



UvA-DARE (Digital Academic Repository)

Forecasting US Commercial Property Price Indexes Using Dynamic Factor Models

van de Minne, A.; Francke, M.; Geltner, D.

DOI

[10.1080/08965803.2020.1840802](https://doi.org/10.1080/08965803.2020.1840802)

Publication date

2022

Document Version

Final published version

Published in

Journal of Real Estate Research

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

van de Minne, A., Francke, M., & Geltner, D. (2022). Forecasting US Commercial Property Price Indexes Using Dynamic Factor Models. *Journal of Real Estate Research*, 44(1), 29-55. <https://doi.org/10.1080/08965803.2020.1840802>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)



Forecasting US Commercial Property Price Indexes Using Dynamic Factor Models

Alex van de Minne, Marc Francke & David Geltner

To cite this article: Alex van de Minne, Marc Francke & David Geltner (2022) Forecasting US Commercial Property Price Indexes Using Dynamic Factor Models, Journal of Real Estate Research, 44:1, 29-55, DOI: [10.1080/08965803.2020.1840802](https://doi.org/10.1080/08965803.2020.1840802)

To link to this article: <https://doi.org/10.1080/08965803.2020.1840802>



© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 09 Dec 2021.



Submit your article to this journal [↗](#)



Article views: 1162



View related articles [↗](#)



View Crossmark data [↗](#)

Forecasting US Commercial Property Price Indexes Using Dynamic Factor Models

Alex van de Minne^a, Marc Francke^{b,c}  and David Geltner^d

^aDepartment of Finance, School of Business, University of Connecticut, Storrs, CT, USA; ^bFinance Group, Amsterdam Business School, University of Amsterdam, Amsterdam, The Netherlands; ^cOrtec Finance, Amsterdam, The Netherlands; ^dCenter for Real Estate, Massachusetts Institute of Technology, Cambridge, MA, USA

ABSTRACT

The general purpose of a dynamic factor model (DFM) is to summarize a large number of time series into a few common factors. In this paper we explore several DFMs on 80 granular, non-overlapping commercial property price indexes in the US, quarterly from 2001Q1 to 2017Q2. We examine the nature and the structure of the factors and the index forecasts that can be produced from the DFMs. We consider specifications of one to four common factors. As a major motivation for the use of DFMs is their ability to improve out-of-sample forecasting of systems of numerous related series, we apply the DFM estimated factors in an Autoregressive Distributed Lag (ARDL) model to forecast individual market index returns. We compare for four markets the forecasts to those from a benchmark univariate autoregression. The results show that the DFM & ARDL model predicts the crisis and subsequent recovery really well, whereas the benchmark model typically extrapolates the past price trend.



KEYWORDS

Asset price dynamics; co-movement; EM algorithm; state space model

Introduction

Having reliable commercial property price index forecasts is obviously important from a practical perspective. Forecasts are important for tactical level portfolio management, where and what to buy and sell in the intermediate term. If (or when) property price derivatives become important, forecasting is essential for pricing real estate derivatives (Geltner & Fisher, 2007). Reliable forecasts are also essential for policymakers given the amount of wealth invested in commercial real estate, for example by pension funds.

It is therefore no surprise that previous literature has taken interest in forecasting property prices. In most cases, forecasts are based on a vector autoregressive model (VAR) or an autoregressive distributed lag (ARDL) model. The main difference between a VAR and ARDL model, is that a VAR model can explicitly model reverse-causality, unlike an ARDL model. In both cases the price index is regressed on its own lags, and lags of

CONTACT Marc Francke  m.k.francke@uva.nl  Finance Group, Amsterdam Business School, University of Amsterdam, Plantage Muidergracht 12, 1018 TV Amsterdam, P.O. Box 15953, Amsterdam 1001, The Netherlands
This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

other explanatory variables. The choice of which explanatory variables to include in such models varies widely across studies and is highly debatable. Explanatory variables that are commonly used in real estate literature include: Inflation, mortgage interest rates, unemployment rate, real per capita income, real construction costs, tax rate and new supply of real estate. One additional issue with the choice of explanatory variables, is that one cannot select too many of them, especially with small T . The loss of degrees of freedom can deteriorate the forecasts, *even if* the (many) explanatory variables are all drivers of the forecasted variable in itself. Examples of ARDL and VAR models used in real estate literature can be found in Malpezzi (1999), Case and Shiller (1990), and Anglin (2006).

This paper presents a different approach to forecasting, which circumvents altogether the discussion on what explanatory variables to use, by first estimating a dynamic factor model (DFM). Such models summarize a large number of related time series into a small number of factors common to the original series. A DFM describes temporal variation in a set of N observed time series that are related or reflect a common system—like GDP, interest rates, employment and such macro-economic variables, or in our case granular commercial property price indexes (CPPIs) themselves—by linear combinations of $M \ll N$ common latent factors which are identified and estimated from the data. As a result we can summarize big quantities of N time series, into a small number of M common factors.¹

Subsequently, these estimated common factors can be used to forecast the original time series themselves, as the common factors may be used as explanatory variables in an ARDL model. Such a model is referred to as a factor augmented ARDL. However, it should be noted that previous literature found conflicting results on whether factor augmented ARDLs actually improve forecasting or not (for a meta-analysis, see Eickmeier and Ziegler (2008), although real estate has not been explored yet).

In our application the granular commercial property price indexes which are our raw material represent non-overlapping commercial property asset submarkets reflecting different property types in different metropolitan areas. In this context, we theorize that DFMs *should* improve forecasting. As noted earlier, DFMs extract a few common—but unobserved—factors from large datasets. The basic idea is that granular real estate markets or segments tend to co-move with each other (Geltner et al., 2014, pp. 554–556). This co-movement may reflect an aspect of the nature of risk in property asset markets. On the one hand there exists idiosyncratic price movement specific to individual assets or granular market segments, largely reflecting the space markets. But just as important, asset-valuation risk reflects changes over time in the capital market that cause changes in the opportunity cost of capital. Time variation in the discount rate causes at least as much volatility in prices (see for example Geltner & Mei, 1995). Such capital market based price movements may have more common elements across space market based granular market segments (van Dijk et al., 2020). We thus believe that the common latent factors should add explanatory power to model real estate property prices.

Real Capital Analytics provided us with the 80 non-overlapping quarterly CPPIs between 2001Q1 and 2017Q2. We observe 40 regions and have a commercial (combined office, retail and industrial) and rental apartment index per region. The regions include 26 metropolitan areas (such as Jacksonville, Boston, and Chicago), where many of them split up into central and non-central areas.

Our back-testing results show that the factor augmented ARDL model almost always outperforms simpler autoregressive models. The factor augmented ARDL was especially accurate in predicting the timing of the Great Financial Crisis (GFC) starting in 2007. The factor augmented ARDL also predicted the subsequent recovery, but in most cases too soon (and too much recovery). Interestingly enough, we found that the factor augmented ARDLs that performed best overall were the models that only include one common factor (we allow for up to four). This motivated us to introduce a new forecasting model denoted the Simplified DFM (SDFM). With the SDFM we simply add the national commercial property price index (with lags) to the ARDL. This national index is also provided to us by RCA. The SDFM performs as well, if not better, compared to the DFM specification of the ARDL.

Our findings are of interest not only for the overall forecast performance results, but also for the structural insights they provide about commercial property asset price dynamics. This is of interest in its own right because it provides insights into the extent and nature of commercial real estate market co-movement. It tells which CPPIs tend to move with which factors, which can be of interest to investors for portfolio management and diversification.

This paper proceeds as follows. Section “Model” describes the DFM used in this research, and the forecast models. Section “Data” gives a description of the data. The estimated factors and loadings from the DFM are given and discussed in Section “Results from Dynamic Factor Model.” Section “Results from the Forecasting Models” gives the forecast results and Section “Conclusion” concludes.

Model

The Dynamic Factor Model

The dynamic factor model reduces many (N) observed time series into a few (M) common latent factors, where $M \ll N$. Let \mathbf{y}_t denote the N -vector of observed time series, and $\boldsymbol{\mu}_t$ the $(M \times 1)$ vector of latent factors, both at time $t = 1, \dots, T$. The dynamic factor model is given by

$$\mathbf{Y}_t = \boldsymbol{\Gamma} \boldsymbol{\mu}_t + \boldsymbol{\alpha} + \boldsymbol{\epsilon}_t, \quad (1)$$

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_{t-1}, \quad (2)$$

where $\boldsymbol{\Gamma}$ is a $(N \times M)$ matrix of factor loadings, $\boldsymbol{\alpha}$ is a $(N \times 1)$ vector of constants, and $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\eta}_t$ are independent and identically distributed error terms. We assume that $\boldsymbol{\epsilon}_t \sim N(0, \mathbf{R})$, and $\boldsymbol{\eta}_t \sim N(0, \mathbf{Q})$ with initial condition for the factors $\boldsymbol{\mu}_0 \sim N(\mathbf{m}_0, \mathbf{V}_0)$. The DFM is an example of a linear Gaussian state space model, which is conditional on the unknown parameters easy to process in a Kalman filter framework, see for example Harvey (1989). The unknown parameters in the model are the elements of $\boldsymbol{\Gamma}, \mathbf{R}, \mathbf{Q}, \boldsymbol{\mu}_0$ and \mathbf{V}_0 and are usually referred to as hyperparameters. We use the Kalman filter and estimate the hyperparameters by the Expectation-Maximization (EM) algorithm. The advantage of this methodology is that it does not require long time series. It also converges relatively fast for large numbers of time series, compared to some of the other methodologies.

In our application we have the log price index returns as observed time series, so $Y_t = \Delta \mathbf{p}_t$, where \mathbf{p}_t is the vector of log price indexes at time t . We model log index returns to account for non-stationarity; log price indexes are typically integrated of order one.

Identification of the Dynamic Factor Model

In order to identify the parameters in Γ , α and \mathbf{Q} in Eqs. (1)–(2), one need to make additional assumptions (see Harvey, 1989, Chapter 4.4). There is substantial literature on how to identify the model parameters (Aguilar & West, 2000; Aßmann et al., 2016; Bai & Wang, 2015; Forni et al., 2000; Geweke & Zhou, 1996; Harvey, 1989; Stock & Watson, 2002). We use the constraints introduced by Zuur et al. (2003), who largely follows Harvey (1989) (Chapter 5.8.1), but with one crucial difference to make the estimation more robust in the EM framework.

Firstly, we set γ_{ij} in Γ to zero if $j > i$. Thus, the upper right corner of matrix Γ fully consists of zeros,

$$\Gamma = \begin{bmatrix} \gamma_{11} & 0 & \cdots & \cdots & 0 \\ \gamma_{21} & \gamma_{22} & \ddots & & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots \\ \gamma_{M-1,1} & \gamma_{M-1,2} & \cdots & \gamma_{M-1,M-1} & 0 \\ \gamma_{M,1} & \gamma_{M,2} & \cdots & \gamma_{M,M-1} & \gamma_{M,M} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \gamma_{N,1} & \gamma_{N,2} & \cdots & \gamma_{N,M-1} & \gamma^M \end{bmatrix} \quad (3)$$

Secondly, one also have to constrain α in a meaningful way. Harvey (1989) suggests to set the first M values to zero. However, Zuur et al. (2003) found that the EM estimates are not robust to this constraint, and takes long to converge. We therefore simply demean the observed time series, and omit α altogether. We further standardize the time series by dividing by its own standard deviation (Bernanke et al., 2005). This step is not necessary for identification, but does increase the efficiency (and speed of convergence) of the EM algorithm.

Thirdly, we set \mathbf{Q} equal to a diagonal matrix, implying uncorrelated innovation errors. Finally, following Zuur et al. (2003) the initial condition of the state vector is specified by a diffuse prior, $\mu_0 \sim N(0, \kappa \mathbf{I})$, where κ is large, and \mathbf{I} denotes the identity matrix. Implementing these restrictions and initial conditions within the EM algorithm is not trivial (Wu et al., 1996). Please consult Zuur et al. (2003) for the exact algorithm used in this paper, which we will omit here for space.²

Note that by implementing these restrictions a unique solution for the factor loadings exists, with one huge caveat. The caveat is that the factors depend on the ordering of the series; the first factor is determined by the first time series, the second factor by the first two time series, and so on. Other solutions therefore exist by simply ordering the time series differently (Harvey, 1989, p. 450).

However, once the parameters have been estimated for a specific ordering of the time series, a factor rotation can be applied to the estimated factor loadings and factors. Many rotation techniques exist, like oblimin, othomax, and quterartimax (see Basilevsky, 1994; Bernaards & Jennrich, 2005; Browne, 2001; Harman, 1976, for overviews).

We use the varimax rotation. Varimax maximizes the sum of the variances of the squared factor loadings. More specifically, we introduce the $(M \times M)$ rotation matrix \mathbf{H} , and rewrite the DFM as

$$\mathbf{y}_t^* = \mathbf{\Gamma}\mathbf{H}^{-1}\boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t, \tag{4}$$

$$\mathbf{H}\boldsymbol{\mu}_t = \mathbf{H}\boldsymbol{\mu}_{t-1} + \mathbf{H}\boldsymbol{\eta}_{t-1} \tag{5}$$

The varimax rotation subsequently seeks a a rotation matrix \mathbf{H} , that creates the largest difference between the loadings. Varimax is (arguably) the most widely used technique for rotation, and most statistical software packages have built-in functions for the varimax rotation.

To reduce computing time we will only focus on restricted versions of the variance-covariance matrix \mathbf{R} of the error term $\boldsymbol{\epsilon}_t$ in Eq. (4): (1) diagonal and equal, (2) diagonal and unequal, and (3) equal variance and covariance.

Forecasting

When having a large number of time series and/or a small number of time periods, vector autoregressions become infeasible due to the large number of unknown parameters relative to the number of observations to be estimated. In this setting the DFM can effectively be used to forecast individual time series. The general idea of DFMs is that the bulk of variation of many variables can be captured by a small number of factors. DFMs exploit the variables' co-movement and efficiently reduce the dimension of the dataset to just a few factors. More specifically, a relatively small number of factors are entered into a forecasting equation to forecast the individual time series, and only a few parameters need to be estimated.

Unsurprisingly, forecasting time series (like GDP) using DFMs has gained in popularity in the last decade or two. For example Eickmeier and Ziegler (2008) summarize 52 papers in which DFMs are used for forecasting inflation and output for a variety of countries. Most of these (and other) studies use the factors in an Autoregressive Distributed Lag (ARDL) framework (Barhoumi et al., 2013; Eickmeier & Ziegler, 2008). The ARDL is given by

$$y_{i,t}^* = \beta_0 + \sum_{m=1}^M \sum_{k=1}^{r_m} \beta_{mk} \hat{\mu}_{m,t-k} + \sum_{k=1}^{l_i} \phi_k y_{i,t-k}^* + \zeta_{i,t} \tag{6}$$

where $\hat{\mu}_{m,t}$ denote the estimated factors from Eqs. (4) to (5), and the error term is normally distributed, $\zeta_{i,t} \sim N(0, \sigma_\zeta^2)$. Conditional on the number of lags (r_m and l_i), the model can be estimated by ordinary least squares, for each time series separately. The number of lags is usually determined by the (penalized) Akaike Information Criteria (AIC or AICc).

In order to use Eq. (6) as a forecasting model, one needs forecasts of the factors $\boldsymbol{\mu}_t$ as well. We forecast the factor returns using an autoregressive model for each factor individually,

$$\mu_{m,t} = \phi_0 + \sum_{k=1}^{r_m} \phi_k \mu_{m,t-k} + \zeta_{mt}. \tag{7}$$

As noted earlier, this three step forecasting procedure (step 1; estimate DFM, step 2; forecast factors, and step 3; use factors in ARDL framework) has gained in popularity, and its easy to see why. However, there are also some drawbacks to this approach. For example, Boivin and Ng (2006) show that the forecast performance of factor models may worsen if one (or more) of the factors that are included are irrelevant for the variable of interest—also known as the oversampling problem. This is partly resolved by dropping irrelevant factors if the AIC deteriorates. Another issue—and related to the first—is that we have to forecast the factors themselves as well. This creates extra uncertainty in the forecasts.³ It should also be noted that there is evidence that the DFM forecast models improve for large N and T (Bai & Ng, 2002; Stock & Watson, 2002).

We use the ARDL to forecast, because (i) it is widely applied, (ii) its simplicity, (iii) the theoretical foundation (Stock & Watson, 1998), and (iv) the co-movement in CPPIs (Van de Minne et al., 2020).

Finally, it should be noted that it is possible to add explanatory variables in the ARDL model (like interest rates) in combination with the latent common factors, as long as the same variables are included in the first stage (the DFM). Doing this is out of the scope of this research though.

The Simplified Dynamic Factor Model

The strength of the factor augmented ARDL is that it uses co-movements between granular CPPIs to forecast the CPPIs themselves, without losing degrees of freedom like in the VAR model. The estimated co-movements come from the DFM. Alternatively, it is possible to replace the latent factors from the DFM by aggregated CPPIs to forecast the granular CPPIs. Examples of aggregated CPPIs are a national index for all property types, and national indexes per property type. These aggregated indexes also contain information on common movements between real estate prices.

The forecasting procedure is similar to what was described in Section “Forecasting”: Forecast the aggregate index(es) by Eq. (7), and subsequently use these forecasts in an ARDL to forecast the granular indexes by Eq. (6). The advantage of this approach is that there is no need to estimate the DFM, which is very time consuming. The interpretation is also more straightforward.

There might be multiple aggregate CPPIs that can be used as covariate in the ARDL, implying the same issues arise we are trying to avoid: Which indexes (“variables”) to include? The results in Section “Forecasting” suggest that the one factor augmented ARDL performs best overall. We therefore argue that including only one national all property types CPPI should suffice.

Data

We obtained a selection of CPPIs from Real Capital Analytics (RCA). RCA is recognized as one of the most respected commercial real estate data firms in the world. RCA focuses on commercial real estate bought and sold by institutional investors. They capture over 90% of all institutional⁴ commercial real estate in the US, and over 70% on average in other countries. RCA estimates CPPIs worldwide using the repeat sales methodology

developed in Van de Minne et al. (2020). This repeat sales model contains a structural time series component and is optimized to reduce revisions in index levels as new data comes in.⁵ Using this methodology Real Capital Analytics is able to estimate reliable indexes even in markets with relatively low numbers of observations.⁶

In this study we use 80 non-overlapping CPPIs, covering commercial real estate in the United States on a granular level. For most metropolitan areas two CPPIs—one for rental apartments and one for combined office, retail and industrial properties (called commercial from now on)⁷—are available. However, some metros have sub-markets indexes, like for Manhattan, the Burroughs and the suburbs in the New York Metro metro area. The indexes are quarterly and run from 2001 to mid 2017. Our total panel data is therefore $N=80$ and $T=66$. For a full overview of available indexes, please consult the website of RCA.

A graphical representation of all indexes is given in [Figure A.1a](#), where the index values in 2001Q1 are equal to 100. [Figure A.1b](#) gives the (log) return of all our indexes, and [Figure A.1c](#) gives the standardized log returns, the left-hand side variables in the DFM. Finally, [Figure A.1d](#) gives a histogram of the average quarterly price growth per market.

Overall, the indexes do seem to co-move. Indeed, the indexes go up prior to the crisis, then crash and subsequently recover starting in approximately 2009. Even though it is not presented here, there is substantial first order autocorrelation in the returns, which indicates predictability. This is not an uncommon finding in real estate literature (see Barkham & Geltner, 1995; Case & Shiller, 1989; Geltner et al., 2003; Geltner & Van de Minne, 2017; Quan & Quigley, 1991, among others). This autoregressive representation (or “inertia”) is inherent to the price formation process in real estate and does not imply arbitrage opportunities. More specifically: (i) Participants in real estate markets have incomplete information about the attributes of the purchase, (ii) some period of costly search must be incurred by both buyers and sellers, due to the heterogeneity of real estate and (iii) trades are decentralized, i.e., market prices are the outcome of pairwise negotiations (Case & Shiller, 1989; Quan & Quigley, 1991). Most of the autocorrelation is gone after a year.

Still, there is a big variety in both exact timing of turning points and in overall growth rate. For example, we find that most markets in the major metro areas have had the highest price growth; New York, Los Angeles and San Francisco had an overall price growth of almost 2% or more per quarter. Also cities like Miami, Portland and Seattle performed relatively well. On the other hand, cities like Las Vegas, Chicago, Dallas, Atlanta and Philadelphia (among others) did poorly. Usually, when apartments had a relative high price growth in a region, the commercial properties did the same.

There are also big differences in volatility. High growth markets also tend to be the most volatile, see for example New York and Los Angeles. More interesting is that some low overall growth markets like Las Vegas and Phoenix also had relative high volatility. Some other markets with low volatility include Raleigh/Durham and Minneapolis.

Results from Dynamic Factor Model

In our application we limit ourselves to a maximum of four factors in the DFM. One consideration is computing time; estimating five factors takes almost a day. Also, it is quite

well established that more factors isn't per se better when it comes to forecasting. The reason is that we need to forecast those factors as well, which adds uncertainty. We also become subject to the aforementioned "over-sampling." Adding more factors also reduces the interpretability of the latent factors. This makes it harder in practice to explain the results, reducing the overall transparency of the (eventual) forecasts. Finally, estimating many factors defeats the purposes of DFM analysis, which is summarizing large quantities of data into a few common factors. In this Section we show the DFM estimation results based on the entire panel of 80 CPPIs, where we use the standardized index returns as dependent variable. We will discuss the factors and corresponding factor loadings.

Table B.1 gives some statistics for DFM(M), where the number of factors M varies between one and four. The results are ordered by model fit (using AICc), from best to worst.⁸

The specification with four factors gives the best overall model fit according to AICc. Interestingly, for forecasting purposes we find that less factors typically works better, see Section "Forecasting." This confirms earlier findings in literature that more factors are not always better for forecasting (Eickmeier & Ziegler, 2008).

Next, we discuss the estimated factor returns $\hat{\mu}_{m,t}$. Note that we model the standardized returns, and not the returns themselves. Since it is difficult to interpret the standardized factors, we first "unstandardize" the factors by

$$r_{m,t} = \frac{1}{N} \sum_{i=1}^N (\hat{\mu}_{m,t} \times \hat{\gamma}_{mi}) \sigma_i, \quad (8)$$

where $r_{m,t}$ is the unstandardized factor return m at time t , and σ_i^2 is the variance of the index returns of market i .

Next, we focus on the factor loadings ($\hat{\gamma}_{mi}$). "Unstandardizing" the factor loadings can be done by performing a ordinary least squares on the market index return $\Delta p_{i,t}$ as dependent variable and the factor indexes $r_{m,t}$ and a constant as independent variables. Figure A.3 gives the kernel distributions of these factor loadings.

In DFM(1) the factor loadings are positive for all index return series, see Figure A.3a. In general most factor loadings are positive, even for higher M . As noted earlier in Section "Data," finding co-movement in property price indexes is quite common. The spatial distribution of the factor loadings can be found in Figure A.9 in Appendix C. In DFM(1) a higher unstandardized factor loading simply means the index has a higher index return volatility. The interpretation of the loadings on one factor is therefore interesting, as it relates to the concept of "beta" in a CAPM model. Indeed, a loading larger (smaller) than 1, simply means that when the common factor increases in value, the target index will increase more (less) than one.⁹ The factor loadings for $M=1$ are relatively high in Florida. Typically, the factor loadings are higher for commercial real estate. Factor loadings for apartments in Boston, Washington DC, California (except for Los Angeles) and Denver are especially on the low side.

One common factor that should drive all property values is the risk-free rate (like T-bills), which determines all discount rate movements to a large extent. The loadings of each market on said common factor is different, because risk-premia are different per market. The common factor from the DFM(1) can therefore be interpreted as the

movement in the risk-free rate plus the *average* risk-premium (i.e., changes in the average discount rate). Unsurprisingly, if we subsequently unstandardize the common factor from the DFM(1) we find an index that is very similar to the all property types national index published by Real Capital Analytics and the average index value per year taken from our data, see [Figure A.4](#).

In DFM(2) the first factor has a high pre-crisis run-up, see [Figure A.2b](#). In contrast, the second factor remains relatively stable, starting at 100 in 2001, and ending at 102 in 2007. For the two factors in percentages the crash is similar. However, the first factor definitely moves first. The second factor experiences a more steep recovery, and shows less volatility overall.

The size of the factor loadings are relatively similar between the two factors, see [Figure A.3b](#). [Figure A.10](#) gives the spatial distribution of these factor loadings. It is evident that the first and second factor are typically a mirror image of each other. Indeed, if the loading is high on the first factor, it is typically low (sometimes even negative) on the second factor, and vice versa. We do never observe a very high loading on both factors. Markets that are mostly impacted by the first factor (red), are almost exclusively apartment markets, like California (excluding San Francisco), Chicago, Boston, New York, Philadelphia, Washington DC and Miami. Most of these are also big cities. For commercial, only the Fort Myers metro area (including Naples and Sarasota) loads almost exclusively on the first factor. The second factor does load onto *both* segments evenly. For apartments we observe Seattle, Denver, Dallas, Raleigh/Durham, Nashville, Washington DC and Atlanta. For commercial we observe Seattle, Denver, Boston and Dallas. Thus, if the second factor loads onto commercial in a specific market, so does apartments in some cases.

The DFM(2) seems to specifically single out apartments in bigger cities on one of the factors (the red line in [Figure A.2b](#)). Larger cities have become more and more popular among younger households in recent decades (Florida, 2005). Combined with relative stringent supply constraints (Saiz, 2010) in said cities could result in the more severe boom-bust cycle observed in this factor.

The factors of DFM(3) are shown in [Figure A.2c](#). The first (third) factor is reminiscent of the second (first) factor of the DFM(2). However, there are some differences. The first factor goes down more in the pre-crisis period compared to the second factor in DFM(2). The second factor is new. Interestingly, the shape is a very typical boom-bust cycle, one that we also observe in the DFM(1). The third factor does go down a full year before the other two. The timing of the crash is similar for the second and third factor though.

The factor loadings are similar in size, see [Figure A.3c](#). Markets typically do not load highly on all factors, but rather on one or sometimes two. Apartments almost never load highly on the second factor (except Charlotte). In contrast, the second factor is the most dominant one for commercial real estate in most markets. Apartments load more aggressively on the low volatile first factor. Almost none of the markets load mainly on the third factor, apart from apartments in North-Carolina and Portland (OR).

DFM(4) gives similar results, see [Figure A.2d](#). We see the typical boom-bust factor (factor 2) and the factor going down at the beginning of the sample, but also recovering more swiftly (factor 3). What is new is that the third factor from DFM(3) is split up into two separate factors (factor 1 and 4). They are very similar (which can be interpreted as

a hint to stop increasing M). The main exception is that the fourth factor shows a bit of the boom-bust cycle, whereas the first does not.

The second factor (typical boom-bust cycle) is relevant for commercial properties only. Again, the third factor is the most dominant factor for apartments. Although quite surprisingly, some commercial property markets also load highly on this factor: Boston, Denver, Seattle and Dallas. The factors are more equally loaded on the markets. Although, the fourth factor seems to impact apartments slightly more. One clear limitation of using higher factor DFM models is that any interpretation of the latent factors becomes moot. This is problematic not only for lack of transparency, but also for possible biases inherited by the algorithms and data itself (Guidotti et al., 2018), which may lead to unfair or wrong decisions. We therefore believe that selecting the optimal number of latent factors should take “interpretability” into account as well.

As the factors are estimated simultaneously, it cannot be said which factor is dominating. Still, a feel for this can be developed by comparing the different factors in [Figure A.2](#), and see which pattern prevails. Arguably, there are two factors that are dominant. The typical boom-bust cycle, which we observe in DFM(1), and the second factor in both DFM(3) and DFM(4), is one. We find that commercial properties typically load on this factor. The second factor we find multiple times is factor 2 in DFM(2), the first factor in DFM(3), and the third factor in DFM(4). This factor slows down at the beginning of the sample, then drops during the Global Financial Crisis, followed by a steep recovery. Apartments typically load on this one.

Results from the Forecasting Models

In this section we use the estimated factors from the DFM in an ARDL model to forecast individual index returns, and compare the performance to a benchmark autoregression and our simplified DFM. We apply our forecasting models to all 80 markets.

We will forecast 8 quarters out-of-sample (2 years). First, we forecast 2 years ahead “pretending” that we are in the first quarter of 2007, using different lag lengths for all the models (see discussion below). Starting in 2007 is interesting as it is just before the GFC, while still preserving at least some prior observations for estimation ($T=22$). It is important to note that we re-estimate the DFM using observations only up to that point in time. The factors and loadings are not discussed here for the sake of brevity, but are available upon request. We subsequently compare the forecasted returns with the actual realized returns to compute the Mean Absolute Percentage Error (MAPE) for all models and markets. We keep on re-estimating the models quarter-by-quarter until 2015Q2 (because our data ends in 2017Q2).

We need to select the best benchmark and (S)DFM & ARDL forecasting model. The benchmark autoregression model depends on the number of lags for the individual market index returns. In our application we use a maximum of 10 lags for all models. Some more choices need to be made in order to keep the estimation feasible in terms of computation time. The DFM depends on the number of factors ($M = 1, \dots, 4$), and the structure of the variance-covariance matrix \mathbf{R} (diagonal and equal, diagonal and unequal, and equal variance and covariance), in total 10 different specifications. In the ARDL model we have to select the number of lags for both the factors and individual market index

returns, r_m and l_i in Eq. (6). We impose the same restriction on the lags of the individual market index returns as for the benchmark model. Moreover, to reduce computation time we require the number of lags for the factors (in case of the DFM and SDFM) to be the same as for the individual market returns, so $r_m = l_i$. Thus, in total we end up with 120 different specifications of the DFM, ARDL and SDFM models combined for every quarter for every market. As a result, we estimate a total of 80 (markets) $\times 58$ ($T - 22$ periods) $\times 120$ (models) $\approx 550,000$ forecasts. (In addition, we also have to estimate our 4 DFM models in the first step per quarter.) In all cases, the “best” forecasting model will be selected by using the penalized Akaike Information Criterion (AICc) (see Bai & Ng, 2002; Breitung & Pigorsch, 2013; Forni et al., 2000, among others).

Table B.2 gives the top five specifications for every model. For the factor augmented ARDL we find that the model with one factor and two lags performed best overall. Four out of the top five models in terms of model-fit use only one factor. As noted earlier, this is partly what motivated us to introduce the simplified dynamic factor model. Note that this finding is contrary to the model-fit of the DFM itself in Section “Results from Dynamic Factor Model,” where we found that higher M gives better model-fit. However, this result is not that surprising. Previous literature found this as well (Eickmeier & Ziegler, 2008), mostly due to the loss of degrees of freedom when using many factors, and some factors might not add much information about the underlying indexes (i.e., “oversampling”). The benchmark model with only two lags fits the data best overall. In general the lag length should not be too long. The ARDL model augmented with the national index (Simple DFM) fits the data best with three lags.

Table B.3 gives some descriptive statistics for all the models. Both the DFM and SDFM outperform the benchmark model by 17% and 9% respectively on a MAPE-basis. Based on the standard deviation, the new models perform 20% and 13% better, respectively. We can thus conclude that adding common trends to ARDL forecasting models boost the prediction precision. In Figure A.5 we give the annualized MAPE (upper panel) and the fraction of times the sign was correct for all three models averaged per year. One first (and unsurprising) observation is that generally the residuals (sign) tend to reduce (improve) as the sample increases in size. At the start of the sample the average MAPE is over 10%, where it is closer to 4% in 2015. It is also clear from Figure A.5 that the SDFM outperforms the factor augmented ARDL. Only in the first year of the sample does the factor augmented ARDL outperform the SDFM, but only marginally so (100 bps reduction in MAPE, or 9%). More noticeably, the factor augmented ARDL seems to struggle in 2009 and 2010, as even the benchmark model performs better during this period.

In order to analyze why certain models perform better at certain points in time we need to zoom in on the market forecasts. Given that we cannot show 80 different market forecasts for 66 different periods, we highlight a few markets to get the general point across. Other market results are available upon request. We focus on three periods, (i) the GFC, (ii) the subsequent recovery and (iii) finally the boom period after the recovery. All three periods can tell us something about the model performance during very specific (turning) points in time.

First we look at the two year forecasts between 2007Q4 and 2008Q4 for four separate markets (Atlanta and Manhattan apartment, and Las Vegas and Sacramento commercial) in Figure A.6. It is clear from Figure A.6 that the factor augmented ARDL (blue line)

predicted the turning point relatively early. (The black line gives the actual realized index.) In Las Vegas and Sacramento the SDFM (green line) predicted similar declines early on, similar to the factor augmented ARDL model starting mid-2008. However, by 2008Q4 (the rightmost figures) the factor augmented ARDL shows heavy declines in all markets, whereas it is a mixed bag for the SDFM. In all the cases the baseline model (red line) is ignoring the turning point altogether. Past price growth is extrapolated to the future using such models, resulting in such delays.

In [Figure A.7](#) we give similar forecasts for the recovery period, between 2009Q3 and 2010Q3. We use again four markets (Chicago and Dallas apartment, and Houston and San Francisco commercial in this case). As shown earlier in [Figure A.5](#) the factor augmented ARDL (blue line) under-performed compared to the other models during this period. [Figure A.7](#) provides some context for this under-performance. In many (but not all) cases, the factor augmented ARDL predicted a recovery, but too early and arguably too much. Hence the sign, and the MAPEs are relatively far off. The SDFM and benchmark models perform relatively similar, which was also apparent from [Figure A.5a](#).

Finally we focus on the period characterized by continuous price growth, see [Figure A.8](#). We analyze four markets again, in this case we provide; Jacksonville and Tampa commercial, and Inland Empire and New York suburbs apartment, starting in 2013Q3 and finishing in 2014Q3. Most forecasts are relatively accurate in predicting the continuous price growth. We do find (not reported here) that the benchmark model has a consistent negative bias during this period, meaning it underestimates the price growth during this period. This is also evident from [Figure A.8](#). The benchmark model seems to be (zero) mean reverting after a year or so in some markets.

Future research could improve prediction accuracy adding explanatory variables like local GDP and interest rates which are typically used to forecast real estate returns or capitalization rates (Chervachidze & Wheaton, 2013), although these variables need to be forecasted themselves. Also, in this paper we only do model selection based on penalized AIC, even though other selection criteria have proven to improve prediction accuracy in some studies (see Bai & Ng, 2002; Breitung & Pigorsch, 2013; Forni et al., 2000, among others). Finally, because the prediction residuals are not perfectly correlated between the three models, we could utilize a models-in-models approach, like model-stacking (Breiman, 1996). Obviously, the results are based on the history of the CPPIs produced by RCA, and there is no guarantee that we will be able to forecast the next big crash. Still, the results are impressive, considering our relative simple setup and lack of any explanatory variables. Both the factor augmented ARDL and the simple version of the DFM predicted (out-of-sample) the crash and subsequent recovery relatively early in the cycle.

Conclusion

The growth in the available amount of economic and financial data has prompted econometricians to develop or adapt new methods enabling them to summarize efficiently information contained in large databases containing many time series. Among these methods, dynamic factor models have seen rapid growth and are popular among macro-economists.

These models can be used to summarize the information contained in a large number of time series into a small number of factors common to the original set of time series.

In this study we apply a dynamic factor model to commercial property price index returns. This is the first time that we can do such a comprehensive study in commercial real estate, due to the recent availability of granular commercial property price indexes. We use 80 different quarterly indexes over the period from 2001Q1 to 2017Q2 in the US, provided by Real Capital Analytics. We estimate the dynamic factor model containing one to four factors. Typically, we find two factors that keep re-emerging. The first factor represents a typical boom-bust cycle. The second factor is already going down pre-crisis. Its recovery is also more swift. Commercial real estate (combined office, retail and industrial properties) tends to load highly onto the second factor, whereas rental apartments are more affected by the first one. We found that adding more common factors made the interpretation of the latent trends moot. This hurts the transparency of the results. We consider this a general limitation of higher order factor DFM models.

In a second step we use the estimated factors in an Autoregressive Distributed Lag (ARDL) model to forecast eight quarters of individual price index returns in four markets. Our findings suggest that the factor augmented ARDL outperforms the benchmark model in almost all instances. The benchmark model almost always misses the crisis and subsequent recovery, whereas the factor augmented ARDL is more accurate. Our results indicate that the forecasts generally improve when the factors are used. The success of the factor augmented ARDL in commercial real estate price index forecasting is interesting, given that this setup is not always successful in other applications, such as Eickmeier and Ziegler (2008). One possible explanation is that some co-movement between commercial property price indexes has been exploited in this setup, perhaps reflecting capital flows that move in an integrated fashion across locations whose space markets may diverge (see van Dijk et al., 2020).

We also introduce a simplified version of the dynamic factor model. In stead of estimated common latent trends, we “simply” use a national index for commercial real estate in the ARDL as an explanatory variable. Using the national index is partly motivated by the fact that we find that the forecasts are typically best when only using one common latent trend. This simplified model performs as good—if not better—then the factor augmented DFM model.

Improving forecasts is important in principle in asset valuation, and in practice is used in asset-allocation models, and in theory for the pricing of real estate derivatives. Apart from forecasting, the DFM analysis also provides useful insight about the structure of commercial real estate index returns, in particular regarding co-movements across markets, something that is important for portfolio diversification.

The goal of our forecasting exercise was not per se to find the “best” forecasting model. For example, we have not used explanatory variables. This can be seen as one of the advantages of the DFM framework: One does not need explanatory variables. However, one can still add explanatory variables in the DFM and ARDL, which could improve forecasts. Also, we find that using the AICc—a penalized version of AIC—would typically, not pick the “best” model (based on recursive forecast residuals). Other information criteria could do better, but this is left for further research. Given that the forecast residuals are not perfectly correlated, we could also employ a models-in-models

approach to boost predictive power. On the data-side we are also aware improvements can be made. At the moment, indexes on the most granular *market* level “lump” together office, retail and industrial into one segment, called commercial. Separate segment indexes are available on a more aggregate market level (say state). We hope that future research does not have to compromise between segments and markets as more and better data is collected conjoined with better index methodologies.

Notes

1. Summarizing large datasets using DFMs can be widely found in practice. A famous example is the Chicago FRB National Activity Indicator, based on 85 monthly series describing the US economy, covering production, income, employment, personal consumption, housing, sales, inventories and orders. Applications of DFMs are abundant in the empirical economic literature as well. A few examples include (Bernanke et al., 2005; Del Negro & Otrok, 2007; Favero et al., 2005; Stock & Watson, 2005).
2. Zuur et al. (2003) also show how to handle missing values, and how to enter covariates in the DFM. However, both are not needed in our research.
3. Alternatively one can add (autoregressive) components in the state Equation (5) and forecast factors within the DFM, see for example the “Factor Augmented Vector Autoregressive Model” (Bernanke et al., 2005). These models are out of the scope of this research.
4. The definition of “Institutional” according RCA is that the property sold at least once in its history for \$2.5 M in the US. For other countries they use different cut-off boundaries.
5. Revisions occur because new (repeat sales) pairs are formed of which the first sale lies somewhere in the past.
6. In total RCA recorded over 50,000 (after filtering) repeated observations in 2017, an average of approximately 650 observations per market. However, there are large differences between markets. For example, the largest market observation-wise (commercial real estate in Los Angeles proper) has over 2,000 observations, whereas the smallest market (apartments in Oakland) has “only” 150 observations. However, as noted before, the threshold for including markets in the index product is based on how stable the estimated indexes are, and not so much on how many observations there are. Although both concepts are obviously related.
7. RCA does provide indexes for the separate segments on a more aggregate market level. However, this is out of the scope of this research.
8. We use AICc, penalized AIC, because it is arguably the most widely used. However, many other criteria have been developed (Bai & Ng, 2002; Breitung & Pigorsch, 2013; Forni et al., 2000), some of them specifically for DFMs.
9. Obviously, we only look at asset returns, and not total returns, so the comparison is not completely accurate. However, we do know that most of the real estate “risk” comes from asset market, and not the rental market, see Geltner and Mei (1995).

Acknowledgments

MIT/CRE Price Dynamics Research Platform Working Paper. This paper was funded by the Real Estate Research Institute. We would also like to thank Real Capital Analytics for providing to us the data.

ORCID

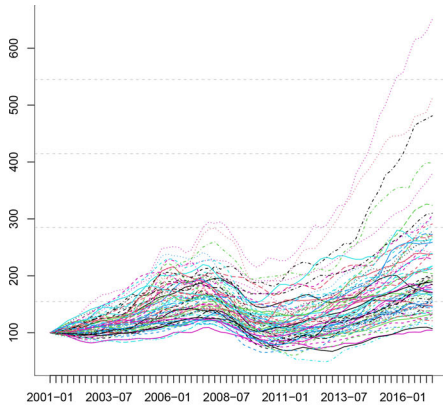
Marc Francke  <http://orcid.org/0000-0002-7239-4868>

References

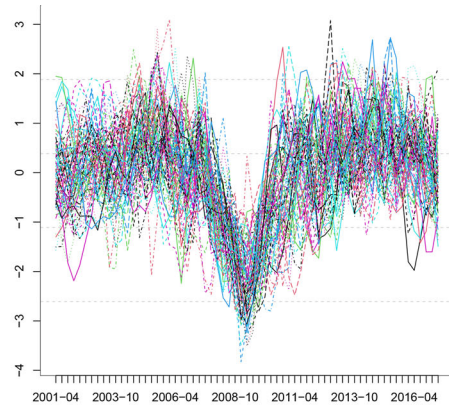
- Aguilar, O., & West, M. (2000). Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics*, 18, 338–357. <https://doi.org/10.2307/1392266>
- Anglin, P. (2006). Local dynamics and contagion in real estate markets. In *The International Conference on Real Estates and Macro Economy*, 19–24.
- Aßmann, C., Boysen-Hogrefe, J., & Pape, M. (2016). Bayesian analysis of static and dynamic factor models: An ex-post approach towards the rotation problem. *Journal of Econometrics*, 192, 190–206. <https://doi.org/10.1016/j.jeconom.2015.10.010>
- Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70, 191–221. <https://doi.org/10.1111/1468-0262.00273>
- Bai, J., & Wang, P. (2015). Identification and Bayesian estimation of dynamic factor models. *Journal of Business & Economic Statistics*, 33, 221–240. <https://doi.org/10.1080/07350015.2014.941467>
- Barhoumi, K., Darné, O., & Ferrara, L. (2013). Dynamic factor models: A review of the literature. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 2013, 73–107.
- Barkham, R., & Geltner, D. M. (1995). Price discovery in American and British property markets. *Real Estate Economics*, 23, 21–44. <https://doi.org/10.1111/1540-6229.00656>
- Basilevsky, A. (1994). *Statistical factor analysis and related methods: Theory and applications*. Wiley-Blackwell.
- Bernaards, C. A., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, 65, 676–696. <https://doi.org/10.1177/0013164404272507>
- Bernanke, B., Boivin, J., & Eliasziw, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, 120, 387–422. <https://doi.org/10.1162/qjec.2005.120.1.387>
- Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132, 169–194. <https://doi.org/10.1016/j.jeconom.2005.01.027>
- Breiman, L. (1996). Stacked regressions. *Machine Learning*, 24, 49–64. <https://doi.org/10.1007/BF00117832>
- Breitung, J., & Pigorsch, U. (2013). A canonical correlation approach for selecting the number of dynamic factors. *Oxford Bulletin of Economics and Statistics*, 75, 23–36. <https://doi.org/10.1111/obes.12003>
- Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36, 111–150. https://doi.org/10.1207/S15327906MBR3601_05
- Case, K. E., & Shiller, R. J. (1989). The efficiency of the market of single-family homes. *The American Economic Review*, 79, 125–137.
- Case, K. E., & Shiller, R. J. (1990). Forecasting prices and excess returns in the housing market. *Real Estate Economics*, 18, 253–273. <https://doi.org/10.1111/1540-6229.00521>
- Chervachidze, S., & Wheaton, W. (2013). What determined the great cap rate compression of 2000–2007, and the dramatic reversal during the 2008–2009 financial crisis? *The Journal of Real Estate Finance and Economics*, 46, 208–231. <https://doi.org/10.1007/s11146-011-9334-z>
- Del Negro, M., & Otrok, C. (2007). 99 luftballons: Monetary policy and the house price boom across US states. *Journal of Monetary Economics*, 54, 1962–1985. <https://doi.org/10.1016/j.jmoneco.2006.11.003>
- Eickmeier, S., & Ziegler, C. (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, 27, 237–265. <https://doi.org/10.1002/for.1056>
- Favero, C. A., Marcellino, M., & Neglia, F. (2005). Principal components at work: The empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics*, 20, 603–620. <https://doi.org/10.1002/jae.815>
- Florida, R. (2005). *Cities and the creative class*. Routledge.
- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and Statistics*, 82, 540–554. <https://doi.org/10.1162/003465300559037>

- Geltner, D., & Mei, J. (1995). The present value model with time-varying discount rates: Implications for commercial property valuation and investment decisions. *The Journal of Real Estate Finance and Economics*, 11, 119–135. <https://doi.org/10.1007/BF01098657>
- Geltner, D. M., & Fisher, J. (2007). Pricing and index considerations in commercial real estate derivatives. *Journal of Portfolio Management*, 33, 1–21.
- Geltner, D. M., MacGregor, B. D., & Schwann, G. M. (2003). Appraisal smoothing and price discovery in real estate markets. *Urban Studies*, 40, 1047–1064. <https://doi.org/10.1080/0042098032000074317>
- Geltner, D. M., Miller, N. G., Clayton, J., & Eichholtz, P. M. A. (2014). *Commercial real estate, analysis & investments* (3rd ed.). OnCourse Learning.
- Geltner, D. M., & Van de Minne, A. M. (2017). Do different price points exhibit different investment risk and return in commercial real estate? *The Journal of Portfolio Management*, 43, 105–119. <https://doi.org/10.3905/jpm.2017.43.6.105>
- Geweke, J., & Zhou, G. (1996). Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies*, 9, 557–587. <https://doi.org/10.1093/rfs/9.2.557>
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys (CSUR)*, 51, 1–42. <https://doi.org/10.1145/3236009>
- Harman, H. H. (1976). *Modern factor analysis*. University of Chicago Press.
- Harvey, A. (1989). *Forecasting structural time series models and the Kalman filter*. Cambridge University Press.
- Malpezzi, S. (1999). A simple error correction model of house prices. *Journal of Housing Economics*, 8, 27–62. <https://doi.org/10.1006/jhec.1999.0240>
- Quan, D. C., & Quigley, J. M. (1991). Price formation and the appraisal function in real estate markets. *The Journal of Real Estate Finance and Economics*, 4, 127–146. <https://doi.org/10.1007/BF00173120>
- Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125, 1253–1296. <https://doi.org/10.1162/qjec.2010.125.3.1253>
- Stock, J. H., & Watson, M. W. (1998). *Diffusion indexes* [Technical report, working paper 6702]. National Bureau of Economic Research.
- Stock, H. J., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20, 147–162. <https://doi.org/10.1198/073500102317351921>
- Stock, J. H., & Watson, M. W. (2005). *Implications of dynamic factor models for VAR analysis* [Technical report, working paper 11467]. National Bureau of Economic Research.
- Van de Minne, A. M., Francke, M. K., Geltner, D. M., & White, R. M. (2020). Using revisions as a measure of price index quality in repeat-sales models. *The Journal of Real Estate Finance and Economics*, 60, 514–553. <https://doi.org/10.1007/s11146-018-9692-x>
- van Dijk, D. W., Geltner, D. M., & van de Minne, A. M. (2020). The dynamics of liquidity in commercial property markets: Revisiting supply and demand indexes in real estate. *The Journal of Real Estate Finance and Economics*. Advance online publication. <https://doi.org/10.1007/s11146-020-09782-5>
- Wu, L. S., Pai, J. S., & Hosking, J. R. M. (1996). An algorithm for estimating parameters of state-space models. *Statistics & Probability Letters*, 28, 99–106. [https://doi.org/10.1016/0167-7152\(95\)00098-4](https://doi.org/10.1016/0167-7152(95)00098-4)
- Zuur, A. F., Fryer, R. J., Jolliffe, I. T., Dekker, R., & Beukema, J. J. (2003). Estimating common trends in multivariate time series using dynamic factor analysis. *Environmetrics*, 14, 665–685. <https://doi.org/10.1002/env.611>

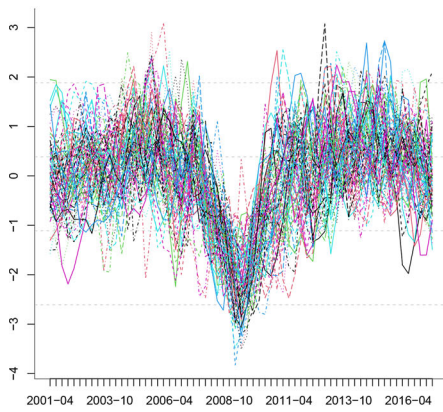
Appendix A. Figures



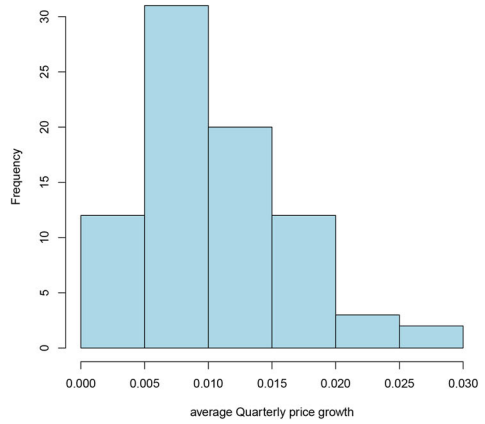
(a) Price levels.



(b) Price returns.

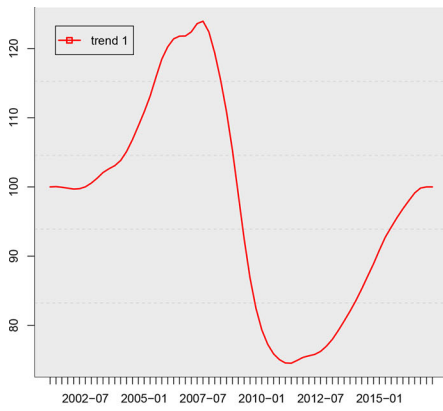


(c) Standardized returns.

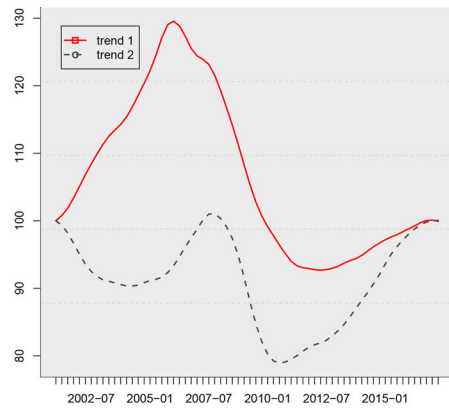


(d) Price Growth.

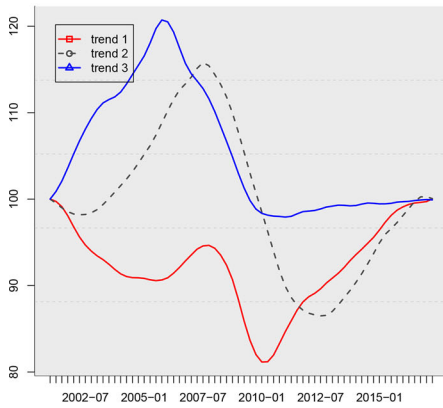
Figure A.1. Price indexes and returns.



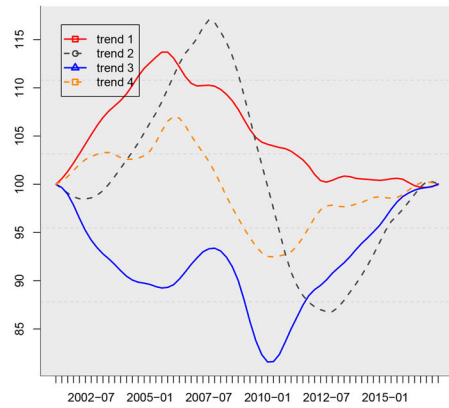
(a) Factor index, $M = 1$.



(b) Factor index, $M = 2$.

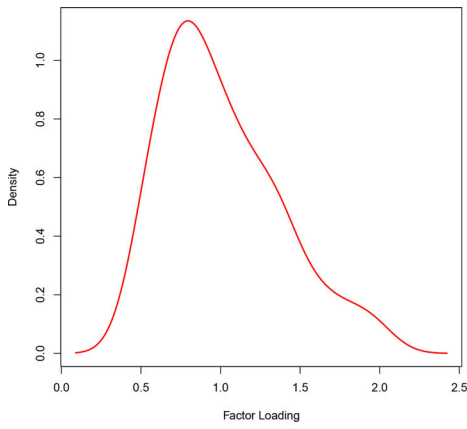


(c) Factor index, $M = 3$.

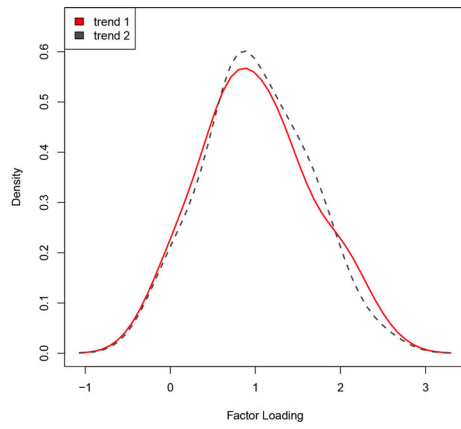


(d) Factor index, $M = 4$.

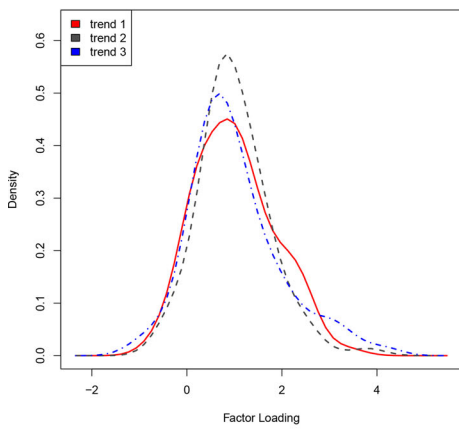
Figure A.2. Unstandardized factor indexes with diagonal and unequal variance-covariance structure.



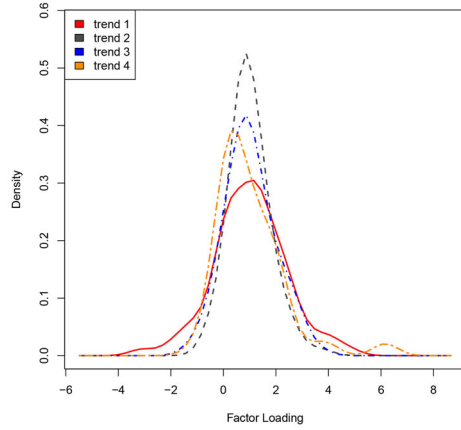
(a) Factor loadings, $M = 1$.



(b) Factor loadings, $M = 2$.



(c) Factor loadings, $M = 3$.



(d) Factor loadings, $M = 4$.

Figure A.3. Kernel distributions of the (unstandardized) factor loadings with diagonal and unequal variance-covariance structure.

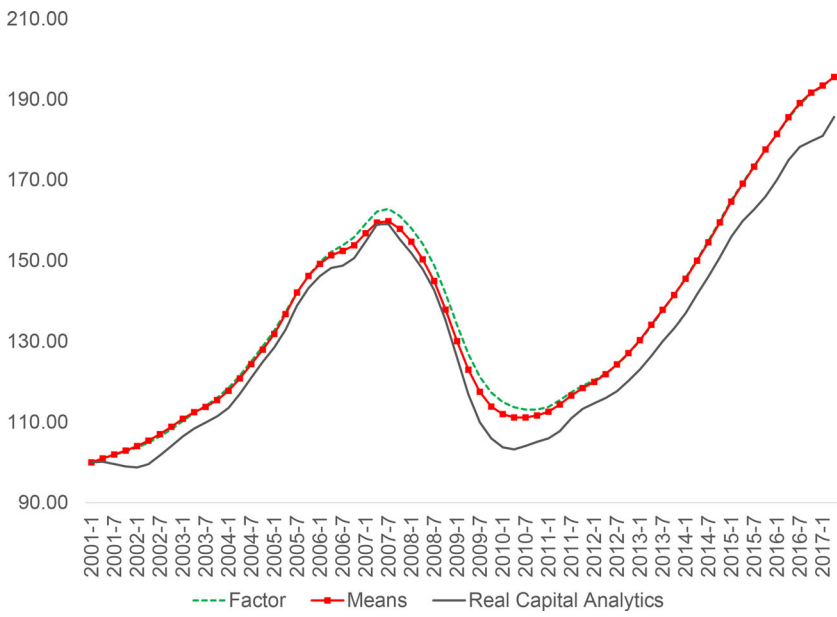
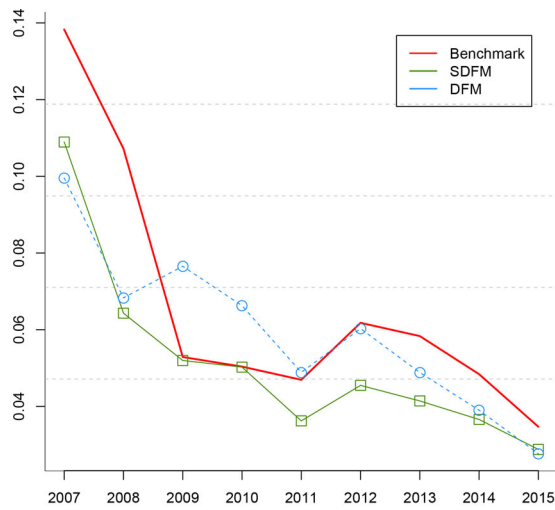
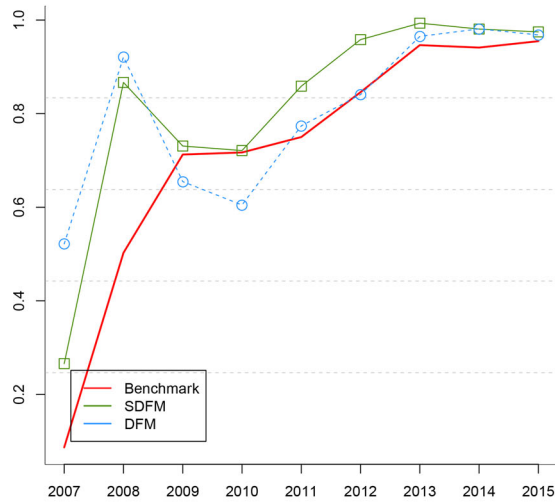


Figure A.4. Factor “index” with $M = 1$ (Factor), the mean index values of our data (Means), and the “all types US index” published by Real Capital Analytics.



(a) MAPE.



(b) Correct sign.

Figure A.5. The upper panel gives the Mean Absolute Percentage Error (MAPE) of the three models averaged per year. The lower panel gives the percentage of times the residual had the correct sign, also averaged per year.

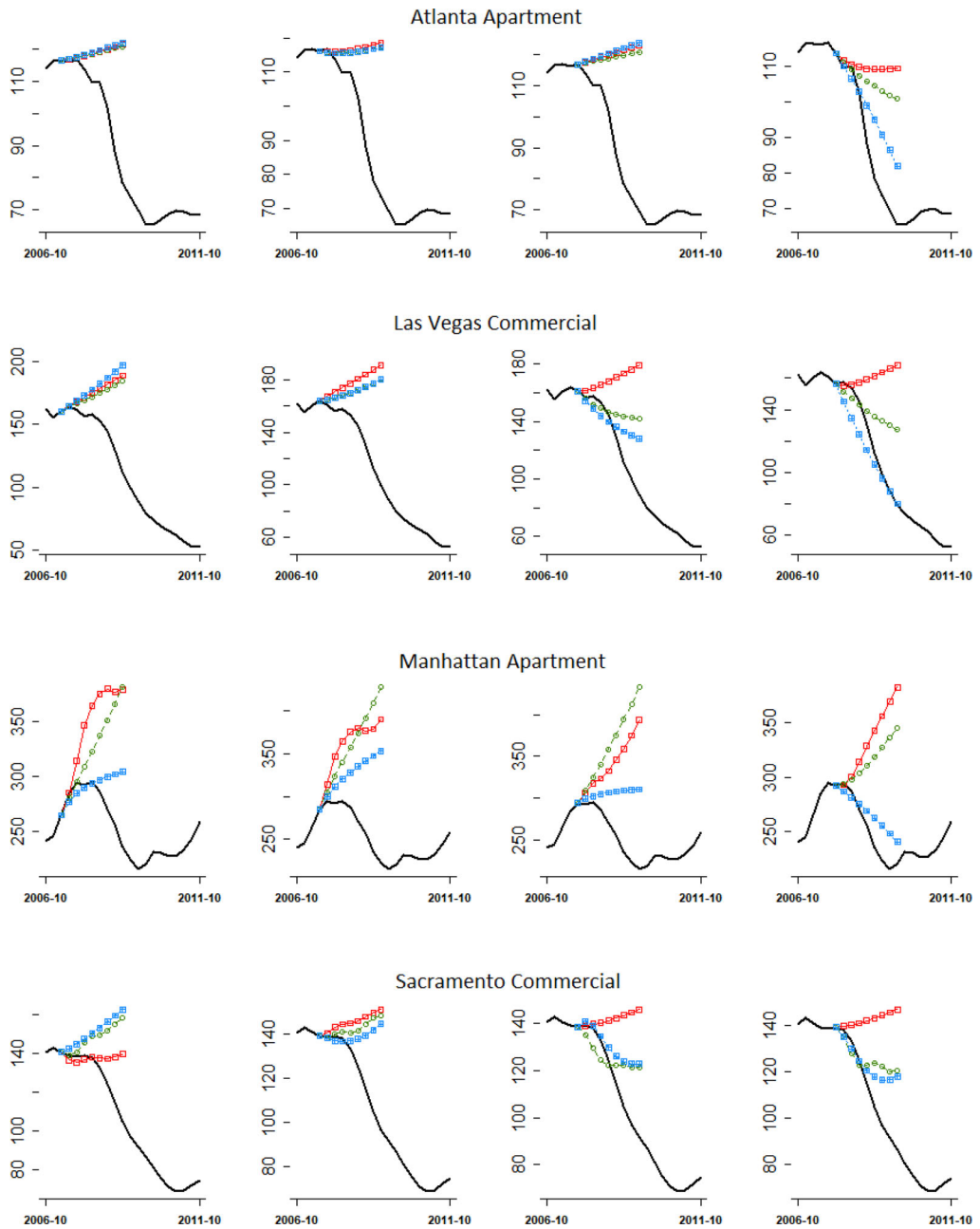


Figure A.6. Forecasts during the Great Financial Crisis. Blue line is the prediction according to the DFM, red gives the benchmark forecast, and green provides the forecasts of the Simplified DFM. Four historic predictions are shown; 2007Q4 through 2008Q4, for four markets.

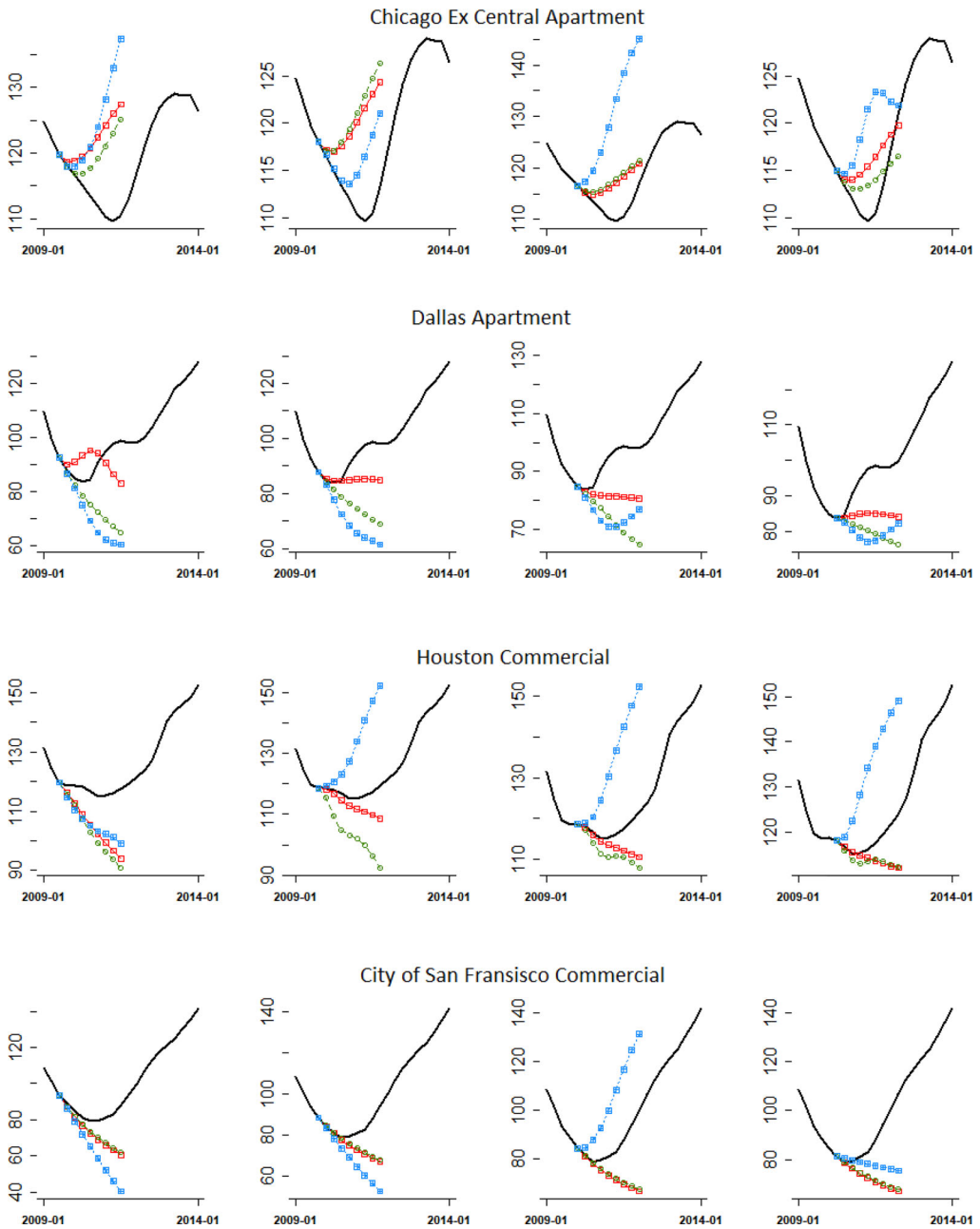


Figure A.7. Forecasts during the recovery. Blue line is the prediction according to the DFM, red gives the benchmark forecast, and green provides the forecasts of the Simplified DFM. Four historic predictions are shown; 2009Q3. through 2010Q3, for four markets.

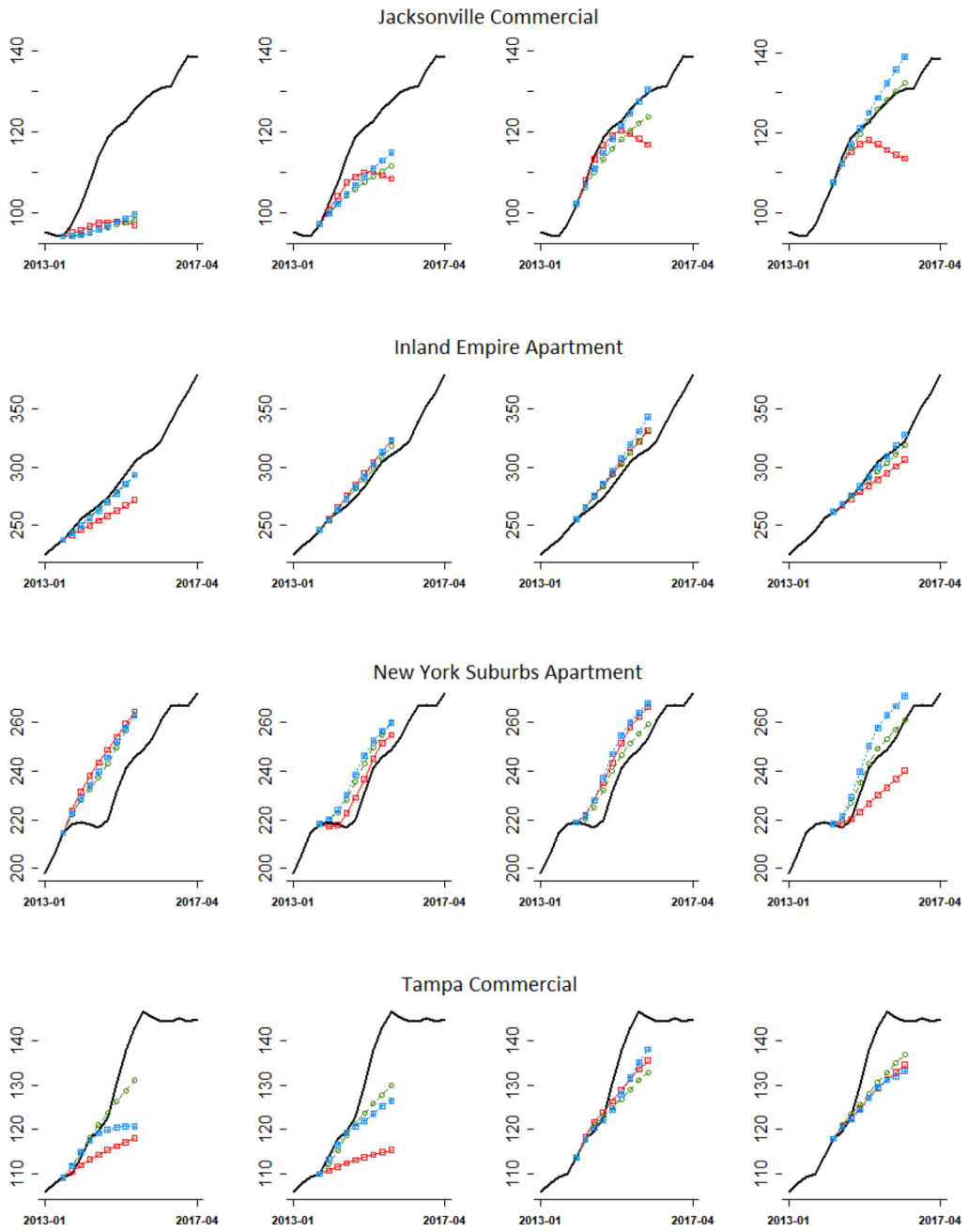


Figure A.8. Forecasts during the boom period. Blue line is the prediction according to the DFM, red gives the benchmark forecast, and green provides the forecasts of the Simplified DFM. Four historic predictions are shown; 2013Q3 through 2014Q3, for four markets.

Appendix B. Tables

Table B.1. Model fit of different specifications of the DFM, ordered by model fit.

R	M	logLik	K	AICc
Diagonal and equal	4	-4,992.7	315	10,656.2
Diagonal and equal	3	-5,130.8	238	10,760.5
Diagonal and equal	2	-5,347.5	160	11,025.3
Diagonal and equal	1	-5,588.4	81	11,341.5

R gives the design of the variance-covariance matrix of the observation errors, *M* is the number of factors, logLik denotes the log likelihood, *K* is the number of parameters, AICc is the weighted or penalized AIC.

Table B.2. Top 5 best models according to the penalized AIC.

	Factor augmented ARDL			Benchmark Lags	Simple DFM Lags
	R	M	Lags		
1	Equal variance/covariance	1	2	2	3
2	Diagonal and equal	1	3	3	4
3	Diagonal and unequal	1	2	1	5
4	Diagonal and equal	4	4	4	2
5	Diagonal and equal	1	4	5	6

R gives the design of the variance-covariance matrix of the observation errors, *M* is the number of factors.

Table B.3. (Annualized) residual statistics.

	DFM	Benchmark	SDFM
MAPE	0.058	0.064	0.053
Median	-0.028	-0.029	-0.027
sd	0.065	0.075	0.060
Min.	-0.178	-0.131	-0.134
Max.	0.121	0.177	0.135
<i>Correlation</i>			
Benchmark	0.640		
SDFM	0.854	0.805	

DFM gives the descriptive statistics of the residuals from the factor augmented ARDL, *Benchmark* gives the descriptive statistics of the residuals from the benchmark model, *SDFM* from the ARDL with the national index. *sd* = standard deviation of the residuals, *MAPE* = mean absolute percentage error.

Appendix C. Spatial Distribution of Factor Loadings

The lowest factor loading per factor index is given a 0, and the highest factor loading is given a 1. Whenever a factor loading is between the min and max, the relative position between the two is given. Some markets are omitted from these figures, as they are a "rest" category, like "South East Rest." Also, if we have multiple markets within 1 metro we take the average. This is the case for Chicago, New York, San Francisco, Miami and Los Angeles.

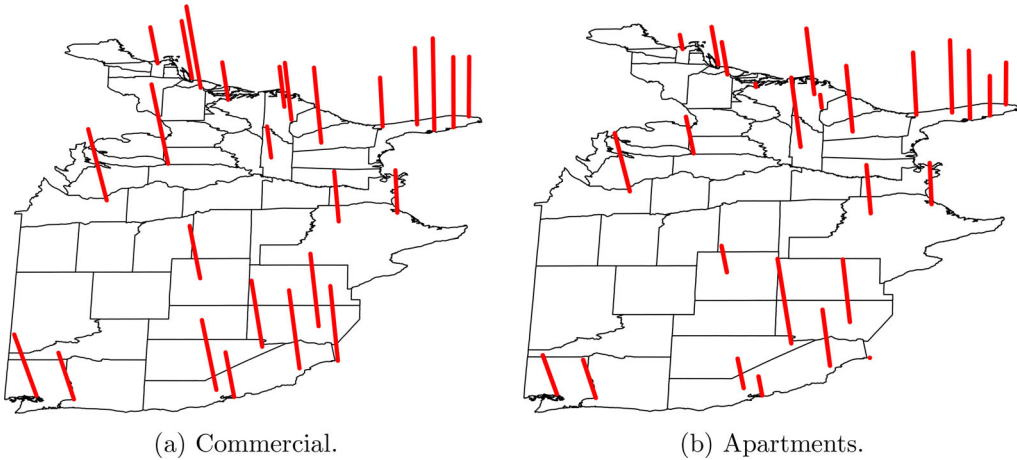


Figure A.9. Factor loadings for $M = 1$.

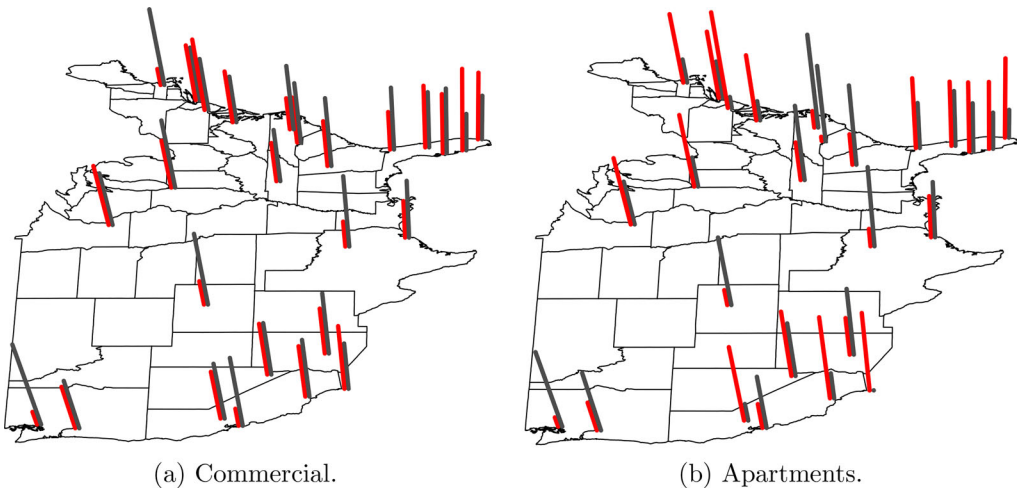


Figure A.10. Factor loadings for $M = 2$.

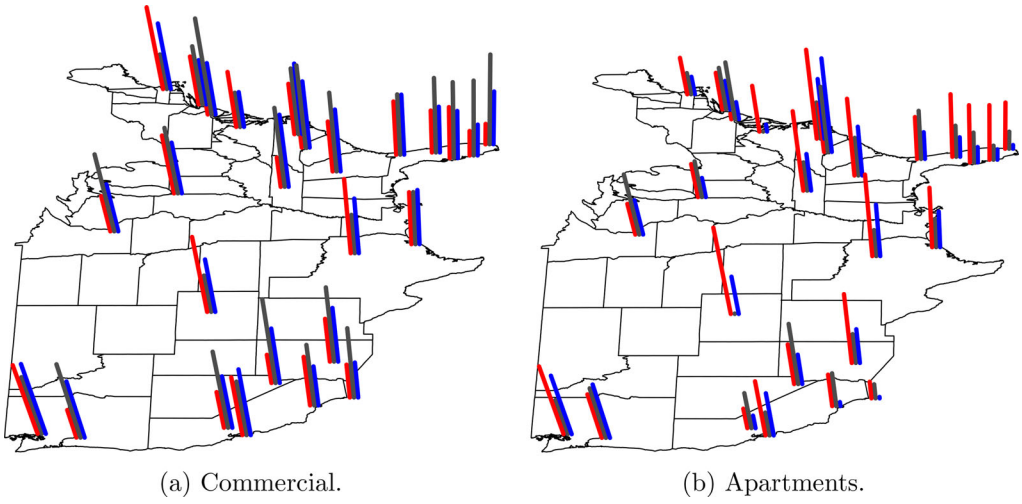


Figure A.11. Factor loadings for $M=3$.

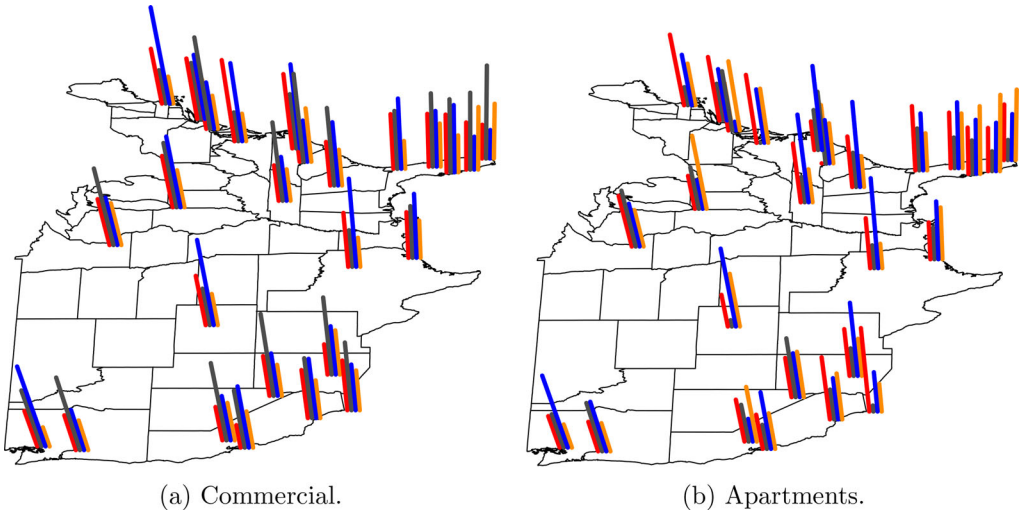


Figure A.12. Factor loadings for $M=4$.