Shrinking Networks: A Spatial Analysis of Bank Branch Closures

Anna Tranfaglia
Federal Reserve Bank of Philadelphia Payment Cards Center
Shrinking Networks: 
A Spatial Analysis of Bank Branch Closures

Anna Tranfaglia*
Federal Reserve Bank of Philadelphia

March 2018

Abstract

As more consumers take advantage of online banking services, branch networks are declining across the country. Limited attention has been given to identifying any possible spatial patterns of branch closures and, more importantly, the community demographics where branches close their doors. This analysis uses an innovative spatial statistics concept to study financial services: Using data from 2010 to 2016, a random labelling test is conducted to understand branch closure clustering in the Philadelphia, Chicago, and Baltimore metropolitan statistical areas (MSAs). Additionally, spatial autocorrelation is tested, and an MSA-level spatial regression analysis is done to see if there is a pattern to branch closures in metropolitan areas. I find evidence of branch closure clusters in the Chicago and Philadelphia MSAs; however, this spatial pattern is only observable within the suburbs, not the primary city itself. Using a random labelling test is a methodological innovation in regional economic studies and propels our understanding of banking deserts and underserved neighborhoods.

Keywords: branch closures, GIS, spatial autocorrelation, marked point process, random labelling test, Philadelphia, Chicago, Baltimore

JEL codes: G21, C21, R12

*anna.tranfaglia@phil.frb.org

Disclaimer: This Philadelphia Fed working paper represents preliminary research that is being circulated for discussion purposes. The views expressed in these papers are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. No statements here should be treated as legal advice. Philadelphia Fed working papers are free to download at https://philadelphiafed.org/research-and-data/publications/working-papers/.
1. Introduction

Since the onset of the Great Recession, many banks have reduced the size of their branching networks. This reduction is due to multiple factors, including increased merger and acquisition activity following the Great Recession, changes to firm strategy, and consumers’ increased reliance on mobile technology, which have weakened the value of physical banking locations for many consumers and decreased the demand for bank branches. However, recent research suggests that proximity, or lack thereof, to a bank branch still matters. Limited physical bank access negatively affects specific populations, such as small businesses and those living in disadvantaged neighborhoods (Nguyen, 2014). Evidence suggests that consumers who had poor exposure to banks as children experience worse financial health than their peers later in life (Brown, Cookson, and Heimer, 2016). Examining the geographic distribution of branch closures is therefore necessary to understand the equality of physical bank access for different neighborhoods. Additionally, studying the geography of branching is central to understanding banking competition, especially in urban landscapes. The goal of this project is twofold: (1) to explore possible spatial patterns, with a specific focus on clustering, in bank branch closures within the Philadelphia, Chicago, and Baltimore metropolitan areas, three regions that recently experienced substantial branching network contractions (Taylor et al., 2017) and (2) if spatial closure patterns emerge, do the census tracts that experience at least one bank branch closure share any especial demographic characteristics?

The U.S. bank branch network is forecast to continue shrinking (Gensler, 2016; Morgan, Pinkovsky, and Yang, 2016). This has served as the genesis for a growing body of literature on consumers’ interactions with the banking system and the role of geographic proximity to financial services, both informal and formal (Smith, Smith, and Wackes, 2008). However, to the author’s knowledge, there is a lack of analysis at the city or metropolitan level on the geographic distribution of branch closures. Understanding these distributions and identifying clusters of branch closures help researchers and policymakers who are interested in the role geographic proximity plays within financial inclusion.

Recent literature underscores the importance of physical proximity with respect to lending. When lending is information intensive, bank closures have a negative effect on local credit supply, and this contraction is more austere in low-income and high-minority neighborhoods (Nguyen, 2014). Spatial price discrimination is another concern; previous research has found that loan rates increase with the distance between the firm (borrower) and the next competing bank (Degryse and Ongena, 2005). Additionally, the distance to a bank does not affect household behavior equally: Low- and moderate-income (LMI) households with access to a nearby bank branch experience larger increases in the likelihood of owning a bank account than non-LMI households (Goodstein and Rhine, 2017; Celerier and Matray, 2017). Exploring the substitution effects of branch closures,
researchers recently found that during 2010–2015, regions that experienced at least a 5 percent reduction in bank branches received 40 percent of LendingClub’s consumer loans (Jagtiani and Lemieux, 2017). There is also evidence that alternative financial service providers are disproportionately concentrated in neighborhoods that are not traditionally served by banks (Smith et al., 2009). These articles and others explore the relationship between credit supply (both formal and from alternative services) and the proximity to a bank. Analysis with respect to proximity to bank branches exists; however, there is a lack of research on the proximity to declining bank access or how branch closures affect consumers.

Additionally, spatial autocorrelation has been analyzed in many social science contexts, including crime (Troy, Grove, and O’Neil-Dunne, 2012), home values (Dubin, 1992; Cohen and Coughlin, 2008), and mortgage prices (Zou, 2014). Within the bank branch literature, however, very little attention is given to this statistical concept. Banks do not decide which branches to close randomly. Rather, current and forecast credit demand and profitability, as well as mergers and acquisitions, help to determine whether a specific branch remains open. Therefore, branch closure locations within a city or market may be spatially dependent for multiple reasons: People with similar levels of income, educational attainment, and race/ethnicity tend to live near one another. These factors are associated with credit demand. Neighborhoods also tend to be composed of housing units with similar values, and evidence exists of high concentrations of subprime mortgages within certain urban neighborhoods (Hwang, Hankinson, and Brown, 2014).

This reality needs to be considered for future research. The presence of spatial autocorrelation affects assumptions required for statistical inference. If data are spatially dependent, a value at a specific location can be predicted solely from a nearby observation; this violates independence. Spatial dependence introduces over- and undercounting bias in ordinary least squares (OLS) regressions, which is indicative of a misspecified model. Therefore, we need to determine whether spatial autocorrelation is present and whether it is vital for analyzing a possible correlation between branch closures and socioeconomic traits. The main contributions of this paper are incorporating spatial autocorrelation into the bank networks context and the original technical use of the random labelling test. This technique, used to test for spatial clustering, has not been widely employed in social science research. Incorporating this spatial statistic allowed for the discovery of varying spatial clustering patterns of branch closures within the MSAs, which would not have been possible using more conventional tools.

Two null hypotheses are tested in this paper: (1) The locations of branch closures that occurred in 2010–2016 in the Philadelphia, Baltimore, and Chicago MSAs are not significantly more clustered than the distribution of all branches within each MSA,
respectively, and (2) tract characteristics are randomly distributed among census tracts that experienced a branch closure.

2. Data and Methodology

2.1 Study Area

The study area for this analysis includes the following MSAs: Philadelphia-Camden-Wilmington, PA-NJ-DE-MD; Baltimore-Columbia-Towson, MD; and Chicago-Naperville-Elgin, IL-IN-WI. The Philadelphia MSA consists of five counties in Pennsylvania (Bucks, Chester, Delaware, Montgomery, and Philadelphia), four in New Jersey (Salem, Gloucester, Camden, and Burlington), one (New Castle) in Delaware, and one in Maryland (Cecil). The Baltimore MSA consists of seven Maryland counties (Howard, Baltimore, Anne Arundel, Baltimore City, Carroll, Queen Anne’s, and Hartford). The Chicago MSA consists of nine counties in Illinois (Cook, DuPage, DeKalb, Will, Grundy, Kendall, Kane, McHenry, and Lake), four (Jasper, Porter, Newton, and Lake) in Indiana, and one (Kenosha) in Wisconsin.

The Philadelphia and Chicago metropolitan regions are both among the top 10 largest (population) in the country. All three geographies rank in the top 10 MSAs that experienced the greatest share of bank branch closures in the past decade (Taylor et al., 2017). Additionally, these MSAs also experienced large numbers, not just rates, of branch closures. The unit of analysis for the cluster analysis is the bank branch, and the unit of analysis for the distribution of neighborhood characteristics is the census tract, a relatively small geographic unit with an approximate population of 4,000. The spatial size of census tracts exhibits a considerable variation within an MSA because the size is dependent on population density. Tracts in the urban center are very small, whereas tracts near the border of the MSA generally are much larger. In this paper, a branch closure is defined as a bank branch that was open in June 2010 and closed any time before June 2016 (these midyear endpoints are explained in detail in the following section). Owing to the specific techniques required for the cluster analysis, I was not able to add branch closures for bank branches that opened after June 2010. For example, if a bank opened a branch in a Chicago tract in 2011 and closed the branch two years later, it was not included in this study.

2.2 Data

The geocoded branch locations come from SNL’s Branch Analytics tool. This tool uses the Federal Deposit Insurance Corporation (FDIC)’s Summary of Deposits (SOD) annual data sets (2010–2016) but is cleaned and geocoded for SNL’s use. SOD data provide a snapshot of the nation’s banks on June 30 every year. For example, the 2016 SOD data capture any changes to the branch network that occurred from July 1, 2015, until June 30, 2016. The SOD data also attempt to keep the branch ID consistent in the event that a bank branch relocates nearby. If a branch moving across the street resulted mistakenly in a new branch
ID, branch closures (identified by firm ID) were checked against any branch openings (by
firm ID) within the tract by the author.

U.S. banks and thrifts increased their branching networks through 2009; however, since
then, more branches have closed each year than opened (Taylor et al., 2017). As a result,
the original set of branch locations for all three MSAs is derived from the 2010 SOD. The
2010 data illustrate the banking landscape within the U.S. as it was on June 30, 2010. The
set of closed branches is therefore any branch that was open on June 30, 2010, and closed
before June 30, 2016 (Table 1).

| Table 1. Branch Closures Since 2010 |
|-------------------------------------|
| MSA      | Bank | Savings Bank | Thrift | Credit Union | Total |
| Baltimore | 150  | 2            | 29     | 0            | 181   |
| Philadelphia | 235 | 82            | 31     | 35           | 383   |
| Chicago   | 586  | 9            | 34     | 0            | 629   |

Source: S&P Global Market Intelligence, SNL Branch Analytics

While the SNL data included the census tract, I reassigned census tract codes to each
branch location as a robustness check using Esri’s ArcGIS Desktop 10.3. For the census
tract/neighborhood characteristics analysis, I generated a count of branch closures within
the tract because this analysis is done at the tract level. Tract-level attributes are from the
U.S. Census’s 2011–2015 American Community Survey 5-Year Estimates. The data for this
study include demographic measures: total tract population, median age, percentage of
tract population less than 24 years of age, distance from tract centroid to the central
business district (CBD), and race/ethnicity. Socioeconomic variables include median
household income, unemployment rate, and the percentage of tract population living below
the federally defined poverty level.

For the tract characteristic analysis, only tracts that had an open bank branch in 2010 are
included. This is to guarantee that the only comparison is between neighborhoods that
experienced a branch closure with neighborhoods that contain branches but didn’t
experience any closures. The main goal is to avoid comparing tracts with a branch closure
to tracts that did not house a branch and did not have any banking services to lose. Table 2
displays the summary statistics for these two groups of tracts. Even though tracts that did
not have a bank branch are not included, both Chicago and Philadelphia tracts that did not
experience a branch closure have a significantly higher percentage of residents living
below the federally defined poverty line. Chicago branch closure tracts are significantly

1 City Hall is the CBD for all three MSAs.
farther from the CBD and have a significantly smaller percentage of Hispanic residents and a larger percentage of white residents than tracts without a branch closure. Similarly, a significantly smaller percentage of black residents live within Philadelphia MSA tracts that contained a branch closure, and a significantly larger share of the population is unemployed in tracts that did not experience any branch closures compared with those tracts that lost at least one branch during the period of study. The Baltimore MSA did not share these trends; tracts with a branch closure had populations with a significantly smaller percentage of whites, a significantly larger percentage of Hispanics, and a greater share of younger individuals (younger than 24 years of age) than those tracts without a closure.

This brief comparative analysis is important because it reinforces that there is no single unifying trend among all three MSAs before any regression analysis is conducted.
Table 2. Neighborhood Characteristics by Branch Closure Status (of Census Tracts that Had Open Branch(es) ≥ 1 in 2010)

| Neighborhood Characteristics | Philadelphia |  |  | Chicago |  |  | Baltimore |  |  |  |
|-----------------------------|--------------|---|---|--------|---|---|-----------|---|---|---|
|                             | Experienced  | Did Not | P-Value | Experienced | Did Not | P-Value | Experienced | Did Not | P-Value |  |
|                             | Closure | Experience | Closure | Closure | Experience | Closure | Closure | Experience | Closure |  |
| Median income ($)           | 74,643      | 69,393 | 0.0159 | 73,706 | 66,856 | 0.0001 | 75,984 | 80,199 | –   |
| Black (%)                   | 14.01       | 16.91  | 0.0719 | 12.07  | 13.66  | –      | 24.80  | 20.61  | –   |
| Hispanic (%)                | 7.16        | 7.41   | –      | 16.99  | 21.49  | 0.0002 | 5.51   | 4.51   | 0.0557 |
| Asian (%)                   | 6.35        | 5.30   | 0.0188 | 6.96   | 6.33   | –      | 5.48   | 5.19   | –   |
| White (%)                   | 74.73       | 72.67  | –      | 72.79  | 69.82  | 0.0304 | 65.17  | 70.34  | 0.0818 |
| Younger than 24 years of age (%) | 24.94 | 25.58 | –      | 26.18  | 25.88  | –      | 25.86  | 24.04  | 0.0902 |
| Total population            | 4387        | 4360   | –      | 5050   | 4673   | 0.0024 | 4678   | 4445   | –   |
| Unemployment rate (%)       | 8.38        | 9.17   | 0.0349 | 8.73   | 9.25   | –      | 7.00   | 7.28   | –   |
| Population below poverty line (%) | 10.76 | 12.40 | 0.0574 | 11.97  | 13.46  | 0.0127 | 10.95  | 10.01  | –   |
| Distance to CBD             | 0.2879      | 0.2805 | –      | 0.3871 | 0.3579 | 0.0410 | 0.2155 | 0.2298 | –   |
| No. of tracts, N            | 309         | 544    | –      | 509    | 714    | –      | 141    | 196    | –   |

Source: 2011–2015 American Community Survey, U.S. Census Bureau. CBD=central business district.

Note: Only p-values corresponding to variables that have significantly different values in closure and nonclosure tracts are listed.
2.3 Spatial Co-occurrence/Clustering of Bank Branch Closures

The fundamental idea of this second-order analysis is that this approach identifies branch closure clustering conditional on the bank branch locations in 2010, a year with relatively robust branching networks. The key question is whether any concentrations of branch closures are greater than would be expected given the branch locations within the Chicago, Baltimore, and Philadelphia MSAs in 2010.

2.3.1 K-Functions

The K-function is a measure of the average number of events within a given radius for any randomly selected event (Dixon, 2002). Specifically, a circle with a given distance is drawn around every closure event in each MSA data set, and then the average number of branch closures is calculated from each circle within that given radius. This process is repeated for many different radii. For Ripley’s K-function, or the homogeneous point process, this can be defined as:

\[ K(r) = \frac{1}{\lambda} E \text{(number of events with distance } r \text{ of each event).} \]  

(1.1)

The parameter \( \lambda \) is the rate, or density; therefore, it is calculated as the number of events per unit area. Lambda has a constant value for the case of an unlabelled, homogeneous point processes. Ripley’s K-function is unsuitable in this analysis, however, as it assumes homogeneity across the entire area of study when calculating Euclidean distances between point data. Bank branch closures cannot occur spontaneously; branch closures occur at the same location where an open branch existed in a previous period. Therefore, a test based on the inhomogeneous K-function is used because it takes into account that the set of branch closure locations is conditional on existing within the set of open bank branches in a previous period (Buzard et al., 2017). However, in this inhomogeneous example, the K-function needs to be adjusted slightly: The parameter instead becomes a function, \( \lambda(x) \), where \( x \) is the location of a randomly selected event (branch closure).

\[ K_{ii}(r) = E \left( \frac{\text{number of type } i \text{ events with radius } r \text{ of type } i \text{ events}}{\lambda(x_i)} \right) \]  

(1.2)

The inhomogeneous intensity function is represented by \( \lambda(x_i) \). Again, this function conveys the likelihood of a given point at location \( x_i \). The inhomogeneous K-function (Equation 1.2) is the basis of the statistical test for the random labelling hypothesis. Since only one marked point pattern is observed (for each MSA), the main challenge in testing the null hypothesis is the lack of any information regarding the true distribution of branch closures in each metropolitan area. To overcome this missing information, estimating the K-functions under the null hypotheses is achieved through Monte Carlo simulation.
2.3.2 Random Labelling Test for Marked Point Patterns

To test for clusters of branch closures, the inhomogeneous distribution of bank branch locations needs to be controlled. The random-labelling test of marked point data accomplishes this data necessity. Random labelling is the result of independent assignment of marks to an original (already existing) point pattern. Bank branch closures can be considered a marked point pattern. A point pattern, \( X^0 = \{x^0_i: i = 1, \ldots, n\} \), is simply the distribution of \( n \) events on a surface (Baddeley, 2008). In addition to focusing on a spatial pattern of points alone, the label- or mark-associated with each point can be of interest. The marked pattern is therefore a stochastic process that generates pairs of joint locations and labels (Smith, 2016). The labels represent data that are qualitatively different. The operating status (open or closed) is the mark in this analysis. So each bank branch location within the set, \( i = 1, \ldots, n \), can be described by the pair \((s_i, j_i)\), where \( s_i \in R \) is the location of the bank branch within the MSA, and \( j_i \in \{0,1\} \) is a mark, or label, indicating whether the bank is open or closed.

The marked random labelling test supplies the correct null hypothesis for analyzing bank branch closures because the marking process (branch operating status) occurs after a point process (decision where a branch is located) has already been established. Therefore, the decision by firms of which branches to close is a secondary process that acts on existing branches within the metro area’s banking networks.

Figures 1–3 illustrate all the branch locations and their respective statuses in June 2016 within the Philadelphia, Baltimore, and Chicago MSAs.
Figure 1. Philadelphia MSA: Total Set of Bank Branches
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD MSA

Source: S&P Global Market Intelligence, SNL Branch Analytics

Figure 2. Baltimore MSA: Total Set of Bank Branches
Baltimore-Columbia-Towson, MD MSA

Source: S&P Global Market Intelligence, SNL Branch Analytics
2.3.3 Random Labelling Hypothesis Test

To determine whether branch statuses are assigned independently of branch locations, the random labelling hypothesis was tested. The underlying assumption for this test is that branch closures are assigned via an inhomogeneous Poisson process and therefore have varying intensity. In other words, the occurrence of branch closures is determined by a Poisson distribution conditional on the spatial distribution of bank branch locations. This test’s null hypothesis implies that each event label (bank branch status) is not influenced by its location, that is:

$$\Pr((m_1, \ldots, m_n) \mid (s_1, \ldots, s_n)) = \Pr(m_1, \ldots, m_n)$$  \hspace{2cm} (1.3)

for all locations $s_1, \ldots, s_n \in R$ and labels $m_1, \ldots, m_n \in \{0,1\}$. Here $\Pr(m_1, \ldots, m_n)$ represents the marginal distribution of branch closure locations, and $\Pr((m_1, \ldots, m_n) \mid (s_1, \ldots, s_n))$ denotes the conditional distribution of marks (branch operation status) given their locations within each MSA. Therefore, the set of observed branches in 2010 does not provide information about whether that branch is now open or closed, just that this location is occupied by a
branch with one status or the other. This does not mean that these two states are equally likely to occur; the majority of bank branches remained open in all three metro areas studied. The distribution of labels, \( \Pr(m_1, ..., m_n) \), captures the likelihood of closed branches.

For random labelling hypothesis tests, the distribution of marked point patterns under the null hypothesis is required. Before testing, there is one more necessary assumption: The marginal distribution of labels does not depend on the order in which unique events are labelled. In other words, the likelihood of the branch statuses within a set, \( (m_1, ..., m_n) \) does not depend on which bank location contains the “1” subscript. More formally, this condition that all points within a set must be exchangeable requires that for all permutations \( (\pi_1, ..., \pi_n) \) of the mark subscripts \( (1, ..., n) \),

\[
\Pr(m_{\pi_1}, ..., m_{\pi_n}) \equiv \Pr(m_1, ..., m_n)
\]

(1.4)

This condition combined with Equation (1.1) implies that all the possible labellings of events are equally likely when generating replicate point patterns. This is necessary for making an exact sampling distribution of marked point patterns under the null hypothesis. For each study area, or MSA, sample marked point patterns are produced by holding the point locations fixed within the study region and assigning the marks by permutation from the observed mark distribution from each MSA.

The null hypothesis of no spatial clustering is tested using Diggle and Chetwynd’s (1991) modified K-function methodology,

\[
\hat{R}_{11}(r) = |A| \{n_i(n_i - 1)\}^{-1} \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} w_{ij} \delta_{ij}(r)
\]

(1.5)

\[
\hat{R}_{22}(r) = |A| \{n(n - 1)\}^{-1} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \delta_{ij}(r)
\]

(1.6)

where \( A \) is the region of study, \( r \) is the length of the radius drawn around a randomly chosen branch in the MSA, and \( \delta_{ij} \) is the indicator that the event \( d_{ij} \leq r \). \( w_{ij} \) is a weight specified for every ordered pair \((i, j)\) within the MSA set. Its value is inversely proportional to the circumference of a circle centered on point \( i \) that passes through point \( j \). \( n_i \) is the number of closed branches, and \( n \) is the number of all branches within \( A \) (the MSA).

\( \hat{R}_{11} \) and \( \hat{R}_{22} \) are comparative with tallies of closed branches and all branches, respectively, within distance \( r \) of a randomly chosen bank branch. From Equations (1.1) and (1.2) and under the random labelling null hypothesis, it follows that

\[
\hat{R}_{11}(r) = \hat{R}_{22}(r).
\]

(1.7)

Equation (1.7) suggests that the most useful way to inspect any divergences from the null is to inspect the differences between the estimates for both K-functions. The test statistic for
the random labelling test is therefore defined by $D(r) = \hat{R}_{11}(r) - \hat{R}_{22}(r)$. The expected value of $D(r)$ with randomly labelled points is zero. If type 1 events (closures) have a greater degree of spatial aggregation with respect to the level of spatial aggregation of type 2 events (open branches), then $D(r)$ is positive and is an indication of clustered labelled points.

The testing procedure is therefore straightforward: Create simulations of branch closure patterns by randomly relabelling the marks while the locations of events remain unchanged. If the observed spatial distribution of closures occurred randomly within each MSA, then the observed $D(r)$ is similar to the $D(r)$ derived from the simulations. Using Philadelphia as an example, each permutation always has 383 “closed” and 1,289 “open” points. The significance test of $D(r)$ is achieved through Monte Carlo simulations. Under the random labelling hypothesis, the probability of obtaining a difference value as large as $D(r)$ is estimated by

$$P_{\text{clustered}}(r) = \frac{m_0}{N+1}. \quad (1.8)$$

Therefore, a two-sided, 95 percent point-wise critical bounds can be estimated from generating 39 random simulations of branch closures (Cronie and van Lieshout, 2016; Dixon, 2002; Smith, 2016). The interpretation of the plotted results is therefore that if any one of the observed difference functions strays outside the envelopes, the null hypothesis would be rejected at distance $r$.

In addition to better representing the set of branch closures than Ripley’s (homogeneous) $K$-function method, the random labelling test has several benefits: By conditioning events on a set of existing locations, the edge effects problem disappears. Also, modeling events as marked point processes is useful when the location process is complex: The set of feasible locations for banks, as well as most physical structures, has a horde of both observed and unobserved restrictions, such as land-use constraints. By modeling open and closed bank branches as a bivariate marked point process, the location process of branches is separated from the distribution of the event, whether the bank was opened or closed in 2016.

2.4 Spatial Autocorrelation

The second question within this analysis considers the relationship between branch closure locations and the underlying tract characteristics. Specifically, do census tracts that experience at least one bank branch closure have similar socioeconomic and demographic characteristics? Both bivariate and multivariate statistical techniques are used to shed light on this question. To explore any possible correlation, one important issue must be addressed: Spatial autocorrelation is a common concern whenever analysis is conducted at the census tract level. Since tracts are defined by neighborhood boundaries, nearby tracts tend to have more similar characteristics than relatively remote tracts. OLS estimation
requires the assumption that observations are independent and identically distributed; errors of the residual are uncorrelated. However, for bank branch closures, tract characteristics are potentially dependent on the levels within neighboring tracts. In this case, OLS estimation can produce biased and inconsistent parameter estimates.

To test for spatial autocorrelation, the global Moran’s I statistic is used to identify any spatial dependence among census tracts with exposure to branch closures (represented by a dummy variable) and eight socioeconomic and geographic variables (median family income, percentage of black, percentage of Asians, percentage of Hispanics, percentage of whites, unemployment rate, percentage of tract population below 24 years of age, and percentage below Federal poverty line). Global Moran’s I is a comprehensive measure of spatial autocorrelation (Bailey and Gatrell, 1995). In this case, Global Moran’s I tests whether tract-level characteristics are randomly distributed or whether neighboring values are more comparable than nonneighboring values. The spdep package for RStudio was used to calculate the Moran’s I statistics and associated Z-scores. While this global statistic was originally proposed by Moran, the spdep packages uses a formula for Moran’s I presented by Cliff and Ord (1981) and Bivand and Piras (2015):

\[
I = \frac{n}{W} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} z_i z_j}{\sum_{i=1}^{n} z_i^2}
\]

where \( n \) is the number of census tracts, \( W \) is the sum of weights \( w_{ij} \) for all census tract pairs, and \( z_i = x_i - \bar{x} \), where \( x \) is the value of the variable at location \( i \) and \( \bar{x} \) is the mean value of the variable in question. This tool tests for homogeneity in tract-level attribute values between each tract and its neighboring tracts. Therefore, a significantly high Moran’s I statistic would be the basis for rejecting the null hypothesis that the attribute is randomly distributed among census tracts within each MSA.

Since a strong spatial autocorrelation is detected between census tracts with respect to these eight characteristics (results are not included but can be provided by the author), OLS is unsuitable to predict unbiased estimates. To generate the most accurate parameter estimates, and therefore conclusions, a spatial regression model is used to control for spatial dependence. A spatial regression model simply adds an additional error term to the linear probability model to account for spatial autocorrelation (Anselin, Syabri, and Kho, 2006).

Even though the results from the Global Moran’s I test indicate that OLS estimates are not suitable, it is still necessary to perform a Lagrange multiplier test on OLS residuals to determine which spatial regression model specification — either spatial error or spatial lag — is appropriate for the regressions. For all three MSAs, the Lagrange multiplier test statistic indicated that a spatial lag model is required. A spatial lag model assumes that the
explanatory variables alone cannot explain the spatial dependence because the spatial structure alters the dependent variable (Anselin, 2004; Zou, 2014). Theoretically, the spatial lag model is designed for use in instances in which externalities or spillover effects are common (Troy et al., 2012; Anselin 2004). In this analysis, socioeconomic characteristics and local credit conditions can explain these effects. The spatial lag models (one unique model per MSA) are identified as:

\[ y = \rho Wy + \sum_{k} \beta_k x_k + \lambda s_1 + \lambda c_2 + u \]

where \( y \) represents whether the tract (which had an open retail bank branch in June 2010) experienced at least one branch closure, \( Wy \) is a \( n \times 1 \) vector of the spatially lagged dependent variable, \( \beta \) denotes the slope of each explanatory variable, \( x \) represents the explanatory variables, \( \rho \) denotes the spatial autoregressive coefficient, \( \lambda s_1 \) represents state fixed effects, \( \lambda c_2 \) represents county fixed effects, and \( u \) denotes a spatially independent error term.

For all MSAs, a k-nearest neighbors spatial weights matrix, \( W \), is used in this study. A spatial weights matrix of \( n \) units is a nonnegative matrix \( W = (w_{ij}: i, j = 1, \ldots, n) \), and each spatial weight, \( w_{ij} \), represents the spatial influence of unit \( j \) on unit \( i \). For a given census tract, \( i \), this weighting structure simply ranks all the other census tracts \( j \), in terms of centroid distance from tract \( i \). The set \( N_k(i) = \{j(1), j(2), \ldots j(k)\} \) contains the k-closest units (census tracts) to unit \( i \). Formally, the k-nearest neighbor spatial weights matrix is defined by:

\[ w_{ij} = \begin{cases} 1, & j \in N_i(k) \\ 0, & \text{otherwise} \end{cases} \]

Census tracts are, of course, not all the same size and shape. Census tracts are smaller in city centers and relatively larger in suburban areas. Therefore, this paper uses row standardizations of the distance weights matrices. To standardize the original spatial weights matrix, the nonnegative weights \( (w_{ii}: i \neq j) \) are normalized to have a unit sum.

\[ \sum_{j=1}^{n} w_{ij} = 1, \quad i = 1, \ldots, n \]

The interpretation of each weight within the matrix is now “the fraction of all spatial influence on unit \( i \), from unit \( j \)” (Smith, 2016).

All regression models and corresponding statistics were run in RStudio. The original linear probability model and the spatial model of best fit are reported in Table 4.

---

2 When regressions are run with a bank fixed effect, the results (not shown) are qualitatively and quantitatively the same.
3. Results

3.1 Cluster Analysis Results

The association between bank branch status (open or closed) and branch location was explored first in this analysis. As can be seen in Figures 4 and 7, there is evidence that the distribution of branch closures in both the Chicago and Philadelphia MSAs are clustered at relatively short distances (at roughly 2–4 km), conditional on the 2010 retail bank branch networks. However, Figure 10 illustrates that within the Baltimore MSA, the estimated difference, $\tilde{R}_{11}(r) - \tilde{R}_{22}(r)$, remains within the Monte Carlo envelopes for all radii, $r$. Therefore, there is no evidence of clustering within the Baltimore MSA.

The majority of all branches closed during this period of study occurred outside the major city’s limits in all MSAs analyzed: Only 16.57 percent of all Baltimore MSA branch closures occurred in Baltimore; Philadelphia (city) experienced 19.32 percent of all branches lost in the MSA, and Chicago experienced 24.16 percent of its MSA's closures. As a result, the same analysis is run next on the primary city itself (Chicago, Philadelphia, and Baltimore, respectively) as well as the suburban region (the entire MSA minus the primary city itself). Similar to the results shown in Figure 10, the random labelling tests failed to produce significant evidence of clustering within Baltimore (city) or within the noncity portion of the Baltimore MSA. However, a pattern emerged in the Philadelphia and Chicago MSAs: Figures 6 and 9 show that the respective differences, $\tilde{R}_{11}(r) - \tilde{R}_{22}(r)$, for the city alone, remain well within the constructed 95 percent of critical bands. Therefore, there is no evidence of clustering among retail bank branch closures in either city (conditional on the branching networks in 2010). So we conclude that 2010–2016 retail bank branch closures do not appear to be clustered and do not have any significant spatial pattern in either primary city, given the locations of retail bank branches in 2010.

The results for the suburban areas do, however, appear to indicate the presence of a spatial pattern. These results are shown in Figures 5 and 8. The estimated difference functions both fail to be completely contained by the 95 percent confidence bands generated from the Monte Carlo simulations. The null hypothesis is rejected in both the noncity regions of Philadelphia and Chicago with relatively smaller values of $r$. Therefore, we conclude that there is evidence of retail branch closures occurring in clusters in the noncity regions of both the Chicago and Philadelphia MSAs. Additionally, the null hypothesis is also rejected at greater distances (approximately when $r = 17 – 20$ km) within the Philadelphia MSA (Figure 5).
Figures 4–10: The plot displays the difference between the summary K-function obtained from the random permutations and the K-function constructed from observed branch closures. The gray area represents the 95 percent confidence interval found by the Monte Carlo simulations; it is constructed from envelopes of 39 random labellings of the MSA branch closure data. The dotted (red) line is the average of the simulations generated under the null hypothesis.
Figure 5. Chicago MSA (Noncity): Observed Closed Branches versus All Branch Locations

Figure 6. Chicago MSA (City Only): Observed Closed Branches versus All Branch Locations
Figure 7. Philadelphia MSA: Observed Closed Branches versus All Branch Locations

Figure 8. Philadelphia MSA (Noncity): Observed Branch Closures versus All Branch Locations
Figure 9. Philadelphia MSA (City Only): Observed Branch Closures versus All Branch Locations

Figure 10. Baltimore MSA: Observed Branch Closures versus All Branch Locations
Bivariate Analysis

The bivariate analysis confirmed the trend illustrated in the summary statistics (Table 2) that Chicago and Philadelphia branch closure census tracts are much more similar than the Baltimore tracts. Table 3 contains these findings. There is a positive and significant relationship between median income and branch closures as well as a significant negative relationship between the unemployment rate and branch closures in both Chicago and Philadelphia. It is also important to note that the signs of the estimated Spearman’s correlation coefficients are not uniform across cities. The percent of the population below the federally defined poverty line is negatively related to experiencing a branch closure in Chicago but positively related within the Baltimore MSA. Additionally, the share of the tract that is white is negatively related to branch closures in Baltimore but has the opposite relationship in the Chicago MSA. These correlations varied greatly across MSAs.

Despite what drives branch closures, there are patterns of branch closures with respect to socioeconomic and demographic variables. The bivariate analysis results indicate that 2010–2016 branch closures were not equally distributed among all tracts that had at least one bank branch in these three metro regions. Results suggest that with respect to median income, poverty, and race, branch closures did not occur uniformly.

Table 3. Spearman’s Correlation Results

|                                      | Philadelphia | Chicago   | Baltimore  |
|--------------------------------------|--------------|-----------|------------|
| Median income ($)                    | 0.0813**     | 0.1087*** | -0.0647    |
| Black (%)                            | -0.0423      | 0.0055    | 0.1426***  |
| Hispanic (%)                         | 0.0017       | -0.0826***| 0.0922*    |
| Asian (%)                            | 0.0786**     | 0.0450    | 0.0139     |
| White (%)                            | 0.0286       | 0.0618**  | -0.1337**  |
| Population younger than 24 years of age (%) | -0.0599*     | 0.0117    | 0.0629     |
| Total population                     | 0.0499       | 0.1016*** | 0.0774     |
| Unemployment rate (%)                | -0.0872**    | -0.0517*  | 0.0146     |
| Population below poverty line (%)    | -0.0509      | -0.0887***| 0.0922*    |
| Distance to CBD (degrees)            | 0.0224       | 0.3871    | -0.0277    |
| No. of tracts, N                     | 309          | 509       | 141        |

*p < 0.1; **p < 0.05; ***p < 0.01. CBD=central business district.

3.2 Multivariate Analysis

Table 4 contains the results of tracts experiencing at least one branch closure — represented by a dummy (0,1) — regressed on multiple neighborhood socioeconomic and
demographic variables. The first column for all metropolitan statistical areas holds the OLS estimation results from the base linear probability model. The second column within each MSA displays the results from spatial lag model. The three spatial lag models consistently produced lower Akaike information criterion (AIC) values compared with the OLS estimates derived from the linear probability model. The AIC value is commonly used to compare a spatial error model with a baseline model; a model with the lower AIC value is interpreted as being “most valid” (Anselin, 2004). This suggests that the spatial lag (SLAG) models produce more robust results compared with the baseline linear probability models. The SLAG models do not noticeably change the coefficients in terms of sign or magnitude; however, in some cases, differences emerge with respect to statistical significance. This result suggests that there is a problem of efficiency with OLS, and this is lessened with the use of the SLAG model.

It is important to note that there is no tract characteristic that has a consistent, significant relationship to experiencing a branch closure across all three MSAs from 2010 to 2016. The percent of unemployed has a significant negative relationship with branch closures in the Baltimore MSA. Additionally, the SLAG model resulted in tract distance to the CDB having a significant negative relationship with Baltimore MSA branch closures. Both Chicago and Philadelphia have a significant positive relationship between median family income and branch closures, whereas no such relationship exists in the Baltimore MSA (Table 3). Within Chicago, the share of tract population that is Hispanic is negatively related to branch closures.

### Table 4. Associations of Census Tract Characteristics with Exposure to at Least One Bank Branch Closure While Accounting and Not Accounting for Spatial Autocorrelation

|                      | Baltimore OLS   | Baltimore SAR (SLAG) | Chicago OLS   | Chicago SAR (SLAG) | Philadelphia OLS | Philadelphia SAR (SLAG) |
|----------------------|----------------|---------------------|---------------|-------------------|----------------|------------------------|
| Median family income ($1,000) | -0.0016       | -0.0015             | 0.0016***     | 0.0016***         | 0.0013*        | 0.0012*                |
| Black (%)            | 0.2628         | 0.2808              | -0.0765       | -0.0767           | -0.2130        | -0.2358                |
| Hispanic (%)         | 0.8591         | 0.8907              | -0.1034       | -0.1223*          | 0.0526         | 0.0290                 |
| Asian (%)            | 0.1728         | 0.2298              | 0.1557        | 0.1568            | 0.2779         | 0.2160                 |
| White (%)            | -0.8094        | -0.7276             | 0.0547        | 0.0551            | -0.2117        | -0.2331                |
| Population younger than 24 years of age (%) | 0.0058         | 0.0061              | 0.0026        | 0.0027            | -0.0020        | -0.0018                |
| Total population     | 0.0000         | 0.0002              | 0.00054**     | 0.00053***        | 0.0001         | 0.0001                 |
| Unemployed (%)       | -0.0226**      | -0.0210**           | 0.0063        | 0.0064            | -0.0025        | -0.0020                |
| Population below poverty line (%) | 0.0007       | 0.0002              | 0.0018        | 0.0018            | -0.0006        | -0.0007                |
| Distance to CDB      | -0.5961        | -0.6205*            | 0.0513        | 0.0507            | 0.0222         | 0.0097                 |
| Moran’s I statistic on residuals | -0.0577      | -0.0034             | 0.0119        | 0.0091            | 0.0307         | 0.0064                 |

*p < 0.1; **p < 0.05; ***p < 0.01. CBD=central business district; OLS=ordinary least squares.
The OLS and SAR results also include the Moran's I test statistics derived from the residuals in both types of models. While comparing the two isn't the formal test for the presence of spatial autocorrelation, evaluating the test statistics generated from the residuals of both models is useful. The Moran's I statistic values for all three SAR models are much lower than the Moran's I in the OLS models. These Moran's I values indicate that spatial autocorrelation has been drastically reduced with the use of the spatial autoregressive models.

4. Discussion and Conclusion

This study revealed evidence of branch closure clustering at small distances (2–5 km) in both the Chicago and Philadelphia MSAs. There was no evidence of any spatial patterns of branch closures in the Baltimore MSA during 2010–2016. Additionally, these spatial clustering patterns seem to be driven by the branch closure locations in quasi-urban areas, or branches within the MSA but not within the primary city itself. This result complements recent literature and trade press articles that indicate banks are prioritizing urban branches at the expense of quasi-urban and rural branches. Also, these results are consistent with the hypothesis that banks are incentivized by the Community Reinvestment Act (CRA) and make an effort to keep branches open in CRA-eligible census tracts. Additionally, while this trend was evident in the Chicago and Philadelphia MSA results, not all metropolitan regions studied in this paper experienced a clear, spatial pattern with respect to branch closures. This finding is important because it supports the hypothesis that banks have different retail branching strategies both across metropolitan regions and within metro areas.

While this paper is primarily concerned with the branch closure processes — not the original branching location decisions — future research should aim to analyze both. When closures are determined by merger and acquisition activity, the primary spatial process — or the decision where to open a bank — becomes more important. That is why the next steps will include analyzing multiple firm decisions, not just closure locations.

This study also attempts to reveal whether there are socioeconomic imbalances associated with the distribution of branch closure locations experienced in three metropolitan regions. Again, it is important to stress that this analysis tried to identify relationships from the set of census tracts that housed an open bank branch in mid-2010. Results from both OLS and spatial regression models indicate a slight evidence of association between experiencing a branch closure and socioeconomic variables (at the census tract or neighborhood level). This evidence suggests that other variables may be much better predictors of branch closures. While other census-tract level characteristics not used in this
paper need to be explored in future work, firm- and branch-level characteristics also need to be included in future branch closure studies.

The results also reveal that making use of spatial models is a clear improvement over more common regression models that cannot control for spatial autocorrelation. There was significant spatial autocorrelation detected in all three samples and the use of spatial models improved estimation precision. This knowledge is useful for upcoming inquiries. Bank branch proximity research reliant on OLS models may introduce unintentional inefficiency. This could lead to imprecise conclusions. For example, in Chicago, the OLS model indicates that there is no relationship between the share of the tract population that is Hispanic and branch closures. However, the spatial lag model shows a significant negative relationship between the two. Other social science fields commonly employ spatial regression models in their research; changes to retail bank branch networks is just one example in which there is much to be gained with the application of spatial models.

Lastly, identifying patterns between tracts that experienced a branch closure versus tracts that contained a branch that did not close — instead of all census tracts within the MSA — removes a great deal of variation from this analysis. In essence, it removes all census tracts that never had any branches to lose. Figures 11–13 illustrate this point. Large regions in both Chicago and Philadelphia have poor access to formal financial institutions. Future bank branching and financial inclusion research need to determine if the scope of the research includes only geographies that recently had retail bank branches but lost them or all regions that do not have access to bank access. This paper used the former approach. However, researchers must decide if all banking deserts should be treated the same or if there are truly fundamental differences between relatively new banking deserts and enduring banking deserts.
Figure 11. Chicago Census Tracts with a Branch (Blue)

Figure 12. Philadelphia Census Tracts with a Branch (Blue)
In summary, this exploratory study uncovers spatial clustering of branch closures within two out of the three metropolitan statistical areas analyzed. Spatial autoregressive models then provided better model fitness for all regions analyzed. Among the limitations of this study is the use of the branch location data. Although the Branch Analytics Tool uses cleaned SOD data, this data set has well-documented challenges with respect to consistent formatting and incomplete address information. Second, this study was cross-sectional. Adding a temporal aspect to this analysis would be very informative.

Examining the spatial distribution of branch closures is crucial to our understanding of financial inclusion, especially our access to the formal financial sector. However, this picture is not complete without including branch openings. Therefore, future research will need to model both branch closures and openings. Shrinking branch networks are expected to continue in the foreseeable future. Determining the spatial structure of either the closures or the remaining branches is the first step in understanding the potential impact of branch closures. Future research is needed; however, this research must take into account the spatial structure common to these data, and the increasing use of spatial models is encouraged.
References

Anselin, L. (2004). Exploring spatial data with GeoDa™: a workbook, Urbana, 51(61801), 309.

Anselin, L., I. Syabri, and Y. Kho (2006). “GeoDa: An Introduction to Spatial Data Analysis,” Geographical Analysis, 38(1), 5–22.

Baddeley, A. (2008). “Analysing Spatial Point Patterns in R. Technical Report,” CSIRO, 2010. Version 4. URL https://research.csiro.au/software/r-workshop-notes.

Bailey, T. C., and A. C. Gatrell (1995). Interactive Spatial Data Analysis (413). Essex: Longman Scientific & Technical.

Bivand, R., et al. (2017). spdep: Spatial Dependence: Weighting Schemes, Statistics and Models. R package version 0.6.15.

Bivand, R. S., J. Hauke, and T. Kossowski (2013). “Computing the Jacobian in Gaussian Spatial Autoregressive Models: An Illustrated Comparison of Available Methods,” Geographical Analysis, 45(2), 150–179.

Bivand, R., and G. Piras (2015). “Comparing Implementations of Estimation Methods for Spatial Econometrics,” Journal of Statistical Software, 63(18), 1–36. http://www.jstatsoft.org/v63/i18/.

Brown, J. R., J. A. Cookson, and R. Heimer “Growing Up Without Finance” (September 8, 2016), 7th Miami Behavioral Finance Conference 2016. https://ssrn.com/abstract=2809164.

Buzard, K., G. A. Carlino, R. M. Hunt, J. K. Carr, and T. E. Smith (2017). “The Agglomeration of American Research and Development Labs,” Federal Reserve Bank of Philadelphia Working Paper 17-18, Journal of Urban Economics, (forthcoming), https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2017/wp17-18.pdf.

Celerier, C., and A. Matray (2017). “Bank Branch Supply and the Unbanked Phenomenon,” HEC Paris Research Paper FIN-2014-1039, https://ssrn.com/abstract=2392278.

Cliff, A.D., and J. K. Ord (1981). Spatial Processes: Models and Applications. Pion Limited, London.

Cohen, J. P., and C. C. Coughlin (2008). “Spatial Hedonic Models of Airport Noise, Proximity, and Housing Prices,” Journal of Regional Science, 48(5), 859–878.
Cronie, O., and M. N. M. van Lieshout, M. N. M. (2016). “Summary Statistics for Inhomogeneous Marked Point Processes,” *Annals of the Institute of Statistical Mathematics*, 68(4), 905–928.

Degryse, H., and S. Ongena (2005). “Distance, Lending Relationships, and Competition,” *The Journal of Finance*, 60(1), 231–266.

Diggle, P. J., and A. G. Chetwynd, A. G. (1991). “Second-Order Analysis of Spatial Clustering for Inhomogeneous Populations,” *Biometrics*, 47(3), 1155–1163.

Dixon, P. M. (2002). “Ripley's K function,” *Encyclopedia of Environmetrics*. 3, 1796–1803.

Dubin, R. A. (1992). “Spatial Autocorrelation and Neighborhood Quality,” *Regional Science and Urban Economics*, 22(3), 433–452.

Gensler, Lauren (2016). “Get Ready for a Lot Fewer Bank Branches Around,” *Forbes Online*, https://www.forbes.com/sites/laurengensler/2016/07/18/bank-branch-closures-more-to-come/#44ead6df5e04.

Goodstein, R. M., and S. L. Rhine (2017). “The Effects of Bank and Nonbank Provider Locations on Household Use of Financial Transaction Services,” *Journal of Banking & Finance*, 78, 91–107.

Hwang, J., M. Hankinson, and K. S. Brown (2014). “Racial and Spatial Targeting: Segregation and Subprime Lending Within and Across Metropolitan Areas,” *Social Forces*, 93(3), 1081–1108.

Jagtiani, J., and C. Lemieux (2017). “Fintech Lending: Financial Inclusion, Risk Pricing, and Alternative Information.” Federal Reserve Bank of Philadelphia Working Paper 17-17. https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2017/wp17-17.pdf.

Morgan, D., M. Pinkovsky, and B. Yang (2016). “Banking Deserts, Branch Closings, and Soft Information,” *Liberty Street Economics*, Federal Reserve Bank of New York.

Nguyen, H. L. Q. (2014). *Do Bank Branches Still Matter? The Effect of Closings on Local Economic Outcomes*. Cambridge: Department of Economics, Massachusetts Institute of Technology.

Smith, T. (2016). “Comparative Analysis of Point Patterns,” Spatial Data Analysis with GIS Applications, Philadelphia: University of Pennsylvania.
Smith, T. E., M. M. Smith, and J. Wackes (2008). “Alternative Financial Service Providers and the Spatial Void Hypothesis,” *Regional Science and Urban Economics*, 38(3), 205–227.

SNL Unlimited. (2017). “Branch Analytics,” *S&P Global Market Intelligence*.

Taylor, J., B. Mitchell, J. Franco, and Y. Xu (2017). “Bank Branch Closures from 2008–2016: Unequal Impact in America’s Heartland,” [http://www.ncrc.org/conference/wp-content/uploads/2017/05/NCRC_Branch_Deserts_Research_Memo_050517_2.pdf](http://www.ncrc.org/conference/wp-content/uploads/2017/05/NCRC_Branch_Deserts_Research_Memo_050517_2.pdf).

Troy, A., J. M. Grove, and J. O'Neil-Dunne (2012). “The Relationship Between Tree Canopy and Crime Rates Across an Urban–Rural Gradient in the Greater Baltimore Region,” *Landscape and Urban Planning*, 106(3), 262–270.

U.S. Census Bureau/American FactFinder. “B02001: Race.” *2011–2015 American Community Survey*. U.S. Census Bureau’s American Community Survey Office, 2015, Web, 7 September 2017, [http://factfinder2.census.gov](http://factfinder2.census.gov).

—— “S0101: Age and Sex.” *2011–2015 American Community Survey*, U.S. Census Bureau’s American Community Survey Office, 2015, Web, 7 September 2017, [http://factfinder2.census.gov](http://factfinder2.census.gov).

—— “S1701: Poverty Status in the Past 12 Months.” *2011–2015 American Community Survey*, U.S. Census Bureau’s American Community Survey Office, 2015, Web, 7 September 2017, [http://factfinder2.census.gov](http://factfinder2.census.gov).

—— “S1901: Income in the Past 12 Months (In 2016 Inflation-Adjusted Dollars).” *2011–2015 American Community Survey*, U.S. Census Bureau's American Community Survey Office, 2015, Web, 7 September 2017. [http://factfinder2.census.gov](http://factfinder2.census.gov).

—— “S2301: Employment Status.” *2011–2015 American Community Survey*, U.S. Census Bureau’s American Community Survey Office, 2015, Web, 7 September 2017, [http://factfinder2.census.gov](http://factfinder2.census.gov).

Zou, Y. (2014). “Analysis of Spatial Autocorrelation in Higher-Priced Mortgages: Evidence from Philadelphia and Chicago,” *Cities*, 40, 1-1.
Appendix A. Selected Covariates

Median Family Income: Philadelphia

Source: S&P Global Market Intelligence, SNL Branch Analytics
Baltimore

Median Family Income: Baltimore

Source: S&P Global Market Intelligence, SNL Branch Analytics
Percent Black: Baltimore

Chicago
Median Family Income: Chicago
Appendix B. Additional Baltimore Spatial Random Labelling Test Results

Test of Random Labelling: Baltimore City Only

![Graph showing test results](image-url)
Test of Random Labelling: Baltimore Non-City Only

$K_{\text{observed}}(r) - K(r)$

Average of Monte Carlo Simulations

95% of Monte Carlo Simulations

$r$ (kilometers)
Appendix C. Branch Closure Counts

Figure 1. Philadelphia MSA: Branch Closures by Census Tract

Branch Closure Count: Philadelphia

Source: S&P Global Market Intelligence, SNL Branch Analytics

Figure 2. Baltimore MSA: Branch Closures by Census Tract

Branch Closure Count: Baltimore
Figure 3. Chicago MSA: Branch Closures by Census Tract

Source: S&P Global Market Intelligence, SNL Branch Analytics
Figure 4. Chicago MSA County Selection: Branch Closures by Census Tract
Branch Closure Count: Cook & DuPage Counties, IL

Source: S&P Global Market Intelligence, SNL Branch Analytics