Which Malls Close

By
William Wheaton
MIT, CBRE Econometric Advisors
Morgan Fleischman[1]
MIT Center for Real Estate
July 2020

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 change 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. National Association of Real Estate Investment Trusts (Nareit) estimates that retail real estate accounts for 13.6 billion square feet and is valued at $2.4 trillion[3].

There has been a lot of discussion that the industry is under significant pressure. According to a report by LoopNet, an affiliate of CoStar Group, analysts at UBS 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. In 2017, Credit Suisse offered a grim outlook of malls, estimating that nearly one in four of the country's 1,169 malls, or 275 malls, will close in this decade[5]. Malls have not traditionally kept pace with the changing expectations of consumers or their differing needs, but now they are quickly transforming themselves to new kinds of destinations investing heavily in new amenities, experiences, and entertainment to improve the shopping experience. On the other hand, the IHL Group, a research firm, has a more positive view. According to a recent study, for every chain with a net closing of stores, 5.2 new stores are opened[6]. 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.

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 shopping center closures are limited. For instance, John M. Clapp, Katsiaryna Bardos, and Tingyu Zhou (2012)[7] 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 the expansion and downsizing of gross leasable area (GLA). There are other related studies that utilize predictive modeling, spatial competition, and probit regression, but this paper differs in several ways. First, it's a rigorous empirical research focusing on just mall closures in the United States, using hundreds of property-level data points. Second, we compute the distances for each mall relative to each other to estimate the amount of spatial competition each mall faces. Finally, we develop a predictive model to identify the centers at risk of going under today.

---

[1] This paper is based on the 2020 MSRED thesis of Morgan Fleischman, "Sorry We're Closed: What Closes Malls and Community Centers in the United States? An Analysis and Predictive Modeling of Distressed Centers"

Data

In this study, we rely on georeferenced data for all retail centers. The information comes from CBRE, which sources neighborhood community centers (NCS) and malls data from several repositories including CoStar (primary source), LoopNet, 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. Here, we focus just on malls.

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 to make sure they are identical. There are three instances in which malls can show up in the surveys:

- *A mall 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 mall is a new property.

- *A mall in 2010 survey but not in 2020 survey*: If the mall shows up in the earlier 2010 survey but not in the 2020 survey, we assume the center has "died" or closed. This could happen for a variety of reasons: decommissioned, going through enough renovation to be "empty" and closed, demolished, turned into another use, or undergoing some other significant transitioning.

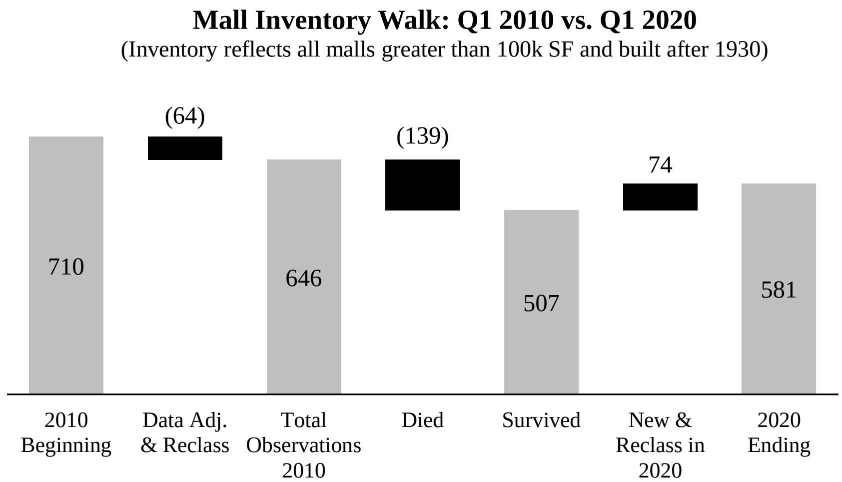
- *A mall in both 2010 and 2020 surveys*: A mall could show up in both surveys. If the center is in same "mall" category, it is a "survived" center. The mall can also transition into a different category between 2010 and 2020. For instance, a mall is repositioned and downgraded to a community center. In the study, we excluded all transitioned centers – all upgrades or downgrades to a different classification – in both surveys.

The study (1) compares survived malls with dead malls and (2) examines only malls. Other centers – such as community, neighborhood, and strip centers – account for the majority of CBRE's inventory but may introduce a lot of random noise in the observations since they are smaller and more ubiquitous. We examine only malls to create greater homogeneity of product.

We also filter out extreme values, particularly centers that were built prior to the 1930's, have flawed input data, and contain net rentable area (NRA) under 100,000 SF. We also remove centers that were either upgraded or downgraded to a different category since we want to look at only dead and survived results. After refining the mall parameters and removing reclassifications, there are 646 observations.

Figure 1 shows the cumulative inventory changes 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. The analysis ignores any new centers that were constructed in 2020 or that were reclassified to a mall from a different center type in 2020. In the past decade, malls expanded, repositioned, and contracted, but ignoring such changes, the close rate was a very significant 22%.

Figure 1: Cumulative Inventory Change for Malls[8]



In the data base, there is only a limited amount of information about each mall at each observation date. Table 1 displays the available data – using the 2010 observation date. Most are self-explanatory. The competition variable was constructed by taking a ring of 5 miles in radius and adding up all of the (other) mall NRA that falls within this. We experimented with alternate ring sizes, but 5-mile radius produced the best results. On average, survived malls tend to be 0.4 million SF larger, have lower 2010 availability rates (by 5.4 percentage points), and face 0.25 million square feet less competition. The survived malls are not much different in age, but are somewhat more likely to have been recently renovated.

Table 1: Descriptive Statistics for Malls

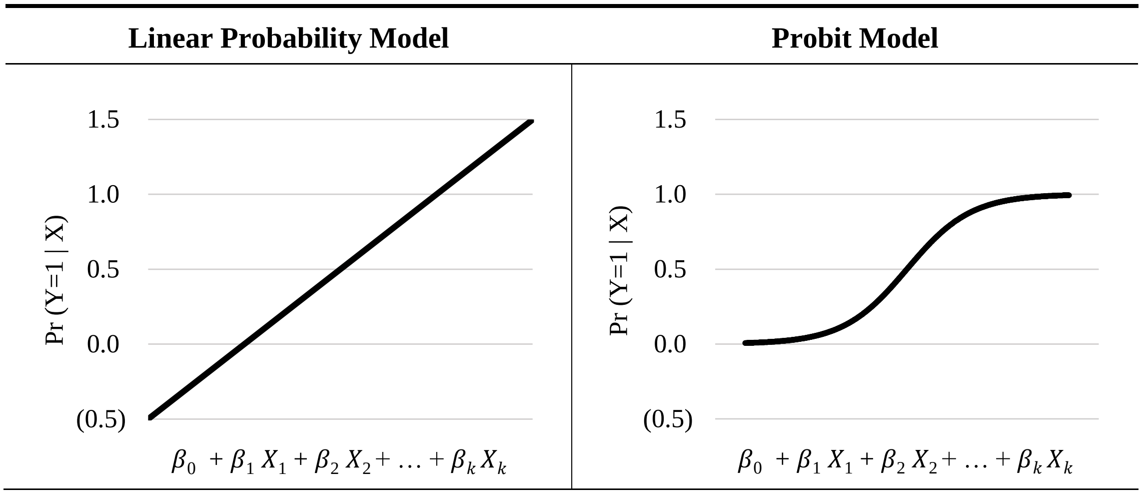
Malls (646 Observations)								
	Survived				Dead			
	Mean	Stdev	Min	Max	Mean	Stdev	Min	Max
Age, Years	43	14	11	78	37	18	11	78
Age Since Renovation, Years	21	6	11	55	23	6	11	36
Availability Rate, %	6.1%	9.3%	0.0%	77.2%	11.5%	15.9%	0.0%	97.2%
NRA, SF	1,076,235	443,304	110,000	2,951,995	631,682	346,705	107,000	1,927,193
Competition NRA within 5.0 miles	743,656	1,030,879	0	5,227,177	1,091,733	1,133,367	0	4,165,617

Methodology

This research uses three steps to experiment and test the predictive modeling of 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 of the probability of closure.

Our first approach to assessing how variables impact the likelihood of a mall's closure is simple OLS regression. When the dependent variable is dichotomous (0 or 1 binary response) this type of equation is called a linear probability model (LPM). A major issue with the ordinary least squares regression and an LPM is that with a binary variable as the dependent variable, the assumptions in the linear regression significance test are violated. It also has the issue that predicted values can easily fall outside of the 0-1 interval that defines "probability." One solution is the probit model, short for "probability unit." It estimates the probability a value will fall into one of the two binary outcomes with a predicted probability that is never below 0 or above 1. The predicted probabilities obey the sigmoid-shaped curve in Figure 2. The vertical axis is a probability, falling between zero and one. The straight line on the left is the linear regression.

Figure 2: Shape of Curve - Linear Probability vs. Probit Model[9]



We used Stata statistical software in this study to estimate a probit function with maximum likelihood estimation (MLE) using the independent variables in Table 1: age, renovation dummy, 2010 availability rate, NRA, and competition NRA within 5.0 miles.

The interpretation of the probit result is essential. The constant in the probit results can be interpreted as a predicted z-score when all independent variables equal zero. The coefficients of the probit model are interpreted as telling us how much the z-score will increase for each one-unit increase in each of the independent variables. However, we need to change the z-score into an actual probability. For this, we use the standard normal distribution (mean of 0 and standard deviation of 1) and ask where in the *cumulative* distribution that z-score stands. If the z-score is exactly "0" the predicted probability is 0.5. If it is "-1" the probability is 0.16, and if it is "+1" it is 0.84. In this approach, all predicted values must lie between 0 and 1, as shown in Figure 2. In the probit model, the impact of a unit change in any independent variables will depend on the values of all the independent variables – or the level of that probability. The effect of a unit increase in an independent variable will be larger around the mean z-values and less when the z-scores are out in the tails. This is what gives the probit its nice "S" shape between 0 and 1.

An important aspect of the estimated OLS regression equation is its ability to most simply predict the effects on Y from a unit change in one or more values of the independent variables X. This is important if the goal is to predict the likelihood of a center's death from all the variables. With the LPM this is straightforward, but to make this assessment with the probit model is more complicated. To calculate a predicted probability for each observation, we first take the value of the independent variables from each observation, multiply these by their probit coefficients and add up, including the constant. This first step results in a predicted z-score value for each observation. We next use the standard normal distribution (mean of 0 and standard deviation of 1) and ask where in the distribution that z-score stands.

Results

The OLS regression results are summarized in Table 2. The age coefficient has no effect according to the t-statistic. One way to interpret this is that older malls might be more deteriorated and more likely to close, but they also happen to be in better locations at the time when they were built. Thus, the results here may indicate there is an offsetting effect between these two notions.

Table 2: Summary of OLS Regression Results

Malls: OLS Regression Results (5.0 Miles)

	Coefficients	**Statistical Significance**
Age	0.000	Insignificant
Renovation Dummy	-0.219	Significant
Availability Rate	0.410	Significant
NRA	-0.282	Significant
Competition NRA	0.039	Significant

Next is the renovation dummy. The coefficient is negative and significant, indicating that renovation decreases the center's likelihood to close by 22 percentage points. Remodeled centers may improve store offerings, enhance shopper's experiences, and create a greater opportunity cost for closure. The availability coefficient is positive and significant with a completely vacant center (in 2010) being 41% more likely to close by 2020 than one completely full. Historic vacancy in 2010 is a good precursor to eventually closing.

The NRA coefficient is perhaps also not surprising. Malls with 1 million greater NRA have a 28% lower closure probability according to the t-statistic. Larger malls are typically owned by institutional firms and REITs. Such owners tend to have deeper pockets, so they can hold on to the centers longer and hope for "better times."

The final result is the level of spatial competition. The significance test reveals that within a 5-mile radius, 1 million more NRA of competing mall space increases the likelihood of mall closure by 4 percentage points. As mentioned above, we experimented with alternate sized rings and 5-mile radius produced the most significant results, which still were always positive on closure likelihood.

In Table 3, we display the results of the probit model noting that the coefficients cannot be interpreted as partial probability effects, as is true with the OLS regression. The signs of the coefficients and their significance however appear virtually the same. With a probit, the coefficients show us how much the variable impacts the z-score. The true partial effect of a unit change in the variable on the closure probability is the product of the probit coefficient with the slope of the sigmoid curve at that point. This of course depends on the values of all the independent variables. When out in the tails of the distribution, the partial effects are quite small, whereas when near to the distribution mean they are much larger.

Table 3: Summary of Probit Regression Results

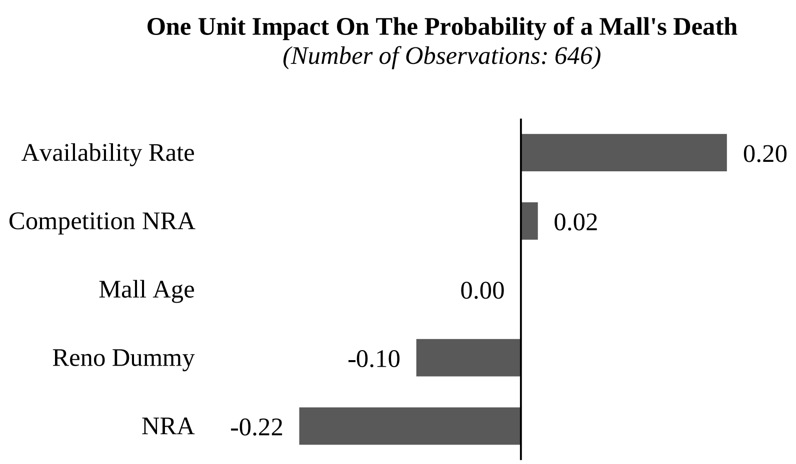
Malls: Probit Regression Results (5.0 Miles)

	Coefficients	**Statistical Significance**
Age	-0.002	Insignificant
Renovation Dummy	-0.703	Significant
Availability Rate	1.384	Significant
NRA	-1.489	Significant
Competition NRA	0.113	Significant

Although we cannot interpret the partial effects in the same way as the OLS model, we can still estimate a unit impact on the probability of a mall dying. The probit results tell us for a unit change in each independent variable, how much the z-score increases or decreases. To get the probability change per unit, we estimate the mean z-score by taking the mean value for each variable multiplied the variables' coefficients from the probit results and then add them up. The result is a predicted z-score for the "mean mall," which is -1.05 in this data. The standardized normal CDF probability of -1.05 is 0.15, which indicates that probability of the "mean mall" failing is 15%. We then take that mean probability of 0.15 multiplied by each coefficient to estimate the percentage point change in the probability of a mall's closure for each unit change in a variable.

Figure 3 summarizes the impact of a unit change on the default probability for each predictor. A unit change is defined differently for each independent variable. For example, a unit change for both "NRA" and "Competition NRA" represents the addition or subtraction of 1 million square feet, since we measured the NRA in millions square feet. The figure shows that there is an outsized effect of the mall's NRA, in 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 other independent variables.

Figure 3: Probit Impact of a Unit Change on the Risk Probability – Malls



In the case of renovation, if the mall gets renovated, then the impact on the probability of a mall's dying decreases by 10 percentage points. The availability variable tells us for every 10 percentage point increase in the availability rate, the impact on the probability of malls going under increases by 2 percentage points. Finally, for every increase of competitive square footage of 1 million the impact on the likelihood of closure also increases by 2 percentage points. The impact of age is muted as there is no statistical significance. Note that these probit impacts on closure probability are about half the magnitude of the OLS impacts in Table 2. This results from the probit model making sure to "squeeze" the predictions into the 0-1 interval. In the OLS model they are allowed to exceed this interval – even though there are never any data points outside of the 0-1 range.

A final application with these models is to examine counter-factuals. How many malls that survived actually had higher predicted value probabilities? Are these in fact more likely to close in the future with the passage of time? And conversely, how many malls that closed had very low predicted probabilities? Perhaps an examination of these would suggest some variables that are missing, which possibly we might see data for with a high predicted value, and a dead center with a low predicted value. In looking over the distribution of predicted closure probabilities, the value of 25% becomes a natural dividing line. 84% of the surviving malls had a predicted closure probability that was less than 25%, while only 24% of the closed malls had predicted values in this range. At the other extreme, virtually no malls had predicted closure rates of 80% or greater. The malls that did close had almost uniformly distributed predicted values throughout the 25%-80% range.

Summary

In today's difficult retail environment, understanding which malls are at risk of becoming distressed is essential to anticipating a possible closure. 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 these investment risks around the country.

Our predictive model provides a framework for retail risk analysis. The statistics of the probit model inform a 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.

Among our findings, the role of renovation, availability in 2010, and spatial competition all make sense and have some precedent in the theoretical literature. Perhaps a more surprising finding was the outsized impact of property size on the probability of closure. In this data, malls in the 2-to-3 million square foot range have a 40% lower probability of closure relative to smaller malls in the 0.5 million range, and this is quite significant statistically.

There are many other factors that can determine a mall failure for which we have no data, such as the tenant mix, transactions history, or management background. But our results hold some important lessons for developing, investing, and managing portfolios and how they measure up against our criteria. First, we know that 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 seems to have no effect on performance.

---

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

[7] 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.
[8] 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.
[9] Dustan, Andrew. *Linear Probability Model vs. Logit (or Probit)*. Department of Agricultural & Resource Economics, University of California, Berkeley.