Sorry We're Closed: What Closes Malls and Community Centers in the United States? An Analysis and Predictive Modeling of Distressed Centers

by

Morgan Fleischman

B.S., Hotel Administration, 2011

Cornell University

Submitted to the Program in Real Estate Development in Conjunction with the Center for Real Estate in Partial Fulfillment of the Requirements for the Degree of Master of Science in Real Estate Development

at the

Massachusetts Institute of Technology

September, 2020

Signature of Author____
Center for Real Estate
July 16, 2020

Certified by____
William Wheaton
Professor, MIT Center for Real Estate
Professor Emeritus, Department of Economics
Thesis Supervisor

Certified by____
Anne Kinsella Thompson
Research Scientist and Visiting Lecturer
MIT Center for Real Estate
Thesis Co-Supervisor

Accepted by____
Professor Dennis Frenchman
Class of 1922 Professor of Urban Design and Planning
Department or Urban Studies and Planning
Director, MIT Center for Real Estate

Sorry We're Closed: What Closes Malls and Community Centers in the United States? An Analysis and Predictive Modeling of Distressed Centers

By

Morgan Fleischman

Submitted to the Program in Real Estate Development in Conjunction with the Center for Real Estate on July 16, 2020 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Real Estate Development

ABSTRACT

The retail industry has transformed markedly in the last decade driven by the confluence of technology, evolving consumer behavior, and innovation. As the sector continues to evolve, retail real estate is coming under significant pressure to keep up with the pace of change. Some centers are poised to quickly adapt and come out stronger, others are left behind and going dark. This paper is particularly interested in examining the geography of distressed retail centers, specifically malls and community centers, understanding what factors lead to their closure, and coming up with a predictive model to measure the properties' probability of defaulting.

We analyze the geography of approximately 4,900 malls and community centers across the United States at two intervals, 2010 and 2020. First, we isolate the centers that have died within the last decade to identify what distinguishes between a dead and survived center. Second, for each dead center we estimate the distance of its competitors to assess impact of spatial competition. Third, we use a linear regression to identify determinants that influence the death of a center. Fourth, we run a probit model on the center's survived-dead response based on each variable. We conclude by developing a predictive model to assess which centers are at greatest risk of underperforming.

Research shows that a property's net rentable area has an outsized impact on the probability of defaulting, with opposite effects for malls and community centers. As malls grow larger, they are less likely to become distressed – whereas the growth of community centers leads to a higher probability of failure. The center's proximity to competition, and the amount of available space both increase the impact on the likelihood of going under. The opportunity to renovate, on the other hand, mitigates the impact.

Thesis Supervisor: William Wheaton
Title: Professor, Center for Real Estate & Professor Emeritus, Department of Economics

Thesis Co-Supervisor: Anne Kinsella Thompson
Title: Research Scientist and Visiting Lecturer

Acknowledgements

This thesis was possible due to several individuals who supported me in this project. So far, I extend my gratitude to my academic and thesis advisor, Professor Bill Wheaton, and to my co-advisor Annie Kinsella Thompson. I thank them for all their help and guidance.

I would also like to thank my friends, classmates, and the MIT CRE faculty and staff for making my graduate school experience memorable. I am grateful for the faculty and staff's unprecedented support during a challenging year dealing with a pandemic.

Finally, I give very special thanks to my parents, family, and Stacie who have been there for me in each step of the way.

Table of Contents

1. Introduction

The retail services industry amounts to well over five trillion dollars annually[1] with at least 1 million retail establishments across the United States. Since 2010, sales have increased nearly 4 percent annually[2]. Although the industry is undergoing an enormous transformation driven by technology, evolving consumer behavior, and innovation, it still plays an important role in shaping the economic, cultural, and social viability of communities across the country. Physical stores account for majority of sales in the market, representing a fundamental asset class that is often well-located with good transport links and parking. National Association of Real Estate Investment Trusts (Nareit) estimates that retail real estate accounts for 13.6 billion square feet in the United States and is valued at $2.4 trillion, approximately 15% of the commercial real estate market share[3].

We often ask ourselves "what will the future of retail real estate look like?" The industry is going through significant changes and evolution, creating a mixed narrative. Some headlines depict retail as a doomed industry, whereas some industry experts set the record straight that the sector is seeing growth. Either way, there is no aspect of the industry that is going to be unaffected.

Starting with the negative sentiment, the consensus belief is that smaller traditional players will be hurt the most. According to a report by LoopNet, an affiliate of CoStar Group, analysts at UBS recently noted that retail stores could take a big hit in the next five years while the share of retail online sales increases from 15% in 2019 to 25% in 2025[4] – a trend that could be further accelerated by the COVID-19 outbreak. The report estimates that 100,000 locations could close by 2025, and closures may be as high as 150,000 locations. In 2017, Credit Suisse offered a grim outlook of malls. The bank estimated that nearly one in four of the country's 1,169 malls, or 275 malls, will close in this decade[5].

The outlook is already evident as prominent retailers have shut down their stores. Last year, Gap Inc. announced closure of 230 stores as sales declined. Payless ShoeSource sought Chapter 11 bankruptcy protection and shut down all 2,100 of its stores. Both Barneys New York and Forever 21 announced closures. In 2020, Macy's plans to shutter hundreds of underperforming stores to focus on flagship

[1] Advance Monthly Retail. *US Census Bureau* Monthly & Annual Retail Trade. https://www.census.gov/retail/index.html; Total Retail sales include motor vehicle and parts.
[2] "State of Retail." National Retail Federation (NRF), https://nrf.com/insights/economy/state-retail.
[3] Thompson, Alexandria. "Total Size of U.S. Commercial Real Estate Estimated Between $14 and $17 Trillion." *Nareit*, https://www.reit.com/news/blog/market-commentary/total-size-us-commercial-real-estate-estimated-between-14-and-17.
[4] Kennedy, Clare. "UBS Estimates 100,000 More Retail Stores Will Close by 2025." *LoopNet*, 27 Apr. 2020, https://www.loopnet.com/learn/ubs-estimates-100000-more-retail-stores-will-close-by-2025/174318906.
[5] Buss, Christian, et al. *Apparel Retail & Brands: Making Sense Of Softlines Following A Tumultuous Twelve Months*. Credit Suisse, May 2017.

locations. Other notable closures and bankruptcies include Sears, Pier 1, Modell's Sporting Goods, J. Crew, Neiman Marcus, and J.C. Penney.

The department store challenges that have dominated the headlines tend to be anchor tenants to many malls. Their underperformance will have an impact on both mall owners and in-line tenants that depend on high-profile brands; in-line tenants contribute a larger share of a mall's net operating income since anchor tenants pay little-to-no rent. Green Street's Advisory & Consulting Group conducted a survey of 950 mall locations and identified that over two-thirds saw a net decrease in the number of national tenants[6]. The advisory group says Class B and C malls were more likely to contract faster and take on significant capital reinvestment. The reasoning is that malls have not kept pace with the changing behaviors of consumers or their needs, which will likely result in some malls being redeveloped to enhance shopper's experiences or repositioned for other purposes. Large players with sophisticated e-commerce operations and supply chains, such as Walmart, Target, and Amazon, may be positioned to come out stronger and seize greater market share.

Needless to say, there are some bright spots in the sector. According to a publication reported by International Council of Shopping Centers (ICSC) and National Retail Federation, both retail trade organizations, for every chain with a net closing of stores, 5.2 new stores are opened[7]. The report was conducted by the IHL Group, a global research and advisory group. The firm reported that the number of chain closings peaked in 2018 and dropped by 66% year-over-year in 2019. The total number of retail chains opening stores increased 56% year-over-year, making the news positive for store openings and closings. Apparel and department store chains saw net closure of 9,651, whereas all other retail segments saw net new openings of 18,226[8].

One common trend is that retailers are adapting for the omnichannel future and rolling out both online and physical channels. Consumers are making purchases online for their convenience, but they are still making their way to visit brick-and-mortar stores for experiences. Casper, Warby Parker, Away, Peloton, and Bonobos – which started as direct-to-consumer platforms – have adopted new sales strategies to reach consumers through experiential testing centers. Malls, on the other hand, have traditionally not kept pace with the changing expectations of consumers or their differing needs. But they are quickly transforming themselves to new kinds of destinations investing heavily in new amenities, experiences, and

[6] Sullivan, Jim, and Otto Aletter. *2017 Mall Tenant Turnover Analysis*. White Paper, Advisory & Consulting Green Street Advisors, 2018, pp. 7.
[7] Holman, Lee, and Greg Buzek. *Retail's Renaissance: The True Story of Store Openings/Closings*. IHL Group, 2019.
[8] Ibid.

entertainment to improve the shopping experience. Anchoring the department stores are pop-up stores, restaurants, bars, cinemas, and specialty fitness centers[9]. Although emerging brands are increasing their physical locations within malls, the issue remains that brands are more aware about mall quality and ownership, so they are opening physical stores at a much slower pace.

There are numerous articles, editorials, and periodicals examining the ongoing trends of stores closures and the outlook of the industry. However, empirical studies on mall and community center closures are limited. John M. Clapp, Katsiaryna Salavei Bardos, and Tingyu Zhou (2012)[10] analyzed the determinants of expansions and contractions of 343 property-level shopping centers in eleven metropolitan areas between 1995 and 2005. Their study does not determine the closure of a center, but merely the expansion and downsizing of the gross leasable area (GLA). Using three types of logistics regressions – ordered, multinomial, and simple logit – they empirically determined that an increase in operating costs and a decrease in revenue per square foot increases the probability of shopping center contraction. The shopping centers were classified as "large" consisting of community centers, and the remaining centers were classified as "small."

We have also come across other studies that utilize methodologies in the retail sector that we are motivated to use including predictive modeling, spatial competition, and the probit model.

In the case of predictive modeling, the works of Mark Eppli and James Shilling (1996)[11] measured the trade-off between store location and retail agglomeration from thirty-eight regional shopping centers. They used retail sales to determine the statistical significance of both retail agglomeration and proximity to competition to the success of a shopping center. Using the Lakshmanan and Hansen (1965) retail gravity model[12], Eppli and Schilling estimated sales of U.S. regional shopping centers and used the estimates as an explanatory variable in an OLS regression to predict actual sales for each center. The study found that retail sales at regional shopping centers are largely determined by center size, and to a lesser degree, by proximity to competition.

---

[9] Petro, Greg. "Shopping Malls Aren't Dying – They're Evolving." Forbes, 5 Apr. 2019, https://www.forbes.com/sites/gregpetro/2019/04/05/shopping-malls-arent-dying-theyre-evolving.
[10] Clapp, John, et al. "Expansions and Contractions of Major US Shopping Centers." *The Journal of Real Estate Finance and Economics*, Aug. 2012, pp. 16–56.
[11] Eppli Mark and Shilling James. How Critical Is a Good Location to a Regional Shopping Center? *Journal of Real Estate Research*, 1996, Vol. 12, No. 2, pp. 459-468.
[12] Lakshmanan, T.R. and W.G. Hansen, A Retail Market Potential Model, *Journal of the American Institute of Planners*, 1965, 31, pp. 134–43.

Our study is also interested in measuring spatial competition of centers. Geoffrey Turnbull and Jonathan Dombrow (2006)[13] test for spatial competition and shopping externality effects on prices and marketing time by using a sample of single-family housing transactions. The variables they use to measure competing houses for sale take into account of the distance between them after mapping all the houses into geographic coordinates. Ming-Long Lee and R. Kelley Pace (2005)[14] examined the spatial distribution of retail sales in Houston using a spatial gravity model. Lee and Pace collected census data and longitude and latitude information of each block-group to calculate the distances among retailers and consumers. They found that both forms of spatial dependence, consumer and retailer, had statistically significant impacts on the estimates of parameters in retail gravity models, providing an opposing view to the findings in Eppli and Shilling's study (1996) that distance parameter may be significantly overstated.

We analyze the empirical study of Tingyu Zhou and John Clapp (2015)[15] to understand the deployment of the probit model. They used probit to predict risks of retail anchor store openings. The model with location fixed effects estimated the probability of openings and closings of anchor stores as a function of the number of competitors. Simon Buechler, Alex van de Minne, and Olivier Schöni's (2020)[16] also used the probit model in "Redevelopment Option Value for Commercial Real Estate" to compute the impact of the redevelopment on the individual prices of different types of commercial real estate properties, including retail. In a related study, Yuling She (2020)[17] built on Buechler's (et al.) methodology to further compute a redevelopment propensity metric on industrial properties. By applying a probit model, we can determine the marginal effects of the mall and community center's age, square footage, and spatial competition on the probability of dying.

Based on these existing empirical studies, our analysis differs in several ways. First, it is a rigorous empirical research on malls and community center closures in the United States. Second, our analysis has thousands of recent property-level data points from CBRE for years 2010 and 2020. Third, we calculate the distances for 4,900 properties relative to each other to estimate the spatial competition based on a radius. Fourth, over a ten-year period we disaggregate malls and community centers into survived and

---

[13] Turnbull, Geoffrey, and Jonathan Dombrow. "Spatial Competition and Shopping Externalities: Evidence from the Housing Market." *The Journal of Real Estate Finance and Economics*, 2006, pp. 391–408.
[14] Lee, Ming-Long, and R. Kelley Pace. "Spatial Distribution of Retail Sales." *The Journal of Real Estate Finance and Economics*, vol. 31, no. 1, Aug. 2005, pp. 53–69.
[15] Zhou, Tingyu, and John Clapp. "Predicting Risks of Anchor Store Openings and Closings." *The Journal of Real Estate Finance and Economics*, Sept. 2015, pp. 449–79.
[16] Buechler, Simon, et al. "Redevelopment Option Value for Commercial Real Estate." *University of Bern, University of Connecticut, Laval University*, 2020.
[17] She, Yuling "Development Option Value for Industrial Property." *Massachusetts Institute of Technology, The Center for Real Estate*, 2020.

dead centers to form our hypothesis. Fifth, we compute a probit model to estimate the probability for centers that died. Finally, we develop a predictive model to identify the centers are at risk of going under today. In the next chapter, we provide an overview of the data. Chapter 3 discusses the research methodology. A review of the results is covered in Chapter 4, and Chapter 5 summarizes our findings.

2. Data and Descriptive Statistics

In this section, we provide context of the data for malls and community centers, identify the changes that took place between two time periods, and introduce the classification of each center. We also discuss the key variables in the data set and describe the fundamental statistics.

2.1 Data Background

In this study, we rely on georeferenced data for all retail centers. The information comes from CBRE, which sources neighborhood community services (NCS) and malls data from several repositories including CoStar (primary source), LoopNet[18], and CBRE proprietary. The data captures an inventory of all centers with availability details allowing brokers to help find lease tenants. The repositories are refreshed, updated, and maintained continuously.

We obtained two slices of data – centers as of Q1 2010 and centers of Q1 2020. The two snapshots, separated by ten years, are compared and matched using latitude and longitude coordinates. There are three instances in which centers can show up in the surveys:

- *A center in 2020 survey but not in 2010 survey*: A center can show up in the survey in 2020 but not in the 2010 survey, indicating it is a new center. In all cases where they show up in the 2020 survey, the year of construction will indicate the center is a new property.
- *A center in 2010 survey but not in 2020 survey*: If the center shows up in the earlier 2010 survey but not in the 2020 survey, we assume the center has "died." It could be "dead" for a variety of reasons, including decommissioned, going through enough renovation to be "empty" and closed, demolished, turned into another use, or undergoing some other significant transitioning.
- *A center in both 2010 and 2020 surveys*: A center could show up in both surveys. If the center is in same category, we assume it is a "survived" center. The center can also transition into a different category between 2010 and 2020. For instance, a mall is mothballed and downgraded to a community center. In the study, we excluded all transitioned centers – all upgrades or downgrades to a different classification – in both surveys.

For the purposes of the study, we are (1) comparing survived centers with dead centers, and (2) examining only malls and community centers. Regarding other centers, strip centers and neighborhood centers account for the majority of CBRE's inventory, but they may introduce lots of random noise in the observations since they are smaller and more ubiquitous. We look at the observations for both malls and

[18] LoopNet is a subsidiary of CoStar Group

community centers individually since each center offers different kinds of shopping experiences, selections of merchandise, types of anchors, and physical layouts.

The number of malls and community centers in CBRE's 2010 inventory are 710 and 5,322, respectively. Malls consist of both regional malls and super regional malls, but we treat them in one group. The data represent centers from 76 core based statistical area (CBSA)[19], a proxy for metropolitan statistical area, shown in Table 2-1. We overlay the latitude and longitude coordinates from the data set onto a CBSA map using ArcGIS to size the number of centers in each area. The data table featuring top fifteen regions contains 5,322 community centers in 75 regions, and 710 malls in 68 regions. We observe that the majority of centers are in three metropolitan areas – Chicago, New York, and Los Angeles – which collectively comprise 17% of the total inventory.

Table 2-1: CBRE Inventory (Malls and Community Centers) by CBSA – Top 15[20]

Distribution of CBRE Inventory By CBSA - Top 15

	Total Inventory			% Mix		
	Total	Malls	Community Centers	Total	Malls	Community Centers
Chicago-Naperville-Elgin, IL-IN-WI	359	29	330	6%	4%	6%
New York-Newark-Jersey City, NY-NJ-PA	345	47	298	6%	7%	6%
Los Angeles-Long Beach-Anaheim, CA	330	49	281	5%	7%	5%
Miami-Fort Lauderdale-Pompano Beach, FL	251	27	224	4%	4%	4%
Dallas-Fort Worth-Arlington, TX	249	21	228	4%	3%	4%
Atlanta-Sandy Springs-Alpharetta, GA	235	20	215	4%	3%	4%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	230	23	207	4%	3%	4%
Houston-The Woodlands-Sugar Land, TX	217	19	198	4%	3%	4%
Phoenix-Mesa-Chandler, AZ	193	21	172	3%	3%	3%
Washington-Arlington-Alexandria, DC-VA-MD-WV	181	27	154	3%	4%	3%
Riverside-San Bernardino-Ontario, CA	174	14	160	3%	2%	3%
Boston-Cambridge-Newton, MA-NH	155	19	136	3%	3%	3%
Orlando-Kissimmee-Sanford, FL	136	10	126	2%	1%	2%
San Francisco-Oakland-Berkeley, CA	135	20	115	2%	3%	2%
Detroit-Warren-Dearborn, MI	127	18	109	2%	3%	2%
All Other	2,715	346	2,369	45%	49%	45%
Total Inventory	**6,032**	**710**	**5,322**	**100%**	**100%**	**100%**
Total CBSA Coverage by Center:	***76***	***68***	***75***			

[19] Core Based Statistical Area (CBSA) is defined by the Office of Management and Budget (OMB). They are defined by: (1) Metropolitan statistical areas (MSA) have at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties, and (2) Micropolitan statistical areas are a new set of statistical areas that have at least one urban cluster of at least 10,000 but less than 50,000 population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.
[20] The full list of CBRE Inventory by CBSA is in Appendix A

Next, for all data we filter out extreme values, particularly centers that were built prior to the 1930's, have erroneous inputs, and contain net rentable area (NRA) under 100,000 SF for malls and 50,000 square feet for community centers. We also remove centers that were either upgraded or downgraded (transitioned) to a different category since we want to look at only dead and survived results. After refining the mall parameters for the study and removing mall reclassifications, we are left with 646 observations for malls and 4,254 observations for community centers.

Figure 2-1 shows the cumulative inventory walk for malls from 2010 to 2020. The study examines 646 mall observations, consisting of 139 malls that died (22%) and 507 malls that survived (78%) from the 2010 survey. The survived malls show up in the 2020 survey, and the malls that died do not show up in the 2020 survey. We ignore any new centers that were constructed in 2020 or that were reclassified to a mall from a different center in 2020.

Figure 2-1: Cumulative Inventory Walk for Malls[21]

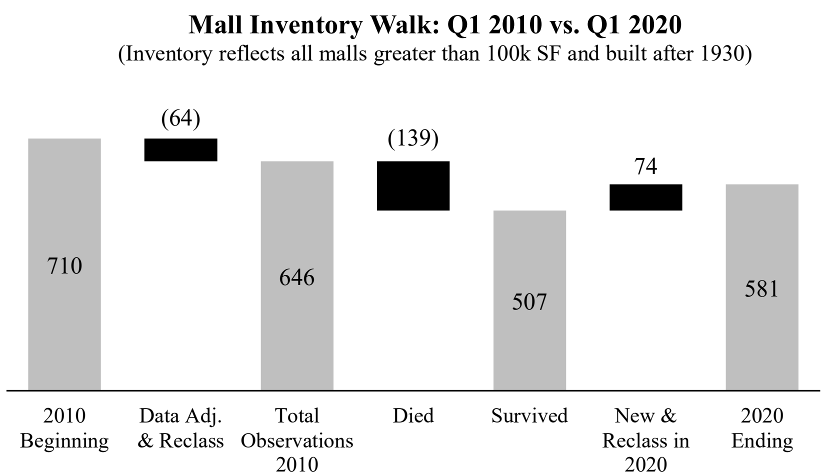
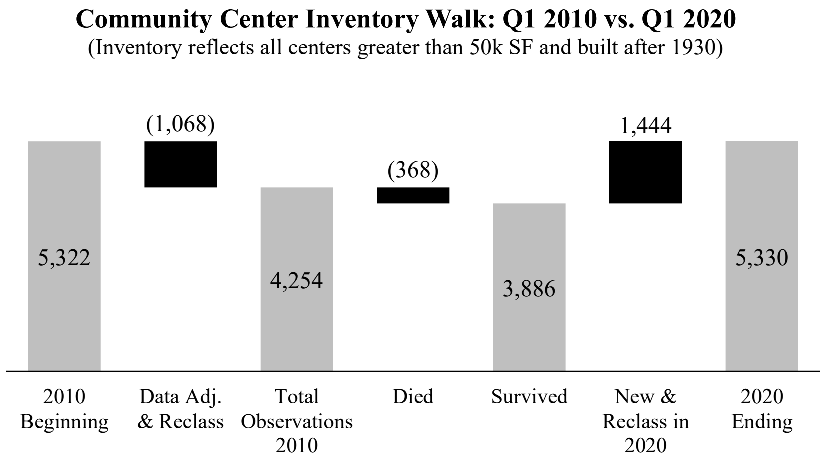


Figure 2-2 is the cumulate inventory walk for community centers from 2010 to 2020. Once we remove all the reclassifications and refine the community center parameters (1,068 properties), we have 4,242 observations in the study, consisting of 3,886 survived centers (91%) and 368 dead centers (9%). The magnitude of the data clean-up is quite considerable, in approximately 69% of the cases, these properties are under 50,000 square feet. 27% of the clean-up have missing dates for when the centers were built, and the rest were centers built prior to 1930.

[21] Data Adj. represents data clean-up including malls that were built prior to 1930's, have erroneous inputs, and contain NRA above 100,000 SF; Reclass stands for reclassification from super regional or regional malls to a different category between the 2010 survey and the 2020 survey, or strip centers, neighborhood centers, or community centers that were upgraded or repositioned to a mall in the 2020 survey.

Figure 2-2: Cumulative Inventory Walk for Community Centers[22]



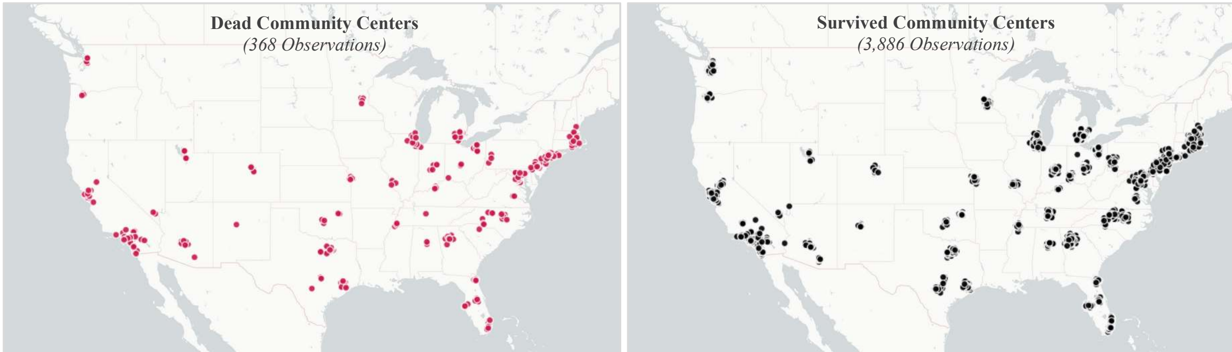
We observe that in the past decade, centers expanded, repositioned, and contracted. The takeaway is that the close rate for malls is higher than the close rate for community centers – more than double the close rate for community centers.

Figures 2-3 and 2-4 display a side-by-side geography of both dead and survived centers in the United States. The data represent the observations that were cleaned up, consisting of 646 malls and 4,254 community centers. For both types of centers, the geography of our sample reinforces the statistics in Table 2-1 that centers are clustered across the major metropolitan areas.

Figure 2-3: Spatial Distribution of Dead vs. Survived Malls in the U.S.[23]

[22] Data Adj. represents data clean-up including community centers that were built prior to 1930's, have erroneous inputs, and contain NRA over 50,000 SF; Reclass stands for reclassification from community centers to a different category between the 2010 survey and the 2020 survey, or strip centers, neighborhood centers, regional malls, and super regional malls that were re-categorized or repositioned to a community center in the 2020 survey.
[23] Hawaii, which contain both malls and community centers, is not shown in Figures 2-3 and 2-4

Figure 2-4: Spatial Distribution of Dead vs. Survived Community Centers in the U.S.[24]



The important definitions for each center in our study is outlined in Table 2-2, which defines the characteristics of regional malls, super-regional malls, and community centers. The classification is defined by the CoStar Group, where the bulk of the data comes from. There are a few distinctions between regional mall and super regional mall classification, but for the purposes of the study, we identify both centers in one category.

Table 2-2: Co-Star Classification of Malls and Community Centers[25]

Center	**Definition**
Regional Mall	Provides shopping goods, general merchandise, apparel, furniture, and home furnishings in full depth and variety. It is built around the full-line department store with a minimum GLA[26] of 100,000 SF, as the major drawing power. For even greater comparative shopping, two, three, or more department stores may be included. In theory a regional center has a GLA of 400,000 SF and may range from 300,000 to more than 1,000,000 SF. Regional centers in excess of 750,000 square feet GLA with three or more department stores are considered super regional.
Super Regional Mall	Similar to a regional mall, but because of its larger size, a super regional mall has more anchors, a deeper selection of merchandise, and draws from a larger population base. As with regional malls, the typical configuration is as an enclosed mall, frequently with multiple levels.
Community Centers	Typically offers a wider range of apparel and other soft goods than neighborhood centers. Among the more common anchors are supermarkets, super drugstores, and discount department stores. Community center tenants sometimes contain value-oriented big-box category dominant retailers. The center is usually configured in a straight line as a strip or may be laid out in an "L" or "U" shape, depending on the site and design. Of all the center types, community centers encompass the widest range of formats. For example, certain centers that are anchored by a large discount department store often have a discount focus. Others with a high percentage of square footage allocated to off-price retailers can be termed as off-price centers. The size of such a center ranges from 100,000 to 350,000 square feet.

[24] Ibid.
[25] *CoStar Glossary*. https://www.costar.com/about/costar-glossary#go_r.
[26] GLA stands for Gross Leasable Area

2.2 Key Variables

The full set of data comes from CBRE with information including the property's ID, name, address, center classification, and longitude and latitude coordinates by pairing centers in 2010 and 2020 at the same location. We are interested in deconstructing the data by center classification to study the differences between properties and come up with a probabilistic prediction to identify which properties are at risk of underperforming. Table 2-3 identifies the key variables collected in this study, consisting of the year when the center was built, the year when the center was renovated, age, renovation dummy, availability rate, net rentable area (SF), and competition within a 2.5-mile and 5.0-mile radius.

Table 2-3: Index for Variables

Label	Description	Unit	Definition
2010_Year_Built	Year Built	Year	The year when the center was built
2010_Year_Renovated	Year Renovated	Year	The year when the center was renovated
Mall_Age	Mall Age	#	Mall age determined by the difference between 2020 and year built
Community_Center_ Age	Community Center Age	#	Community center age determined by the difference between 2020 and year built
Reno_Dummy	Renovation Dummy	1 = Renovated 0 = Not Renovated	Dummy variable indicating if the center has been renovated
2010_AvailabilityRate	Availability Rate	%	Availability rate defined as the ratio of available space to total rentable space[27]
Dead_Mall _Dummy	Dead Mall Dummy	1 = Died 0 = Survived	Dummy variable indicating if the mall has "died"
Dead_Community_ Center_Dummy	Community Center Dummy	1 = Died 0 = Survived	Dummy variable indicating if the community center has "died"
NRA_2010	Net Rentable Area (SF)	SF (in millions)	The net rentable area (square footage) of a center
Competition_ 2.5_Miles_NRA	Competition NRA within 2.5 mile	SF (in millions)	The total net rentable area of the all competition within 2.5 miles radius
Competition _ 5_Miles_NRA	Competition NRA within 5.0 miles	SF (in millions)	The total net rentable area of the all competition within 5.0 miles radius

It is worth mentioning that instead of calculating the number of centers within a radius, we calculated the amount of net rentable area (NRA), a better representation of competition.

2.3 Calculation of Competition

Next, we add a special section to discuss the calculation of the center's competition for both survived and dead centers. As mentioned, the data from CBRE contains longitude and latitude coordinates for every

[27] According to CoStar, availability rate is the percent of space available on the last day of each quarter or the current date in the case of the current quarter: total available SF divided by the total rentable building area (RBA) on the last day of each quarter.

center in our study. In each of the 4,900 centers, we estimate the distance of a center to the other 4,899 properties and ignore any geographical boundaries. This is important because if two malls are near each other and separated by state lines, we ignore the separation and estimate the straight distance.
How did we calculate the distance for thousands of observations? The calculation between two coordinate points is nuanced as they are on a sphere due to the curvature of the Earth. Since we cannot use flat-grid calculations, we employ a trigonometric equation known as the Haversine Formula to get an estimate of a "straight" distance – not the driving or route distance. It is given by[28]:

$$d = 2r\sin^{-1}\left(\sqrt{\sin^2\left(\frac{\phi_2-\phi_1}{2}\right) + \cos(\phi_1)\ \cos(\phi_2)\ \sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right)}\right) \tag{2.1}$$

Where:
r = the earth's radius
d = distance
ϕ_1 and ϕ_2 = latitude of the two coordinate points
λ_1 and λ_2 = longitude of the two coordinate points

We make small adjustments to the trigonometric formula to run our formula in Excel. Since $\sin^{-1}(x) = \arcsin(x)$, we can swap out the function to use ASIN(), the inverse of sine. We also use other functions COS() and SIN() to perform our calculations. Within the equation, we use RADIANS() function to convert longitude and latitude into a decimal value in radians, where in Excel automatically converts the value of a number from degrees to radians. One radian equals π divided by 180 degrees[29]. Finally, "r" represents the earth's radius, or 3,959 miles, or half of the Earth's diameter[30]. With the coordinates converted to radians, we use the final trigonometric to estimate distance in Excel:

$$d = r * 2\mathrm{asin}\left(\sqrt{\sin^2\left(\frac{\phi_2-\phi_1}{2}\right) + \cos(\phi_1)\ \cos(\phi_2)\ \sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right)}\right) \tag{2.2}$$

Note that in the second equation the latitude and longitude in the formula are expressed in radians. Since the earth is not a perfect sphere, the calculated distances are an approximation. After we determine the radius of the competition for each observation of both centers, we add up the total NRA of all competition within the radius, using an Excel logic function.

[28] Morgan, Andrew, et al. *Mastering Spark for Data Science*. Birmingham, UK, 2017.
[29] An entire circle (360 degrees) is $2\pi r$, and half of it is πr (180 degrees); rearranging results in $r = \pi / 180°$
[30] One can substitute 3,959 miles or 6,371 km to return the estimated distance in miles or kilometers

2.4 Descriptive Statistics

In this section, we discuss the important descriptive statistics of the data set to form our hypothesis by dividing up the data into malls and community centers.

Beginning with malls, we have 646 observations, consisting of 507 survived malls (78%) and 139 dead malls (22%). The facts in Table 2-4 make intuitive sense for the malls that survived. On average, survived malls tend to be larger (1.1 million square feet), have lower availability rates (6.1%), and face less competition for both 2.5- and 5.0-mile radii (0.2m SF / 0.7m SF). The survived malls are older (average age of 43 years old), likely to have better locations, and are renovated more recently than the malls that underperformed. Noteworthy, the dead malls are, on average, 445,000 square feet (NRA) smaller than the malls that survived and face, on average, 348,000 square feet (NRA) greater competition than the malls that survived within a 5-mile radius.

Table 2-4: Descriptive Statistics of Data Set for Malls

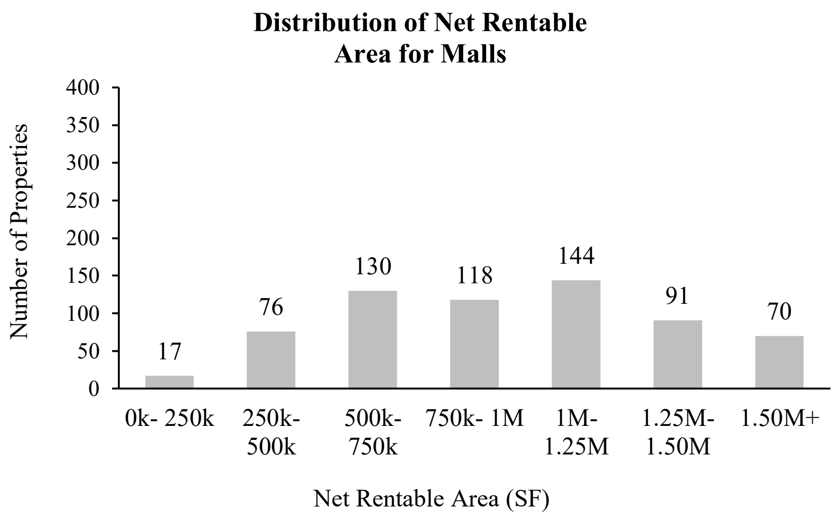
Mall - All Properties	N	Mean	Median	Standard Deviation	Min	Max	10th Percentile	90th Percentile
Age, Years	646	42	45	15	11	78	20	61
Age Since Renovation, Years	646	21	20	6	11	55	14	30
Availability Rate, %	646	7.3%	3.2%	11.3%	0.0%	97.2%	0.0%	19.4%
NRA, SF	646	980,580	953,026	461,857	107,000	2,951,995	430,406	1,529,526
Competition NRA within 2.5 miles	646	262,628	0	566,612	0	3,873,066	0	2,259,774
Competition NRA within 5.0 miles	646	818,552	361,226	1,062,550	0	5,227,177	0	2,259,774

Survived	N	Mean	Median	Standard Deviation	Min	Max	10th Percentile	90th Percentile
Age, Years	507	43	46	14	11	78	22	60
Age Since Renovation, Years	507	21	20	6	11	55	14	29
Availability Rate, %	507	6.1%	2.6%	9.3%	0.0%	77.2%	0.0%	17.4%
NRA, SF	507	1,076,235	1,077,241	443,304	110,000	2,951,995	519,599	1,572,861
Competition NRA within 2.5 miles	507	231,216	0	566,612	0	3,873,066	0	902,194
Competition NRA within 5.0 miles	507	743,656	127,053	1,030,879	0	5,227,177	0	2,126,785

Dead	N	Mean	Median	Standard Deviation	Min	Max	10th Percentile	90th Percentile
Age, Years	139	37	35	18	11	78	14	62
Age Since Renovation, Years	139	23	22	6	11	36	15	33
Availability Rate, %	139	11.5%	6.7%	15.9%	0.0%	97.2%	0.6%	26.5%
NRA, SF	139	631,682	549,393	346,705	107,000	1,927,193	256,318	1,081,133
Competition NRA within 2.5 miles	139	377,205	0	566,612	0	2,370,000	0	1,350,646
Competition NRA within 5.0 miles	139	1,091,733	977,275	1,133,367	0	4,165,617	0	2,772,908

Figure 2-5 shows the distribution of mall size, expressed in net rentable area (NRA), within the data set to provide a scale of these properties. Nearly 60% of the 646 observations are within the 500,000 to 1,000,000 square feet range, which is consistent with CBRE's definition for regional and super regional malls and International Council of Shopping Center's (ICSC)[31] classification for both categories. ICSC assumes 400,000-800,000 square feet for regional mall and 800,000+ for super regional mall (Appendix B). The distribution is bell-shaped with the mean of the observations at 980,580 square feet.

Figure 2-5: Distribution of NRA for Malls



In this study, community centers have 4,254 observations. 91% of the sample, or 3,886 properties, survived between 2010 and 2020, whereas 368 properties closed, or 9% of the observations. Clearly, the close rate is lower for community centers than malls. In Table 2-5, one notable difference from malls is that survived community centers are, on average, smaller than the centers that died (202,083 square feet). This shows that size plays an important role in both property types. This may be partially driven by the challenge of leasing up a large community center, including community centers with different layouts and tenant mix.

From a competition standpoint, the survived centers on average had less competition than dead centers within a 2.5-mile radius but no effect within a 5.0-mile radius; the 2.5-mile radius seems plausible as

[31] ICSC caveats with a disclaimer in its classification report (Appendix B): while every effort is made to ensure the accuracy and reliability of the information contained in the report, ICSC does not guarantee and is not responsible for the accuracy, completeness, or reliability of the information contained in the report.

community centers draw customers from closer distances than malls. Finally, survived community centers are also older (average age of 37 years old) and likely to have better locations.

Table 2-5: Descriptive Statistics of Data Set for Community Centers

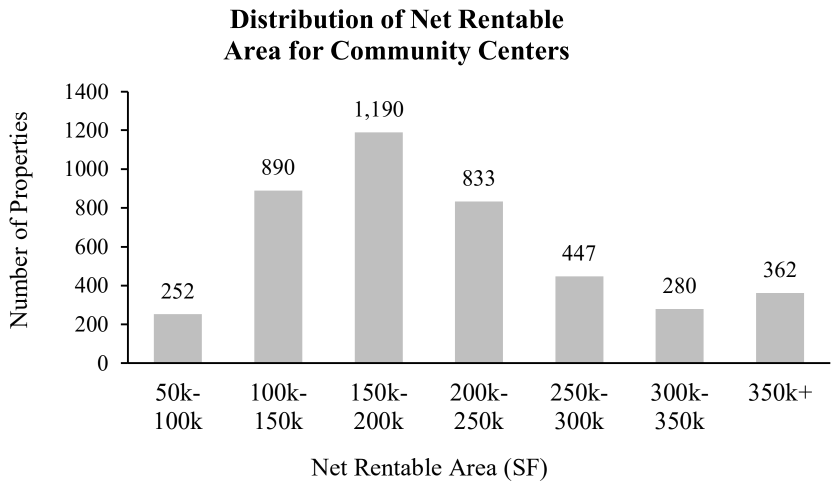
Community Centers - All Properties								
	N	Mean	Median	Standard Deviation	Min	Max	10th Percentile	90th Percentile
Age, Years	4,254	36	34	15	10	88	17	59
Age Since Renovation, Years	4,254	24	23	7	10	55	15	33
Availability Rate, %	4,254	13.3%	7.1%	16.9%	0.0%	100.0%	0.0%	35.0%
NRA, SF	4,254	211,492	189,905	100,180	50,000	1,056,294	111,414	334,274
Competition NRA within 2.5 miles	4,254	581,121	499,344	470,267	0	2,679,390	366,681	3,137,515
Competition NRA within 5.0 miles	4,254	1,655,291	1,501,261	1,066,113	0	5,679,425	366,681	3,137,515

Survived								
	N	Mean	Median	Standard Deviation	Min	Max	10th Percentile	90th Percentile
Age, Years	3,886	37	35	15	10	85	17	59
Age Since Renovation, Years	3,886	24	23	7	10	55	15	33
Availability Rate, %	3,886	13.2%	7.1%	16.9%	0.0%	100.0%	0.0%	34.6%
NRA, SF	3,886	202,083	187,009	100,180	50,000	716,109	112,519	310,557
Competition NRA within 2.5 miles	3,886	573,726	0	470,267	0	3,873,066	0	902,194
Competition NRA within 5.0 miles	3,886	1,655,172	1,495,001	1,066,113	0	5,679,425	361,729	3,137,289

Dead								
	N	Mean	Median	Standard Deviation	Min	Max	10th Percentile	90th Percentile
Age, Years	368	34	31	15	11	88	14	60
Age Since Renovation, Years	368	23	22	7	10	50	15	32
Availability Rate, %	368	13.7%	6.6%	16.9%	0.0%	100.0%	0.0%	38.6%
NRA, SF	368	310,854	280,026	100,180	50,000	1,056,294	93,604	550,827
Competition NRA within 2.5 miles	368	659,211	0	470,267	0	2,370,000	0	1,350,646
Competition NRA within 5.0 miles	368	1,656,542	1,543,122	1,066,113	0	4,722,567	393,278	3,131,356

Figure 2-6 highlights the distribution of community center's NRA within the data set. The shape of the curve has a slight, positive skew where approximately 68% of the 4,254 observations are between 100,000 and 250,000 square feet. According to CBRE, community center ranges between 100,000 and 350,000 square feet, and ICSC's typical range for community centers is between 125,000 and 400,000 square feet. The data are aligned with both classifications. About 90% of the 350,000+ SF range are centers that are between 350,000 and 550,000 SF, adjacent to the lower boundary of mall square footages.

Figure 2-6: Distribution of NRA for Community Centers



2.5 Descriptive Statistics by Geography

The data set also features centers across the United States, and we disaggregate the data by state in Table 2-6. The malls that survived and died across the top 15 states represent 77% of the observations (496 observations), and the top three states California, Texas and Florida contribute one-third of the observations. This is reasonable given the size and population of these markets. As a proxy, Texas, California, and Florida are the top three states for Walmart's total stores in the United States, accounting for a quarter of its physical footprint[32]. Note that the descriptive statistics by state will not tie to Table 2-1 (CBRE Inventory by CBSA) since core based statistical areas may cover more than one state.

Next, for each state, we divided the centers by survived and dead and compared the mean for each variable. The statistical facts for malls in Table 2-6 show similar themes as the facts in Table 2-4. The NRA for all survived malls in all states is bigger, with a couple states featuring survived centers that are almost twice the NRA than a closed center. For closed centers, the competition within a 2.5-mile and 5.0-mile radius is greater, as compared to survived centers, for nearly all states. In some states like Texas, Pennsylvania, Illinois, North Carolina, and Michigan, the competition for dead centers within the 5-mile radius is nearly three times greater than survived ones. In all other top 15 states except for NJ, MD, AZ, and VA, competition is at least 50% greater for dead centers vs. operating centers. For other variables, malls that survived tend to be older and have higher occupancy rates (lower availability rates) across most states.

[32] *Walmart Inc. 2020 Annual Report*. 23 Apr. 2020; Total Wal-Mart stores in the U.S. is 5,355, and Texas, California, and Florida each have following number of stores, respectively: 603, 320, 386.

Table 2-6: Descriptive Statistics of Data Set by Geography – Mall

Mall																		
					Mean													
					Age		Age Since Renovation		Avail. Rate (%)		NRA (SF, M)		Comp. NRA 2.5 Miles (SF, M)		Comp. NRA 5.0 Miles (SF, M)	
Location	N	Survived	Dead	Death Contrib.	Survived	Dead	Survived	Dead	Survived	Dead	Survived	Dead	Survived	Dead	Survived	Dead
CA	106	93	13	2%	47	40	20	22	5.7%	6.8%	1.0	0.7	0.3	0.4	0.9	1.7
TX	50	36	14	2%	40	33	20	22	6.9%	14.7%	1.3	0.7	0.0	0.3	0.5	1.4
FL	49	41	8	1%	39	32	21	18	4.8%	11.4%	1.1	0.9	0.1	0.5	1.0	1.5
PA	30	22	8	1%	44	41	24	24	8.4%	11.3%	1.2	0.5	0.2	0.6	0.4	1.3
NJ	28	22	6	1%	48	40	20	25	2.9%	6.0%	1.2	0.8	0.3	0.4	0.9	0.4
OH	28	20	8	1%	45	40	22	31	8.0%	34.7%	1.1	0.6	0.1	0.4	0.6	0.8
NY	26	20	6	1%	41	59	23	22	6.2%	2.4%	1.1	0.6	0.5	0.7	1.0	1.6
MD	25	22	3	0%	48	41	20	17	7.8%	12.4%	1.0	0.8	0.0	0.2	0.5	0.3
IL	24	23	1	0%	47	20	22	-	2.7%	0.5%	1.2	0.6	0.1	2.3	0.7	3.6
MA	24	19	5	1%	44	37	22	32	4.0%	5.1%	0.9	0.5	0.1	0.1	0.3	0.5
AZ	24	18	6	1%	40	20	20	21	11.4%	7.9%	1.1	0.8	0.1	0.4	0.8	1.1
GA	21	17	4	1%	41	25	23	21	6.2%	7.1%	1.1	0.9	0.2	1.0	0.6	1.4
NC	21	16	5	1%	41	39	24	21	6.9%	7.3%	1.0	0.6	0.2	0.2	0.4	1.7
MI	20	14	6	1%	48	39	23	23	4.5%	12.7%	1.3	0.6	0.3	0.4	0.4	1.4
VA	20	16	4	1%	38	57	20	28	2.7%	5.8%	1.2	0.5	0.2	0.3	0.9	1.1
All Other	150	108	42	7%	41	35	20	22	7.0%	12.5%	1.0	0.6	0.3	0.2	0.8	0.7
Total	**646**	**507**	**139**	**22%**	**43**	**37**	**21**	**23**	**6.1%**	**11.5%**	**1.1**	**0.6**	**0.2**	**0.4**	**0.7**	**1.1**

Table 2-7 features descriptive statistics for community centers that survived and died in the top 15 states, representing 78% of the observations (3,329 observations). The top three states California, Texas and Florida also represent one-third of the sample. There are three main differences between malls and community centers. First, survived community centers are, on average, smaller than the centers that died. Second, average availability rates across states is mixed. Third, competition within a 5-mile radius between both types of centers are muted, but within the 2.5-mile radius, the results are significant: all states except for AZ and MA show dead centers have at least 30% more competition vs. survived centers.

Table 2-7: Descriptive Statistics of Data Set by Geography – Community Centers

Community Center																
					Mean											
					Age		Age Since Renovation		Avail. Rate (%)		NRA (SF, M)		Comp. NRA 2.5 Miles (SF, M)		Comp. NRA 5.0 Miles (SF, M)	
Location	N	Survived	Dead	Death Contrib.	Survived	Dead	Survived	Dead	Survived	Dead	Survived	Dead	Survived	Dead	Survived	Dead
CA	671	610	61	9%	38	37	24	25	9.7%	12.7%	0.2	0.3	0.7	0.8	1.8	1.8
TX	409	365	44	7%	33	33	24	23	16.5%	13.0%	0.2	0.3	0.6	0.7	1.9	1.9
FL	374	355	19	3%	36	33	23	23	13.3%	5.0%	0.2	0.3	0.6	0.6	2.1	2.1
PA	156	147	9	1%	41	37	24	22	13.3%	6.9%	0.2	0.3	0.4	0.8	1.3	1.5
NJ	160	143	17	3%	39	34	23	21	12.1%	11.7%	0.2	0.3	0.3	0.6	1.1	1.5
OH	180	169	11	2%	39	40	24	22	16.9%	24.5%	0.2	0.4	0.5	0.7	1.4	1.9
NY	119	105	14	2%	43	37	24	22	10.8%	6.7%	0.2	0.3	0.4	0.3	1.4	1.6
MD	143	130	13	2%	42	34	25	24	11.8%	7.9%	0.2	0.4	0.6	0.6	1.4	1.3
IL	271	250	21	3%	35	30	24	20	14.4%	19.7%	0.2	0.3	0.7	0.7	2.2	1.9
MA	132	121	11	2%	43	45	24	21	11.4%	6.8%	0.2	0.2	0.3	0.5	0.8	0.9
AZ	150	139	11	2%	31	22	24	12	11.8%	12.9%	0.2	0.2	0.7	0.8	2.2	1.8
GA	195	185	10	2%	34	35	24	27	17.2%	11.5%	0.2	0.4	0.6	0.7	1.6	1.9
NC	156	136	20	3%	34	29	21	20	14.9%	9.6%	0.2	0.4	0.5	0.6	1.3	1.5
MI	99	84	15	2%	39	33	22	28	15.2%	26.9%	0.2	0.3	0.4	0.6	1.2	1.5
VA	114	105	9	1%	37	32	26	27	10.7%	14.8%	0.2	0.4	0.5	0.8	1.3	1.7
All Other	925	842	83	13%	36	31	24	24	13.5%	16.8%	0.2	0.3	0.6	0.6	1.5	1.5
Total	**4,254**	**3,886**	**368**	**9%**	**37**	**34**	**24**	**23**	**13.2%**	**13.7%**	**0.2**	**0.3**	**0.6**	**0.7**	**1.7**	**1.7**

3. Research Methodology

This retail empirical study uses three steps to experiment and test the predictive modeling of retail centers that are at risk of defaulting. The first step is to run an Ordinary Least Squares (OLS) regression, the second step is to perform a probit regression, and the final step is to calculate the predicted values for each observation to determine its normal cumulative distribution function.

3.1 Simple Ordinary Least Squares (OLS) Regression

One of the basic and commonly used tools to predict a value of one variable from another variable is a linear regression. Linear regression, also known as an Ordinary Least Squares (OLS) regression, estimates the relationship between a dependent variable (the variable that we are interested in predicting a value) and one or more explanatory, independent variables. Linear regression works to find the best-fitting straight line through a set of data points and comes up with an equation. The OLS regression formula is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \tag{3.1}$$

Where in the equation, β_0 represents a constant, and the other betas are the coefficients, which are the estimates of the actual sample parameter that the OLS regression estimates. Y is the unit measurement of the center's death dummy, and X is vector of independent variables including age, renovation dummy, availability rate, NRA, competition NRA within 2.5 miles, and competition NRA within 5.0 miles. Finally, the error term ε stands for the variation in the dependent variable that the independent variables do not explain.

The overarching idea of performing an OLS regression is to examine two things. First, does a set of predictor variables help predict the death of a mall or community center? Second, which independent variables, determined by the coefficients, are significant predictors of the dependent variable? The coefficients will explain the magnitude and sign direction of the beta estimates.

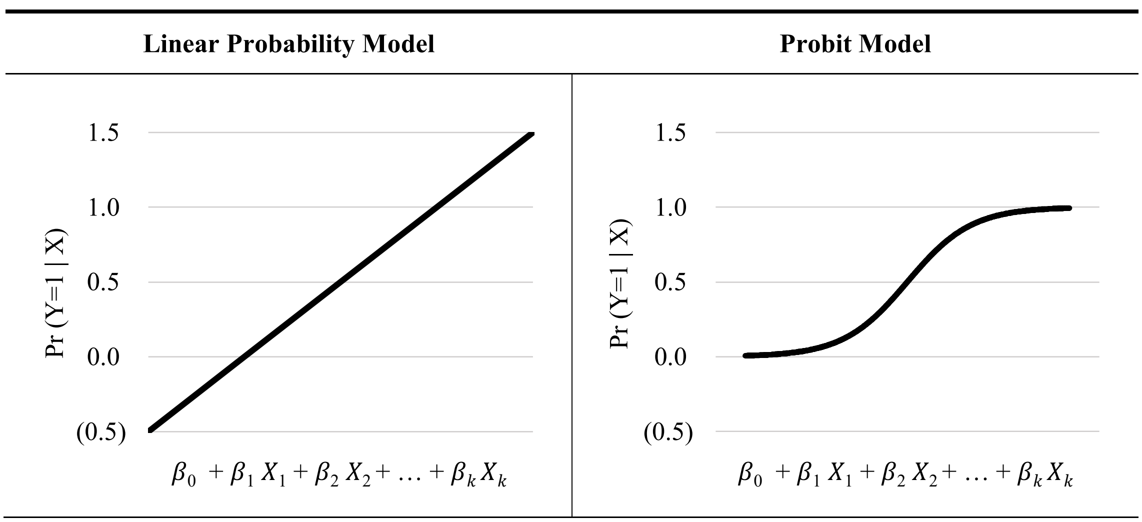
Another important component of the OLS model, which will be discussed in the next chapter, is the statistical significance. The statistical significance reveals that changes in the independent variables correlate with changes in the dependent variable.

3.2 Probit Analysis

A major issue with the ordinary least squares regression is that it uses a binary variable as the dependent variable – either the center died or survived. The assumptions in the linear regression significance test are violated when when there is a dichotomous dependent variable. When a dependent variable is categorical and the explanatory variables are fixed, the linear regression violates the homoskedasticity and normality of errors assumptions, resulting in unreliable significance levels associated with test statistics[33]. One solution is the probit model, or short for “probability unit” – it estimates the probability a value will fall into one of the two binary outcomes. The method was popularized in part due to the work of D.J. Finney (1971)[34].

Technicalities aside, the probit model differs from the linear probability model in that the predicted probability of Y = 1 is never below 0 or above 1. The cumulative distribution function is increasing and continuous, reflected by the sigmoid-shaped curve in Figure 3-1. The vertical axis is a probability, falling between zero and one. The straight line on the left is the linear regression:

Figure 3-1: Shape of Curve - Linear Probability vs. Probit Model[35]



The generalized linear probability model with a single regressor takes the equation when Y is binary:

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \tag{3.2}$$

[33] Noreen, Eric. “An Empirical Comparison of Probit and OLS Regression Hypothesis Tests.” *Journal of Accounting Research*, vol. 26, no. 1, 1988, pp. 119–33. JSTOR.
[34] Finney, D. J. *Probit Analysis*. 3rd Ed, University Press, 1971.
[35] Dustan, Andrew. *Linear Probability Model vs. Logit (or Probit)*. Department of Agricultural & Resource Economics, University of California, Berkeley.

The issue with the OLS model is that it writes the probability of Y=1 as being linear[36]:

$$\Pr(Y=1|X) = \beta_0 + \beta_1 X_1 \tag{3.3}$$

Instead, we want to set the model using the following parameters[37]:

$$0 \leq \Pr(Y=1|X) \leq 1 \text{ for all X, and}$$
$$\Pr(Y=1|X) \text{ to be increasing in X for } \beta_1 > 0$$

The probit model satisfies the above parameters. The regression models the probability that Y=1 and takes the inverse standard normal distribution of the probability:

$$\Phi^{-1}(p) = (\beta_0 + \beta_1 X_1) + \varepsilon \tag{3.4}$$

Where:
Φ = standard normal cumulative distribution function
p = conditional probability
$\beta_0 + \beta_1 X_1 = Z$, the "z-value" of the probit model
ε = error term

Since $p = \Pr(Y=1|X)$, the probit regression model can be re-written as follows:

$$\Pr(Y=1|X) = \Phi(\beta_0 + \beta_1 X_1) + \varepsilon \tag{3.5}$$

A probit regression with multiple regressors will incorporate additional variables. The model is given by:

$$\Pr(Y=1|X) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k) + \varepsilon \tag{3.6}$$

Φ is the normal cumulative distribution function and $Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$ is the "z-value" of the probit. However, β_1 is the effect on the z-score of a unit change in X_1, holding X_2, X_k constant.

We used a statistical software Stata in this study to convert the percent death to probits automatically. The predictor variables of a center are the age, renovation dummy, availability rate, NRA, competition NRA within 2.5 miles for community centers, and competition NRA within 5.0 miles for malls.

[36] *Econ 423: Regression with a Binary Dependent Variable*. Department of Economics, University of Maryland.
[37] Ibid.

The interpretation of the probit result is essential. The constant in the probit results can be interpreted as a predicted z-score, and we can look up the z-score in a table to estimate the probability associated with it. The coefficients of the probit model rarely have direct interpretation as they tell us how much the z-score will increase for each one-unit increase in each of the variables we are regressing. Stated in another way, the probit model tells us how big the marginal effects are of changes in the regressors (x variable) affecting the probability of getting Y = 1, or the death of a center.

3.3 Predictive Modeling

One important aspect of the estimated regression equation is its ability to predict the effects on Y from a change in one or more values of the independent variables X. The goal of this section is to predict the likelihood of a center's death from the other variables. We use the regression equation to make a prediction and use the coefficients from the probit model to define the relationship between each independent variable and the dependent variable.

To calculate a predicted probability for each observation, or the mean value of the dependent variable, we simply take the value of the independent variables from each CBRE observation, multiply by its probit coefficients and add up, including the constant:

$$\widehat{Y} = \hat{\beta}_0 + \hat{\beta}_1(\text{Age}) + \hat{\beta}_2(\text{Reno}) + \hat{\beta}_3(\text{Avail. Rate}) + \hat{\beta}_4(\text{NRA}) + \hat{\beta}_4(\text{Comp\_NRA}) \tag{3.7}$$

We need to see whether there is a significant relationship between the variables. The regression prediction is more accurate if the observed values are closer to the predicted values. The further the spread of the predicted values from the observed values, the less the predictions can provide significant information. The coefficient of determination (R-Squared) is a good measure of prediction and always lies between 0 and 1.

After we determine the predicted z-score value for each observation, we determine the standard normal cumulative distribution of it with a mean of 0 and standard deviation of 1. In a standard cumulative distribution function (CDF), the Y-axis of the probit model in Figure 3-1 is between 0 and 1. The interval is a probability value and represents the probability of a value from our normal distribution being less than or equal to a given value. We can use the normal CDF to determine which centers will have a higher average value relative to surviving ones.

4. Results

In this section, we will go over the results of the OLS regression, probit model, and the predicted values for estimating the cumulative distribution function.

4.1 Simple OLS on Dead Center Dummy

To recap, the simple OLS regression model from Chapter 3 assumes:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \tag{3.1}$$

Where β_0 represents is the constant, and X is vector of independent variables including age, renovation dummy, availability rate, NRA, competition NRA within 2.5 miles, and competition NRA within 5.0 miles. The regression results are summarized in 4-1 for radii 2.5 and 5.0 miles, divided up by malls and centers. Tables 4-2 and 4-3 are the full results of the OLS regression divided up into two different radius measurements.

Starting with the age coefficient in Table 4-1 for malls and community centers with competition in 2.5- and 5.0-mile radii. When we look at the age result, it has no effect according to the t-statistic. One way to interpret this is, it is likely that older malls are more deteriorated and more likely to close, but they also happen to be in better locations at the time when they were built. Perhaps at the time, the malls were more compact since they were the first players in the market. Seeing that they are likely in better locations, older malls will be able to hang on longer. Thus, the results here may indicate there is an offsetting effect between these two notions.

Table 4-1: Summary of OLS Regression Results

	2.5 Mile Competition Radius			5.0 Mile Competition Radius		
	Coefficients:		Statistical Significance:	Coefficients:		Statistical Significance:
	Mall	Community		Mall	Community	
Age	0.000	0.000	Insignificant: Both	0.000	0.000	Insignificant: Both
Renovation Dummy	-0.219	-0.035	Significant: Both	-0.219	-0.035	Significant: Both
Availability Rate	0.406	0.055	Significant: Both	0.410	0.058	Significant: Both
NRA	-0.287	0.868	Significant: Both	-0.282	0.872	Significant: Both
Competition NRA	0.038	0.024	Insignificant: Mall Significant: Community	0.039	0.000	Significant: Mall Insignificant: Community

Next is the renovation dummy. We find the betas are negative and significant in all instances which indicate that renovation decreases the center's likelihood to close, as remodeled centers may improve store offerings and shopper's experiences. The availability coefficients are positive and significant across all variables in both regressions. Any unit increase in available space could make it more challenging for a center to lease up vacant space and compete.

We are surprised by the NRA coefficients between centers and malls. The NRA, which has significant effect according to the t-statistic, is largely negative for malls but largely positive for community centers. This is an important finding since the opposite coefficients can be explained by the number of properties in the United States and by type of ownership. The mall results indicate that for every 1 million square feet increase, the mall is less likely to be distressed. Malls are typically owned by institutional firms and REITs such as Simon Property Group and Brookfield. Owners tend to have deeper pockets, so they can hold on to the centers longer and make them "too big to fail."

For community centers, it is the opposite, as they tend to be much more prolific since there are thousands of properties as opposed to hundreds of them. Because of greater inventory, it makes sense for these centers to be smaller. The coefficients are positive, so for every 1 million square feet increase, the centers face higher risk of default. Larger community centers are essentially properties that are almost competing with malls, so these centers need to work harder to be fully occupied. The results suggest that community centers may have been overbuilt, so the supply of community centers is likely to have a negative effect. They can't be too big because if they are too big, then it more difficult to lease up since they have a limited market area and can only draw customers from a certain range of distances.

Another important result is the spatial competition. The significance test reveals that within a 5-mile radius, mall spatial competition matters but it does not matter for community centers. For malls that have competition in a 5.0-mile radius, it makes sense that competition is deadlier since malls are located further away from one another. In a 2.5-mile radius, spatial competition for malls has no effect according to the t-statistic, but it matters for community centers since these centers are more ubiquitous and located more closely together. In either instance, where there is significance, any unit increase in competition will drive up the likelihood of property underperformance.

Table 4-2: OLS Regression Results for Malls

Malls with 2.5 Mile Competition Radius

Regression Statistics	
Multiple R	0.47
R Square	0.23
Adjusted R Square	0.22
Standard Error	0.36
Observations	646

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.620	0.054	11.503	0.000	0.514	0.725	0.514	0.725
Mall_Age	0.000	0.001	-0.004	0.997	-0.002	0.002	-0.002	0.002
Reno_Dummy	-0.219	0.042	-5.191	0.000	-0.301	-0.136	-0.301	-0.136
2010_AvailabilityRate	0.406	0.130	3.124	0.002	0.151	0.661	0.151	0.661
NRA_2010	-0.287	0.032	-8.846	0.000	-0.351	-0.224	-0.351	-0.224
Competition_Mall_2.5_Miles_NRA	0.038	0.026	1.491	0.137	-0.012	0.088	-0.012	0.088

Malls with 5.0 Mile Competition Radius

Regression Statistics	
Multiple R	0.48
R Square	0.23
Adjusted R Square	0.23
Standard Error	0.36
Observations	646

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.601	0.054	11.135	0.000	0.495	0.707	0.495	0.707
Mall_Age	0.000	0.001	-0.168	0.867	-0.003	0.002	-0.003	0.002
Reno_Dummy	-0.219	0.042	-5.236	0.000	-0.302	-0.137	-0.302	-0.137
2010_AvailabilityRate	0.410	0.129	3.168	0.002	0.156	0.664	0.156	0.664
NRA_2010	-0.282	0.032	-8.726	0.000	-0.345	-0.218	-0.345	-0.218
Competition_Mall_5_Miles_NRA	0.039	0.014	2.868	0.004	0.012	0.066	0.012	0.066

Table 4-3: OLS Regression Results for Community Centers

Community Centers with 2.5 Mile Competition Radius

Regression Statistics	
Multiple R	0.32
R Square	0.10
Adjusted R Square	0.10
Standard Error	0.27
Observations	4,254

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.091	0.015	-5.990	0.000	-0.121	-0.062	-0.121	-0.062
Community_Center_Age	0.000	0.000	-1.264	0.206	-0.001	0.000	-0.001	0.000
Reno_Dummy	-0.035	0.010	-3.534	0.000	-0.055	-0.016	-0.055	-0.016
2010_AvailabilityRate	0.055	0.024	2.239	0.025	0.007	0.102	0.007	0.102
NRA_2010	0.868	0.041	21.104	0.000	0.787	0.948	0.787	0.948
Competition_Community_Center_2.5_Miles_NRA	0.024	0.009	2.802	0.005	0.007	0.042	0.007	0.042

Community Centers with 5.0 Mile Competition Radius

Regression Statistics	
Multiple R	0.32
R Square	0.10
Adjusted R Square	0.10
Standard Error	0.27
Observations	4,254

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.078	0.016	-4.969	0.000	-0.109	-0.047	-0.109	-0.047
Community_Center_Age	0.000	0.000	-1.286	0.199	-0.001	0.000	-0.001	0.000
Reno_Dummy	-0.035	0.010	-3.499	0.000	-0.054	-0.015	-0.054	-0.015
2010_AvailabilityRate	0.058	0.024	2.393	0.017	0.011	0.106	0.011	0.106
NRA_2010	0.872	0.041	21.188	0.000	0.791	0.952	0.791	0.952
Competition_Community_Center_5_Miles_NRA	0.000	0.004	-0.070	0.944	-0.008	0.007	-0.008	0.007

Moving forward, as the results indicated, we will simplify the results to run the probit using 2.5-mile radius for community centers and a 5.0-mile radius for malls since the t-statistic for spatial competition for the respective radii are significant. Next, we will divide up the sample into survivors and dead to examine their differences.

4.2 Probit Model on Dead Center Dummy

To recap, we use the non-linear regression probit model since we have an indicator variable with two outcomes, dead vs. survived centers. The probit model shows us where the results fall along the s-curve, but the coefficients cannot be interpreted as partial effects. If we have a low probability, then we have a low marginal effect, whereas, a high probability has a larger marginal effect. The marginal effects of the regressors show how much the probability of the outcome variable changes when we adjust the value of a regressor, holding equal all other variables.

The probit results for community centers and malls in a 2.5-mile and 5.0-mile competition radius are summarized in Table 4-4. The signs of the coefficients are the same for both OLS and probit regressions. However, the magnitude of the coefficients is amplified. For instance, the beta of a mall's NRA turns out to be five times the size of the linear probability model coefficient, which would indicate that the probit model has placed more weight on that factor than the linear model. In Table 4-5 showing the mall's full probit results, the t-statistics suggest that NRA is greater in the probit model that – once in a cumulative distribution function curve – the NRA has a higher effect in the probit model than the linear model. For all other variables, the coefficient is about three times larger than the linear model.

Table 4-4: Probit Regression Results Summary

	Coefficients		**Statistical Significance**
	Mall	Community	
	5.0 Miles	*2.5 Miles*	
Age	-0.002	-0.002	Insignificant: Both
Renovation Dummy	-0.703	-0.244	Significant: Both
Availability Rate	1.384	0.360	Significant: Both
NRA	-1.489	3.857	Significant: Both
Competition NRA	0.113	0.168	Significant: Both

In Table 4-4, we can see that the mall's NRA coefficient has the greatest magnitude. All else being equal, the average marginal effect of having an additional 1 million square feet makes it less likely for a center to default. If you have a huge amount of NRA in a mall, it is likely not going to close.

Table 4-5: Probit Regression Results for Malls and Community Centers[38]

Probit Summary for Malls (Competition Within 5.0 Mile Radius)

```
Probit regression                               Number of obs     =        646
                                                LR chi2(5)        =     168.69
                                                Prob > chi2       =     0.0000
Log likelihood = -252.04346                     Pseudo R2         =     0.2507

------------------------------------------------------------------------------
    Dead_Dum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Age |  -.0023889   .0052209    -0.46   0.647    -.0126217     .007844
  1.Reno_Dum |  -.7025547   .1752708    -4.01   0.000    -1.046079   -.3590304
       Avail |   1.383974   .4999253     2.77   0.006      .404138    2.363809
         NRA |   -1.48898   .1824284    -8.16   0.000    -1.846533   -1.131427
     Comp_5M |    .113102   .0556839     2.03   0.042     .0039634    .2222405
       _cons |   .8365232   .2464906     3.39   0.001     .3534105    1.319636
```

Probit Summary for Community Centers (Competition Within 2.5 Mile Radius)

```
Probit regression                               Number of obs     =      4,254
                                                LR chi2(5)        =     299.42
                                                Prob > chi2       =     0.0000
Log likelihood = -1102.5851                     Pseudo R2         =     0.1195

------------------------------------------------------------------------------
    Dead_Dum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Age |  -.0019872   .0022109    -0.90   0.369    -.0063205    .0023461
  1.Reno_Dum |  -.2439208   .0736552    -3.31   0.001    -.3882823   -.0995593
       Avail |   .3597405   .1644656     2.19   0.029     .0373939    .6820871
         NRA |   3.857287   .2450712    15.74   0.000     3.376956    4.337618
    Comp_25M |   .1675041   .0596848     2.81   0.005      .050524    .2844843
       _cons |  -2.273236   .1070315   -21.24   0.000    -2.483014   -2.063458
```

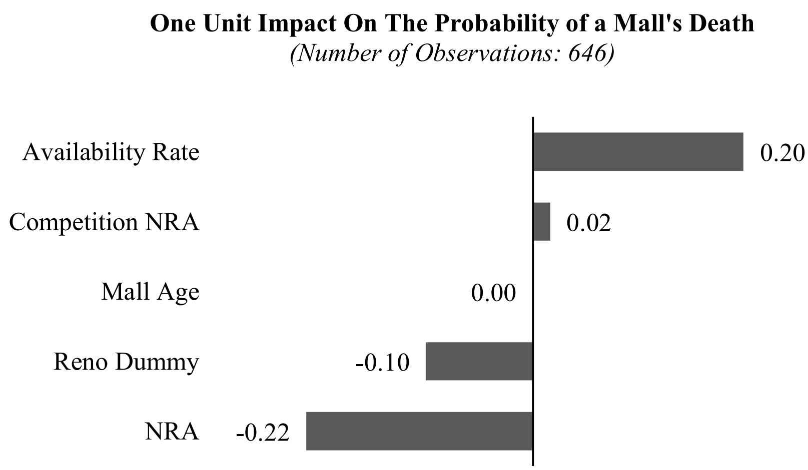
Although we cannot interpret the partial effects in the same way as the OLS model, we estimate a unit impact on the probability of a mall dying. The probit results tell us for every unit change, the z-score increases or decreases. To get the percentage change for per unit, we estimate z-score by taking the mean value for each variable multiplied the variable's coefficients from the probit results, given by:

$$Y = \beta_0 + \beta_{Age}(\mu_{Age}) + \beta_{Reno}(\mu_{Reno}) + \beta_{Avail.}(\mu_{Avail.}) + \beta_{NRA}(\mu_{NRA}) + \beta_{Comp\ NRA}(\mu_{Comp\ NRA}) \qquad (4.1)$$

[38] Regression calculated in Stata; results from the probit model is also in Appendix C-F

Where μ represents the mean for each variable. Based on the mean values, the mall's predicted z-score is -1.05. The standardized normal CDF probability of -1.05, with a mean 0 and standard deviation of 1, is 0.15 which indicates that probability of a mall failing is 15% based on the mean values. We take the standard probability of 0.15 multiplied by each coefficient to estimate the percentage change in the probability of a mall's death for each variable. Figure 4-1 summarizes the impact of a unit change on the default probability for each predictor.

Figure 4-1: Impact of a Unit Change on the Risk Probability – Malls

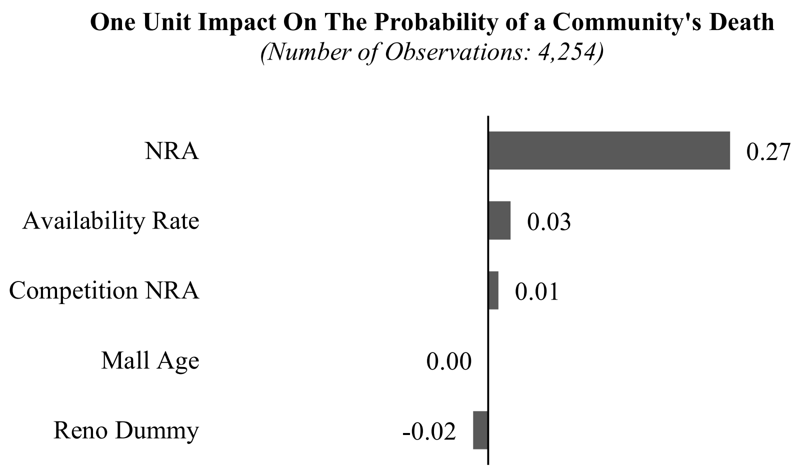


There is an outsized effect of the mall's NRA, which adding 1 million square feet (a unit change) reduces probably for malls dying by 22 percentage points. Note that these unit impacts are not constant – they are only valid for mean values of all the variables.

In the case of renovation, if the mall gets renovated, then the impact on the probability of a mall's dying decreases by 10%. The availability variable tells us for every available square foot, the impact on the probability of malls going under increases by 20 percentage points. What is consistent in our study is that if the mall adds 1 million square feet and increases the availability rate, then these two effects nearly wash each other out. Finally, for every increase of competitive square footage, the impact on the likelihood of underperformance increases by 2%, less significantly than the other variables.

One surprising but an important result from the study is the NRA coefficient's swing from malls to community centers in Table 4-4, from -1.489 to 3.857. The model has placed great weight on that factor where NRA has a huge, negative effect on community centers. Figure 4-2 reveals that adding 1 million square feet to community centers increases the effect on the likelihood of defaulting by 27 percentage points. If the center adds 1 million square feet and increases the unit of the other variables, the marginal effects of NRA will be enough to overshadow the combined effects of other variables.

Figure 4-2: Impact of a Unit Change on the Risk Probability – Community Centers



The renovation variable indicates that a renovation lowers the probability of a mall's dying by 2%. The availability variable tells us for every square foot that becomes available, the impact on the probability increases by 3%. Spatial competition's marginal effects are even lower than availability rate, that every increase in competition's NRA, the impact on the probability of a center being distressed increases by 1 percentage point. The impact of age is muted as there is no statistical significance.

4.3 Results of the Predictive Model

This section will discuss the model for determining predictions from the estimates. In section 3.3, we introduce the methodology for computing the effects for all the predictors using probit results, adding them up, and transforming to a standard normal cumulative distribution function. The normal CDF is the range of the predicted probability.

How did we transform the z-values to the standard normal cumulative distribution? Recall that $\Phi^{-1}(p)$ is the inverse standard normal distribution function which is later rewritten to get our expected values. The inverse function $\Phi(z)$ represents the standard normal cumulative distribution function, where z is the predicted value. In Excel, we use the NORM.S.DIST(z, TRUE) formula for $\Phi(z)$, where TRUE is equal to cumulative distribution function[39].

[39] *Probit Regression | Real Statistics Using Excel*. http://www.real-statistics.com/logistic-regression/probit-regression.

Figure 4-3 shows the characteristics of the normal cumulative distribution function. The horizontal axis is the predicted values, or the z-scores, and the vertical axis is the cumulative probability, between 0 and 1. We see a steeper-sloped sigmoid curve ("s-curve") for community centers, suggesting less risk of the distribution. The s-curve for malls stretches to lower predicted values – a wider curve may indicate the higher the standard deviation or dispersion of risk. When disaggregating the curves into dead and survived centers, the shape of the curve between malls and community centers is noticeable. For dead malls, the model predicts that 42% of the observations are at greater risk of failing, or above 50% CDF, whereas 7% of the dead community centers are modeled above 50%. For survived malls, 95 percent of the observations fall below 50% CDF, and 99.8 percent of the community centers fall below the 50% CDF threshold.

Figure 4-3: Standard Cumulative Distribution Curves for Malls & Community Centers

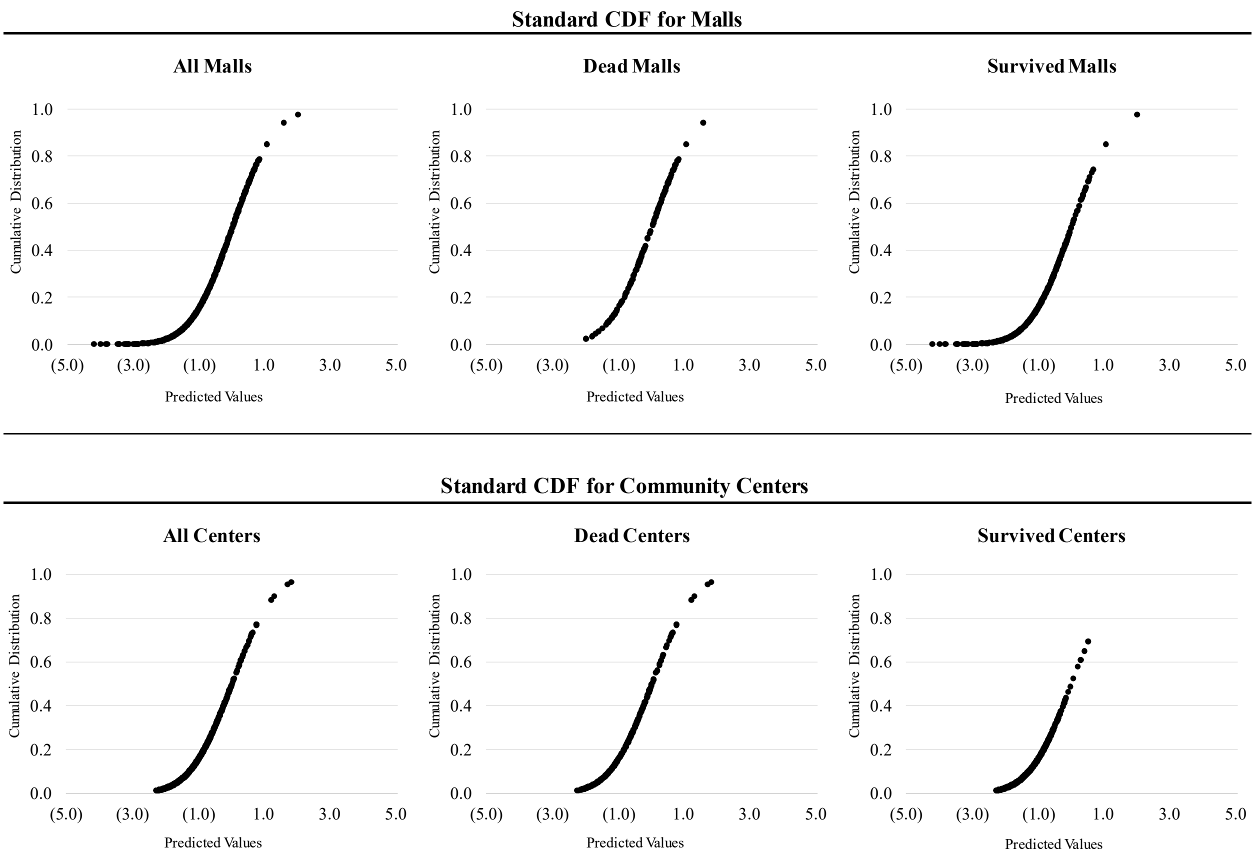


Table 4-6 compares survived malls and dead malls and summarizes the mean values for each predictor variables based on a range. In both results summary, the left column is the normal cumulative distribution function range in tabular format, from 0% to 100%. The column represents the Y-axis of the

CDF graph, and the range from 0% to 100% is deconstructed into 10% increments. The second column is the number of observations in each range, expressed as a percent of total observations. The other five columns on the right are mean values of each predictor variables for all observations in that range. The predictor variables are availability, age, renovation dummy, NRA, and competition NRA.

Table 4-6: CDF Tabular Summary for Malls

Survived Malls - 5.0 Miles
(507 Observations)

Normal CDF	% Observed	Mean				
		Avail. Rate	Age	Reno Dum.	NRA	Comp. NRA
					Million	*Million*
90%+	0%	77%	29	0.0	0.3	5.2
80-90%	0%	54%	30	0.0	0.4	1.2
70-80%	1%	10%	13	0.0	0.5	3.9
60-70%	2%	13%	32	0.1	0.4	1.0
50-60%	2%	14%	29	0.1	0.6	0.5
40-50%	4%	18%	40	0.7	0.5	1.3
30-40%	7%	7%	42	0.7	0.6	1.3
20-30%	11%	7%	42	0.8	0.7	1.0
10-20%	23%	6%	42	0.7	1.0	0.7
0-10%	50%	4%	46	1.0	1.4	0.5
Total	**100%**	**6%**	**43**	**0.8**	**1.1**	**0.7**

Dead Malls - 5.0 Miles
(139 Observations)

Normal CDF	% Observed	Mean				
		Avail. Rate	Age	Reno Dum.	NRA	Comp. NRA
					Million	*Million*
90%+	1%	80%	48	0.0	0.2	0.0
80-90%	1%	45%	45	0.0	0.2	0.0
70-80%	11%	15%	26	0.1	0.3	1.9
60-70%	15%	11%	27	0.0	0.4	0.8
50-60%	14%	11%	29	0.2	0.5	0.7
40-50%	7%	13%	32	0.4	0.6	0.8
30-40%	21%	12%	46	0.9	0.5	1.5
20-30%	10%	8%	47	0.7	0.8	1.1
10-20%	12%	4%	48	0.9	0.9	0.9
0-10%	9%	12%	36	0.8	1.4	0.6
Total	**100%**	**12%**	**37**	**0.5**	**0.6**	**1.1**

The survived malls that that have low predicted values – bottom half of the table from 0% to 50% where nearly 95% of the observations lie – indicate the same story as the regression results. On average, these centers have lower availability rate, are more likely to be renovated, are larger in size, and face lower competition. Note that the mean values for these centers are older, suggesting that the properties are well-located, but regression results determined no statistical significance for age.

We observe that in the dead mall table, the malls that have a higher probability of going under – top half of the table from 50% to 100% where 42% of the observations lie – have opposite effects compared to the survived malls. These observations indicate, on average, higher availability rates, are more likely not to be renovated, and are smaller in NRA. We also observe that the two dead centers at greatest risk are among the oldest, reflecting the hypothesis that older centers may go under because of deterioration. Still, the results indicate that age has no statistical significance.

It is important to note that the results indicate we have a mismatch between a survived mall with a high predicted value, and a mismatch between a dead mall with a low predicted value. Through qualitative research, we determined that a mall's actual status differed from its expected status because of timing of when the mall was is in transition, repositioning, or facing new nearby competition.

- *Malls that survived despite high predicted values*: In Table 4-6, there are four survived malls with a high predictive value. We observed that some of these malls were open but going through transitioning:
    - The mall with the highest predictive value is in Nevada with an NRA of roughly 0.3M square feet. Ranked the sixth smallest in the list of dead centers, the mall has an availability rate of 77%. We uncover that the mall's anchor tenants – Macy's, Sears and JC Penney – are all vacant. The center announced in mid-2019 that it was going through a $30 million repositioning starting in summer of 2020 and completing in 2021.
    - The second mall is based in New York, with approximately 0.4M square feet of NRA and availability rate of 31%. The mall was bought out by a new owner in 2016 and announced a moderate upgrade. In the same year, a nearby mall was closed and replaced by a $275 million development project taking place in the subsequent years that included outdoor dining areas along with more than 70,000 square feet of retail space.
    - The next mall is a 0.3M SF retail center in Kentucky with 17% of the space available. In 2020, the owners announced a redevelopment project that would add community greenspace and bring in a new mix of tenants. The announcement came after Barnes & Noble and several of its primary restaurant tenants closed in 2019.
    - The last one is 0.5M SF mall in New York with 27% availability rate. It went through transitioning in late 2018 and 2019, following the closure of Toys "R" Us and Babies "R" Us. The space was repurposed to a large food retail market, taking up nearly 0.1M SF.
- *Malls that died despite low predicted values*: There are approximately twelve malls at the lowest predicted value range in Table 4-6. We find out that these dead properties are not necessarily

"dead." In many cases, these centers have been renovated or repositioned, yet parts of the property remained opened, including these examples:

- A 1.4M SF center in Florida that is nearly occupied (2% availability rate) went through a $20 million transformation to bring in new retailers and restaurants, and a development of a multifamily rentals. The new retail tenants will fill in Macy's and Sears empty spaces.
- A 1.2M SF mall in California announced a rebranding in 2015 and completed renovations in 2018. During this time, parts of the center were opened.
- A 1.1M SF mall in Maryland announced in 2017 the $108 million redevelopment to a new outdoor shopping center, anchored by Costco. The constructed began the following year and completed by 2019.
- A 1.9M SF mall in Texas announced a multimillion-dollar renovation of the property in 2015. As part of the transformation, the mall gave itself a new name and rebranded the center. It also added a new landscaping, lighting, and a new public art program.

Next, we compare the results of dead and survived community centers in Table 4-7. The table yields an interesting finding that very few survived centers have high predicted values. From our probit results, we know that the NRA coefficient has the greatest effect for community centers. We find that the size of the survived community centers become smaller as we move down the ranges from high predicted probability to low predicted probability. Community Centers that died and have high predicted probabilities, namely in the 60-70% range, are on average bigger in size, endure significant competition, and have high availability rates.

Table 4-7: CDF Tabular Summary for Community Centers

Survived Community Centers - 2.5 Miles

(3,886 Observations)

Normal CDF	**% Observed**	**Mean**				
		Avail. Rate	**Age**	**Reno Dum.**	**NRA**	**Comp. NRA**
					Million	*Million*
90%+	0%	-	-	-	-	-
80-90%	0%	-	-	-	-	-
70-80%	0%	-	-	-	-	-
60-70%	0%	18%	45	0.5	0.7	1.2
50-60%	0%	26%	47	0.0	0.6	0.5
40-50%	0%	11%	37	0.2	0.5	0.7
30-40%	1%	13%	33	0.3	0.5	1.0
20-30%	3%	14%	31	0.2	0.4	0.8
10-20%	18%	16%	33	0.2	0.3	0.7
0-10%	78%	13%	38	0.4	0.2	0.5
Total	**100%**	**13%**	**37**	**0.4**	**0.2**	**0.6**

(cont.)

Dead Community Centers - 2.5 Miles (368 Observations)						
Normal CDF	**% Observed**	**Mean**				
		Avail. Rate	**Age**	**Reno Dum.**	**NRA**	**Comp. NRA**
					Million	*Million*
90%+	1%	14%	49	0.5	1.0	1.4
80-90%	1%	15%	53	0.5	0.9	0.8
70-80%	2%	8%	20	0.0	0.7	0.7
60-70%	2%	14%	32	0.2	0.7	0.9
50-60%	2%	10%	36	0.3	0.6	0.8
40-50%	6%	10%	23	0.0	0.5	0.7
30-40%	12%	12%	31	0.3	0.5	0.7
20-30%	13%	9%	30	0.4	0.4	0.7
10-20%	16%	15%	34	0.3	0.3	0.7
0-10%	46%	16%	37	0.3	0.1	0.6
Total	**100%**	**14%**	**34**	**0.3**	**0.3**	**0.7**

Since the magnitude of the NRA coefficient was substantial for community centers, the dead community centers with the highest predicted values are centers with the largest NRA and face significant competition – two effects that exacerbate the performance of community centers.

Like malls, we discover a mismatch between a dead community center with a low predicted value, and a survived community center with a high predicted value. The explanation for why a community center's actual status differed from its expected status is because of timing of transitioning and repositioning.

- *Community centers that survived despite high predicted values*: There are a couple of community centers that survived with high predicted values. We find the mismatch driven in part by the exit of large tenants:
  - A 0.7M square-foot center in Texas announced at the beginning of 2020 the closure of two anchor tenants, including Macy's and Sears.
  - A 0.6M square-foot center in Maryland announced in 2019 the closure of two large tenants, A.C. Moore and Dressbarn.
- *Community Centers that died despite low predicted values*: Nearly half of the dead centers have low predicted values. In many cases, these centers went through repositioning or redevelopment:
  - A 0.1M SF retail center in Kansas announced a redevelopment in 2017. The demolition took place in 2019 with plans to tear down all buildings, except for two anchor restaurants Olive Garden and Red Lobster, to reposition the site to add specialty stores, health and fitness centers, offices, and a multifamily building.

- A declining shopping center with 0.1M SF in California went through a renovation and adaptive reuse to transform the center into a retail, government services, and senior housing hub. The project was completed in 2017.
- A 0.1M SF community center in Maryland closed few notable tenants including Barnes & Noble, Bahama Breeze, and Pier 1 a few years ago to make way for a new $30 million development that will open in 2020. The development will include more than 20 new retailers and a new multifamily property.
- A retail center in California with roughly 0.1M SF underwent a $3 million renovation in 2017 to reposition into an open-air marketplace; in 2019, the center revamped roster of restaurant tenants.

In this section, we discussed the results for our retail risk modeling and identified the risk profiles for both malls and community centers. We also highlighted the mismatch between the predicted and actual values for centers that died or survived despite low or high predicted values. In the next section, we summarize our findings.

5. Conclusion

In today's tough retail environment, understanding which malls or community centers are at risk of becoming distressed is essential to anticipating a possible death. Centers can lose much of their value to investors and the economic well-being of the communities they serve. This analysis draws on the trove of data accumulated by CBRE and identifies the factors that lead to investment risks and opportunities around the country.

Our predictive model provides a framework for retail risk analysis where risk is defined by probabilities of mall or community center's death. The statistics and the probit model inform our range of probabilities, enabling investors and operators to rethink their investment thesis and develop scenarios about which centers are attractive or which may need to be sold or repositioned.

A key finding is that a property's net rentable area (NRA) has an outsized impact on the probability of failing. Malls with very large spaces are less likely to go under, whereas community centers with large spaces are at increased risk of contraction. We saw the coefficient of the mall's NRA was five times the size of the linear probability model coefficient, placing more weight on our probit results, while the coefficients for other variables – renovation, availability, and competition – were about three times bigger than the linear model. We determined that the potential to die grows as retail spaces become increasingly available, and when malls and community centers are close to competition within 5.0 miles and 2.5 miles, respectively. When the centers have an opportunity to renovate, the potential to become distressed for both centers decreases. Finally, age has no statistical effect, and can be said to have a yin yang effect, since the older centers could be at risk because of deterioration or may perform successfully as they are in prime locations.

We also uncovered that centers have a mismatch between a survived property with high predicted values as well as a mismatch between a dead property with low predicted values. In these cases, we identified a center's actual status differing from its expected status is driven by timing of when a center may be in transition, repositioning, or undergoing significant changes.

Our results hold some important lessons for developing, investing, and managing centers. First, size does matter. Second, competition, availability rate, and renovation trigger changes in the probabilities of defaulting but not as materially as the size of the center. Finally, age has no effect on performance.

6. References

Advance Monthly Retail. *US Census Bureau* Monthly & Annual Retail Trade. https://www.census.gov/retail/index.html.

Buechler, Simon, et al. "Redevelopment Option Value for Commercial Real Estate." *University of Bern, University of Connecticut, Laval University*, 2020.

Buss, Christian, et al. *Apparel Retail & Brands: Making Sense Of Softlines Following A Tumultuous Twelve Months*. Credit Suisse, May 2017.

Clapp, John, et al. "Expansions and Contractions of Major US Shopping Centers." *The Journal of Real Estate Finance and Economics*, Aug. 2012, pp. 16–56.

CoStar Glossary. https://www.costar.com/about/costar-glossary#go_r. Accessed June 2020.

DiPasquale, Denise, and William C. Wheaton. *Urban Economics and Real Estate Markets*. Prentice Hall, 1996.

Dustan, Andrew. *Linear Probability Model vs. Logit (or Probit)*. Department of Agricultural & Resource Economics, University of California, Berkeley.

Econ 423: Regression with a Binary Dependent Variable. Department of Economics, University of Maryland.

Eppli Mark and Shilling James How Critical Is a Good Location to a Regional Shopping Center? *Journal of Real Estate Research*: 1996, Vol. 12, No. 2, pp. 459-468.

Finney, D. J. *Probit Analysis*. 3rd ed, University Press, 1971.

Holman, Lee, and Greg Buzek. *Retail's Renaissance: The True Story of Store Openings/Closings*. IHL Group, 2019.

Kennedy, Clare. "UBS Estimates 100,000 More Retail Stores Will Close by 2025." *LoopNet*, 27 Apr. 2020, https://www.loopnet.com/learn/ubs-estimates-100000-more-retail-stores-will-close-by-2025/174318906.

Lakshmanan, T.R. and W.G. Hansen, A Retail Market Potential Model, *Journal of the American Institute of Planners*, 1965, 31, 134–43.

Lee, Ming-Long, and R. Kelley Pace. "Spatial Distribution of Retail Sales." *The Journal of Real Estate Finance and Economics*, vol. 31, no. 1, Aug. 2005, pp. 53–69.

Morgan, Andrew, et al. *Mastering Spark for Data Science*. Birmingham, UK, 2017.

Noreen, Eric. "An Empirical Comparison of Probit and OLS Regression Hypothesis Tests." *Journal of Accounting Research*, vol. 26, no. 1, 1988, pp. 119–33. JSTOR.

Petro, Greg. "Shopping Malls Aren't Dying – They're Evolving." Forbes, 5 Apr. 2019, https://www.forbes.com/sites/gregpetro/2019/04/05/shopping-malls-arent-dying-theyre-evolving.

Probit Regression | Real Statistics Using Excel. http://www.real-statistics.com/logistic-regression/probit-regression.

She, Yuling "Development Option Value for Industrial Property." *Massachusetts Institute of Technology, The Center for Real Estate*, 2020.

"State of Retail." National Retail Federation (NRF), https://nrf.com/insights/economy/state-retail

Sullivan, Jim, and Otto Aletter. *2017 Mall Tenant Turnover Analysis*. White Paper, Advisory & Consulting Green Street Advisors, 2018, pp. 7.

Thompson, Alexandria. "Total Size of U.S. Commercial Real Estate Estimated Between $14 and $17 Trillion." *Nareit*, https://www.reit.com/news/blog/market-commentary/total-size-us-commercial-real-estate-estimated-between-14-and-17.

Turnbull, Geoffrey, and Jonathan Dombrow. "Spatial Competition and Shopping Externalities: Evidence from the Housing Market." *The Journal of Real Estate Finance and Economics*, 2006, pp. 391–408.

Walmart Inc. 2020 Annual Report. 23 Apr. 2020.

Zhou, Tingyu, and John Clapp. "Predicting Risks of Anchor Store Openings and Closings." *The Journal of Real Estate Finance and Economics*, Sept. 2015, pp. 449–79.

APPENDIX A: CBRE INVENTORY BY CBSA[40]

Distribution of CBRE Inventory By CBSA

	Total Inventory			% Mix		
	Total	Malls	Community Centers	Total	Malls	Community Centers
Chicago-Naperville-Elgin, IL-IN-WI	359	29	330	6%	4%	6%
New York-Newark-Jersey City, NY-NJ-PA	345	47	298	6%	7%	6%
Los Angeles-Long Beach-Anaheim, CA	330	49	281	5%	7%	5%
Miami-Fort Lauderdale-Pompano Beach, FL	251	27	224	4%	4%	4%
Dallas-Fort Worth-Arlington, TX	249	21	228	4%	3%	4%
Atlanta-Sandy Springs-Alpharetta, GA	235	20	215	4%	3%	4%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	230	23	207	4%	3%	4%
Houston-The Woodlands-Sugar Land, TX	217	19	198	4%	3%	4%
Phoenix-Mesa-Chandler, AZ	193	21	172	3%	3%	3%
Washington-Arlington-Alexandria, DC-VA-MD-WV	181	27	154	3%	4%	3%
Riverside-San Bernardino-Ontario, CA	174	14	160	3%	2%	3%
Boston-Cambridge-Newton, MA-NH	155	19	136	3%	3%	3%
Orlando-Kissimmee-Sanford, FL	136	10	126	2%	1%	2%
San Francisco-Oakland-Berkeley, CA	135	20	115	2%	3%	2%
Detroit-Warren-Dearborn, MI	127	18	109	2%	3%	2%
Denver-Aurora-Lakewood, CO	119	8	111	2%	1%	2%
Baltimore-Columbia-Towson, MD	106	15	91	2%	2%	2%
Tampa-St. Petersburg-Clearwater, FL	105	7	98	2%	1%	2%
Seattle-Tacoma-Bellevue, WA	103	23	80	2%	3%	2%
San Diego-Chula Vista-Carlsbad, CA	99	10	89	2%	1%	2%
Sacramento-Roseville-Folsom, CA	96	9	87	2%	1%	2%
Las Vegas-Henderson-Paradise, NV	95	11	84	2%	2%	2%
St. Louis, MO-IL	92	14	78	2%	2%	1%
Kansas City, MO-KS	84	8	76	1%	1%	1%
Minneapolis-St. Paul-Bloomington, MN-WI	84	20	64	1%	3%	1%
Columbus, OH	82	8	74	1%	1%	1%
San Antonio-New Braunfels, TX	79	10	69	1%	1%	1%
Cincinnati, OH-KY-IN	75	12	63	1%	2%	1%
Indianapolis-Carmel-Anderson, IN	72	10	62	1%	1%	1%
Louisville/Jefferson County, KY-IN	71	5	66	1%	1%	1%
Portland-Vancouver-Hillsboro, OR-WA	71	7	64	1%	1%	1%
Pittsburgh, PA	71	16	55	1%	2%	1%
Jacksonville, FL	69	5	64	1%	1%	1%
San Jose-Sunnyvale-Santa Clara, CA	66	10	56	1%	1%	1%
Nashville-Davidson--Murfreesboro--Franklin, TN	64	10	54	1%	1%	1%
Cleveland-Elyria, OH	64	10	54	1%	1%	1%
Raleigh-Cary, NC	64	4	60	1%	1%	1%
Charlotte-Concord-Gastonia, NC-SC	61	7	54	1%	1%	1%
Memphis, TN-MS-AR	61	8	53	1%	1%	1%
Providence-Warwick, RI-MA	59	8	51	1%	1%	1%
Richmond, VA	57	9	48	1%	1%	1%
Austin-Round Rock-Georgetown, TX	56	6	50	1%	1%	1%
Birmingham-Hoover, AL	49	5	44	1%	1%	1%
Oklahoma City, OK	45	8	37	1%	1%	1%
Salt Lake City, UT	44	4	40	1%	1%	1%
Oxnard-Thousand Oaks-Ventura, CA	41	5	36	1%	1%	1%
Tulsa, OK	33	4	29	1%	1%	1%
Albuquerque, NM	33	3	30	1%	0%	1%
Greensboro-High Point, NC	33	5	28	1%	1%	1%
Worcester, MA-CT	33	3	30	1%	0%	1%
All Other	279	39	240	5%	5%	5%
Total	**6,032**	**710**	**5,322**	**100%**	**100%**	**100%**

[40] CBSA: Core Based Statistical Area

APPENDIX B: ICSC 2020 SHOPPING CENTER CLASSIFICATIONS

ICSC

U.S. Shopping-Center Classification and Typical Characteristics*

Type of Shopping Center	Concept	Typical GLA Range (Sq. Ft.)	Acres	# of Anchors	% Anchor GLA	Typical Number of Tenants	Typical Type of Anchors	Trade Area Size
General-Purpose Centers								
Super-Regional Mall	Similar in concept to regional malls, but offering more variety and assortment.	800,000+	60-120	3+	50-70%	N/A	Full-line department store, mass merchant, discount department store, fashion apparel store, mini-anchor, cineplex or other large-scale entertainment attraction, and food-and-beverage service cluster.	5-25 miles
Regional Mall	General merchandise or fashion-oriented offerings. Typically, enclosed with inward-facing stores connected by a common walkway. Parking surrounds the outside perimeter.	400,000-800,000	40-100	2+	50-70%	40-80 stores	Full-line department store, mass merchant, discount department store, fashion apparel store, mini-anchor, cineplex or other large-scale entertainment attraction, and food-and-beverage service cluster.	5-15 miles
Community Center ("Large Neighborhood Center")	General merchandise or convenience-oriented offerings. Wider range of apparel and other soft goods offerings than neighborhood centers. The center is usually configured in a straight line as a strip, or may be laid out in an L or U shape, depending on the site and design.	125,000-400,000	10-40	2+	40-60%	15-40 stores	Discount store, supermarket, drug, large-specialty discount (toys, books, electronics, home improvement/furnishings or sporting goods, etc.)	3-6 miles
Neighborhood Center	Convenience-oriented.	30,000-125,000	3-5	1+	30-50%	5-20 stores	Supermarket	3 miles
Strip/Convenience	Attached row of stores or service outlets managed as a coherent retail entity, with on-site parking usually located in front of the stores. Open canopies may connect the storefronts, but a strip center does not have enclosed walkways linking the stores. A strip center may be configured in a straight line, or have an "L" or "U" shape. A convenience center is among the smallest of the centers, whose tenants provide a narrow mix of goods and personal services to a very limited trade area.	< 30,000	<3	Anchor-less or a small convenience-store anchor.	N/A	N/A	Convenience store, such as a mini-mart.	<1 mile
Specialized-Purpose Centers								
Power Center	Category-dominant anchors, including discount department stores, off-price stores, wholesale clubs, with only a few small tenants.	250,000-600,000	25-80	3+	70-90%	N/A	Category killers, such as home improvement, discount department, warehouse club and off-price stores	5-10 miles
Lifestyle	Upscale national-chain specialty stores with dining and entertainment in an outdoor setting.	150,000-500,000	10-40	0-2	0-50%	N/A	Large-format upscale specialty	8-12 miles
Factory Outlet	Manufacturers' and retailers' outlet stores selling brand-name goods at a discount.	50,000-400,000	10-50	N/A	N/A	N/A	Manufacturers' and retailers' outlets	25-75 miles
Theme/Festival	Leisure, tourist, retail and service-oriented offerings with entertaiment as a unifying theme. Often in urban areas, they may be adapted from older—sometimes historic—buildings, and part of a mixed-use project.	80,000-250,000	5-20	Unspecified	N/A	N/A	Restaurants, entertainment	25-75 miles
Limited-Purpose Property								
Airport Retail	Consolidation of retail stores located within a commercial airport	75,000-300,000	N/A	N/A	N/A	N/A	No anchors; retail includes specialty retail and restaurants	N/A

***Disclaimer: While every effort is made to ensure the accuracy and reliability of the information contained in this report, ICSC does not guarantee and is not responsible for the accuracy, completeness or reliability of the information contained in this report. Use of such information is voluntary, and reliance on it should only be undertaken after an independent review of its accuracy, completeness, efficiency, and timeliness. Criteria used in the definitions above are intended to be only typical of general features, rather than covering all situations.**

APPENDIX C: STATA PROBIT RESULTS FOR COMMUNITY CENTERS (2.5M)

COMM 2.5 Mile(obs=4,254)

```
             |   PropID Dead_Dum      Age Reno_Dum    Avail    NRA Comp_25M       SF
-------------+------------------------------------------------------------------------
      PropID |   1.0000
    Dead_Dum |   0.0998   1.0000
         Age |  -0.4473  -0.0588   1.0000
    Reno_Dum |  -0.3588  -0.0523   0.5078   1.0000
       Avail |  -0.0575   0.0075   0.0722   0.0322   1.0000
         NRA |  -0.0521   0.3053  -0.0303   0.0547  -0.0775   1.0000
    Comp_25M |   0.0111   0.0511  -0.0008   0.0117   0.0517   0.0299   1.0000
          SF |  -0.0521   0.3053  -0.0303   0.0547  -0.0775   1.0000   0.0299   1.0000
```

```
. probit Dead_Dum Age i.Reno_Dum Avail NRA Comp_25M

Iteration 0:   log likelihood = -1252.2946
Iteration 1:   log likelihood = -1103.9792
Iteration 2:   log likelihood = -1102.5854
Iteration 3:   log likelihood = -1102.5851
Iteration 4:   log likelihood = -1102.5851

Probit regression                               Number of obs     =      4,254
                                                LR chi2(5)        =     299.42
                                                Prob > chi2       =     0.0000
Log likelihood = -1102.5851                     Pseudo R2         =     0.1195

------------------------------------------------------------------------------
    Dead_Dum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Age |  -.0019872   .0022109    -0.90   0.369    -.0063205    .0023461
  1.Reno_Dum |  -.2439208   .0736552    -3.31   0.001    -.3882823   -.0995593
       Avail |   .3597405   .1644656     2.19   0.029     .0373939    .6820871
         NRA |   3.857287   .2450712    15.74   0.000     3.376956    4.337618
    Comp_25M |   .1675041   .0596848     2.81   0.005      .050524    .2844843
       _cons |  -2.273236   .1070315   -21.24   0.000    -2.483014   -2.063458
```

APPENDIX D: STATA PROBIT RESULTS FOR MALLS (2.5M)

MALL 2.5 Mile(obs=646

```
             |      _ID Dead_Dum      Age Reno_Dum    Avail      NRA Comp_~5M       SF
-------------+------------------------------------------------------------------------
         _ID |   1.0000
    Dead_Dum |   0.3343   1.0000
         Age |  -0.5511  -0.1718   1.0000
    Reno_Dum |  -0.5196  -0.3017   0.6155   1.0000
       Avail |   0.0296   0.1974  -0.0090  -0.0851   1.0000
         NRA |  -0.2233  -0.3959   0.0873   0.1838  -0.2000   1.0000
   Comp_2_5M |   0.0715   0.1060   0.0170   0.0020   0.0358  -0.1547   1.0000
          SF |  -0.2233  -0.3959   0.0873   0.1838  -0.2000   1.0000  -0.1547   1.0000
```

```
. probit Dead_Dum Age Reno_Dum Avail NRA Comp_2_5M

Iteration 0:   log likelihood = -336.38952
Iteration 1:   log likelihood = -257.78607
Iteration 2:   log likelihood = -253.76604
Iteration 3:   log likelihood = -253.75457
Iteration 4:   log likelihood = -253.75457

Probit regression                                 Number of obs     =        646
                                                  LR chi2(5)        =     165.27
                                                  Prob > chi2       =     0.0000
Log likelihood = -253.75457                       Pseudo R2         =     0.2457

------------------------------------------------------------------------------
    Dead_Dum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Age |  -.0024034   .0052255    -0.46   0.646    -.0126453    .0078385
    Reno_Dum |  -.6803029   .1742321    -3.90   0.000    -1.021792   -.3388143
       Avail |   1.371438   .5012835     2.74   0.006     .3889407    2.353936
         NRA |  -1.530633   .1828133    -8.37   0.000     -1.88894   -1.172325
   Comp_2_5M |   .0818166   .1001701     0.82   0.414    -.1145131    .2781463
       _cons |   .9372838   .2434597     3.85   0.000     .4601116    1.414456
```

APPENDIX E: STATA PROBIT RESULTS FOR COMMUNITY CENTERS (5.0M)

COMM 5 Mile(obs=4,254)

```
             |   PropID Dead_Dum      Age Reno_Dum    Avail      NRA  Comp_5M       SF
-------------+------------------------------------------------------------------------
      PropID |   1.0000
    Dead_Dum |   0.0998   1.0000
         Age |  -0.4473  -0.0588   1.0000
    Reno_Dum |  -0.3588  -0.0523   0.5078   1.0000
       Avail |  -0.0575   0.0075   0.0722   0.0322   1.0000
         NRA |  -0.0521   0.3053  -0.0303   0.0547  -0.0775   1.0000
     Comp_5M |  -0.0145   0.0004   0.0313   0.0160   0.0439   0.0047   1.0000
          SF |  -0.0521   0.3053  -0.0303   0.0547  -0.0775   1.0000   0.0047   1.0000
```

```
. . probit Dead_Dum Age i.Reno_Dum Avail NRA Comp_5M

Iteration 0:   log likelihood = -1252.2946
Iteration 1:   log likelihood = -1107.7838
Iteration 2:   log likelihood = -1106.4696
Iteration 3:   log likelihood = -1106.4694
Iteration 4:   log likelihood = -1106.4694

Probit regression                               Number of obs     =      4,254
                                                LR chi2(5)        =     291.65
                                                Prob > chi2       =     0.0000
Log likelihood = -1106.4694                     Pseudo R2         =     0.1164

------------------------------------------------------------------------------
    Dead_Dum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Age |  -.0020155   .0022038    -0.91   0.360    -.0063349     .002304
  1.Reno_Dum |  -.2389715    .073556    -3.25   0.001    -.3831386    -.0948044
       Avail |   .3889689   .1636819     2.38   0.017     .0681582    .7097795
         NRA |   3.870965   .2446013    15.83   0.000     3.391556    4.350375
     Comp_5M |    -.00036   .0276232    -0.01   0.990    -.0545005    .0537805
       _cons |  -2.178184   .1086515   -20.05   0.000    -2.391137   -1.965231
```

APPENDIX F: STATA PROBIT RESULTS FOR MALLS (5.0M)

MALL 5 Mile(obs=646)

```
             |   PropID Dead_Dum      Age Reno_Dum    Avail      NRA  Comp_5M       SF
-------------+------------------------------------------------------------------------
      PropID |   1.0000
    Dead_Dum |   0.3343   1.0000
         Age |  -0.5511  -0.1718   1.0000
    Reno_Dum |  -0.5196  -0.3017   0.6155   1.0000
       Avail |   0.0296   0.1974  -0.0090  -0.0851   1.0000
         NRA |  -0.2233  -0.3959   0.0873   0.1838  -0.2000   1.0000
     Comp_5M |   0.0551   0.1347   0.0805   0.0431   0.0197  -0.1344   1.0000
          SF |  -0.2233  -0.3959   0.0873   0.1838  -0.2000   1.0000  -0.1344   1.0000


. probit Dead_Dum Age i.Reno_Dum Avail NRA Comp_5M

Iteration 0:   log likelihood = -336.38952
Iteration 1:   log likelihood = -256.12656
Iteration 2:   log likelihood = -252.04796
Iteration 3:   log likelihood = -252.04346
Iteration 4:   log likelihood = -252.04346

Probit regression                               Number of obs     =        646
                                                LR chi2(5)        =     168.69
                                                Prob > chi2       =     0.0000
Log likelihood = -252.04346                     Pseudo R2         =     0.2507

------------------------------------------------------------------------------
    Dead_Dum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Age |  -.0023889   .0052209    -0.46   0.647    -.0126217     .007844
  1.Reno_Dum |  -.7025547   .1752708    -4.01   0.000    -1.046079   -.3590304
       Avail |   1.383974   .4999253     2.77   0.006      .404138    2.363809
         NRA |   -1.48898   .1824284    -8.16   0.000    -1.846533   -1.131427
     Comp_5M |    .113102   .0556839     2.03   0.042     .0039634    .2222405
       _cons |   .8365232   .2464906     3.39   0.001     .3534105    1.319636
```