

Noname

X

X

Matt Larriva, CFA

Abstract Purpose

To explain and forecast both national and market-level cap rate expansions as simply a binary interaction between Gross Domestic Product and Consumer Price Index.

Design/Methodology/Approach

We use a binary logistic regression with binary independent variables, trained on a minority oversampling set, to explain cap rate expansion at the national level and at 20 US metropolitan areas.

Results

Both the national model and the market models have accuracy near 70% and capture about 40% of all cap rate expansion periods with a robust confusion matrix, such that implementing the recommendations of the models would result in significant outperformance versus a buy-and-hold strategy.

Originality

This model contributes to the existing corpus by 1) establishing a statistically significant relationship between the interaction effect of GDP and CPI versus cap rates which 2) holds explanatory power at the national and market levels, by 3) mapping the ground truth from a non-discrete to a binary variable.

Practical Implications

The accuracy of the model suggests a simpler and more robust explanation for understanding cap rates and navigating residential property markets in the present environment of rapidly fluctuating interest rates, not only at the national level but also at the market level.

Keywords Binary Logistic Regression · Cap Rates · US Real Estate Markets · Multifamily Cap Rate · Apartment Cap Rate

Address(es) of author(s) should be given

1 Introduction

At 6 trillion dollars (NMHC, 2022), the value of apartment stock in the US constitutes over 1% of the total wealth of the world (Williams, 2021). Accurately projecting and explaining the value of that apartment stock has implications for direct investors, owners, renters, and anyone who owns part of the 255 billion dollars worth of securitized residential real estate that exists in public REITs (NAREIT, 2022).

Despite the importance of explaining and forecasting the near and long-term value of this real estate, there is little to guide investors as to the pricing, and therefore risk, of the asset class.

There has long been a supposed relationship between interest rates and cap rates: that the two move together. Proponents would cite historical examples; they would reason that interest rates relate to debt rates which influence real estate values; others would claim that inflation controlled interest rates and inflation was negative for real estate—eroding the value of its bond-like cashflows. While such arguments have been empirically voided and academically challenged (Larriva and Linneman, 2021), the cap rate and interest rate dynamics of 2022 offer yet another counterpoint.

This is not to say that interest rates are independent of cap rates. Indeed, much research has been dedicated to the opposite point. But the inconsistency of the relationship implies a missing component. Models have historically tried to solve for this component by introducing earnings expectations, demand-side factors, debt variables and spreads, and other germane variables. They often end with highly complex models that require regime-switching dynamics to account for the tail events (COVID-19, The Great Recession).

We revisit the question of cap rate expansion from a simpler lens which treats cap rate compression as one binary outcome and cap rate expansion the other. While it is helpful to know the degree to which real estate values will increase or decrease, it is almost as helpful to simply know if the value will be up or down. Real estate is, after all, still an alternative asset class, which means it is marked by illiquidity. Contrasted with public equities, one does not trade in and out of an apartment building daily, thus has less interest in the degree of price movement and more interest simply in the direction. Should a commercial real estate manager be a net buyer this year? Should a multifamily operator expect higher or lower NOI next year? Answering “yes” or “no” is often as useful as “yes, by exactly 267 basis points.”

We pair the simpler binary dependent variable with a novel binary input variable which illustrates a deceptively elegant principal of multifamily real estate. Namely, owners can often pass through higher expenses to tenants, if the tenants’ wages increase in proportionate to the expense. If an owner has to increase rents by 5% to compensate for the increased cost of, say, maintenance personnel, then he can pass this through (or more) to his tenant base, if their wages have increased somewhat proportionally. Thus multifamily real estate values should be determined by the ability of renters to pay inflation-like increases. Conversely, if the expenses of operations increase but the tenants’ wages do not, then rents cannot be increased, and cash yield will decrease, which will reduce the value of the asset. We propose that gross domestic product (GDP) is a sufficient proxy for wages, and the consumer price index (CPI) is a sufficient proxy for inflation. We model a binary situation: is GDP increasing faster (decreasing more slowly) than CPI? If so, cap rates should compress. If not—if CPI is increasing faster (decreasing more slowly) than GDP—then cap rates should expand.

Examining multifamily real estate at the national and market level, we evaluate the simple yet elegant thesis that the interplay between consumer price index changes (CPI) and gross

domestic product changes (GDP) explains cap rate compression and expansion. Testing this on 20 metropolitan statistical areas (MSAs) and the national series yield statistically significant models with high levels of accuracy and recall. Signals for cap rate expansions are conservatively issued, and acting upon them outperforms a buy-and-hold strategy at the national and MSA level, on average.

The remainder of the paper is outlined as follows: we review the literature surrounding cap rate forecasting and explanation as a function of GDP or growth metrics; CPI or inflationary metrics; or both variables together. Next we review the input data we used and the transformations we made. We then review our motivation for choosing a binary dependent variable and how we reduced the input variables to binary themselves. We specify a model and explain its validity by goodness of fit. The next section examines the results of our model, both at the national and MSA level. We review the results and show the strength of the model versus a buy-and-hold strategy before showing the cases (at the MSA-level) where the model works particularly well. Finally we conclude with suggestions for further research.

2 Literature Review

The literature on cap rate forecasting is extensive, and it can be viewed well through two different lenses: variable selection and methodology selection.

The first framework is illustrated in Ghysels et. al. Forecasting Real Estate Prices (Ghysels et al., 2013) which divides the literature by variable selection. The authors enumerate three camps: those models which use lagged return (or price change) as a variable, those which use ratios such as rent-to-price or price-to-income, and those which use more granular property or regional data.

The second lens by which to view the research is through methodology selection, as documented by Larriva and Linneman (Larriva and Linneman, 2021). The authors divide the research into simple time series methods, multivariate timeseries, and machine learning methods. The division is roughly chronological.

Neither lens quite suits this study's purpose, as the variables used in this work—raw GDP and CPI—are not often found in the research. More common is a spread to an inflation metric or a derivative of GDP. Separately, the level of granularity we seek to explain is somewhat uncommon (MSA level). More common is either property-level or national. Finally, the method we use is new to the research, to our review. The following section, then will seek to outline research that contains either growth or inflationary metrics, or their derivatives, regardless of method or granularity.

In the mid-1990s, research was concentrated on either single-variate autoregressive methods (Gau, 1984), (Gau, 1985), (Linneman, 1986) or variations on OLS that used multi-stage estimation (Case and Shiller, 1990), (Abraham and Hendershott, 1994). The variables used in these methods of forecasting tended to fall into classes that line up with the components of a cap rate: risk-free rate, real interest rates, relative return expectations, and risk. Many models' variables included a government long-bond to capture the risk-free and real rates, an economic spread (credit or pricing-based) to capture the relative expense or return of real estate, and macroeconomic factors to capture the growth prospects and risk in the marketplace. Some models included a term to capture forward estimates including either sentiment or expected rent or NOI growth. This lines up with the Gordon growth model of pricing cap rates as the discount rate less the perpetual NOI growth rate. Thus, the value of macroeconomic factors, or their derivatives, was established.

2.1 The Relationship between Inflation and Cap Rates

Inflation has long been researched with regard to cap rates (Froland, 1987); (Sivitanides et al., 2001); (Chandrashekaran and Young, 2000). Supporters suggest that inflation rates are inversely related to cap rates either via real rates, the weighted average cost of capital, or increased borrowing costs. While there are myriad reasons why the relationship should hold, there are just as many observable exceptions—cases when cap rates increase when interest rates decline, or the reverse. Understandably then, the research into the topic reaches mixed conclusions.

2.1.1 In Support of the Interest Rate and Cap Rate Relationship

Froland (Froland, 1987) reports negative correlations between cap rates and inflationary expectations. Sivitanides et. al. (Sivitanides et al., 2001) further supported and quantified this relationship at the MSA level. The authors use the average cap rates over the past 16 years of four property types across 14 metropolitan markets to examine how cap rates behave. And the statistical specification was estimated using a dual Time-Series Cross-Section method which corrects for cross section correlations and group-wise heteroskedasticity. Inflation was found to be a major driver of cap rates. In fact, Sivitanides et al. found that an expected increase in economy wide inflation of one percent annually lowers office cap rates by 46 basis points, multi-family cap rates by 40 basis points, retail cap rates by 54 basis points, and by 20 basis points in industrial cap rates.

Research from Wheaton et. al. (Wheaton et al., 2001) further substantiates the above work using a VAR model to show that inflation is a significant variable in determining cap rates at the MSA level.

2.1.2 In Opposition of the Interest Rate and Cap Rate Relationship

In contrast to the above works, Chandrashekaran and Young (Chandrashekaran and Young, 2000) find that there is not a statistically significant relationship between inflation and cap rates. To reach this conclusion Chandrashekaran and Young use two regression models: one with macroeconomic variables and one with lagged cap rates. The model with lagged cap rates uniformly performed better. The authors concluded that their attempt to predict cap rates using macroeconomic variables, such as inflation, were unsuccessful.

Gimpelevich (Gimpelevich, 2011) supports Chandrashekaran and Young using Monte Carlo simulations of real estate returns called the Simulation-Based Excess Return Model (SERM). The simulation results in poor correlation between inflation rates and cap rates (Gimpelevich, 2011).

More recently, Larriva and Linneman (Larriva and Linneman, 2021) used multivariate timeseries analysis to show the superiority of a model which did not use return expectations or inflationary metrics. Their model used granger causality to select variables and establish a highly robust explanatory and predictive VECM model, without the use of interest rates.

Although there are exceptions, different conclusions in this space can often be attributed to differences in data and methodology. Research which finds that inflation is a good predictor of future cap rates typically use time series data for both macroeconomic variables and cap rates while research which dismisses the value of inflation rates, generally, use cross sectional econometric methods.

We propose herein that perhaps the reason for the disagreement in the impact of inflation rates to cap rates is because only some models are including the vital second part of the relationship: GDP.

2.2 GDP

The growth or contraction of a nation or MSA's gross domestic product is likely the single most important indicator of economic health. That cap rates would be closely related to this metric is not surprising. And there is ample research which substantiates and explores this.

Granger causality between GDP and real estate markets was established in 1997 by Green (Green, 1997) who found that, under a wide variety of time series specifications, residential investment causes GDP, while non-residential investment is caused by GDP. The phrase "caused by" is used in the Granger sense in this context.

Quan and Titman (Quan and Titman, 1997) build on the Granger causality and seek to quantify the relationship. The team utilized time series regressions to examine the effects of changes in macroeconomic variables (including GDP) on real estate values and rents. To begin Quan and Titman explore the connection between stock and real estate market returns, and after establishing their connection the duo explore factors and develop regressions to explain this connection. The first theory they explore is that real estate and stock prices are both guided by future macroeconomic expectations such as GDP growth. The second is that commercial real estate prices rise and fall because of changing political and economic fundamentals, and under this theory the relationship between the stock and real estate market will be much weaker. After controlling for macroeconomic variance Quan and Titman find that the correlation between real estate and stock prices are primarily because of economic fundamentals, and that rental rates are strongly correlated with GDP growth.

Once this connection was established, research turned to its implication on an international scale: does worldwide GDP growth or local GDP growth have a larger effect on property values and cap rates? Case, et. al. (Case et al., 2000) explore whether or not correlations across global real estate markets are due to world changes in GDP, and estimate the value of local economic performance in real estate markets. This is an important question in real estate because all real estate is essentially local. And if local growth has a stronger effect on property values than national or global growth investors can better leverage local growth patterns to maximize return. The researchers found that that international property returns move together in dramatic fashion due to the effects of changes in GNP. Specifically, correlations of real estate are due in part to common exposure to fluctuations in the global economy, as measured by an equal-weighted index of international GDP changes. Some markets, such as Asia, are more effected by local changes rather than world changes in GDP. In other words, even though real estate is fundamentally local changes in the global economy carry enough weight to have significant effects in local real estate markets.

2.3 GDP and Inflation together

Studies that focus on macroeconomic variables often include both GDP (or its derivative) and Inflation (or its derivative). The results often suggest that one variable is significant while the other is not.

The aforementioned Quan and Titman research used cross sectional analyses to show that changes in GDP are very strongly related to movements in real estate values. But despite the significance of GDP, inflation and interest rates had relatively little effect on real estate values. Approaching the question with a timeseries analysis corroborated this conclusion. Quan and Titman found real estate values and rents were still significantly affected by changes in GDP. Inflation, again, appeared to be not significant.

Aizenman and Jinjarak make a somewhat contrary finding. In their 2009 work, (Aizenman and Jinjarak, 2009) they determine that the most economically significant variable in accounting for changes in real estate valuation is lagged real estate valuation appreciation (defined as real estate inflation minus CPI inflation), followed in importance by lagged declines of Current Account divided by GDP. Thus they note that an inflation derivative is more important than a GDP derivative.

2.4 Similar Models and Novelty

The study that is most comparable to our own is that of Duca and Ling (Duca et al., 2017) whose research finds that cap rates are positively correlated with inflation via risk premia, and negatively correlated with rent growth expectations via GDP. Specifically, risk premia and real Treasury rates drive cap rates. They explain this theoretically via the discount factor (i.e. required rate of return).

Upon review, we note the frequency with which research disagrees on the significance (ordinal or absolute) of GDP metrics and inflation metrics. Our work makes three significant contributions to the existing body of work.

First, we assert that neither GDP nor inflation variables are sufficient alone to explain cap rates. This can be seen by the above review. Simply using both as independent variables is not sufficient either. It is their interaction, we assert, that explains the most about cap rate directions.

Second, we posit that we can learn more about cap rate determinants if we view cap rates as binary instead of continuous. While, of course, cap rate moves are non-discrete, this is not an useful way to analyse them, as real estate is still highly illiquid. The relevant question then, is not “will cap rates increase by, say, 0.0125 next year?” but instead “will cap rates increase next year?” Thus we model cap rate moves as random variables generated by a Bernouilli distribution as opposed to random variables generated by, say the Normal distribution.

Finally, we propose a model that functions at the MSA and the national level. The more significant of these two is the former, because it determines the later. But the fact that the same data series are available at the nation and MSA and that they both corroborate the statistical findings is further attestation to the validity of the model.

3 Data

For the national model there are three data series that are used: the Consumer Price Index, the Gross Domestic Product, and the nominal Apartment cap rate series.

3.1 National

The Gross Domestic Product is a quarterly, unadjusted (not de-seasoned) series from the St. Louis Fed's economic data page (FRED). It is important to use a nominal (versus real) series, as we are comparing this to the consumer price index, which is a close proxy to the interest rate series used to convert the GDP from nominal to real. That is to say, we would be double counting the impact of inflation.

The Consumer Price Index series is sourced to the Bureau of Labor Statistics, and it is also a non-seasonally adjusted series. In both cases, we did not seek to alter the series with seasonal adjustments because the raw data itself contains information which we do not wish to strip. Knowing if the economy grew faster than inflation, even if it was in 4Q, and even if it was due to holiday sales, is valuable information. Such a situation suggests a very different environment from its seasonally adjusted counterpart which may show an economy in which 4Q growth is similar to say the prior quarter.

Finally, cap rate series at the national level are sourced to Green Street. Green Street defines its cap rate series as the next-twelve-months' NOI divided by the spot asset value.

3.2 Metropolitan

MSA-level data follows the conventions of the national data. Non-seasonally adjusted series are selected both for MSA GDP and the MSA's CPI. The former is from FRED while the latter is from the Bureau of Labor Statistics.

As national CPI is presented monthly, and national GDP is presented quarterly, we proceeded with quarterly data for the national model, pairing the CPI release most close to the GDP release.

More challenging was the MSA-level data wherein GDP is presented only annually and CPI is presented with varying frequencies depending on the region. Atlanta is published on even months. Boston is published on odd months, and New York is published monthly. Some MSAs are published bimonthly. The Bureau of Labor statistics only offers CPI measures for 22 Core Based Statistical Areas (CBSA). Two of these are "Urban Alaska" and "Urban Hawaii", but these have no equivalent GDP series, so they were discarded. The remaining 20 CBSAs have GDP series. But the GDP series are not necessarily for the exact same areas i.e., some GDPs are published for an MSA, some are published for a CBSA. Adding to this confusion: the BLS revised its reporting in 2018 and started offering CBSA measurements instead of MSA measurements, thus revising the coverage areas.

We attempt to reconcile this by matching MSA to CBSA where needed (between the BLS and FRED). And we average the measurements of CPI over a year before matching them to the GDP metrics. In this way, we end up with 21 geographies: 20 MSAs and 1 national geography. Each geography has a cap rate, CPI, and GDP.

As many practitioners assert that cap rates are functions of real interest rates, it is worth addressing the question, “is this simply ‘real interest rates’ by a different name?” It is not. Mathematically, the series have very low correlation. Practically, real rates contain an expectation of returns and an expectation of inflation, whereas GDP and CPI are merely spot empirical values.

4 Model Construction

4.1 Model Selection

While most analyses forecasting cap rates allow the random variable to be generated from a normal distribution, we opted instead to represent the variable as generated from a Bernoulli distribution. While this choice might be unorthodox—modelling a continuous variable as a binary variable—we argue that it is well suited for two reasons.

First, it is practical. Real estate is not an equity, it is an alternative asset. As such, it is illiquid. This lack of liquidity means an investor cannot enter and exit the market readily. As such, a would-be buyer is somewhat ambivalent between caring if a cap rate will go up by 40 basis points next year and if it will simply go up. In either case, next year’s prices will offer a better value. Similarly a would-be seller should be ambivalent between knowing if next year a cap rate will go up by 10 basis points or if it will go up by 100. In either situation, next year is a worse time to sell than this year. Contrast this with an equities trader who can easily place stop-losses or “cost-average” into a position. Such a trader cares very much whether his holdings will increase 10% next year, hence estimating liquid asset pricing is well suited by a continuous variable. But estimating illiquid asset pricing is, we argue, better done with a discrete variable.

Second, it maps to the binary nature of the motivation of the model: will GDP growth (contraction) exceed CPI growth (contraction)? It matters less the exact amount one exceeds or does not exceed the other. We argue that the question is simply: can a landlord pass on rising expenses to tenants or not? In the case where GDP is rising faster than inflation, we argue the tenants will be wealthier than the growth in expenses, and thus able to absorb higher rents, creating an asset worth more. Conversely, when inflation outpaces the growth of the economy, we argue the tenant is not in a position to absorb higher costs, leaving the landlord to bear them, creating an asset worth less.

Thus a logistic regression with a single binary input lent itself well to our analysis, with cap rate expansion as the response variable, and GDP change > CPI change as the independent variable.

4.2 Model Specification

We preprocessed the GDP and CPI series differently for the national versus MSA model because the national data was more frequent. In both cases, the preprocessing included differencing the series and then creating a simple trailing average. The goal in smoothing the series was to prevent an overly sensitive signal if, say, one period had inflation that quickly reversed.

CPI and GDP percentage changes at t_0 were compared to trailing average historical percentage changes. A signal was generated when 1) the t_0 GDP change was less than its trailing average changes and 2) the t_0 CPI change was greater than its trailing averages. The signal suggests cap rate expansion in the next period.

The cap rate series were simply differenced and their sign taken to be a boolean variable such that cap rate expansion is TRUE and cap rate non-expansion (and contraction) is FALSE.

To establish the validity of the model for predictive power, we used a training and testing data set.

For the MSA model, we trained five separate models (Year = 2015, 2016....2020) which were specified using only data prior to that date. The each of the five models were used to predict the next years' (Years 2016-2021, 2017-2021,...2021) cap rate expansions and contractions based on the signal series outlined above. Each of these five models were specified as one—that is the data from each MSA was stacked into one matrix, absent the name of the MSA, and used to specify the model. The model then predicted the future values for each MSA based only on that MSA's historical data. Thus the same estimators were used to forecast the cap rate activity of the out-of-sample periods. Only the test data changed (was limited to only one MSA).

National models were created every two years from 2000, as data is both more frequent and longer lived than MSA-level data.

Then, to establish the validity of the models for explanatory power, we trained both the MSA model and the national models using all time periods available (in-sample).

The ground truth of cap rate expansion does not present a balanced set, as cap rate compressions are far more frequent than cap rate expansions (roughly two to three times more years see compressions than expansions). The risk in modelling an unbalanced set is the risk of converging to an useless model that always forecasts the more common occurrence (i.e., always forecasting cap rate compression independent of the input). We overcame the imbalance through a Synthetic Minority Over-sampling Technique as created by (Chawla et al., 2002). This balanced the frequency of cap rate expansions and contractions to prevent a the logistic regression from estimating all compressions (which would produce a low-variance high-bias estimator).

Thus our model can intuitively be understood in terms of the GDP and CPI historical series

$$GDP = \{g_i\}_{i=0}^N$$

$$CPI = \{c_i\}_{i=0}^N$$

...and their differenced series

$$GDP' = \{g_i - g_{i-1}\}_{i=1}^N$$

$$CPI' = \{c_i - c_{i-1}\}_{i=1}^N$$

...and their respective n-period trailing averages

$$h_i = \frac{1}{n} \sum_{j=i-n}^i GDP'_j$$

$$d_i = \frac{1}{n} \sum_{j=i-n}^i CPI'_j$$

...such that comparing the current delta to prior averages determines cap rate expansion in the following way

$$\text{Cap Rate Expansion Variable}_{t_i+1} = \begin{cases} 0 & \text{if } GDP'_i - h_i > CPI'_i - d_i \\ 1 & \text{if } GDP'_i - h_i < CPI'_i - d_i \end{cases}$$

Where

$$\text{Cap Rate Expansion} \sim \mathcal{B}(n = 1, p = 0.77)$$

And p is given by the solution (coefficient) of the logistic regression:

$$p = \frac{1}{1 + e^{(-\beta_1)(CREvariable)}}$$

$$p = \frac{1}{1 + e^{(-1.25)(1)}}$$

$$p \approx 0.77$$

4.3 Model Goodness of Fit

In both the National and the MSA models, the logistic regressions displayed consistent coefficients, inline with our hypothesis. The coefficients had low standard errors relative to the coefficients' magnitude, and the 2.5% to 97.5% bounds did not include zero in any case.

Where a year is listed in the OOS in Year column, the data was segmented into training and test sets. The training periods were all prior to the split year, and the test set was the split year through 2021. When “all-time” is listed under the Split-Year, the model was trained on all data and evaluated on all available data.

The columns, “Coef.”, “Std.Err.”, “Z”, “ $P > |z|$ ”, and the 2.5% to 97.5% bounds are all populated with the parameters from the trained models using the training data. The columns to the right, “accuracy” through “specificity” are based on the test data.

It is worth noting that the coefficient decreased consistently as time advances, suggesting that perhaps the signal is not as robust as it once was. However the standard error decreases correspondingly, producing a Z score which is as high as the earlier periods. The decrease in the coefficient (corresponding to a lower probability of cap rate expansion for the CRE variable) could be due to the difficulty of trying to incorporate the GFC and the unorthodox interest rate behavior of the COVID-19 era.

Goodness of Fit: National Model

OOS in Year	Coef.	Std.Err.	Z	$P > z $	[0.025	0.975]
2000	2.2	1.05	2.08	0.0371	0.13	4.26
2004	1.7	0.77	2.22	0.0266	0.2	3.21
2008	1.25	0.57	2.21	0.0271	0.14	2.36
2012	1.34	0.46	2.93	0.0034	0.45	2.24
2016	1.19	0.43	2.76	0.0059	0.34	2.04
all-time	1.25	0.36	3.49	0.0005	0.55	1.96

Table 1: Relevant parameters and metrics from the logistic regression trained on national data, both in and out of sample

Goodness of Fit: MSA-level Model

OOS in Year	Coef.	Std.Err.	Z	$P > z $	[0.025	0.975]
2015	1.13	0.26	4.3	0.0000	0.62	1.65
2016	1.12	0.26	4.22	0.0000	0.6	1.63
2017	1.02	0.25	3.99	0.0001	0.52	1.52
2018	0.83	0.22	3.85	0.0001	0.41	1.25
2019	0.85	0.21	4.07	0.0000	0.44	1.26
in-sample	0.77	0.18	4.3	0.0000	0.42	1.12

Table 2: Relevant parameters and metrics from the logistic regression trained jointly on all MSA data, both in and out of sample

5 Results

Upon implementing the models, results were evaluated using a confusion matrix, as the period we are attempting to capture—cap rate expansion—is far less frequent than cap rate contraction.

Accuracy ($TP+TN/(TP+FP+TN+FN)$) can be understood as the portion of time the model correctly identified a cap rate expansion or contraction. At the national level, cap rate expansion has happened 29.5% of quarters since 2000. The national model then is no more accurate than forecasting constant cap rate contraction, which would have yielded an accuracy of 70.5%. This is not true at the MSA level, where cap rates have expanded 27% of the time (compressed 73% of the time) and the model yielded only 68% accuracy. That is, it would have done better by forecasting constant cap rate compression (yielding an accuracy of 73%).

Those subtleties are why we further examine precision, recall, and specificity. Precision ($TP/(TP+FP)$) is the ratio of correctly predicted positive observations to total predicted positive observations. The national model achieves 67% precision, meaning two-thirds of the time a signal is generated, it is accurate. The jointly trained MSA model had relatively weak precision, out-of-sample, but surged to 68% in-sample.

Recall ($TP/(TP+FN)$) is the ratio of correctly identified cap rate expansions to the total number of cap rate expansions. In many MSAs there were only three or six years of cap rate expansions. Compare this to the national model which had far more quarters of cap rate expansion. The recall of the national model was high at 0.41, meaning 41% of the times when cap rates expanded, the model correctly forecasted them. The recall of the MSA model was equal.

The F1-score combines the precision and recall into their harmonic mean, and is valuable to consider precision and recall together, though it is less intuitive.

Finally, the specificity ($TN/(TN+FN)$) of both models was most noteworthy. A high specificity value suggests that the periods where cap rate compression was forecasted were highly accurate. This is especially useful in a series where cap rate compressions are profitable and are more common than expansions. These periods should not be missed, and indeed, the forecasts generated by both models would allow an investor to participate in upward of 80% of the cap rate compressions experienced.

Confusion Matrix: National Model

OOS in Year	accuracy	precision	recall	F1-score	specificity
2000	0.76	0.55	0.5	0.52	0.85
2004	0.75	0.53	0.5	0.51	0.84
2008	0.77	0.54	0.54	0.54	0.85
2012	0.78	0.56	0.56	0.56	0.85
2016	0.75	0.5	0.6	0.55	0.8
all-time	0.7	0.67	0.41	0.51	0.88

Table 3: A confusion matrix showing the values of the national model with both train/test split, and in-sample data

Confusion Matrix: MSA Model

OOS in Year	accuracy	precision	recall	F1-score	specificity
2015	0.65	0.16	0.25	0.19	0.73
2016	0.63	0.16	0.33	0.21	0.68
2017	0.61	0.11	0.33	0.16	0.65
2018	0.67	0.06	0.17	0.09	0.72
2019	0.62	0.07	0.33	0.12	0.65
in-sample	0.68	0.46	0.4	0.43	0.79

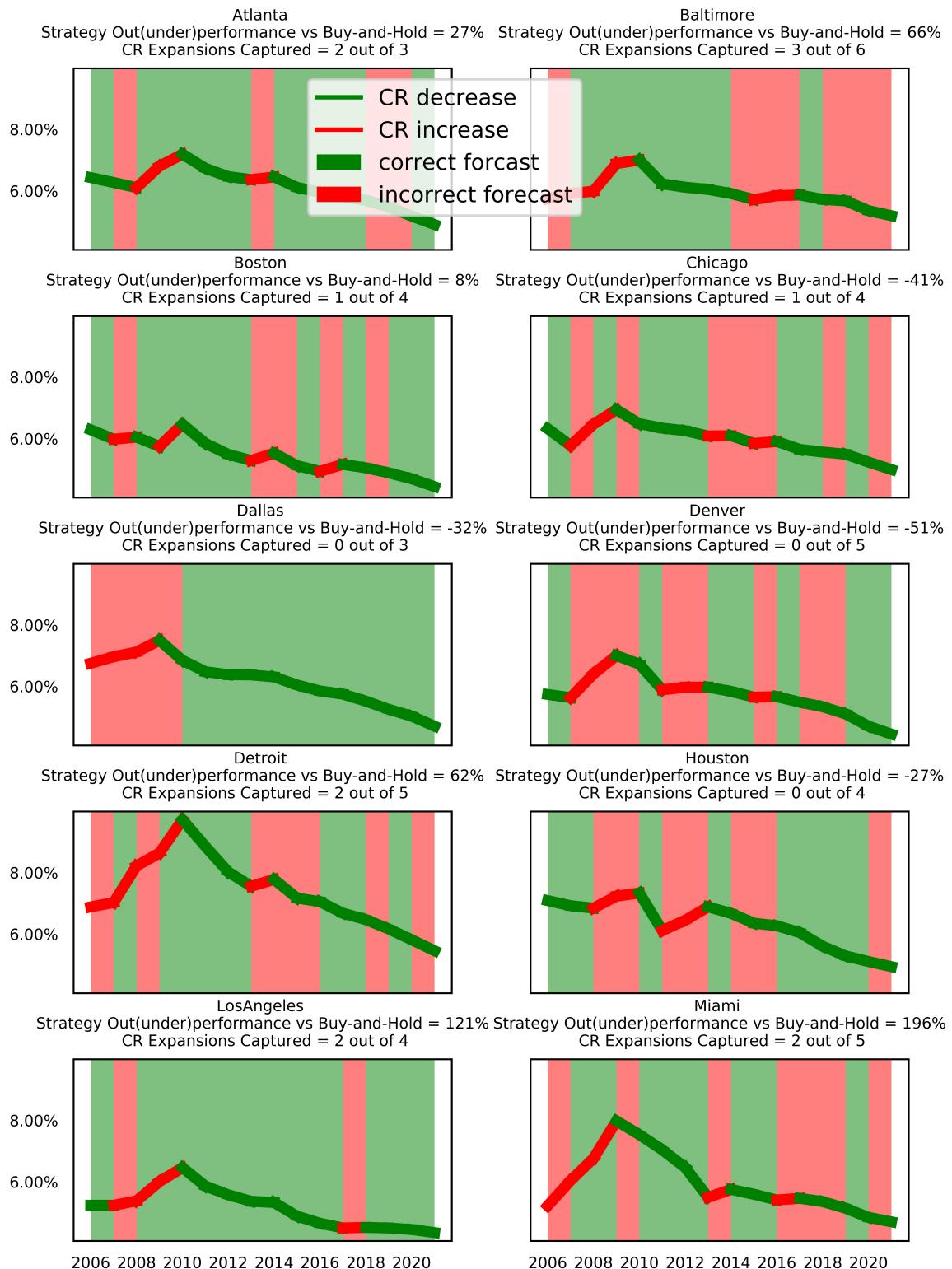
Table 4: A confusion matrix showing the values of the MSA-level model with both train/test split, and in-sample data

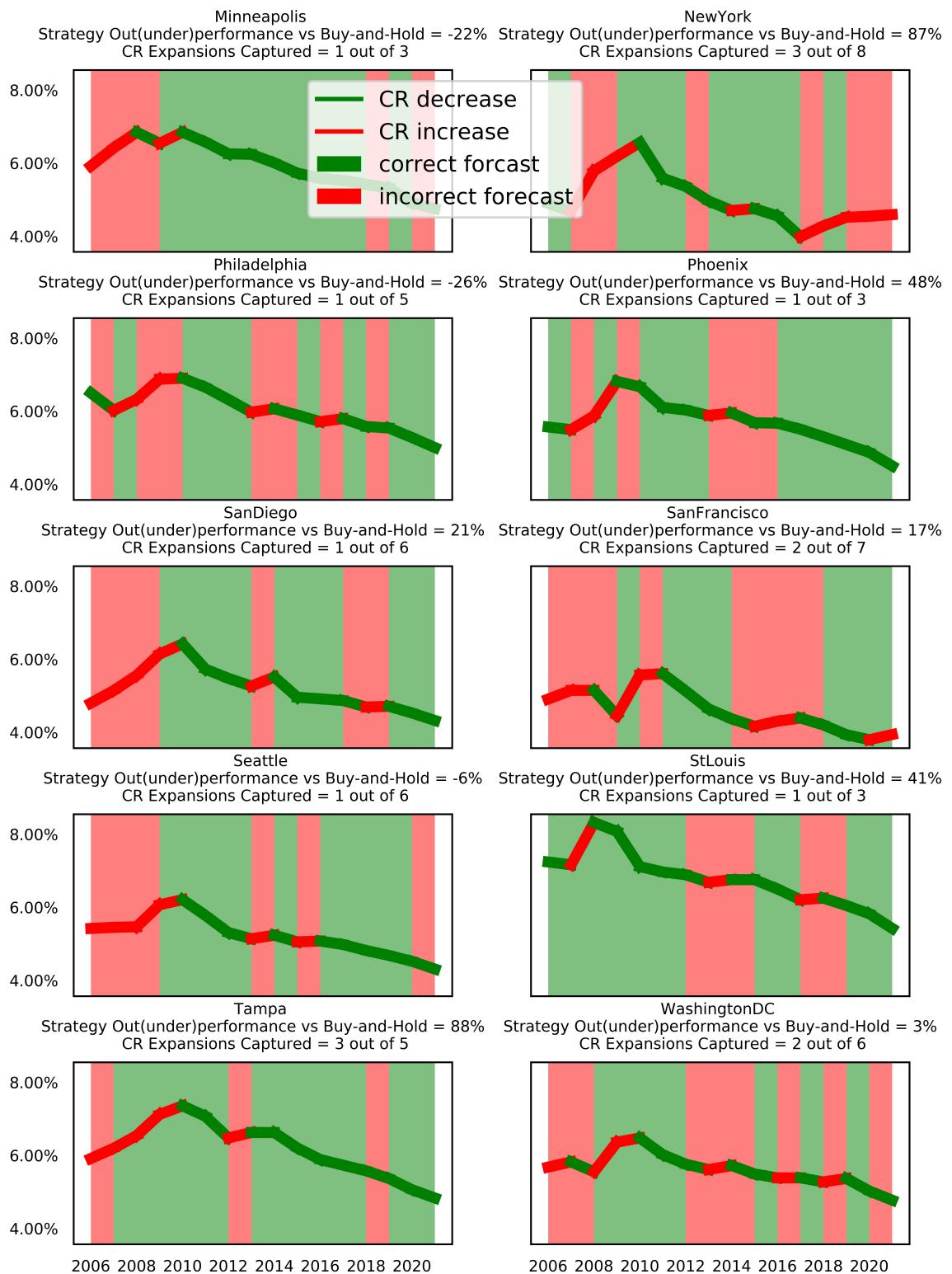
Still, despite the four metrics of the confusion matrix, there remains the question of efficacy in practice. Because entering or exiting real estate holdings is not done easily or without great transactions costs, these models would only be useful to the extent they significantly outperformed a buy-and-hold strategy.

To examine this we implemented the national and MSA models' recommendations and yielded the following results.

At the national level, implementing the recommendations made by the model would have yielded cap rate compression 13% greater than a buy-and-hold strategy. This would have happened in part because of the avoidance of sharp periods of cap rate expansion of the 2005-2009 period, and the late '80s.

At the MSA-level, implementing the recommendations would have resulted in significant outperformance of 29% relative to a buy-and-hold strategy. But the variance is quite high. In Miami, implementing the model would have led to outperformance of 196%, while in Baltimore, implementing the model would have resulted in underperformance of 66%. Understanding these results is the focus of the next section.







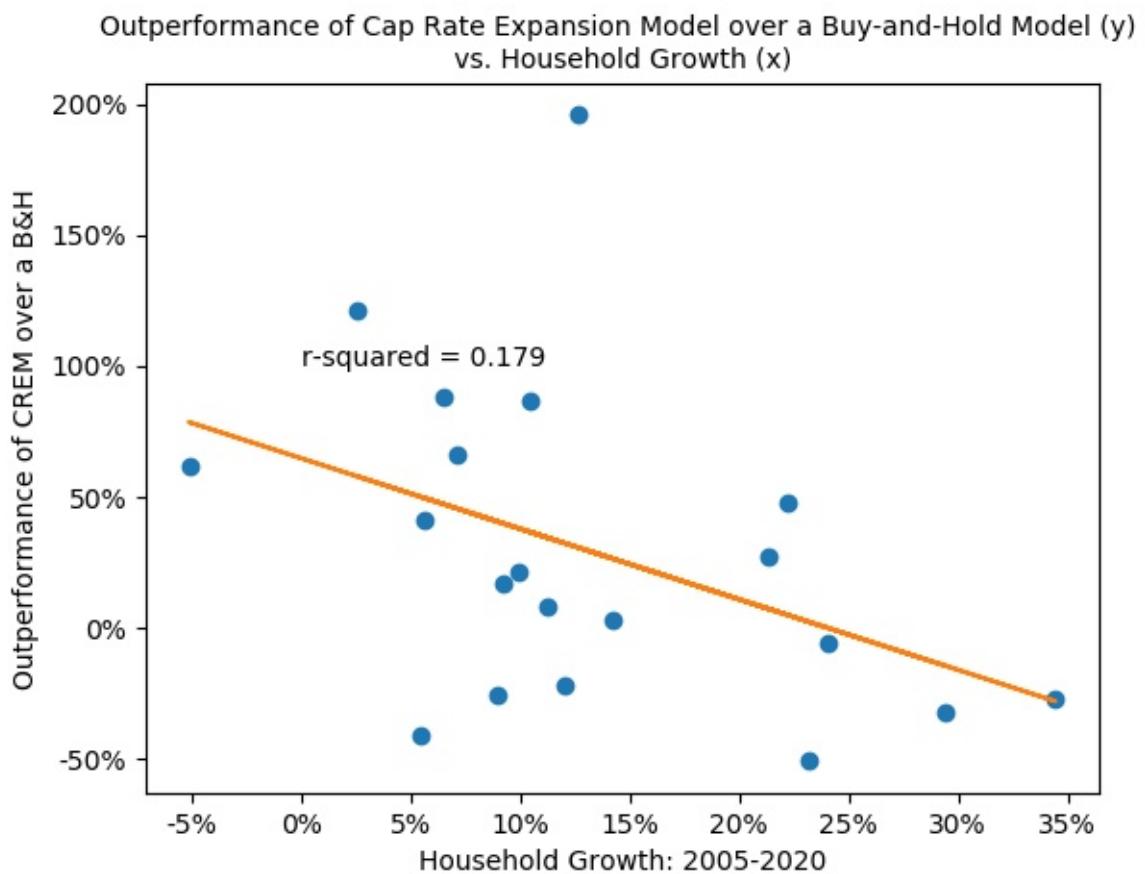
5.1 Analysis of Results

The more robust predictor, in all cases, was the national model. This model, both in and out of sample, managed to forecast cap rate expansions accurately, with few false negatives and false positives. The model benefited from additional data points: both more history and more frequent readings.

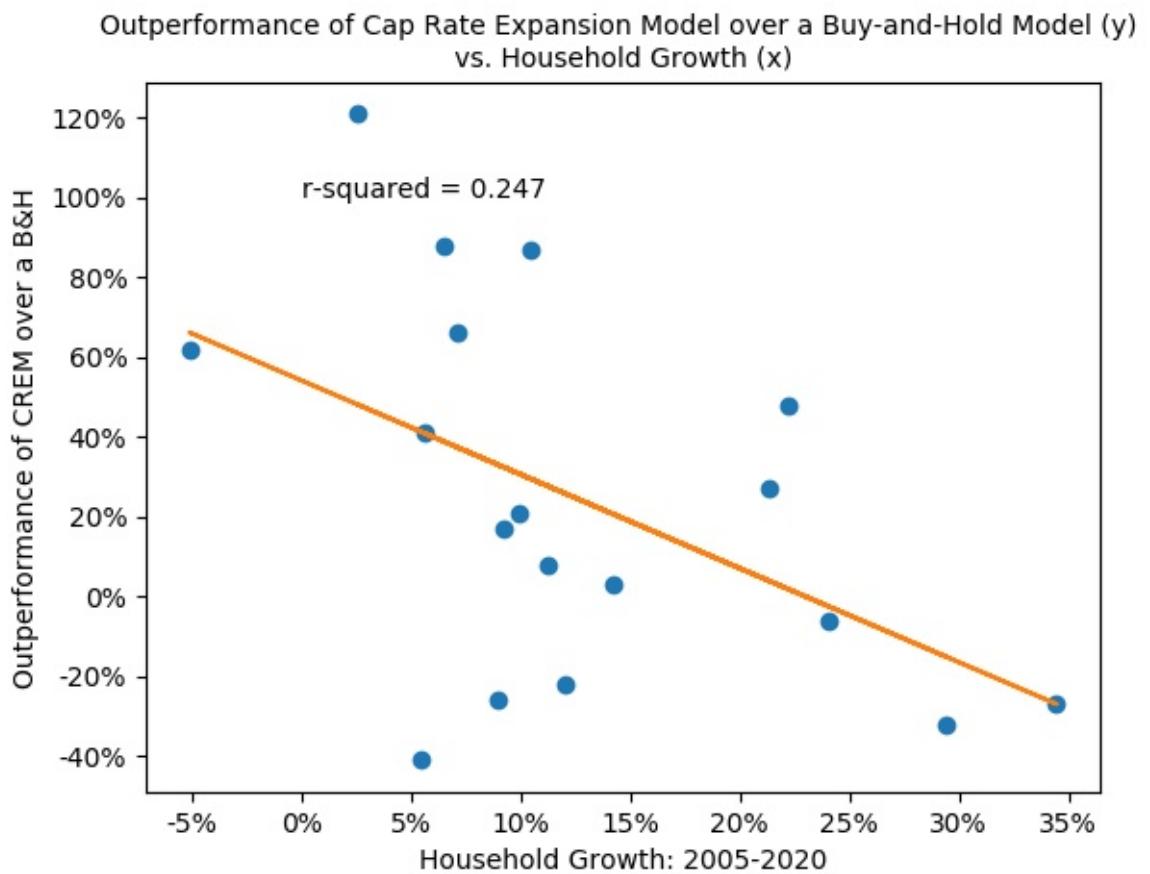
The MSA model that was jointly trained was relatively weak, which is likely owing to two factors. First, there is simply not enough data available. There were only 20 years of GDP and CPI data available consistently, and they were only reported annually. The model, however should not be disregarded, as data is becoming more widely distributed, with greater frequency, and with higher accuracy.

Upon further review, we noted that it seemed the MSAs wherein the model performed well were all established or low-population-growth cities, while the MSAs where the model performed poorly seemed to generally have higher population growth.

Viewing these results in a regression, the results are more pronounced and show that household growth is significantly correlated with model underperformance, thus suggesting this model is very well suited for MSAs with national-average-like population growth.



Removing the leverage point (Miami 196% outperformance and 13% household growth) yields an even tighter relationship.



And this makes intuitive sense: that a market with average-to-low organic growth would have its real estate values dictated by the interplay between inflation growth and GDP growth. Conversely, an area which was experiencing sharp population growth would have these dynamics overridden by the outsized demand for real estate.

6 Conclusion

Declarations

6.1 Funding

This research is a part of one author's role as VP of Research and Data Analytics at a Real Estate Private Equity firm.

6.2 Conflict of interest

One author works for a Real Estate Private Equity firm which has ownership interest in many office and multifamily assets throughout the US.

6.3 Availability of data and material

Data available upon request.

6.4 Code availability

Code available upon request.

6.5 Authors' contributions

Each of the authors confirms that this manuscript has not been previously published and is not currently under consideration by any other journal.

References

- Abraham, J. and Hendershott, P. (1994). Bubbles in metropolitan housing markets. *National Bureau of Economic Research*.
- Aizenman, J. and Jinjarak, Y. (2009). Current account patterns and national real estate markets. *Journal of Urban Economics*, 66(2):75–89.
- Case, B., Goetzmann, W. N., and Rouwenhorst, K. G. (2000). Global real estate markets-cycles and fundamentals.
- Case, K. and Shiller, R. (1990). Forecasting prices and excess returns in the housing market. *Real Estate Economics*.

- Chandrashekaran, V. and Young, M. S. (2000). The predictability of real estate capitalization rates. In *Annual Meeting of American Real Estate Society, California*.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357.
- Duca, J. V., Hendershott, P. H., and Ling, D. C. (2017). How taxes and required returns drove commercial real estate valuations over the past four decades. *National Tax Journal*, 70(3):549–584.
- Froland, C. (1987). What determines cap rates on real estate? *The Journal of Portfolio Management*, 13(4):77–82.
- Gau, G. W. (1984). Weak form tests of the efficiency of real estate investment markets. *The Financial Review*, 19(4):301–320.
- Gau, G. W. (1985). Public information and abnormal returns in real estate investment. *Real Estate Economics*, 13(1):15–31.
- Ghysels, E., Plazzi, A., Valkanov, R., and Torous, W. (2013). Forecasting real estate prices. *Handbook of Economic Forecasting*, page 509–580.
- Gimpelevich, D. (2011). Simulation-based excess return model for real estate development: A practical monte carlo simulation-based method for quantitative risk management and project valuation for real estate development projects illustrated with a high-rise office development case study. *Journal of Property Investment Finance*, 29:115–144.
- Green, R. K. (1997). Follow the leader: how changes in residential and non-residential investment predict changes in gdp. *Real estate economics*, 25(2):253–270.
- Larriva, M. and Linneman, P. (2021). The determinants of capitalization rates: Evidence from the us real estate markets. *Journal of Property Investment & Finance*.
- Linneman, P. (1986). An empirical test of the efficiency of the housing market. *Journal of Urban Economics*, 20(2):140–154.
- NAREIT (2022).
- NMHC (2022). Quick facts: Investment returns on apartments.
- Quan, D. C. and Titman, S. (1997). Commercial real estate prices and stock market returns: An international analysis. *Financial Analysts Journal*, 53(3):21–34.
- Sivitanides, P., Southard, J., Torto, R. G., and Wheaton, W. C. (2001). The determinants of appraisal-based capitalization rates. *Real Estate Finance*, 18(2):27–38.
- Wheaton, W. C., Torto, R. G., Sivitanides, P. S., Southard, J. A., Hopkins, R. E., and Costello, J. M. (2001). Real estate risk: a forward-looking approach. *Real Estate Finance*, 18(3):20–28.
- Williams, O. A. (2021). World's wealth hits half a quadrillion dollars.