000)
HPI Annualized rate	-0.228***	-0.110***
	(0.000)	(0.000)
HPI Standard Deviation	0.039***	0.025***
	(0.000)	(0.000)
% Minority	0.018***	-0.004***
	(0.000)	(0.000)
Median income (in $1,000)	-0.006***	0.005***
	(0.000)	(0.000)
% Urban	-0.006***	0.001
	(0.000)	(0.134)
% Vacant	0.013***	-0.002
	(0.000)	(0.291)
Median Home Age (2010-year built)	0.004***	0.003***
	(0.001)	(0.001)
Current HPI over Average Origination HPI	0.027***	-0.009***
	(0.000)	(0.000)
Refinance Rate	-2.045***	-1.502***
	(0.000)	(0.000)
Subprime Default Rate	0.075***	0.064***
	(0.000)	(0.000)
Fixed Effects	Time & CBSA	Time & CBSA
Constant	0.930***	1.979***
	(0.000)	(0.000)
Observations	168,428	168,428
Adjusted R-squared	0.706	0.548

Notes: This table reports 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, $\sum_{k=1}^{t-1} Sub_{i,t-1}$ 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 (column 1) and the change in default rates from the average default rate in 2003 (column 2.) *** p<0.01, ** p<0.05, * p<0.1

Table 5: Two-stage Least Squares Regression

VARIABLES	Stage 1 Subprime Rate	Stage 2 90+ Day Default Rate
Subprime: Sum of past Subprime Rates		0.019***
1 Quarter Lag		(0.000)
1 Quarter change in Unemployment	-2.55***	
	(0.000)	
HPI Annualized rate	-4.31***	
	(0.000)	
HPI Standard Deviation	1.38***	
	(0.000)	
% Minority	0.64	
	(0.000)	
Median income (in $1,000)	-0.75***	
	(0.000)	
% Urban	-0.11***	
	(0.000)	
% Vacant	-0.26***	
	(0.000)	
Median Home Age (2010-year built)	-0.09***	
	(0.000)	
Current HPI over Average Origination HPI	0.97***	
	(0.000)	
Refinance Rate	42.85***	
	(0.000)	
Subprime Default Rate	2.43***	0.059***
	(0.000)	(0.000)
Constant	-44.64***	0.183***
	(0.000)	(0.000)
Observations	168,428	168,428
R-squared	0.496	0.390

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

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

$$\delta_{i,t} = \gamma + \gamma_1 \sum_{k=1}^{t-1} Sub_{i,t-1} + \gamma_2 Sub\delta_{i,t} + u_{i,t}$$

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 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. *** p<0.01, ** p<0.05, * p<0.1