# Geographic Patterns of Serious Mortgage Delinquency: Cross-MSA Comparisons

**Lariece M. Brown**
Federal Deposit Insurance Corporation

**Hui-Chin Chen**
**Melissa T. Narragon**
Freddie Mac

**Paul S. Calem**
Federal Reserve System

*The views expressed in this article are those of the authors and do not necessarily reflect those of the Federal Deposit Insurance Corporation, the Federal Housing Finance Agency, the Federal Reserve Board, or Freddie Mac.*

## Abstract

*This article examines the distribution of impaired mortgages across neighborhoods, defined at the ZIP Code level, in 91 metropolitan areas as of the fourth quarter of 2008, well into the recent U.S. mortgage crisis. We catalogue serious mortgage delinquency patterns by metropolitan area based on features of the geographic distribution, including measures of dispersion across neighborhoods and of spatial autocorrelation. The findings are potentially informative for assessing local and neighborhood consequences of the mortgage crisis and for selecting and implementing strategies to ameliorate the effects of foreclosure.*

# Introduction

The tremendous volume of mortgage delinquencies and foreclosures since 2007 is an ongoing national crisis, but fashioning an appropriate policy or private-sector response requires assessing the local manifestations of the crisis. That the appropriate response depends on the neighborhood distribution of seriously delinquent mortgages in a metropolitan area—the extent to which such mortgages are concentrated in high-foreclosure neighborhoods and whether the latter are sparse or numerous, and are clustered together, dispersed, or isolated—has become increasingly clear.

For example, Goldstein (2010) introduced a data-based tool labeled "Market Value Analysis" that can be used to target public-sector and nonprofit neighborhood stabilization funds.[1] The author emphasized that "targeting places where the problem is manageable and the surrounding markets have strength is critical to success" (Goldstein, 2010: 73). An illustrative application to the city of Philadelphia identified neighborhoods where vacancy and foreclosure were geographically confined so that interventions are likely to succeed.

This article surveys and classifies the variety of spatial patterns of serious delinquency observed across U.S. metropolitan areas. The article's primary objectives are to highlight important differences in the spatial distribution of mortgage delinquency across metropolitan areas and to promote discussion of what public- and private-sector strategies are most suitable in each context. In particular, our typology may facilitate information sharing among cities with similar circumstances.

Secondarily, the article examines some housing market and economic conditions associated with the different spatial patterns. Although overall delinquency rates are highest in cities with large house price declines or high unemployment rates, this examination highlights how most other cities have high-delinquency pockets, mostly because of subprime lending concentrations.

Specifically, this article examines the mortgage delinquency distribution across neighborhoods, defined at the ZIP Code level, within U.S. metropolitan statistical areas (MSAs) as of the fourth quarter of 2008, well into the mortgage crisis. The results classify metropolitan areas into six groups:

1. Low-to-moderate mean and high spatial autocorrelation: a modest number of high- or moderately high-delinquency neighborhoods that are clustered together.
2. High mean and standard deviation: wide variation across neighborhoods, with most delinquencies occurring in distressed neighborhoods.
3. High positive skewness: mostly multiple high-delinquency neighborhoods, some with extremely high delinquency rates.
4. Low-to-moderate mean, high positive skewness, and steep gradient around the peak delinquency neighborhood: a modest number of neighborhoods distinguished by high delinquency rates, including at least one spatial outlier.

[1] The analytical approach constructs a set of neighborhood indicators, such as foreclosure and vacancy rates, assessed at the census block-group level, and uses them to cluster neighborhoods into categories reflecting dimensionality and degree of distress.

5. Steep gradient around the peak delinquency neighborhood, indicating at least one spatial outlier: in general, isolated problem neighborhoods.

6. All other cities: somewhat more varied, but generally exhibiting moderate mean and low-to-moderate standard deviation of spatial delinquency.

This article contributes to a developing literature analyzing foreclosures and REO (Real Estate Owned) properties from a geographic perspective and deriving implications for neighborhood stabilization strategy.[2] Immergluck (2009) classified metropolitan areas based on the level and change in density of REO properties from 2006 through 2008 and compared REO accumulation across central city and suburban locations. The analysis highlighted three types of metropolitan areas: (1) areas with low-to-moderate initial REO densities and stable prices, (2) those with initially high REO density and either stable prices or declines in value and increases in REO density from 2006 through 2008, and (3) "boom and bust" areas characterized by steep declines in home values accompanied by rising REO density over this period. The latter category tended to have higher REO concentrations in suburban areas.[3] The author emphasized that "understanding the accumulation of REO inventories across and within metropolitan areas is important for formulating policies and informing community development practice regarding how to stabilize communities and neighborhoods that have been affected by surging foreclosures and vacant properties" (Immergluck, 2009: 28).[4]

Immergluck (2010a) revisited the subject, drilling down to the neighborhood (ZIP Code) level to investigate factors affecting REO accumulation from 2006 through 2008. The analysis indicated that the locations of high-risk lending activity and rapid housing development explain most of the urban-versus-suburban distribution of REO accumulation across metropolitan areas.[5] Edmiston (2009) examined factors associated with foreclosure rate differences across census tracts within the 10th Federal Reserve District as of year-end 2008.[6] The analysis found that concentrations of foreclosures in lower income areas are explained by concentrations of subprime mortgages.

The analysis in this article proceeds as follows. We first calculate distributional moments of the ZIP Code-level delinquency rates, and several measures of their spatial distribution across ZIP Codes. We next conduct a cluster analysis (using the principal component measures) to determine metropolitan area groupings based on common geographical patterns. Finally, we conduct a principal components regression analysis, exploring the relationship of these distributional moments and spatial measures (reduced to their principal components) to subprime lending patterns and economic factors.

[2] REO properties are those that have been acquired by lenders via foreclosure.

[3] The analysis also indicates that among suburban ZIP Codes, those with long commute times experienced larger REO increases over the November 2006-to-2008 period than those with shorter commute times.

[4] For instance, the paper suggests, as an implication of disproportionate REO shares in ZIP Codes with long commute times, that "it may be unwise to spend scarce resources attempting to redevelop residential patterns that may not be highly sustainable in the context of more conservative mortgage markets or higher long-term energy and transportation costs" (Immergluck, 2009: 28).

[5] Immergluck (2010b) examined both levels of and changes in REO activity from August 2006 through August 2008 across metropolitan areas, particularly in relation to changes in home values and the legal environment affecting foreclosures.

[6] The 10th Federal Reserve District consists of Colorado, Kansas, Nebraska, Oklahoma, Wyoming, and parts of western Missouri and northern New Mexico.

The article is organized correspondingly. The next section describes our data sources. The section on geospatial characterization follows with the calculation of the distributional moments and spatial measures and their principal components. The section following that presents the cluster analysis, emphasizing the implications of the results for developing appropriate policy responses. The principal component regression analysis precedes the concluding section.

## Data Sources

We draw data for the study from several sources. We obtain estimates of prime and subprime mortgage delinquency rates as of October 2008 by ZIP Code, using the CoreLogic TrueStandings Servicing® online data analytics tool.[7] This online business intelligence platform accesses the prime and subprime mortgage databases of CoreLogic. These databases provide current information on the payment status of active mortgages serviced by the top mortgage-servicing institutions or sold to Fannie Mae or Freddie Mac.[8] Historical information for both paid-off and active loans is also available, by origination month, as are the state, county, and ZIP Code location of the financed property. We restrict our attention to first-lien, conventional mortgages. For this article, we define delinquency as 60 or more days past due.[9]

The CoreLogic databases do not provide a full count of all active mortgage loans in all ZIP Codes, because not all institutions that service mortgages contribute to these databases. Therefore, we adjust the active loan counts from the CoreLogic servicing data based on an estimate of the undercount in each ZIP Code.[10] Specifically, we measure the undercount by comparing the number of 2005 and 2006 mortgage originations in the CoreLogic data against the number reported to federal regulatory authorities in Home Mortgage Disclosure Act (HMDA) data. The procedure is discussed in greater detail in the following section.

[7] Information about TrueStandings Servicing® is available at http://www.corelogic.com.

[8] The loans CoreLogic assigns to its subprime database are either serviced by institutions that specialize in servicing subprime loans or identified as subprime by the servicing institution. Despite the recent demise of most subprime-specializing institutions, the subprime database continues to track active subprime loan performance because the servicing of these loans has largely transferred to other institutions that contribute to the database. In contrast to CoreLogic's more commonly used, loan-level subprime securities database, the subprime-servicing database provides information on loans retained in bank portfolios as well as those in securities.

[9] Although adverse neighborhood effects generally are associated with properties in later stages of foreclosure and REO, we favor including all loans 60 or more days past due in our analysis of delinquency patterns, for several reasons. First, foreclosure moratoria and loan modification programs have artificially slowed the transition through foreclosure into REO, so that our measure may be a better indicator of actual "facts on the ground." Second, our measure is somewhat forward looking, because most loans in early stages of delinquency as of the analysis date will move into later stages of foreclosure and REO, given the relatively low cure rates associated with the mortgage crisis. Third, early stages of delinquency are relevant when considering effective policy responses. Moreover, because the 60-plus-days-past-due measure is dominated by longer term delinquent loans that are in foreclosure or REO, and because neighborhoods with lower delinquency rates in general will also have higher cure rates, we would not expect classifications based on longer term delinquency to be much different from those arising from our cluster analysis.

[10] Although CoreLogic takes steps to eliminate duplication, some duplicate reporting of loans may occur in the data obtained from Fannie Mae and Freddie Mac by the servicers of these loans. In some ZIP Codes, we observe excess counts and adjust these counts, as well.

Thus, we develop estimates of active loan counts, prime and subprime, as of October 2008, by ZIP Code. We aggregate the estimated prime and subprime delinquency rates and active loan counts to obtain estimates of overall mortgage delinquency rates by ZIP Code.

We use additional data sources to obtain explanatory variables for the regression analysis of metropolitan-area delinquency characteristics. We use 2005, 2006, and 2007 HMDA data to construct variables descriptive of the mortgage market in a metropolitan area, such as share of home purchase loans by occupancy type (owner *vs.* nonowner occupied). We rely on Economy.com for data describing local economic and housing market conditions from 2005 through 2008, including annual house price appreciation rates, annual changes in housing starts, affordability index, and unemployment rates by MSA.

## Estimating Active Loan Counts by ZIP Code

As discussed previously, we adjust the active loan counts from the CoreLogic servicing data by comparing 2005 and 2006 origination counts in the CoreLogic data with origination counts from HMDA data. Because the CoreLogic data provide the state, county, and ZIP Code associated with a mortgage, whereas HMDA data indicate the state, county, and census tract, not the ZIP Code, we first map state, county, and census tract into ZIP Code(s).[11] We apply separate adjustments to prime and subprime loan counts, associating high-cost mortgages in HMDA data (those with a reported above-prime rate spread) with subprime.

Let $n_j$ denote the number of originations reported in the CoreLogic subprime servicing data, and let $N_j$ denote the number of subprime (high-cost) originations in HMDA data, for ZIP Code $j$ in 2005 through 2006. Our adjustment factor is then the ratio $\alpha_j = n_j/N_j$. We multiply the 2008 active loan count in the CoreLogic subprime servicing data by $\alpha_j$ to obtain the estimated active subprime loan count for ZIP Code $j$. We apply the analogous procedure to estimate active prime loan counts. ZIP Codes with fewer than 50 estimated total (prime plus subprime) active loans are excluded from the study.[12]

Note that this procedure assumes that the within-ZIP delinquency rates observed for subprime loans included in the CoreLogic subprime servicing data are representative of the aggregate (observed and unobserved) within-ZIP delinquency rate; we make the same assumption regarding the prime data. Likewise, this procedure assumes that the servicing databases are representative with respect to within-ZIP proportions of 2005-to-2006 originations that remain active in 2008. Although assessing the accuracy of these assumptions is not possible, the fact that we are holding constant both geographic (ZIP Code) location and risk category (prime versus subprime) provides some assurance that the observed quantities will be reasonable approximations. At the least, correcting for the undercounts is preferable to not doing so.

[11] Where a census tract traversed more than one ZIP Code, we allocated the mortgages across the ZIP Codes in proportion to the loan counts observed in Freddie Mac internal data.

[12] We also exclude ZIP Codes where $\alpha_j$ is implausibly large or small. In addition, we apply consistency checks for the prime active counts using Freddie Mac internal data. For instance, if the estimated active prime loan count for a ZIP Code is less than the number of active loans in Freddie Mac data, we use the active loan count and delinquency rate from Freddie Mac data instead.

### MSA Selection

As defined by the Office of Management and Budget, 371 MSAs were in the United States, as of December 2006.[13] To limit the scope of this study to major cities and to ensure the statistical relevance of the measures calculated at the ZIP Code level, we select the 88 MSAs with at least 50 ZIP Codes or 100,000 active mortgages in our data. We include an additional 3, marginally smaller MSAs (Knoxville, Tennessee, Boise, Idaho, and Sioux Falls, South Dakota) to achieve better geographic representation. In Appendix A, we provide the complete list of selected MSAs and the number of ZIP Codes and active mortgages in each.

Large MSAs usually contain several cities along with the suburban areas around the cities. For simplicity, we abbreviate the full name of an individual MSA in the following text by referring to the major city in the MSA. For example, we refer to the New York-Northern New Jersey-Long Island MSA as “New York.” Note that we include as part of an MSA any ZIP Codes that extend beyond the MSA boundary into adjacent non-MSA areas.

## Geospatial Characterization

In this article, we address how delinquent loans, as of October 2008, in individual MSAs were distributed in relation to neighborhood delinquency rate, and whether any generalized patterns emerge across MSAs. Using the ZIP Code-level data described previously, we calculate eight MSA distributional statistics to quantify the patterns in a standardized way. These distributional statistics become the basis of cross-MSA comparisons and analysis.

Note that the focus is the distribution of delinquent loans in relation to neighborhood delinquency rate, not the distribution of the overall population of mortgage borrowers, homeowners, or households in relation to neighborhood delinquency rate. Although these distributions will tend to be similar, we view the former as more relevant for policy analysis addressing the mortgage crisis. For example, the share of a city’s delinquent mortgages contained in high-delinquency neighborhoods is a more important consideration for judging the relevance of the neighborhood dimension than the share of the city’s population located in these neighborhoods.

From a policy perspective, characterizing the shape of the distribution is of interest; for example, knowing whether neighborhoods with extremely high delinquency rates comprise a long tail may be important. Initially, we attempted to fit metropolitan-area delinquency distributions to two-parameter lognormal or beta functional forms. In many cases, however, the data do not conform to these distributions and require greater flexibility in fitting the mean, standard deviation, and shape characteristics (skewness and kurtosis) of the distributions. Therefore, we calculate four descriptive statistics characterizing how the delinquent mortgages in an MSA are distributed in relation to the neighborhood delinquency rate: mean, standard deviation, skewness, and kurtosis.

These moments characterize the delinquent loan distributions across individual ZIP Codes but have no spatial component. The extent to which high-delinquency neighborhoods are spatially isolated, dispersed, or clustered is also of interest from a policy perspective. For example,

[13] See OMB (2006) for more detail.

clustering may imply that delinquency problems are contained (or containable) within a limited geographical area and likely require neighborhood-specific responses. Therefore, we also calculate four gradient and spatial autocorrelation measures, which indicate spatial aspects of the neighborhood delinquency distribution.

We calculate the *mean* as the mean neighborhood delinquency rate for all the delinquent loans in the MSA. Because our data are at the ZIP Code level, we represent neighborhood by ZIP Code and calculate the mean as the weighted average ZIP Code delinquency rate, weighting by number of delinquent loans in the ZIP Code. Note that this is *not* equivalent to the overall measured delinquency rate for the MSA, which we would obtain by weighting by number of active loans.

We use the same weighting concept to calculate standard deviation, skewness, and kurtosis. Note that the standard deviation from this calculation is small because each loan in the same ZIP Code is assigned the same delinquency rate. Therefore, the deviation among delinquent loans in the same ZIP Code is 0; the measure captures only the deviation among the ZIP Codes.

We compute weighted skewness (tendency of the deviations to be larger in one direction) as[14]

$$Skewness = \frac{n}{(n-1)(n-2)} \sum w_i^{3/2}((x_i - \bar{x})/\hat{\sigma})^3, \qquad (1)$$

where $n$ is the number of ZIP Codes, $w$ is the number of delinquent loans in a ZIP Code, $x$ is the ZIP Code delinquency rate, and $\sigma$ is the standard deviation of the delinquency rate distribution. We compute weighted kurtosis (heaviness of the tail of the distribution) as

$$Kurtosis = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum w_i^2((x_i - \bar{x})/\hat{\sigma})^4 - \frac{3(n-1)^2}{(n-2)(n-3)}. \qquad (2)$$

MSAs with high kurtosis usually have high skewness as well; details are provided in the following paragraphs.

*Gradient.* We calculate two measures of gradient—greatest rate of change in delinquency rate between the ZIP Code with the highest (peak) delinquency rate and neighboring ZIP Codes.[15] When we restrict attention to ZIP Codes directly adjacent to the peak-delinquency ZIP Code, we obtain what we call the “first-layer gradient.” We obtain the “second-layer gradient” by focusing on those ZIP Codes adjacent to the directly adjacent ZIP Codes (those that touch the boundaries of the first layer). Specifically,

$$FirstLayerGradient = \underset{i=1....n}{Max}(D_i - D_{Max})/D_{Max}, \qquad (3)$$

$$SecondLayerGradient = \underset{j=1....k}{Max}(D_j - D_{Max})/D_{Max}, \qquad (4)$$

where $D_{Max}$ is the highest ZIP Code delinquency rate in the MSA, $D_i$ is the delinquency rate of the $n$ ZIP Codes adjacent to the ZIP Code with the highest delinquency rate, and $D_j$ is the delinquency rate of the $k$ ZIP Codes adjacent to the $n$ first-layer ZIP Codes.

[14] See “The Univariate Procedure—Descriptive Statistics” from SAS 9.1.3 Online Documentation (The SAS Institute, 2003) at http://support.sas.com/onlinedoc/913/docMainpage.jsp.

[15] In calculus, the gradient of a vector field is the vectors that point in the direction of the greatest rate of increase, with magnitude equal to the greatest rate of change.

A steep gradient suggests that high-delinquency neighborhoods are more isolated or extreme. An MSA with flat first- and second-layer gradients is likely to have a broad region of high-delinquency neighborhoods. An MSA without any high-delinquency-rate areas will have low gradient measures.[16]

## Spatial Autocorrelation

Spatial autocorrelation refers to the degree to which observations from nearby locations (in our context, nearby ZIP Codes) are more likely to have similar magnitude (similar delinquency rate) than by chance alone (Fortin, Dale, and ver Hoef, 2002). We calculate two spatial autocorrelation measures: Moran's I and Geary's C.[17]

Moran's I measures autocorrelation with respect to deviations between individual values of the spatial variable and the mean value:

$$I = \frac{\sum\sum W_{ij}(X_i - \bar{X})(X_j - \bar{X})/W}{\sum (X_i - \bar{X})^2/n}, \tag{5}$$

where $X$ stands for the ZIP Code delinquency rate, $i$ and $j$ represent any pair of ZIP Codes in the MSA, $W_{ij}$ is a weight that declines with the distance between the centroids of ZIP Codes $i$ and $j$; $W$ is the sum of $W_{ij}$, and $n$ is the total number of ZIP Codes in the MSA. Moran's I takes a value between -1 and 1, where 1 means perfectly positive correlation and 0 represents random distribution.

Geary's C measures autocorrelation directly with respect to differences between individual values of the spatial variables. Geary's C is computed as

$$C = \frac{\sum\sum W_{ij}(X_i - X_j)^2/2W}{\sum (X_i - \bar{X})^2/(n-1)}. \tag{6}$$

It varies from 0 for perfect positive autocorrelation to about 2 for a strong negative autocorrelation. If correlation is absent, the expected value equals 1.

A low value of Geary's C corresponds to a high value of Moran's I, both indicating a high degree of spatial autocorrelation. Moran's I is a global indicator, whereas Geary's C is more sensitive to local differences across neighborhood pairs. In general, Moran's I and Geary's C will agree on the existence of spatial autocorrelation, but not necessarily on the magnitude.

Exhibit 1 shows the summary statistics for the eight analysis variables. The mean value across the 91 MSAs of the MSA mean variable is about 0.08, and the mean skewness is about 1.6, consistent with substantial positive skewness for most MSAs.

[16] The gradient measures apply only to the neighborhoods surrounding the ZIP Code with the highest delinquency rate. If a large MSA has multiple pockets of high-delinquency areas, the gradient measures will describe only one of them. Also, the ZIP Code size may affect the gradient measure, as does the delinquency rate differential across neighborhoods; for instance, larger ZIP Codes may mask substantial within-ZIP variation. Nevertheless, the results of our cluster analysis that follows suggest that the gradient measure is an effective tool for identifying metropolitan areas where high-delinquency neighborhoods tend to be more isolated.

[17] Much of our discussion of these spatial autocorrelation measures is drawn from Fortin, Dale, and ver Hoef (2002) and Lembo (2008).

**Exhibit 1**

Summary Statistics of Analysis Variables

| ZIP Code Variable | Mean | Standard Deviation |
|---|---|---|
| Mean | 0.079 | 0.034 |
| Standard deviation | 0.005 | 0.002 |
| Skewness | 1.566 | 0.921 |
| Kurtosis | 5.711 | 7.170 |
| First-layer gradient | 0.261 | 0.181 |
| Second-layer gradient | 0.329 | 0.198 |
| Moran's I | 0.142 | 0.113 |
| Geary's C | 0.946 | 0.185 |

The mean values of the spatial autocorrelation measures (0.14 for Moran's I and 0.95 for Geary's C) suggest that spatial autocorrelation in each city, in general, is not high. These values may be somewhat misleading, however, because we define neighborhoods rather broadly, at the ZIP Code level. Thus, the measure does not reflect the degree of homogeneity of smaller neighborhood units sharing the same ZIP Code. A higher degree of spatial autocorrelation likely would be observed with a narrower definition of neighborhood, such as the census tract level.[18]

Appendix B reports the mean and skewness statistics and the Moran's I spatial autocorrelation measure for each of the 91 MSAs. Although we do not use it in classifying metropolitan-area delinquency patterns, appendix B also includes the Gini coefficient, a measure of inequality in the spatial distribution of delinquent loans.[19] It is calculated using the following formula, where X is the cumulated portion of active loans and Y is the cumulated portion of delinquent loans across ZIP Codes, ordered by number of delinquent loans:

$$G = 1 - \sum_{k=1}^{n} (X_k - X_{k-1})(Y_k + Y_{k-1}) . \qquad (7)$$

A Gini coefficient equal to 0 indicates that delinquent loans in the MSA are distributed exactly in proportion to active loans. The greater the Gini coefficient, the more likely some ZIP Codes contain a disproportionate share of delinquent loans relative to active loans.[20]

## Principal Component Analysis

The multidimensionality associated with the full set of descriptive statistics introduced in the previous sections would confound an effort to analyze delinquency patterns or to draw intuitively meaningful comparisons across metropolitan areas. Moreover, a high correlation exists among these measures, especially among those that quantify related, but not identical, aspects of the

[18] Moreover, the spatial autocorrelation measures indicate the overall degree of spatial autocorrelation, not specifically the degree to which neighborhoods in the high-delinquency tail of a distribution are clustered. Nonetheless, they may be useful as relative measures for comparing spatial patterns across metropolitan areas.

[19] The Gini coefficient is commonly used along with the Lorenz Curve to measure income distribution inequality (Litchfield, 1999).

[20] Among the 91 selected MSAs, the Gini coefficient has a mean value of 0.20 and a standard deviation of 0.09. San Francisco has the highest Gini coefficient (0.47).

distribution (exhibit 2). These correlations complicate the description of delinquency patterns and impart redundancy to an analysis conducted using the full set of descriptive measures.

For example, the mean and standard deviation have a 0.72 correlation, whereas skewness and kurtosis are 92-percent correlated. The variables associated with the spatial aspects of the distribution also are highly correlated with each other and with the standard distribution moments. At the 5-percent significance level, 11 pairings of the 8 focus variables have significant Pearson correlation measures.

Therefore, in anticipation of conducting classification (cluster) and regression analyses of geographic delinquency patterns, we reduce dimensionality by applying principal component analysis (PCA). PCA is often applied in the economic geography literature to reduce the number of variables used to describe and group cities or places along a number of socioeconomic dimensions without losing all the information contained in the numerous variables of interest (Vicino, Hanlon, and Short, 2007).[21] In this application, we use PCA to reduce the number of variables used to describe a metropolitan area from eight measures to four principal components. The principal components essentially are indices that enable us to describe delinquency rate patterns that vary across metropolitan areas, reducing the dimensions of analysis without significant loss of information contained in the original set of analysis variables.

PCA identifies components that are linear combinations of the original variables (factors), orthogonal, and ordered with respect to proportion of variance in the data that is explained. Using PCA with eight variables results in eight components. We reduce the number of components used in the analysis to four by examining the eigenvalues of each component and the proportion of variance explained by each component. We can also provide an interpretation of each component by examining the loadings (weights) on each factor. We choose components that account for a greater variance than any single variable contributes, as captured by an eigenvalue greater than 1. Otherwise, the component accounts for less variance than what is attributed to individual variables.

**Exhibit 2**

Correlation Coefficients of Analysis Variables

| | Mean | Standard Deviation | Skewness | Kurtosis | First-Layer Gradient | Second-Layer Gradient | Moran's I | Geary's C |
|---|---|---|---|---|---|---|---|---|
| Mean | 1.00 | | | | | | | |
| Standard deviation | 0.72* | 1.00 | | | | | | |
| Skewness | (0.16) | (0.10) | 1.00 | | | | | |
| Kurtosis | (0.11) | (0.13) | 0.92* | 1.00 | | | | |
| First-layer gradient | (0.39)* | (0.19) | 0.07 | 0.08 | 1.00 | | | |
| Second-layer gradient | (0.28)* | (0.13) | 0.27 | 0.24* | 0.55* | 1.00 | | |
| Moran's I | 0.08 | 0.26* | (0.08) | (0.12) | (0.15) | (0.17) | 1.00 | |
| Geary's C | 0.06 | 0.01 | 0.36* | 0.31* | (0.15) | 0.05 | (0.30)* | 1.00 |

** Significant at the 5-percent level.*

[21] Researchers have also used PCA, for example, to develop neighborhood quality indices as a function of neighborhood characteristics (Can, 1992) and to include a composite measure of neighborhood quality in house price index construction (Can and Megbolugbe, 1997).

Exhibit 3 reports the eigenvalues associated with each component. An inspection of the eigenvalues shows that the first four components have eigenvalues greater than 1.00 and the fifth component's eigenvalue is only 0.59.

Exhibit 3 also shows the proportion of the variance in the data that each component captures. The first two components account for more than one-half of the variation, and the third and fourth components account for nearly one-third of the variation. Each remaining component accounts for less than 10 percent of the total variation in the data, and omitting them is consistent with the analysis of the eigenvalues.

Examining the factor coefficients (loadings) can yield a high-level interpretation for each of the first four components. The fact that each component has at least two significant loading variables, whereby the variables with the largest coefficients are conceptually related, facilitates interpretation. In the first component, the skewness and kurtosis—both measures related to the tails of the delinquency distribution—have coefficients near 0.50. Thus, this component is viewed as a skewness/kurtosis component. With the second component, the mean and standard deviation are the most relevant coefficients; they contribute 0.52 and 0.59, respectively. The third component is most related to the spatial gradient measure, with our two variations of the gradient having the largest coefficients, about 0.55 each. The fourth component is the measure of autocorrelation captured by Moran's I and Geary's C. The coefficients of C and I have opposite signs, consistent with the negative and statistically significant correlation between the two measures.

**Exhibit 3**

Principal Component Analysis Results

| Component | Eigenvalue | Percent of Variance | Cumulative Percent |
|---|---|---|---|
| 1 | 2.59 | 32 | 32 |
| 2 | 1.85 | 23 | 55 |
| 3 | 1.18 | 15 | 70 |
| 4 | 1.08 | 14 | 84 |
| 5 | 0.59 | 7 | 91 |
| 6 | 0.41 | 5 | 96 |
| 7 | 0.22 | 3 | 99 |
| 8 | 0.07 | 1 | 100 |

## Cluster Analysis

Each of the four principal components from the PCA has a specific value, or "component score," that equals the weighted sum of the original eight distributional measures, whereby the weights are the factor loadings. To classify MSAs based on delinquency patterns, we conduct a k-means cluster analysis of these component scores (Derudder et al., 2003).

Examining the clusters obtained under alternative specifications of number of groupings, we find that six clusters are most satisfactory. Appendix C lists the metropolitan areas by cluster.

Broadly speaking, the groupings suggested by the cluster analysis reflect the degree to which an MSA's delinquent mortgages are concentrated in high-foreclosure neighborhoods and the spatial pattern of those neighborhoods: congregated, dispersed throughout the MSA, or relatively few and isolated.

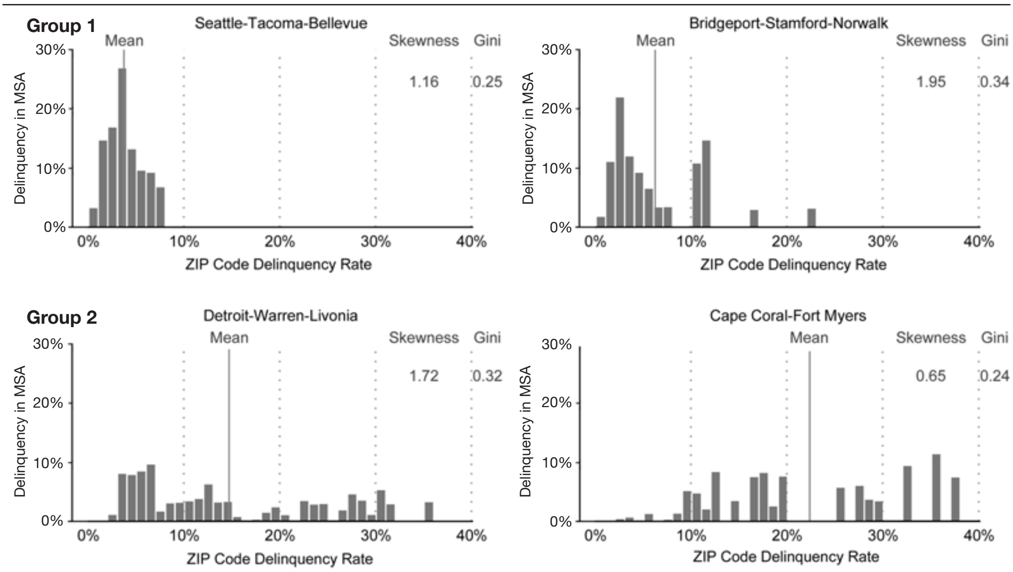
To help characterize the clusters and enable us to visualize how delinquency patterns vary across metropolitan areas in relation to component scores, we create density plots selected as examples for each cluster. The density plots presented in exhibit 4 show how delinquent mortgages in each MSA are distributed in relation to the ZIP Code delinquency rate. They provide a visual reference for components 1 and 2, which are closely associated with this distribution.

The bars in each chart represent the proportion of delinquent loans associated with each neighborhood delinquency rate band, whereby we apply a 1-percentage-point bandwidth. For example, about 5 percent of the delinquent loans in Miami are located in ZIP Codes with a delinquency rate of between 9 and 10 percent, whereas about 10 percent are associated with a neighborhood delinquency rate between 17 and 18 percent.

We also created and examined density maps highlighting the range of ZIP Code delinquency rates through color coding. These maps provide a visual reference for spatial autocorrelation and gradient (closely associated with components 3 and 4). We are unable to reproduce them here, but note some of our observations in the following discussion.[22]

**Exhibit 4**

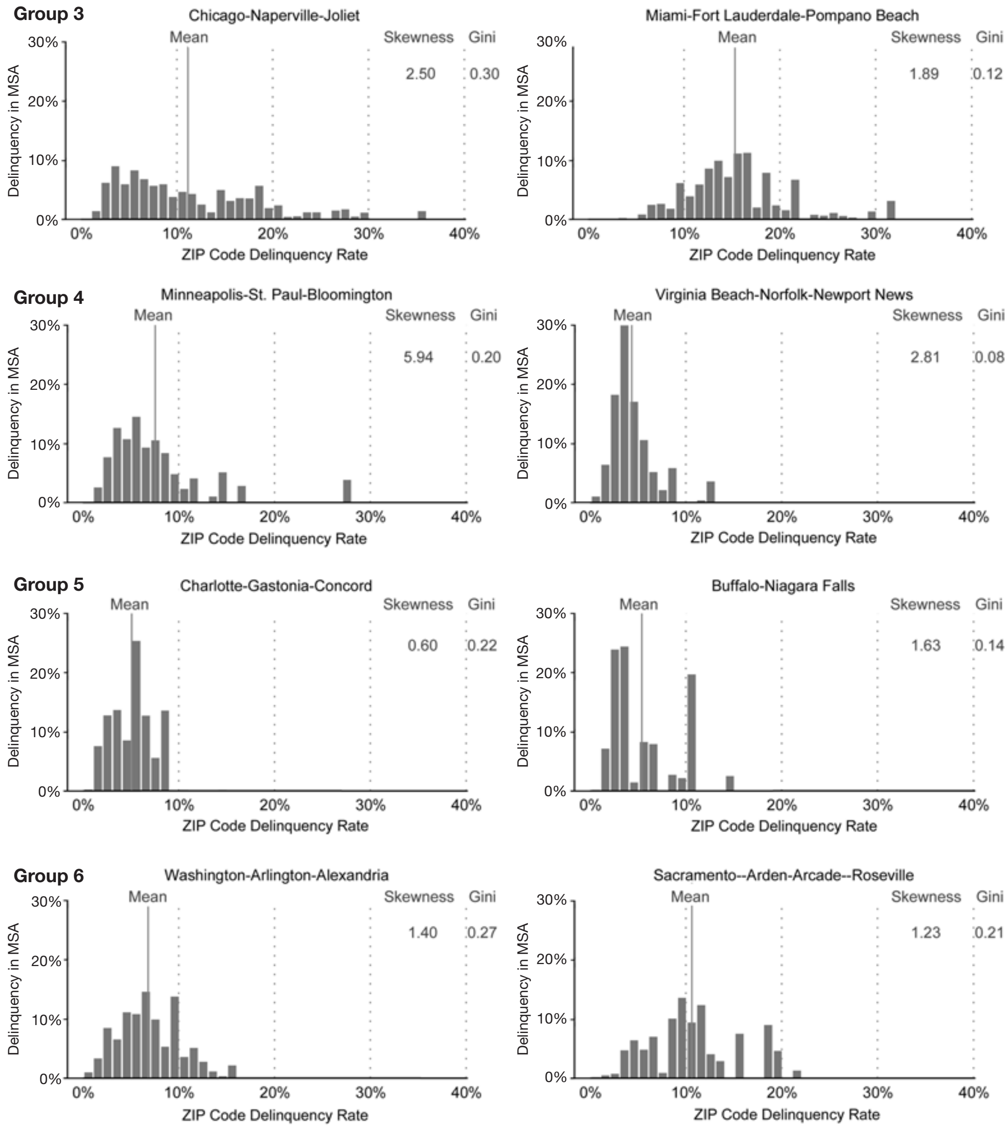
Loan-Level Delinquency Density by ZIP Code Delinquency Rate (1 of 2)




[22] The authors will provide the maps on request.

## Exhibit 4

Loan-Level Delinquency Density by ZIP Code Delinquency Rate (2 of 2)




*MSA = metropolitan statistical area.*

## Group 1

The first cluster analysis grouping consists of MSAs with high spatial autocorrelations and low or moderate delinquency rate means. These MSAs contain a modest number of high- or moderately high-delinquency neighborhoods that are clustered together or comprise a distinct pocket of neighborhoods within the MSA. Examples include Austin, Raleigh, and (as of the third quarter of 2008 analysis date) Seattle.[23]

The density plots associated with this group, as illustrated by those of Bridgeport and Seattle, are relatively compact, with most of the mass in low-delinquency neighborhoods. The distinguishing characteristic of this group, spatial clustering of the higher delinquency neighborhoods, is not evident from the density plots but is observable in density maps. For example, for the Bridgeport metropolitan area, we observe a distinct, concentrated pocket of high delinquency in the urban core. Throughout the remainder of the MSA, we observe lower delinquency rates.

In general, these metropolitan areas have relatively stable housing market and economic environments overall; foreclosure rates in the higher delinquency neighborhoods may or may not rise to a level of concern. Neighborhood effects of delinquency and foreclosure, to the extent they are a concern, would be limited to the higher delinquency pockets, which should then receive particular attention.

We would advise first assessing the potential for effects on house values in adjacent neighborhoods that could cause the foreclosure problem to expand, and taking countermeasures as needed.[24] Targeted use of Neighborhood Stabilization Program (NSP) funds to acquire and rehabilitate properties close to the boundaries of the high-foreclosure area is a possible containment strategy.

In many, if not most, cases, the high-foreclosure pocket will consist of neighborhoods where subprime lending was concentrated (the regression analysis in the following section provides some empirical support for this statement). Thus, strategies to prevent foreclosure, such as loan modification to reduce the payment burden on households with high-cost subprime loans, could help stem neighborhood decline. The high-delinquency pocket may also need to be the focus of efforts to mitigate adverse neighborhood effects of REO and vacant properties, applying the kinds of strategies discussed at length in Fleischman (2010), Ryan (2010), and others in the same volume.[25]

[23] Housing values in Seattle declined substantially and unemployment rose after the third quarter of 2008. As a result, the current delinquency distribution for Seattle likely is different from that in our data, with mean delinquency higher.

[24] Negative externalities associated with foreclosures include lower prices for nearby properties, reduced local property tax base, and high crime rates. Kingsley, Smith, and Price (2009) include a survey of the literature regarding the effect of foreclosures on families and communities.

[25] These strategies include (1) use of public- and nonprofit-sector resources to acquire and rehabilitate foreclosed properties, including bulk acquisitions; (2) partnerships of public-sector and nonprofit agencies with mortgage lenders and servicers to facilitate the sale of REO properties to owner occupants, particularly first-time homebuyers, or to existing occupants (tenants or former owners); (3) partnerships of public-sector and nonprofit agencies with mortgage lenders, servicers, and investors to develop viable REO rental or rent-to-own options for former owners or for existing or new tenants; (4) property code enforcement to mitigate the adverse neighborhood effects of vacancy and abandonment, and legal strategies to facilitate lien transfers to parties willing to perform maintenance or rehabilitation; and (5) demolition of vacant properties and planning for long-term reuse and redevelopment of vacant lots.

In some cases, particularly if the high-foreclosure pocket is an area where overdevelopment led to severe home value declines, market-driven recovery may be the best option. Home value declines may suffice to bring homebuyers back into the community as owner occupants, or to attract private investors who see an opportunity to rehabilitate properties for rental or resale.[26]

## Group 2

The second grouping from the cluster analysis exhibits a high mean and standard deviation for delinquency rates. These cities have wide variation across neighborhoods, with most delinquencies occurring in distressed neighborhoods.

Metropolitan areas in this group include Cape Coral, Detroit, Memphis, Palm Beach, and Stockton. Although they may have some spatial concentrations, high means and very high-delinquency-rate areas in the right tail of the distribution are their most prominent features, as illustrated by the density plots for Cape Coral and Detroit. Widespread occurrence of moderate-to-high delinquency rates characterizes the density maps for these metropolitan areas.

The large number and broad swath of neighborhoods affected by high and very high delinquency necessitate a citywide or regional planning perspective, in contrast with the neighborhood focus associated with Group 1. Strategies to address foreclosure and REO, such as developing viable REO rental or rent-to-own options for former owners or for existing or new tenants, will have to be scalable. Using public- and nonprofit-sector resources directly to acquire and rehabilitate foreclosed properties is unlikely to be an effective strategy, given the scale and scope of the problem. Redevelopment plans may need to incorporate demolition of vacant and abandoned properties and planning for long-term reuse and redevelopment of vacant lots, a strategy that is being used effectively in Cleveland, for example.

The cities in Group 2 contain neighborhoods with extremely high delinquency rates, where as many as one out of every three or four borrowers is delinquent. Restoring stability to these neighborhoods will be a special challenge, requiring particularly intensive or imaginative strategies.

These metropolitan areas also are characterized by steep declines in house prices or by high unemployment. The extent of market-driven recovery will be tied to long-term population and employment prospects for the city or region.

## Group 3

A third grouping is distinguished by a highly positively skewed, long, or fat-tailed distribution. Most metropolitan areas in this group have mean delinquency rates in the moderate range and multiple high-delinquency neighborhoods, which may be clustered together. Delinquency rate dispersion is more one-sided than for Group 2, as reflected in the measures of skewness and kurtosis. This group includes Atlanta, Baltimore, Chicago, Miami, and New York.

[26] Local economic trends, property age and condition, and long-run neighborhood conditions such as high vacancy rates before the mortgage crisis will influence the extent or pace of market-driven recovery, as emphasized by Newburger (2010).

Miami is an example of a metropolitan area with both widespread high delinquency and substantial positive skewness. It has by far the highest mean delinquency rate among cities in Group 3 and is closer to Group 2 in this respect. Chicago is more typical of Group 3. A delinquency map of Chicago shows many high-delinquency neighborhoods, mostly clustered on the city's south side and into neighboring areas southeast of the city, including Gary, Indiana. Chicago has a kurtosis value of 12 and Miami's measured kurtosis is 7.9, both well above the 5.7 sample average or the 3 associated with a normal distribution.

In the case of Miami, where high delinquency rates are widespread throughout the city and its environs, a regional perspective is required, as with the cities in Group 2. In a city more typical of Group 3, such as Chicago, the focus can be on the neighborhoods constituting the high-delinquency tail of the distribution.

Most metropolitan areas in this group have numerous high-delinquency-rate neighborhoods, requiring a planning perspective that encompasses sizable sections of the city or region. In these cases, strategies to address foreclosure and REO will have to be scalable, as discussed for Group 2.[27] The cities in Group 3, like those in Group 2, contain neighborhoods with extremely high delinquency rates, presenting a special challenge.

Often, the higher delinquency neighborhoods will be those where subprime lending was concentrated. Thus, strategies to prevent foreclosure, such as loan modification to reduce the payment burden on households with high-cost subprime loans, could help stem neighborhood decline.

As with Group 1, assessing the potential for spillover effects that could cause the foreclosure problem to expand into adjacent neighborhoods, and taking countermeasures as needed, would be advisable. Again, targeted use of NSP funds to acquire and rehabilitate properties is a possible containment strategy.

## Group 4

Group 4 consists of metropolitan areas with low-to-moderate mean delinquency rates, high positive skewness, and steep gradient around the peak-delinquency neighborhood. Low-to-moderate delinquency neighborhoods predominate in these MSAs. As reflected in the skewness measure, however, some neighborhoods will have distinctly higher delinquency, and at least one spatial outlier neighborhood is characterized by a high gradient value.

In general, metropolitan areas in Group 4 have fewer and less extreme high-delinquency neighborhoods than those in Group 3. They tend to have more high-delinquency pockets, or more spatial separation of high-delinquency neighborhoods, in comparison with Group 1. They also are distinguished by the outlier neighborhood having a much higher delinquency rate than neighboring ZIP

[27] A few cities in Group 3 (Hartford, Oklahoma City, Rochester, and Syracuse) have a relatively low mean delinquency rate. Thus, although the neighborhood delinquency rate distribution is positively skewed, relatively few neighborhoods have high or very high delinquency. From a policy perspective, these cities more closely resemble those in Group 4, although they lack the gradient or spatial outlier aspect.

Codes, which suggests that the high-delinquency pockets are relatively self-contained (spillover is limited). In many, if not most, cases, the high-delinquency neighborhoods reflect concentrations of subprime lending.

Areas in this group include Albany, Barnstable, Minneapolis-St. Paul, and Virginia Beach. The delinquency map for Minneapolis, for example, shows two distinct high-delinquency pockets, one on the east side of St. Paul and another in northwest Minneapolis, extending north over the city boundary into the lower suburbs. They are relatively self-contained, largely surrounded by areas with much lower delinquency rates.

As with Group 1, neighborhood effects of delinquency and foreclosure would be limited to the higher delinquency pockets, which should then receive particular attention. The policy considerations noted for Group 1 apply to Group 4, with two nuances. First, the spatially separated high-delinquency pockets that are more characteristic of Group 4 may not be amenable to the same responses. Second, the high gradient measure suggests that containing the foreclosure problem may be of less concern.

## Group 5

Group 5 is dominated by the gradient measure. Unlike the cities in Group 4, the cities in Group 5 have more or less symmetrical delinquency rate distributions, but, like those cities, they have a high gradient measure. All metropolitan areas slotted to this group have low-to-moderate delinquency means except for Riverside-San Bernardino, which may more appropriately belong to Group 2 (high-mean-delinquency cities), and fell into this group only because of an extreme outlier ZIP Code.

The large gradient suggests that the ZIP Code with the highest delinquency rate is isolated from other problematic ZIP Codes. It is possible that this ZIP Code is one of several problematic neighborhoods that are not near each other or that the MSA does not have many neighborhoods with very high delinquency rates.

The density plots for Buffalo and Charlotte, selected for exhibit 4, suggest that policy implications for this group vary, depending on the nature of the outlier ZIP Code and on potential effects of foreclosure in other, higher delinquency neighborhoods. For example, the Buffalo MSA has a single outlier ZIP Code that is associated with the large gradient but, reflecting the distribution's symmetry, also has substantial mass in neighborhoods with delinquency rates above 10 percent. A delinquency map of Buffalo indicates that the outlier ZIP Code is in the Niagara Falls area, where the delinquency rate is higher than in the other portions of the MSA beyond central Buffalo, whereas a large portion of urban Buffalo has moderately high delinquency rates. Thus, if the neighborhood foreclosure rates are considered problematic, Buffalo may require scalable strategies analogous to those discussed for Group 3. Charlotte, on the other hand, has no ZIP Code with a delinquency rate of 9 percent or more, which suggests that effects on neighborhood stability may not be a concern.

## Group 6

The sixth group is the largest cluster. Group 6 consists of metropolitan areas that have low-to-moderate scores for all components; examples include Philadelphia, Pittsburgh, Sacramento, and Washington, D.C. Most cities in this group have moderate mean and skewness. A few, such as

Sacramento and Tampa, have high means but are distinguished from cities in Groups 2 and 3 by lower standard deviation and skewness; that is, less heterogeneity of neighborhood default rates, without the extremes associated with Groups 2 and 3.[28]

The density plots for Sacramento and Washington, D.C., selected for exhibit 4, illustrate the relatively compact, mildly skewed delinquency distributions that characterize most cities belonging to this group. A delinquency map of Washington, D.C., shows scattered high-delinquency neighborhoods, mostly adjacent to and east of the city or in outlying suburbs to the southwest of the city. A delinquency map of Sacramento illustrates the different case of widespread high delinquency rates through much of the metropolitan area.

Policy implications for these cities vary with the share of delinquencies in high-delinquency neighborhoods. The more typical metropolitan areas in this group, such as Washington, D.C., have a moderate delinquency mean and some scattered high-delinquency neighborhoods, largely tied to subprime concentrations. As with Group 1, neighborhood effects of delinquency and foreclosure would be limited to the higher delinquency pockets, which should then receive particular attention. Cities in this group, such as Sacramento, with high delinquency means and widespread high delinquency rates, require a citywide or regional perspective, comparable with that of Group 2.

## Regression Analysis of Spatial Characteristics

Although this article's primary objective is to classify cities according to spatial characteristics of mortgage delinquency, a secondary goal is to examine the housing market and economic conditions that influence these characteristics. As we emphasized in the preceding section, understanding these factors is important for developing appropriate policy responses. For example, a high foreclosure rate in a lower income neighborhood that is a consequence of concentrations of high-risk lending to vulnerable borrowers might require a different response than would a spike in foreclosures in a far suburb resulting from overbuilding during the housing boom.

In this section, we develop an exploratory, multivariate regression analysis relating the spatial characteristics to economic and housing market conditions across metropolitan areas. The analysis highlights the contribution of subprime lending patterns and identifies the aspects of a metropolitan area's delinquency patterns that are most closely tied to the housing market cycle and to economic conditions. This analysis is a preliminary attempt to identify some basic relationships; it is not intended to be comprehensive.

First, we introduce a set of potential explanatory variables that we classify into three groups: (1) subprime lending spatial distribution measures, (2) housing market factors, and (3) other economic factors. Next, we estimate regression equations for each of the four principal components characterizing the spatial distributions.[29]

[28] The cluster classifications for these cities are robust to using predicted values for their components from the regression analysis in place of actual values.

[29] To select efficiently among the large number of potential explanatory variables, we initially use a stepwise regression procedure for each of the four principal components. Because stepwise regression may generate some arbitrary selections, we subsequently evaluate and test the robustness of the resulting variable selections to the inclusion of omitted variables and adjust the specifications as appropriate. We dropped a few variables where the selection was questionable because of marginal statistical significance and colinearity or redundancy with other included variables.

## Subprime Spatial Distribution

Because subprime loans are disproportionately represented among delinquent mortgages, we expect that distributional moments and spatial patterns of previous subprime lending activity in a metropolitan area influence mortgage delinquency patterns. We describe the characteristics of subprime lending across ZIP Codes using measures analogous to those used for mortgage delinquency: mean, standard deviation, kurtosis, and skewness for percent of active loans that are subprime (weighting by active subprime count), as well as spatial autocorrelation and gradient measures. Exhibit 5 provides summary statistics for the eight analysis variables.

Applying a PCA again reduces the set of descriptive measures to four principal components, again closely associated with (1) skewness and kurtosis, (2) gradient measures, (3) mean and standard deviation, and (4) autocorrelation measures, respectively. Exhibit 6 shows the results of the PCA of subprime spatial characteristics. We include these four principal components as explanatory variables for the regression analysis.

**Exhibit 5**

Summary Statistics of Analysis Variables (Subprime)

| ZIP Code Variable | Mean | Standard Deviation |
|---|---|---|
| Mean | 0.208 | 0.058 |
| Standard deviation | 0.011 | 0.005 |
| Skewness | 1.238 | 0.709 |
| Kurtosis | 3.661 | 4.841 |
| First-layer gradient | 0.199 | 0.165 |
| Second-layer gradient | 0.274 | 0.154 |
| Moran’s I | 0.148 | 0.077 |
| Geary’s C | 0.930 | 0.177 |

**Exhibit 6**

Principal Component Analysis Results (Subprime)

| Component | Eigenvalue | Percent of Variance | Cumulative Percent |
|---|---|---|---|
| 1 | 2.43 | 30 | 30 |
| 2 | 2.17 | 27 | 58 |
| 3 | 1.23 | 15 | 73 |
| 4 | 1.13 | 14 | 87 |
| 5 | 0.41 | 5 | 92 |
| 6 | 0.29 | 4 | 96 |
| 7 | 0.26 | 3 | 99 |
| 8 | 0.07 | 1 | 100 |

## Housing Market Variables

We expect distributional moments and spatial patterns of mortgage delinquency to be tied to housing market activity. For example, delinquency rates will be higher in cities with more rapidly depreciating home values during 2007 and 2008.

Variables associated with the housing market boom and bust considered in the regression analysis include (1) annual home price appreciation rate from the third quarter of 2005 through the third quarter of 2006 and from the third quarter of 2006 through the third quarter of 2008 in each MSA, (2) annual change in MSA housing starts over these periods, (3) the percentage of MSA home purchase loans in 2005 and 2006 that were for nonprimary residence (investment property or second home), and (4) the National Association of REALTORS® housing affordability index for the third quarter of 2005 and the third quarter of 2006.

We also construct measures of the spatial distribution of housing market activity for inclusion in our regression equations. Specifically, we calculate the distributional moments of percent change in home purchase loan originations from the third quarter of 2005 through the third quarter of 2006 and from the third quarter of 2006 through the third quarter of 2007 in each MSA: mean, standard deviation, kurtosis, and skewness (weighting by *ex ante* number of originations), along with the spatial autocorrelation measures. Spatial patterns of home purchase lending activity during the housing boom or at the beginning of the downturn may help differentiate neighborhoods where the market "overheated," as reflected in subsequent mortgage delinquency patterns.

We include, as a potential control variable for each regression equation, the size of the MSA housing market, measured as the log of the total number of active loans as of October 2008. In the equation for the gradient component, we include the ratio of maximum to median ZIP Code delinquency rate in the MSA to control for the potential effect of an outlier neighborhood.[30]

## Economic Factors

Deteriorating economic conditions also affect delinquency patterns. For instance, we expect higher average delinquency rates, on average, in cities with more rapidly increasing unemployment during 2008, or in cities with higher unemployment levels.

Spatial patterns of mortgage delinquency in an MSA may reflect the spatial distribution of borrowers' incomes. We describe the spatial characteristics of borrowers' median incomes within an MSA using the distributional moments—mean, standard deviation, skewness, and kurtosis (weighting by active loan count)—along with the two spatial autocorrelation measures.[31]

## Regression Results

Appendix D lists the economic and housing market variables that we ultimately selected for inclusion in one or more of the regression equations based on consideration of statistical significance and robustness.[32] The mean value and standard deviation of each variable across the 91 metropolitan areas included in the study also appear in appendix D.

[30] The neighborhood delinquency gradient factor may reflect idiosyncratic factors that determine the maximum neighborhood delinquency rate, rather than economic or housing market conditions affecting the broader metropolitan area.

[31] We calculate the borrowers' median income for each ZIP Code relative to MSA median income from pooled 2005, 2006, and 2007 HMDA data.

[32] We employed stepwise regression as a first pass to develop baseline specifications, which we then evaluated for robustness by testing each variable excluded by the stepwise procedure.

Exhibit 7 summarizes the regression results. Results for the skewness, mean, gradient, and spatial autocorrelation components appear in columns 1 through 4, respectively.

**Exhibit 7**

Stepwise Regression Results for Delinquency Spatial Characteristics (Coefficient Estimates, With t-Values in Parentheses)

| Explanatory Variable | Dependent Variable | | | |
|---|---|---|---|---|
| | (1) Skewness Component | (2) Mean Component | (3) Gradient Component | (4) Spatial Auto-Correlation Component |
| Intercept | 2.7935*** (2.8) | – 0.0158 (1.2) | 4.4175*** (3.0) | – 1.2038*** (9.6) |
| Subprime skewness component | 0.7753*** (10.7) | | | – 0.2168*** (4.5) |
| Subprime gradient component | | | 0.1847** (2.1) | 0.1153** (2.3) |
| Subprime mean component | – 0.1886*** (2.7) | 0.5491*** (9.3) | | – 0.1497*** (3.1) |
| Subprime autocorrelation component | | | – 0.1661* (1.9) | 0.6060*** (12.4) |
| Home price percent change, third quarter 2006 to third quarter 2007 | | – 0.1507*** (11.9) | | |
| Home price percent change, third quarter 2005 to third quarter 2006 | | | 0.0416** (2.5) | |
| Affordability index, third quarter 2006 | | | – 0.0048** (2.2) | |
| Unemployment percent change, third quarter 2007 to third quarter 2008 | – 0.0128*** (2.7) | 0.0081** (2.2) | | |
| 2008 unemployment rate | | | 0.1146* (1.9) | |
| Home purchase percent nonprimary residence 2006 | – 2.8750*** (2.6) | | | |
| Ratio, maximum-to-median ZIP Code delinquency rate | | | 0.7109*** (5.7) | |
| Log of MSA active loan count | | | – 0.5333*** (4.7) | |
| Spatial autocorrelation of home purchase lending percent change, third quarter 2006 to third quarter 2007 | | | | 1.2024*** (10.4) |
| Skewness of 2006 ZIP Code median borrower relative to MSA median family income | | | – 0.1861* (1.9) | |
| R-squared | 0.597 | 0.711 | 0.442 | 0.798 |
| Adjusted R-squared | 0.578 | 0.701 | 0.387 | 0.786 |

*MSA = metropolitan statistical area.*

** Significant at the 10-percent level. ** Significant at the 5-percent level. *** Significant at the 1-percent level.*

One broad conclusion that emerges from the analysis is that the shape (skewness and kurtosis) of the neighborhood delinquency rate distribution and the spatial autocorrelation of neighborhood delinquency rates are closely tied to spatial patterns of subprime lending activity during 2005 and 2006.[33] In the regression equations for the skewness/kurtosis and autocorrelation components, the estimated coefficient of the subprime lending counterpart of the dependent variable is the strongest explanatory variable. Thus, the regression analysis supports our previous contention that high-delinquency pockets in metropolitan areas characterized by high positive skewness or high spatial autocorrelation will often be neighborhoods with high subprime concentrations.

A second, broad, and not particularly surprising, conclusion is that economic conditions are at least as important as the subprime share mean/standard deviation component in influencing the delinquency rate mean/standard deviation component. This conclusion is consistent with our previous observation that metropolitan areas in Group 2 experienced harsher housing market or economic declines.

Third, spatial autocorrelation of neighborhood delinquency rates is strongly influenced by spatial autocorrelation of market decline during the "bust" period. Specifically, neighborhood delinquency spatial autocorrelation is positively related to the spatial autocorrelation of percent change in home purchase loan originations from the third quarter of 2006 through the third quarter of 2007.

The gradient component exhibits a somewhat eclectic set of associations. It is positively related to the subprime gradient component and inversely related to subprime spatial autocorrelation. In addition, the neighborhood delinquency gradient component exhibits a positive association with MSA house price appreciation from the third quarter of 2005 through the third quarter of 2006 and an inverse association with housing affordability as of the third quarter of 2006. The latter relationships are consistent with rapidly rising house prices triggering overdevelopment that subsequently generated high-foreclosure pockets.[34]

## Conclusion

We first classified metropolitan areas into six groupings distinguished by their geographic patterns of serious mortgage delinquency. Understanding these patterns and their contributing factors may be informative for assessing local and neighborhood effects of the mortgage crisis and for developing appropriate strategies to mitigate the effects on communities.

[33] We also estimated a regression equation for the Gini coefficient and found that it is very closely tied to the Gini coefficient of subprime lending (relative to total lending) activity.

[34] We observe various additional results specifically for the gradient component. It is inversely related to the log of MSA active loan count, indicating that steeper gradients occur in smaller cities. It is positively related to the ratio of maximum-to-median ZIP Code delinquency rate, confirming the importance of controlling for idiosyncratic neighborhood effects. The stepwise regression for the delinquency gradient component also yields three variables that are statistically significant at the 10-percent level in the equation: the subprime autocorrelation component, the third quarter 2008 unemployment rate, and the skewness of 2006 borrowers' median income relative to MSA median family income across ZIP Codes. F-tests indicate joint and pairwise significance at the 5-percent level for these three variables.

Second, we examined some housing market and economic conditions associated with the different spatial patterns. Although overall delinquency rates are highest in cities with large home price declines or high unemployment, the examination in this article highlights how most other cities have high-delinquency pockets, mostly because of subprime lending concentrations.

The first cluster consists of MSAs with high spatial autocorrelation and low- or moderate-delinquency rate means. These MSAs contain a modest number of high- or moderately high-delinquency neighborhoods that are clustered together or comprise a distinct pocket of neighborhoods within the MSA. The second grouping exhibits a high mean and standard deviation for delinquency rates. These MSAs have wide variation across neighborhoods, with most delinquencies occurring in distressed neighborhoods.

A third grouping is distinguished by a highly positively skewed, long, or fat-tailed distribution. Metropolitan areas in the fourth cluster are characterized by low-to-moderate mean delinquency rates, high positive skewness, and a steep gradient around the peak delinquency neighborhood, whereas those in the fifth cluster are distinguished specifically by their steep gradient. The sixth group consists of metropolitan areas that have low-to-moderate scores for all components.

These classifications are potentially useful for understanding the effects of the mortgage crisis on the dynamics of housing market decline and recovery. For instance, home prices appear to be stabilizing in some metropolitan areas despite little reduction in the inventory of foreclosed properties. Most likely, the foreclosures are concentrated in specific neighborhoods that are lagging behind the overall market recovery, as negative spillover effects tend to diminish with distance.[35]

We believe the analysis has practical applications for selecting or adapting appropriate strategies and policy responses to stabilize neighborhoods and contain foreclosure spillover effects. For example, NSP funds might be most effective for reversing or containing problems associated with foreclosure when spatially targeted to neighborhoods detached from or on the perimeters of broader areas of elevated delinquency and foreclosure. Metropolitan areas with low-to-moderate delinquency means and highly skewed delinquency distributions (Groups 3 and 4, and some cities in Group 6) are those where strategic deployment of NSP funds could be particularly effective at containing neighborhood decline.

Finally, we recognize that this study relies on data from 2008 and that housing markets have further deteriorated in many cities since then, so some cities may need to be reclassified. Although looking back has value, we wish to emphasize the role of this study as an example or template for ongoing analysis.

[35] See Frame (2010) and Lee (2008) for reviews of the literature on price-related spillover effects.

# Appendix A

### Exhibit A-1

Metropolitan Areas Included in the Study (1 of 2)

| Metropolitan Area | CBSA Code | Number of ZIP Codes | Number of Active Loans |
|---|---|---|---|
| Akron, OH | 10420 | 45 | 106,942 |
| Albany-Schenectady-Troy, NY | 10580 | 89 | 95,775 |
| Albuquerque, NM | 10740 | 30 | 110,559 |
| Allentown-Bethlehem-Easton, PA-NJ | 10900 | 65 | 124,184 |
| Atlanta-Sandy Springs-Marietta, GA | 12060 | 186 | 952,203 |
| Austin-Round Rock, TX | 12420 | 74 | 183,871 |
| Bakersfield, CA | 12540 | 37 | 105,919 |
| Baltimore-Towson, MD | 12580 | 140 | 424,934 |
| Barnstable Town, MA | 12700 | 50 | 62,961 |
| Baton Rouge, LA | 12940 | 51 | 83,673 |
| Birmingham-Hoover, AL | 13820 | 97 | 144,998 |
| Boise City-Nampa, ID | 14260 | 31 | 95,176 |
| Boston-Cambridge-Quincy, MA-NH | 14460 | 227 | 553,793 |
| Bridgeport-Stamford-Norwalk, CT | 14860 | 46 | 154,159 |
| Buffalo-Niagara Falls, NY | 15380 | 68 | 101,719 |
| Cape Coral-Fort Myers, FL | 15980 | 28 | 143,676 |
| Charleston-North Charleston, SC | 16700 | 37 | 109,212 |
| Charlotte-Gastonia-Concord, NC-SC | 16740 | 69 | 314,654 |
| Chicago-Naperville-Joliet, IL-IN-WI | 16980 | 353 | 1,275,160 |
| Cincinnati-Middletown, OH-KY-IN | 17140 | 137 | 310,996 |
| Cleveland-Elyria-Mentor, OH | 17460 | 93 | 308,361 |
| Columbia, SC | 17900 | 37 | 103,838 |
| Columbus, OH | 18140 | 90 | 247,042 |
| Dallas-Fort Worth-Arlington, TX | 19100 | 238 | 638,709 |
| Dayton, OH | 19380 | 62 | 113,818 |
| Deltona-Daytona Beach-Ormond Beach, FL | 19660 | 26 | 104,561 |
| Denver-Aurora, CO | 19740 | 117 | 432,145 |
| Des Moines-West Des Moines, IA | 19780 | 53 | 82,061 |
| Detroit-Warren-Livonia, MI | 19820 | 212 | 748,855 |
| Duluth, MN-WI | 20260 | 50 | 33,292 |
| Fresno, CA | 23420 | 42 | 123,920 |
| Grand Rapids-Wyoming, MI | 24340 | 53 | 110,170 |
| Hartford-West Hartford-East Hartford, CT | 25540 | 84 | 161,788 |
| Honolulu, HI | 26180 | 28 | 115,861 |
| Houston-Sugar Land-Baytown, TX | 26420 | 210 | 718,052 |
| Indianapolis-Carmel, IN | 26900 | 95 | 263,862 |
| Jacksonville, FL | 27260 | 53 | 216,240 |
| Kansas City, MO-KS | 28140 | 165 | 254,570 |
| Knoxville, TN | 28940 | 44 | 99,447 |
| Las Vegas-Paradise, NV | 29820 | 59 | 326,813 |
| Los Angeles-Long Beach-Santa Ana, CA | 31100 | 341 | 1,544,348 |
| Louisville/Jefferson County, KY-IN | 31140 | 88 | 176,558 |
| Madison, WI | 31540 | 51 | 77,252 |

**Exhibit A-1**

Metropolitan Areas Included in the Study (2 of 2)

| Metropolitan Area | CBSA Code | Number of ZIP Codes | Number of Active Loans |
|---|---|---|---|
| Memphis, TN-MS-AR | 32820 | 65 | 176,957 |
| Nashville-Davidson-Murfreesboro-Franklin, TN | 34980 | 80 | 223,277 |
| New Haven-Milford, CT | 35300 | 40 | 123,831 |
| New Orleans-Metairie-Kenner, LA | 35380 | 64 | 147,791 |
| New York-Northern New Jersey-Long Island, NY-NJ-PA | 35620 | 820 | 1,956,999 |
| Oklahoma City, OK | 36420 | 74 | 125,597 |
| Omaha-Council Bluffs, NE-IA | 36540 | 76 | 90,279 |
| Orlando-Kissimmee, FL | 36740 | 86 | 391,956 |
| Oxnard-Thousand Oaks-Ventura, CA | 37100 | 24 | 127,552 |
| Palm Bay-Melbourne-Titusville, FL | 37340 | 24 | 114,259 |
| Philadelphia-Camden-Wilmington, PA-NJ-DE-MD | 37980 | 317 | 798,019 |
| Phoenix-Mesa-Scottsdale, AZ | 38060 | 129 | 604,605 |
| Pittsburgh, PA | 38300 | 200 | 255,626 |
| Portland-South Portland-Biddeford, ME | 38860 | 72 | 85,125 |
| Portland-Vancouver-Beaverton, OR-WA | 38900 | 114 | 343,125 |
| Poughkeepsie-Newburgh-Middletown, NY | 39100 | 67 | 88,915 |
| Providence-New Bedford-Fall River, RI-MA | 39300 | 99 | 233,860 |
| Raleigh-Cary, NC | 39580 | 44 | 158,985 |
| Richmond, VA | 40060 | 90 | 210,529 |
| Riverside-San Bernardino-Ontario, CA | 40140 | 146 | 713,532 |
| Rochester, NY | 40380 | 84 | 106,879 |
| Sacramento--Arden-Arcade--Roseville, CA | 40900 | 99 | 399,296 |
| Salt Lake City, UT | 41620 | 37 | 127,176 |
| San Antonio, TX | 41700 | 90 | 194,622 |
| San Diego-Carlsbad-San Marcos, CA | 41740 | 94 | 520,128 |
| San Francisco-Oakland-Fremont, CA | 41860 | 153 | 778,902 |
| San Jose-Sunnyvale-Santa Clara, CA | 41940 | 55 | 233,907 |
| Santa Rosa-Petaluma, CA | 42220 | 30 | 100,052 |
| Sarasota-Bradenton-Venice, FL | 42260 | 37 | 139,775 |
| Scranton--Wilkes-Barre, PA | 42540 | 56 | 63,693 |
| Seattle-Tacoma-Bellevue, WA | 42660 | 149 | 576,654 |
| Sioux Falls, SD | 43620 | 30 | 27,949 |
| Springfield, MA | 44140 | 77 | 87,622 |
| St. Louis, MO-IL | 41180 | 183 | 380,268 |
| Stockton, CA | 44700 | 29 | 117,827 |
| Syracuse, NY | 45060 | 61 | 68,873 |
| Tampa-St. Petersburg-Clearwater, FL | 45300 | 122 | 454,621 |
| Toledo, OH | 45780 | 58 | 107,862 |
| Tucson, AZ | 46060 | 33 | 150,494 |
| Tulsa, OK | 46140 | 62 | 99,479 |
| Virginia Beach-Norfolk-Newport News, VA-NC | 47260 | 87 | 246,509 |
| Washington-Arlington-Alexandria, DC-VA-MD-WV | 47900 | 270 | 1,125,879 |
| Wichita, KS | 48620 | 55 | 75,696 |
| Worcester, MA | 49340 | 77 | 120,910 |
| Youngstown-Warren-Boardman, OH-PA | 49660 | 59 | 76,062 |

*CBSA = Core Based Statistical Area.*

# Appendix B

### Exhibit B-1

Selected Distributional Measures (1 of 2)

| Metropolitan Area | CBSA Code | Mean | Skewness | Gini Coefficient | Moran's I |
|---|---|---|---|---|---|
| Akron, OH | 10420 | 0.1011 | 0.7611 | 0.2485 | 0.32132 |
| Albany-Schenectady-Troy, NY | 10580 | 0.0609 | 2.8483 | 0.1480 | 0.02833 |
| Albuquerque, NM | 10740 | 0.0471 | 2.2612 | 0.2120 | 0.09045 |
| Allentown-Bethlehem-Easton, PA-NJ | 10900 | 0.0523 | 2.4191 | 0.1419 | 0.12101 |
| Atlanta-Sandy Springs-Marietta, GA | 12060 | 0.0919 | 3.2221 | 0.2523 | 0.14049 |
| Austin-Round Rock, TX | 12420 | 0.0407 | 2.3014 | 0.1672 | 0.11915 |
| Bakersfield, CA | 12540 | 0.1284 | 1.1880 | 0.1189 | 0.15969 |
| Baltimore-Towson, MD | 12580 | 0.0619 | 2.5160 | 0.2544 | 0.20582 |
| Barnstable Town, MA | 12700 | 0.0525 | 4.1814 | 0.2769 | 0.12282 |
| Baton Rouge, LA | 12940 | 0.0657 | 0.6837 | 0.0359 | 0.03857 |
| Birmingham-Hoover, AL | 13820 | 0.0981 | 0.5870 | 0.2278 | 0.17512 |
| Boise City-Nampa, ID | 14260 | 0.0512 | 0.6044 | 0.1518 | 0.29785 |
| Boston-Cambridge-Quincy, MA-NH | 14460 | 0.0700 | 2.1865 | 0.3684 | 0.11547 |
| Bridgeport-Stamford-Norwalk, CT | 14860 | 0.0633 | 1.9451 | 0.3447 | 0.34573 |
| Buffalo-Niagara Falls, NY | 15380 | 0.0546 | 1.6284 | 0.1355 | 0.06291 |
| Cape Coral-Fort Myers, FL | 15980 | 0.2259 | 0.6535 | 0.2364 | 0.07114 |
| Charleston-North Charleston, SC | 16700 | 0.0504 | 1.0662 | 0.0213 | 0.06203 |
| Charlotte-Gastonia-Concord, NC-SC | 16740 | 0.0516 | 0.6037 | 0.2160 | 0.08647 |
| Chicago-Naperville-Joliet, IL-IN-WI | 16980 | 0.1126 | 2.4996 | 0.3025 | 0.15895 |
| Cincinnati-Middletown, OH-KY-IN | 17140 | 0.0659 | 0.9297 | 0.1037 | 0.08096 |
| Cleveland-Elyria-Mentor, OH | 17460 | 0.1129 | 1.7977 | 0.2763 | 0.22396 |
| Columbia, SC | 17900 | 0.0559 | 0.2479 | 0.0691 | 0.09935 |
| Columbus, OH | 18140 | 0.0861 | 1.5809 | 0.2257 | 0.22189 |
| Dallas-Fort Worth-Arlington, TX | 19100 | 0.0656 | 1.7768 | 0.2015 | 0.13064 |
| Dayton, OH | 19380 | 0.1044 | 1.2386 | 0.2032 | 0.25994 |
| Deltona-Daytona Beach-Ormond Beach, FL | 19660 | 0.1018 | 0.3803 | 0.1537 | 0.02248 |
| Denver-Aurora, CO | 19740 | 0.0628 | 1.4959 | 0.2304 | 0.06629 |
| Des Moines-West Des Moines, IA | 19780 | 0.0683 | 1.5758 | 0.2299 | 0.07425 |
| Detroit-Warren-Livonia, MI | 19820 | 0.1487 | 1.7237 | 0.3200 | 0.22105 |
| Duluth, MN-WI | 20260 | 0.0491 | 1.0778 | 0.0354 | 0.01245 |
| Fresno, CA | 23420 | 0.0951 | (0.7155) | 0.0659 | 0.13991 |
| Grand Rapids-Wyoming, MI | 24340 | 0.0739 | 3.2392 | 0.1880 | 0.02531 |
| Hartford-West Hartford-East Hartford, CT | 25540 | 0.0481 | 2.2628 | 0.2192 | 0.08115 |
| Honolulu, HI | 26180 | 0.0357 | 1.6452 | 0.2797 | 0.20223 |
| Houston-Sugar Land-Baytown, TX | 26420 | 0.0658 | 1.1848 | 0.1085 | 0.08990 |
| Indianapolis-Carmel, IN | 26900 | 0.0920 | 1.4298 | 0.2475 | 0.12805 |
| Jacksonville, FL | 27260 | 0.1071 | 2.0112 | 0.1780 | 0.22384 |
| Kansas City, MO-KS | 28140 | 0.0729 | 1.6895 | 0.2186 | 0.11854 |
| Knoxville, TN | 28940 | 0.0460 | 0.8428 | 0.1424 | 0.07278 |
| Las Vegas-Paradise, NV | 29820 | 0.1247 | 0.7089 | 0.0890 | 0.19760 |
| Los Angeles-Long Beach-Santa Ana, CA | 31100 | 0.0848 | 2.6415 | 0.2944 | 0.08394 |
| Louisville-Jefferson County, KY-IN | 31140 | 0.0915 | 2.6032 | 0.2267 | 0.18207 |
| Madison, WI | 31540 | 0.0324 | 2.4176 | 0.0517 | 0.15087 |
| Memphis, TN-MS-AR | 32820 | 0.1421 | 0.2423 | 0.2622 | 0.05278 |
| Miami-Fort Lauderdale-Pompano Beach, FL | 33100 | 0.1546 | 1.8939 | 0.1163 | 0.10401 |
| Milwaukee-Waukesha-West Allis, WI | 33340 | 0.1119 | 1.6206 | 0.4404 | 0.26654 |

**Exhibit B-1**

Selected Distributional Measures (2 of 2)

| Metropolitan Area | CBSA Code | Mean | Skewness | Gini Coefficient | Moran's I |
|---|---|---|---|---|---|
| Minneapolis-St. Paul-Bloomington, MN-WI | 33460 | 0.0762 | 5.9378 | 0.2029 | 0.06764 |
| Nashville-Davidson-Murfreesboro-Franklin, TN | 34980 | 0.0557 | 1.6972 | 0.1803 | 0.05318 |
| New Haven-Milford, CT | 35300 | 0.0683 | 0.8762 | 0.2614 | 0.18388 |
| New Orleans-Metairie-Kenner, LA | 35380 | 0.0643 | 0.8645 | 0.1011 | 0.15283 |
| New York-Northern New Jersey-Long Island, NY-NJ-PA | 35620 | 0.0857 | 2.8248 | 0.3672 | 0.08619 |
| Oklahoma City, OK | 36420 | 0.0562 | 2.2695 | 0.0327 | 0.11450 |
| Omaha-Council Bluffs, NE-IA | 36540 | 0.0654 | 2.1485 | 0.2275 | 0.07015 |
| Orlando-Kissimmee, FL | 36740 | 0.1409 | 0.8574 | 0.1644 | 0.10881 |
| Oxnard-Thousand Oaks-Ventura, CA | 37100 | 0.0754 | 1.0449 | 0.2381 | 0.89622 |
| Palm Bay-Melbourne-Titusville, FL | 37340 | 0.1071 | 1.3009 | 0.1575 | 0.10468 |
| Philadelphia-Camden-Wilmington, PA-NJ-DE-MD | 37980 | 0.0659 | 1.6754 | 0.2785 | 0.11554 |
| Phoenix-Mesa-Scottsdale, AZ | 38060 | 0.1289 | 1.0739 | 0.3001 | 0.17823 |
| Pittsburgh, PA | 38300 | 0.0675 | 1.6749 | 0.1077 | 0.07618 |
| Portland-South Portland-Biddeford, ME | 38860 | 0.0495 | 1.2519 | 0.0650 | 0.21971 |
| Portland-Vancouver-Beaverton, OR-WA | 38900 | 0.0358 | 0.7176 | 0.1386 | 0.09663 |
| Poughkeepsie-Newburgh-Middletown, NY | 39100 | 0.0562 | 0.4385 | 0.0987 | 0.05572 |
| Providence-New Bedford-Fall River, RI-MA | 39300 | 0.0834 | 2.4916 | 0.2760 | 0.14911 |
| Raleigh-Cary, NC | 39580 | 0.0410 | 1.5304 | 0.2310 | 0.40647 |
| Richmond, VA | 40060 | 0.0526 | 1.6577 | 0.2453 | 0.05181 |
| Riverside-San Bernardino-Ontario, CA | 40140 | 0.1487 | 1.0673 | 0.1543 | 0.11091 |
| Rochester, NY | 40380 | 0.0516 | 2.1316 | 0.1556 | 0.14953 |
| Sacramento--Arden-Arcade--Roseville, CA | 40900 | 0.1072 | 1.2348 | 0.2130 | 0.20979 |
| Salt Lake City, UT | 41620 | 0.0329 | 0.4505 | 0.1421 | 0.04594 |
| San Antonio, TX | 41700 | 0.0564 | 0.9467 | 0.1606 | 0.19861 |
| San Diego-Carlsbad-San Marcos, CA | 41740 | 0.0750 | 0.3835 | 0.2473 | 0.16840 |
| San Francisco-Oakland-Fremont, CA | 41860 | 0.1029 | 1.2504 | 0.4745 | 0.14809 |
| San Jose-Sunnyvale-Santa Clara, CA | 41940 | 0.0790 | 0.9437 | 0.4204 | 0.18254 |
| Santa Rosa-Petaluma, CA | 42220 | 0.0702 | 2.1641 | 0.2843 | 0.00757 |
| Sarasota-Bradenton-Venice, FL | 42260 | 0.1168 | 1.8652 | 0.1700 | 0.08905 |
| Scranton--Wilkes-Barre, PA | 42540 | 0.0639 | 1.0407 | 0.1653 | 0.08902 |
| Seattle-Tacoma-Bellevue, WA | 42660 | 0.0380 | 1.1601 | 0.2524 | 0.26297 |
| Sioux Falls, SD | 43620 | 0.0432 | 1.6651 | 0.1184 | 0.05387 |
| Springfield, MA | 44140 | 0.0922 | 1.6064 | 0.2997 | 0.15298 |
| St. Louis, MO-IL | 41180 | 0.0864 | 1.3642 | 0.2687 | 0.19471 |
| Stockton, CA | 44700 | 0.1568 | 0.6151 | 0.1601 | 0.15657 |
| Syracuse, NY | 45060 | 0.0517 | 2.0942 | 0.1833 | 0.14539 |
| Tampa-St. Petersburg-Clearwater, FL | 45300 | 0.1155 | 1.0958 | 0.1374 | 0.07677 |
| Toledo, OH | 45780 | 0.0960 | 2.2450 | 0.2553 | 0.36648 |
| Tucson, AZ | 46060 | 0.0600 | 0.8474 | 0.2653 | 0.11890 |
| Tulsa, OK | 46140 | 0.0669 | 2.2258 | 0.0745 | 0.09841 |
| Virginia Beach-Norfolk-Newport News, VA-NC | 47260 | 0.0439 | 2.8062 | 0.0805 | 0.04837 |
| Washington-Arlington-Alexandria, DC-VA-MD-WV | 47900 | 0.0687 | 1.4021 | 0.2728 | 0.15820 |
| Wichita, KS | 48620 | 0.0399 | 0.9506 | 0.1035 | 0.01248 |
| Worcester, MA | 49340 | 0.0774 | 1.2448 | 0.2081 | 0.09926 |
| Youngstown-Warren-Boardman, OH-PA | 49660 | 0.0931 | 1.7036 | 0.0273 | 0.11455 |

*CBSA = Core Based Statistical Area.*

# Appendix C

## Exhibit C-1

Groupings From the Cluster Analysis

**Group 1**
Albuquerque, NM
Austin-Round Rock, TX
Boise City-Nampa, ID
Bridgeport-Stamford-Norwalk, CT
Honolulu, HI
Madison, WI
Oxnard-Thousand Oaks-Ventura, CA
Portland-South Portland-Biddeford, ME
Raleigh-Cary, NC
San Antonio, TX
Seattle-Tacoma-Bellevue, WA

**Group 2**
Akron, OH
Bakersfield, CA
Cape Coral-Fort Myers, FL
Cleveland-Elyria-Mentor, OH
Dayton, OH
Deltona-Daytona Beach-Ormond Beach, FL
Detroit-Warren-Livonia, MI
Las Vegas-Paradise, NV
Memphis, TN-MS-AR
Milwaukee-Waukesha-West Allis, WI
Orlando-Kissimmee, FL
Palm Bay-Melbourne-Titusville, FL
Phoenix-Mesa-Scottsdale, AZ
San Francisco-Oakland-Fremont, CA
Sarasota-Bradenton-Venice, FL
Stockton, CA

**Group 3**
Atlanta-Sandy Springs-Marietta, GA
Baltimore-Towson, MD
Boston-Cambridge-Quincy, MA-NH
Chicago-Naperville-Joliet, IL-IN-WI
Columbus, OH
Hartford-West Hartford-East Hartford, CT
Jacksonville, FL
Los Angeles-Long Beach-Santa Ana, CA
Louisville/Jefferson County, KY-IN
Miami-Fort Lauderdale-Pompano Beach, FL
New York-Northern New Jersey-Long Island, NY-NJ-PA
Oklahoma City, OK
Rochester, NY
Springfield, MA
Syracuse, NY
Toledo, OH
Tulsa, OK
Youngstown-Warren-Boardman, OH-PA

**Group 4**
Albany-Schenectady-Troy, NY
Allentown-Bethlehem-Easton, PA-NJ
Barnstable Town, MA
Grand Rapids-Wyoming, MI
Minneapolis-St. Paul-Bloomington, MN-WI
Providence-New Bedford-Fall River, RI-MA
Santa Rosa-Petaluma, CA
Virginia Beach-Norfolk-Newport News, VA-NC

**Group 5**
Baton Rouge, LA
Buffalo-Niagara Falls, NY
Charleston-North Charleston, SC
Charlotte-Gastonia-Concord, NC-SC
Des Moines-West Des Moines, IA
Nashville-Davidson-Murfreesboro-Franklin, TN
New Haven-Milford, CT
Omaha-Council Bluffs, NE-IA
Richmond, VA
Riverside-San Bernardino-Ontario, CA
Scranton--Wilkes-Barre, PA
Sioux Falls, SD
Wichita, KS

**Group 6**
Birmingham-Hoover, AL
Cincinnati-Middletown, OH-KY-IN
Columbia, SC
Dallas-Fort Worth-Arlington, TX
Denver-Aurora, CO
Duluth, MN-WI
Fresno, CA
Houston-Sugar Land-Baytown, TX
Indianapolis-Carmel, IN
Kansas City, MO-KS
Knoxville, TN
New Orleans-Metairie-Kenner, LA
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
Pittsburgh, PA
Portland-Vancouver-Beaverton, OR-WA
Poughkeepsie-Newburgh-Middletown, NY
Sacramento--Arden-Arcade--Roseville, CA
Salt Lake City, UT
San Diego-Carlsbad-San Marcos, CA
San Jose-Sunnyvale-Santa Clara, CA
St. Louis, MO-IL
Tampa-St. Petersburg-Clearwater, FL
Tucson, AZ
Washington-Arlington-Alexandria, DC-VA-MD-WV
Worcester, MA

# Appendix D

**Exhibit D-1**

Economic and Housing Market Variables Selected by the Stepwise Regression Procedure

| Variable | Mean | Standard Deviation |
|---|---|---|
| Subprime Gini coefficient | 0.1711 | 0.0938 |
| Percent owner occupied among 2006 conventional home purchase loans | 0.8320 | 0.0737 |
| Log of MSA active loan count | 12.1988 | 0.8728 |
| Percent change in HPI third quarter 2005 to third quarter 2006 | 6.6987 | 5.9766 |
| Percent change in HPI third quarter 2006 to third quarter 2007 | 0.6482 | 4.8186 |
| Percent change in unemployment rate third quarter 2006 to third quarter 2007 | 1.7955 | 13.0455 |
| Percent change in unemployment rate third quarter 2007 to third quarter 2008 | 31.4344 | 16.1052 |
| Third quarter 2008 unemployment rate | 5.9697 | 1.5235 |
| Third quarter 2006 affordability index | 125.1476 | 48.4117 |
| Home purchase lending percent change third quarter 2007 to third quarter 2008 | – 0.2364 | 0.1168 |
| The mean, avg_income_to_mfi_05 | 0.0014 | 0.0003 |
| Skewness, avg_income_to_mfi_05 | 1.6476 | 1.0766 |
| The mean, avg_income_to_mfi_06 | 0.0015 | 0.0004 |
| Skewness of 2006 ZIP Code median borrower income relative to MSA median family income | 1.6877 | 0.8648 |
| Geary's C for home purchase lending percent change third quarter 2006 to third quarter 2007 | 1.0012 | 0.4399 |
| Ratio of maximum-to-median ZIP Code delinquency rate | 2.8754 | 0.8092 |

*HPI = house price index. MSA = metropolitan statistical area.*

## Authors

Lariece M. Brown is a senior financial economist at the Federal Deposit Insurance Corporation.

Hui-Chin Chen is a former financial analyst at Freddie Mac.

Melissa T. Narragon is an economic research senior at Freddie Mac.

Paul S. Calem is a senior economist at the Federal Reserve System.

## Acknowledgments

The authors thank Calvin Schnure and two anonymous referees for their thoughtful feedback.

## References

Can, Ayse. 1992. "Residential Quality Assessment: Alternative Approaches Using GIS," *The Annals of Regional Science* 26: 97–110.

Can, Ayse, and Isaac F. Megbolugbe. 1997. "Spatial Dependence and House Price Construction," *Journal of Real Estate Finance and Economics* 14: 203–222.

Derudder, B., P.J. Taylor, F. Witlox, and G. Catalano. 2003. "Hierarchical Tendencies and Regional Patterns in the World City Network: A Global Urban Analysis of 234 Cities," *Regional Studies* 37 (9): 875–886.

Edmiston, Kelly. 2009. "Characteristics of High-Foreclosure Neighborhoods in the Tenth District," *Economic Review* 94: 51–75.

Fleischman, Daniel. 2010. "Nonprofit Strategies for Returning REO Properties to Effective Use." In *REO & Vacant Properties: Strategies for Neighborhood Stabilization*, a joint publication of the Federal Reserve Banks of Boston and Cleveland and the Federal Reserve Board: 107–114.

Fortin, Marie-Jos'ee, Mark R.T. Dale, and Jay ver Hoef. 2002. "Spatial Analysis in Ecology." In *Encyclopedia of Environmetrics*, edited by Abdel H. El-Shaarawi and Walter W. Piegorsch. Chichester, United Kingdom: Wiley: 2051–2058.

Frame, W. Scott. 2010. "Estimating the Effect of Mortgage Foreclosures on Nearby Property Values: A Critical Review of the Literature," *Economic Review* 95 (3): 1–9.

Goldstein, Ira. 2010. "Maximizing the Impact of Federal NSP Investments Through the Strategic Use of Local Market Data." In *REO & Vacant Properties: Strategies for Neighborhood Stabilization*, a joint publication of the Federal Reserve Banks of Boston and Cleveland and the Federal Reserve Board: 65–75.

Immergluck, Dan. 2010a. "Neighborhoods in the Wake of the Debacle: Intrametropolitan Patterns of Foreclosed Properties," *Urban Affairs Review* 46 (1): 3–36.

———. 2010b. "The Accumulation of Lender-Owned Homes During the U.S. Mortgage Crisis: Examining Metropolitan REO Inventories," *Housing Policy Debate* 20: 619–635.

———. 2009. Intrametropolitan Patterns of Foreclosed Homes: ZIP-Code-Level Distributions of Real-Estate-Owned (REO) Properties During the U.S. Mortgage Crisis. Community Affair discussion paper No. 01-09. Federal Reserve Bank of Atlanta.

Kingsley, G. Thomas, Robin E. Smith, and David Price. 2009. "The Impacts of Foreclosures on Families and Communities: A Primer." Urban Institute: http://www.urban.org/url.cfm?ID=411909 (accessed August 12).

Lee, Kai-yan. 2008. Foreclosure's Price-Depressing Spillover Effects on Local Properties: A Literature Review. Community Affairs discussion paper, 2008-01. Federal Reserve Bank of Boston.

Lembo, Jr., Arthur J. 2008. "Spatial AutoCorrelation: Moran's I, Geary's C." http://faculty.salisbury.edu/~ajlembo/419/lecture15.pdf (accessed April 3).

Litchfield, Julie A. 1999. "Inequality: Methods and Tools." http://siteresources.worldbank.org/INTPGI/Resources/Inequality/litchfie.pdf (accessed March 2, 2009).

Newburger, Harriet. 2010. "Acquiring Privately Held REO Properties With Public Funds: The Case of the Neighborhood Stabilization Program." In *REO & Vacant Properties: Strategies for Neighborhood Stabilization*, a joint publication of the Federal Reserve Banks of Boston and Cleveland and the Federal Reserve Board: 101–106.

Office of Management and Budget (OMB). 2006. *Update of Statistical Area Definitions and Guidance on Their Uses*. Bulletin No. 07-01. White House website: http://www.whitehouse.gov/sites/default/files/omb/assets/omb/bulletins/fy2007/b07-01.pdf (accessed March 2, 2009).

Ryan, Jr., Jay N. 2010. "REO Disposition and Neighborhood Stabilization: A Servicer's View." In *REO & Vacant Properties: Strategies for Neighborhood Stabilization*, a joint publication of the Federal Reserve Banks of Boston and Cleveland and the Federal Reserve Board: 95–100.

Vicino, Thomas J., Bernadette Hanlon, and John Rennie Short. 2007. "Megalopolis 50 Years On: The Transformation of a City Region," *International Journal of Urban and Regional Research* 31 (2): 344–367.

## Additional Reading

Hatcher, Larry. 1994. *A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute.

Weisstein, Eric W. n.d. "Gradient." MathWorld—A Wolfram Web Resource. http://mathworld.wolfram.com/Gradient.html (accessed March 2, 2009).