An Empirical Analysis of Housing Price Dynamics with Factors

Yongheng Deng Hyungsik Roger Moon Yanyu Wu

November 16, 2008

Abstract

The U.S. housing price indexes are subject to measurement problems that severely impair their ability to capture the true risk. In this paper, we seek alternative methodology of utilizing latent-variable statistical methods and provide new insight to understand housing market dynamics. Housing prices are assumed to respond to external forces as proxies by way of a set of macroeconomic variables and financial indexes. Latent variable models allow us to extract interpretable common information about unobserved real estate returns.

This methodology of this paper is based on the framework in Bai & Ng's papers (2002 &2006). We applies a pure statistical approach to extract the latent factors and, more importantly, examines whether the observed macro factors are exact factors according to the information criteria proposed by Bai &Ng (2006). For robustness, we test both OFHEO repeated sales index and MRAC median home price index. We found geographical pattern of factor loadings for housing price appreciation at MSA level. The results indicate that income, consumption and GDP is a comparatively accurate factor.

1 Introduction and Motivation

It is widely recognized households could reduce and diversify house price risk via a functioning market in house price derivatives such as future and options. The integrity of such markets depends on the accurate measurement of price levels and volatilities. Practically speaking, a broad housing futures market would necessarily be based upon the housing price indexes published by large organizations.

Hedonic, Repeated Sales and Hybrid are three commonly used methods of housing price construction. The most widely used technique for estimating housing price trends is the repeat sales method introduced by Bailey, Muth, and Nourse (1963). The weighted repeat sales model extended by Case and Shiller (1987) is widely used in academic research. The series of state and metropolitan indexes published by the US Office of Housing Enterprise Oversight (OFHEO) are based on a modified version of weighted repeated sale methodology.

The repeated sales model has gained most popularity because it controls for heterogeneity of structural and location characteristics while requiring only transaction prices and sales dates. One more attribute inherent in the repeated sales methodology is that the entire historical path of these indexes is subject to revision as new information is revealed. However, numerous recent studies examine various potential biases in the repeat sales house price indices. Those biases include: (1) “Renovation Bias” due to the inability to account for the possibility of structure changes between two sales; (2) “Hedonic Bias” due to the inability to account for depreciation, maintenance and improvements; (3) “Trading-frequency Bias” caused by relative infrequency of sales causes; (4) “Sample Selection Bias” due to the fact that it only includes properties transacted more than once; (5) “Aggregation Bias” caused by the specific interval employed. (Cho, 1996)

To solve those measurement problems, various alternatives of repeated sales models has been proposed in the recent studies. Among which, hedonic repeated sales models (also termed as hybrid models, Case and Quigley 1991; Quigley 1995; Englund et al 1998) are the most popular one. Englund, Quigley, and Redfearn (1999) combines single sales and repeat sales and utilizes information on all sales, as well as all available information on housing attributes, to estimate trends in housing prices.

Considerable interest has been focused on the implications of the assumptions of constancy of the housing quality. Clapham et al (2006) find that hedonic indexes appear to be substantially more stable than repeated sales indexes and are not prone to the systematic downward revision. Although, the hybrid method takes advantage of the information that is present in repeated sales, but without ignoring information on single sales and represents an obvious improvement over the repeated sales method, it is data intensive and depending crucially on the inclusion of the correct set of properties, and the correct function form for the regression. That is why the repeated sales index is widespread used. Recently, a couple of papers examine the index stability by studying the index revision. (Clapham et al 2006, Deng and Quigley 2008) Methods that are subject to substantial revision raise questions about the viability of derivatives markets.

No matter what indexes we use, they unavoidably subject to intrinsic drawbacks, either unavailability of information or biases. Our paper seeks alternative methods of measuring housing price risks by looking broadly at the financial and macroeconomic dynamics. Understanding the future path of house prices in relation to economic stresses such as oil price shocks, financial market distress, household income, and others is critical to successful strategic planning and risk management. We use factor analysis to obtain a better identification of innovations to housing prices. The methodology follows a two-step procedure. The first step estimates the factor and factor loadings. The second step evaluates the macroeconomic variables and financial indexes using the factor estimated in the first step.

The remainder of the paper proceeds as follows: Section 2 presents the empirical evidence we found in housing markets. Section 3 explains estimation of the empirical models. Section 4 describes dataset. In Section 5 presents and interprets the empirical results. Section 6 concludes.

2 Empirical Evidence: Macroeconomics and Housing

A rising tide of foreclosure is flooding the markets. Many economists believe that the there will be a multi-year slump in national housing market. The troubles in housing precipitated the credit crisis in the summer of 2007 and became a significant drag on the economy. The raising risks of a recession attracts hot discussions in housing markets recently. Take a look at the foreclosure variation across states, we find the evidence that the foreclosure problem appears to be greatest in the West, particularly in Nevada, where the home prices soared in the housing boom and now drop rapidly. It appears that in the most heavily affected states the sales totals lagged behind the number of foreclosure notices. The states with low rates tend to be states that missed the boom in housing prices and now have reasonably good economies, for a typical example, South Dakota. The importance of housing price dynamics to what we call the business cycle attracts great attention in academic recently. House price volatility, although generally lower than the volatility of financial asset prices, can have important effects on economic activity and financial stability.

[Figure 1 insert here]

Figure 2 describes the house price trends and cycles from 1980Q2 to 2007Q4. Real house prices are very volatile and fluctuate over time, with an average standard deviation of the growth rate of around 1 percent per quarter. In recent years, while house prices have been buoyant, their volatility has declined markedly, and this phenomenon is found to be worldwide. (See "World Economic Outlook" Sep 2004 by IMF)

[Figure 2 insert here]

Changes in house prices may be caused by a variety of factors that affect both the supply of and the demand for housing. What are the factors that drive house price dynamics? Important questions have been whether housing market fluctuations are an independent source of shocks or whether they just reflect macroeconomic fluctuations. Figure 3 shows the housing price appreciation (HPA) and key macroeconomic variables[1] over the period since 1975. The strongest relationship seems to be that between house prices and consumption of nondurable goods. It is not surprising that we see a negative relationship between the growth rates of fixed investment and housing price appreciation, while a positive relationship between the growth rates of nondurable goods and housing price appreciation.

[1] All variables have been detrended by taking logs and then regressed on a constant and a linear trend.

[Figure 3 insert here]

The link between housing markets and the rest of the economy operates primarily through the effects of house price fluctuations, as they represent the main source of fluctuations in housing wealth. If that relationship is stable, then fundamentals can explain house prices. Given the importance of housing in household wealth, it also seems reasonable to conjecture that the observed downs and ups in housing prices could have substantial macroeconomic impacts. It would be interesting to explain these movements in the housing market, and to what extent they are related to macroeconomic movement in business cycles.

Those observations prompt our interest to quantify such relationships. There is a large literature that attempts to use regression analysis to link house prices to variables thought to be fundamental determinants of house prices. The fundamentals include real disposable income (per capita, growth), per capita output, housing affordability, interest rates(short term, long term) , real credit (growth), residential investment, stock price, population growth, bank crisis, and demographic variables. However, in those regression analyses, the independent variables are chosen arbitrarily to some extent. We look for more robust methods to quantify those relationships.

3 Methodology

Factor models have been widely used for studying asset returns. Considering the strong cross-sectional correlation among MSAs, assumptions in some methods of panel data analysis are violated.

In traditional simple regression-like analysis, the non-factor return for one asset (e_i) is assumed to be uncorrelated with that of every other (e_j). However, the data show non-negligible correlation among HPAs in different MSAs. High correlations across MSAs are generally associated with a common loading on major factors. As stated previously, the principal component method ensures an optimal fitting of the model.

[Table 1 insert here]

The first goal of this paper is to determine the number of factors (r). Bai and Ng (2006) propose some panel criteria. They show that the number of factors can be consistently estimated under the framework of large cross sections (N) and large time dimensions (T).

Let HPA_{it} (Housing Price Appreciation) be the housing price appreciation of the ith MSA at time t, for $i = 1, ..., N$ (N is the total number of MSAs included in the analysis) and $t = 1, ..., T$ (T is total number of quarters). Suppose that HPA_{it} admits a static factor model representation with r common factors f_t., i.e., assume that the variation of the HPA of all MSAs can be explained by a small set of r unobservable factors contained in the $T \times N$ matrix HPA[2], and related to HPA through a $K \times N$ matrix of factor loadings λ, Equation (1) is a generic representation:

$$HPA_{it} = \lambda_1 f_{1t} + \lambda_2 f_{2t} + ... + \lambda_r f_{rt} + e_{it} \tag{1}$$

where $f_{1t}, f_{2t}, ..., f_{rt}$ are the common factors that determine HPA_{it}; $\lambda_1, \lambda_2, ..., \lambda_r$ are factor loadings associated with and e_{it} is the idiosyncratic component of HPA_{it}. The products, $\lambda_1 f_{1t}, \lambda_2 f_{2t}, ..., \lambda_r f_{rt}$ are the common components of HPA_{it}.

By solving the following minimization problem, we can obtain the estimates of λ^k and F^k. The subscript in λ^k and F^k denotes the number of factors included in the estimation. The solution of the maximization problem (2) can be derived by the principal component method.

$$V(k) = \underset{\Lambda,F}{min}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(HPA_{it} - \lambda_i^k F_t^k)^2 \tag{2}$$

$$s.t. F^{k'} F^k / T = I_k$$

The estimated factor matrix $\widetilde{F}^k$ is $\sqrt{T}$times the eigenvectors corresponding to the k largest eigenvalues of the $T \times T$ matrix $HPA \times HPA'$. Given $\widetilde{F}^k$,the factor loadings are obtained through ordinary least squares, $\widetilde{\Lambda}^{k'} = \widetilde{F}^{k'} \times HPA/T$ is the corresponding matrix of factor loadings.

[2]standardized housing price appreciation

To estimate the true number of factors, r, Bai and Ng (2002) propose to minimize the criterion functions. The simulation results in the paper suggest that are preferred criteria. Therefore, they are considered in this paper.

$$
\begin{aligned}
\widetilde{r}_{PCp1} &= \arg\min_{k\leq k\max} PC_{p1}(k) = \arg\min_{k\leq k\max}\{V(k) + k\widehat{\sigma}^2(\frac{N+T}{NT})\ln(\frac{N+T}{NT})\} \\
\widetilde{r}_{PCp2} &= \arg\min_{k\leq k\max} PC_{p2}(k) = \arg\min_{k\leq k\max}\{V(k) + k\widehat{\sigma}^2(\frac{N+T}{NT})\ln[\min(N,T)]\} \\
\widetilde{r}_{ICp1} &= \arg\min_{k\leq k\max} IC_{p1}(k) = \arg\min_{k\leq k\max}\{\ln V(k) + k(\frac{N+T}{NT})\ln(\frac{N+T}{NT})\} \\
\widetilde{r}_{ICp2} &= \arg\min_{k\leq k\max} IC_{p2}(k) = \arg\min_{k\leq k\max}\{\ln V(k) + k(\frac{N+T}{NT})\ln[\min(N,T)]\}
\end{aligned}
$$

where
$\widehat{\sigma}^2 = V(K\max, \widetilde{F}^{k\max})$is the scaling to the penalty term.

Considering that BIC_3 has good properties in the presence of cross section correlation and gives a result of a smaller number of true factors (the penalty term has a greater weight), the paper includes BIC_3 in the study.

$$
\widetilde{r}_{BIC3} = \arg\min_{k\leq k\max} BIC3(k) = \arg\min_{k\leq k\max}\{V(k) + k\widehat{\sigma}^2(\frac{(N+T-k)\ln(NT)}{NT})\}
$$

Practically, Ahn & Horenstein (2008) argue that Bai & Ng's estimators tend to choose more than the true number of factors, which impairs the predictivity in forecast. They propose eigenvalue ratio estimator (ER) and growth ratio estimator (GR) and prove that they outperform the Bai & Ng estimators when either the number of cross sections or the number of time series is small, which fit our case better.

$$
\begin{aligned}
\widetilde{r}_{ER} &= \arg\max_{k\leq k\max}\{\frac{\widetilde{u}_k}{\widetilde{u}_{k+1}}\} \\
\widetilde{r}_{GR} &= \arg\max_{k\leq k\max}\{\frac{\ln(\widetilde{u}_k^*)}{\ln(\widetilde{u}_{k+1}^*)}\}
\end{aligned}
$$

where,$\widetilde{u}_k$ is kth largest eigenvalue of the sample covariance matrix of standardized HPA, and $\widetilde{u}_k^* = (\sum_{j=k}^T \widetilde{u}_j)/(\sum_{j=k+1}^T \widetilde{u}_j)$.

Table 2 reports the estimation of a number of factors according to minimization of above seven criteria.

The major difficulty in implementing multi-factor models is the identification of common and relevant factors. Above, a set of implicit factors have been derived endogenously by using principal components analysis. However, the difficulty then lies in the economic interpretation of the implied factors. The implicit factors are not easy to interpret. Having gotten these principal components most highly correlated with HPAs, we move on to test whether the observed factors which we believe or know to be relevant are highly correlated to the principal components.

Bai and Ng (2006) propose criteria to evaluate the observed factors via latent factors. The economic intuition is then straightforward. In this paper we use simple criteria to test the fitness.

Denote $Macro_t$ as macroeconomic and financial indexes, and suppose that $Macro_t = \delta' F_t + \varepsilon_t$ where F_t are the derived k latent factors. Then, by running an OLS regresstion, the estimated $Macro$ is,

$$\begin{aligned} \widehat{Macro} &= P_{\tilde{F}} \times Macro. \\ where P_{\tilde{F}} &= \tilde{F}\left(\tilde{F}'\tilde{F}\right)^{-1}\tilde{F}'. \end{aligned}$$

Our first criteria is the R^2

$$R^2 = \frac{\widehat{Macro}'\widehat{Macro}}{Macro'Macro}$$

Another descriptive statistic is the sample correlation between $Macro$ and $\widehat{Macro}$:

$$Corr = \frac{\frac{1}{T}\sum_{t=1}^T Macro_t'\widehat{Macro}_t - \left(\frac{1}{T}\sum_{t=1}^T Macro_t\right)\left(\frac{1}{T}\sum_{t=1}^T \widehat{Macro}_t\right)}{\sqrt{\left(\frac{1}{T}\sum_{t=1}^T Macro_t^2 - \left(\frac{1}{T}\sum_{t=1}^T Macro_t\right)^2\right)}\sqrt{\left(\frac{1}{T}\sum_{t=1}^T \widehat{Macro}_t^2 - \left(\frac{1}{T}\sum_{t=1}^T \widehat{Macro}_t\right)^2\right)}}$$

4 Data

The data used in this research are MSA-level repeat transactions home price indices estimated by the Office of Federal Housing Enterprise Oversight (OFHEO) and MRAC Single Family Residence Home Price Index.

The Office of Federal Housing Enterprise Oversight (OFHEO) has estimated repeat sales price indexes for US census regions, states, and metropolitan areas. Housing prices for 381 different metropolitan housing markets are updated and released quarterly. MSAs are normalized to 100 in the first quarter of 1995. The difference in normalization dates has no impact on appreciation rates obtained from the index. These data exploit variations in the geographical distribution of housing prices among the US MSAs, states and census divisions. The observations start from the first quarter in1975, however, there are a lot of missing data in the early periods because that the OFHEO effort has been undertaken continuously since 1996. The principal component analysis is based on a balanced panel; therefore the paper narrows our sample by requiring a MSA to have a record of continuous observations ending in the fourth quarter of 2003 and beginning no later than the first quarter of 1984. Our second sample selects a balance panel with time horizon starting from the first quarter of 1980 and ending at the fourth quarter of 2007, which is a larger set. The first sample delete the periods with high variation in data across MSAs. This requirement that MSAs should have a continuous record introduces selection bias into the samples.

MRAC Single Family Residence Home Price Index is based on mean or median housing prices. The reason we include this dataset in our study is that it is a balanced panel in which all MSAs has continuous observations starting from the first quarter of 1976 and ending in the second quarter of 2007. This sample doesn't suffer the selection bias, and more satisfies the condition of large T and large N as the methodology in this paper requires. In addition, it is more robust to test both datasets since we agree neither of them can exactly represent house price dynamics.

Policy makers are interested in how macroeconomic indicators react with each other under policy shocks and we are interested in how the macroeconomic indicators help us to understand the housing price construction and movements.

When evaluating the macroeconomic variables, we rely on a panel of quarterly observations[3] on US macroeconomic variables and financial indexes, measuring aggregate consumption, income, investment, interest rates, spreads, stock and bond returns etc. (see Appendix for the descriptions of the macroeconomic indicators and asset class benchmarks included in the study)

5 Empirical Results

5.1 Factors and Factor Loadings: Testing the consistency of HPI at national level and MSA level

In the first part, I factor analyze the MSAs as a single group. We transform the data so that each series is mean zero, variance 1.

Table 2 shows the results of determining the numbers of true factors. k-max is set to be 15. Our first panel excludes the periods with high volatility. The sample period is 1984Q2-2003Q4. The PC and IC criteria choose k-max as the optimal number of latent factors. That means the two PC and two IC criteria select more than 15 factors. This result is not promising and it might be due to lack of dimensionality in cross sections, i.e. the dimension is too small to satisfy the requirement of large N and large T. It also indicates high variation of HPA across MSAs. However, BIC3 shows much lower values. Seven mutually orthogonal principal components, which explain most of the cross-sectional variance of OFHEO HPA, are extracted by the BIC3 criteria. Strikingly but not surprisingly, the ER and GR estimators only capture 1 factor, because the second largest eigenvalue drops significantly.

Given our results in Figure 4, which we will discuss later, there is strong evidence for one common factor presenting in US housing price indics at MSA level.

[Table 2 insert here]

Considering the sample selection bias in OFHEO HPA, we include a comparative study of MRAC HPA. The number of cross sections is almost doubled. It has

[3] If quarterly data are not available, monthly data are transformed to quarterly data by taking the average within quarter. The data are first taken natural log and then detrended.

a balanced panel consisting of all 381 MSAs and 122 quarters. PCp1 and PCp2 choose 7 factors while ICp1 and ICp2 choose only 3 factors. More significantly, BIC3 suggest the presence of only 2 factors! ER and GR again both pick only 1 factor.

Figure 4 shows the first factor extracted from the indics at MSA level and its Hodrick-Prescott (HP) filter, compared to the national index. When we plot them together, we see a rough overlap. The figure shows that the national index tracks the summary indicator closely, meaning that historically the national index has provided a good summary of overall housing price dynamics. This finding provides evidence of consistency in OFHEO housing price indics. There have been occasional divergences between them. The national index is more volatile and it has higher peaks and lower troughs than the first factor in most periods. Still, the correlation between the two series is very high. However, we fail to see such relationship when we compare the national index with the second factor, third factor, and so on. (see Figure 5). This is an indirect but strong evidence of the presence of one common factor.

[Figure 4 insert here]
[Figure 5 insert here]

Figure 6 shows the factor loadings. The values of the factor loadings are indeterminate. We only care about the comparative importance across MSAs. The first factor has strongest effect on census divisions PAC, MA, and NE and it causes the house prices in these divisions move in the same direction, while MT and WSC in the opposite direction. The second factor has its strongest effect on PAC. It causes PAC, MT and WSC move in the same direction. The fourth factor causes WSC, NE and MT move in the same direction. Its strongest effect is on WSC. The third and fifth factors show no strong regional pattern. The sign consistency of factor loadings on NE and MA is not surprising because of location similarity. They represent the east coast and PAC represents the west coast. These three divisions show concentrated characteristics on factor loadings.

[Figure 6 insert here]

5.2 Testing the Fitness of Macroeconomic Variables and Financial Market Indexes in Factor Analysis

I set r to 5 in the analysis for both the panel of OFHEO HPA and MRAC HPA in order to do comparative studies.

Case, Quigley and Shiller (2003) find strong evidence that variations in housing market wealth have important effects upon consumption. Changes in housing price may have substantial macroeconomic effects through private consumption. A long list of factors can be expected to have a depressing effect on consumer consumption with the deflating housing bubble. When house price falls, homeowner's total wealth is declining and the value of mortgage equity is lowered. Consumers also face constraints due to the declines in the stock market and the tightening of lending terms at depository institutions. The resulted rise of the prices of energy, food, and other commodity taxes the disposable incomes of households and holds back consumer spending. Those households with significant mortgage debt may need to adjust non-durable consumption when confronted by a negative, unanticipated economic shock, which is called the "lock-in" effect. That is why we see an almost perfect fit of the movements of non-durable income and house price. All the consumption variables including Personal Consumption (NPC, RPC), Durable Consumption (NDC, RDC), Non-durable Consumption (NNC, RNC) are all significant in both panels of OFHEO HPA and MRAC HPA.

House prices and income are thought to be linked by a stable long-run relationship. If so, then the gap between the two may be a useful indicator of when house prices are above or below their equilibrium values, and therefore a useful predictor of future house-price changes. Our result confirms it. In the panel of MRAC, real disposable income doesn't perform as well as the others. House prices are also found to be significantly influenced by GDP growth rates and provide a good long-term hedge against inflation but a poor year-to-year hedge although the aggregate house price is more volatile than GDP.

Housing markets can be viewed as an extension of capital markets and residential property may be seen as a potential institutional investment class. This is not only due to the fact that housing is an alternative portfolio choice available to investors but also because financing terms available on capital markets have a significant effect on the return on housing. Even so, housing markets diverge from capital markets in a number of ways. They face high governmental intervention, investments are illiquid, indivisible, structurally and locationally heterogeneous. The level of integration between two markets is determined by the extent to which assets in these markets are affected by common economic factors. If the two markets are significantly integrated then it is expected that a large asset substitution will occur, with such substitution having a significant impact on price movements in both markets. However, the result indicates that the contemporaneous relation between quarterly house price changes and stock returns is not statistically significant.

House prices, like other asset prices, are interest rate sensitive and respond to changes in the monetary policy, thereby contributing to the transmission of monetary policy impulses to the economy. The difference between the yield on long-term debt guaranteed by Fannie Mae and that of similar Treasury debt soared in recent years following the subprime crisis, providing a new sign of the nervousness that has affected financial markets. In testing the spreads of asset benchmarks, Spread between 3-month Treasury Bills and Federal Funds Rate (SFYG3), Spread between 6-month Treasury Bills(SFYG6) and Federal Funds Rate(FYFF) are found significant in the panel of OFHEO HPA. However, in the panel of MRAC HPA, the short term assets are no longer significant. 5-year and10-year treasuries become significant instead. Spread of AAA Bond yield and Federal Funds Rate, and spread of BAA Bond yield and Federal Funds Rate are both significant as a measure of risk.

A downside risk is whether a rising unemployment will force some people to sell their houses, creating further downward pressure on housing prices. Civil Employment (CE) is comparatively significant in OFHEO HPA but weak in MRAC HPA.

6 Concluding Remarks

The housing market endures significant cyclical movements and volatility. There are many interesting research topics about the nexus of macroeconomics and housing. This paper presents insights for estimating housing price index via other indices. Factor analysis is used to analyze house price dynamics by exploiting macroeconomic and financial data. The results confirm the importance of many aspects in affecting the course of housing prices. Specifically, this paper examines the one-by-one relationship between macroeconomic and financial indexes and housing price indexes. We found geographical pattern of factor loadings for housing price appreciation at MSA level. The dynamics of income, consumption and GDP can explain housing price dynamics.

Housing price forecast is crucial for both individual investors and mortgage companies. Many organizations apply different methodologies to conduct forecasts. The follow-up research for this paper could be forecasting HPA based on the principle components we derive. Stock and Watson (2002) estimates the indexes and constructs forecasts using an approximate dynamic factor model; the predictors are summarized using a small number of indexes constructed by principal component analysis.

7 References

References

[1] Ahn, S.C. & Horenstein A.R., 2008, "Eigenvalue Ratio Test for the Number of Factors", Working paper, Arizona State University.

[2] Bai, J., & S. Ng, 2002, "Determining the Number of Factors in Approximate Factor Models", Econometrica, 70, 191-221.

[3] Bai, J., & S. Ng, 2006, "Evaluating Latent and Observed Factors in Macroeconomics and Finance", Journal of Econometrics, 131, 507-537.

[4] Bailey, M. J., R. F. Muth, & H. O. Nourse, 1963, "A Regression Method for Real Estate Price Index Construction", Journal of the American Statistical Association, 58(304), 933–942.

[5] Case, B., & J.M. Quigley, 1991, "The Dynamics of Real Estate Prices", The Review of Economic Statistics, 22, 50–58.

[6] Case, K.E., J.M. Quigley, & R.J. Shiller, 2003, "Home buyers, Housing and the Macroeconomy", in: Richards, A., & T. Robinson, "Asset Prices and Monetary Policy," Reserve Bank of Australia, 149–188.

[7] Case, K. E., and R. J. Shiller, 1987, "Prices of Single-Family Homes Since 1970: New Indexes from Four Cities", New England Economic Review, 45-56.

[8] Case, K.E., & R. J. Shiller, 1989, "The Efficiency of the Market for Single Family Houses", American Economic Review, 79, 125-137.

[9] Clapham, E., P. Englund, & J.M. Quigley, C.L. Redfearn, 2006, "Revisiting the Past and Settling the Score: Index Revision for House Price Derivatives", Real Estate Economics 34(2), 275–302.

[10] Deng, Y., & J. M. Quigley, 2008 forthcoming, "Index Revision, House Price Risk, and the Market for House Price Derivatives", The Journal of Real Estate Finance and Economics.

[11] Englund, P., J.M. Quigley & C. Redfearn, 1998, "Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices", Journal of Urban Economics, 44(2), 171–196.

[12] Englund, P., M. Hwang, & J. M. Quigley, 2002, "Hedging Housing Risk", Journal of Real Estate Finance Economics, 24(1/2), 167–200.

[13] Cho, M., 1996, "House Price Dynamics: A Survey of Theoretical and Empirical Issues", Journal of Housing Research, 7(2), 145-72

[14] Quigley, J. M., 1995, "A Simple Hybrid Model for Estimating Real Estate Price Indices", Journal of Housing Economics, 4(1), 1-12.

[15] Stock, J.H., & M.W. Watson, 2002, "Forecasting using principal components from a large number of predictors", Journal of the American Statistical Association, 97, 1167–79

Figure 1 Housing Boom and Foreclosure

The foreclosure problem appears to be greatest in the West, particularly in Nevada, where the home prices soared in the housing boom and drop rapidly now. It appears that in the most heavily affected states the sales totals lagged behind the number of foreclosure notices. The states with low rates tend to be states that missed the boom in housing prices and now have reasonably good economies, for a typical example, South Dakota.

Figure 1.1 HPI of top 5 states with lowest foreclosure rate in 2007

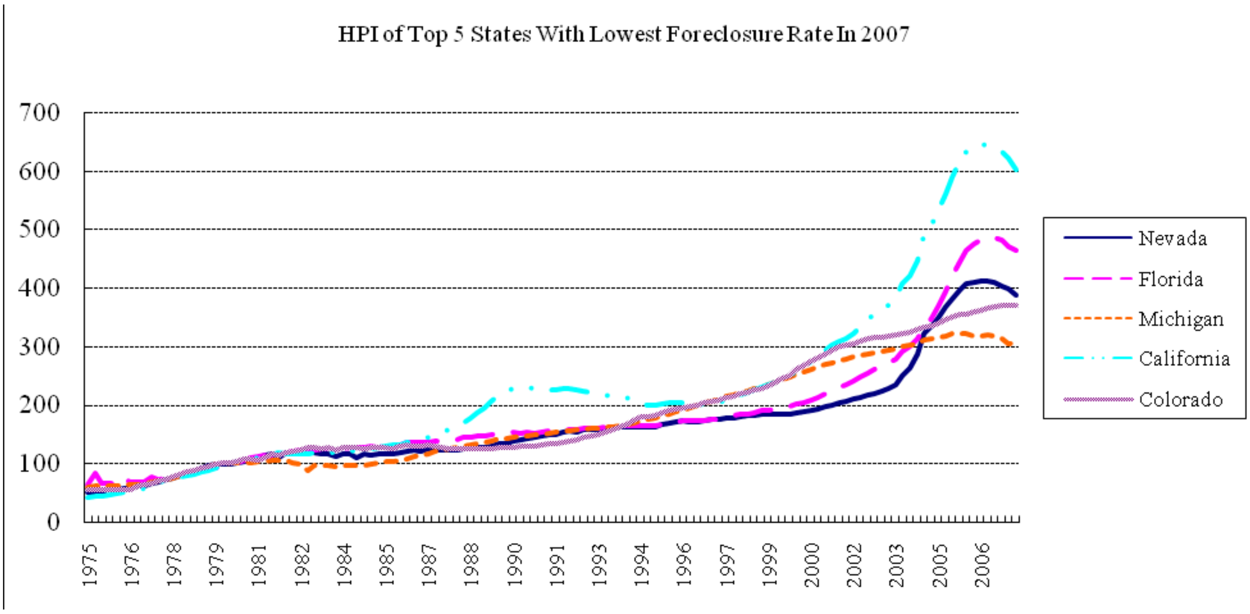


Figure 1.2 HPI of Bottom 5 states with lowest foreclosure rate in 2007

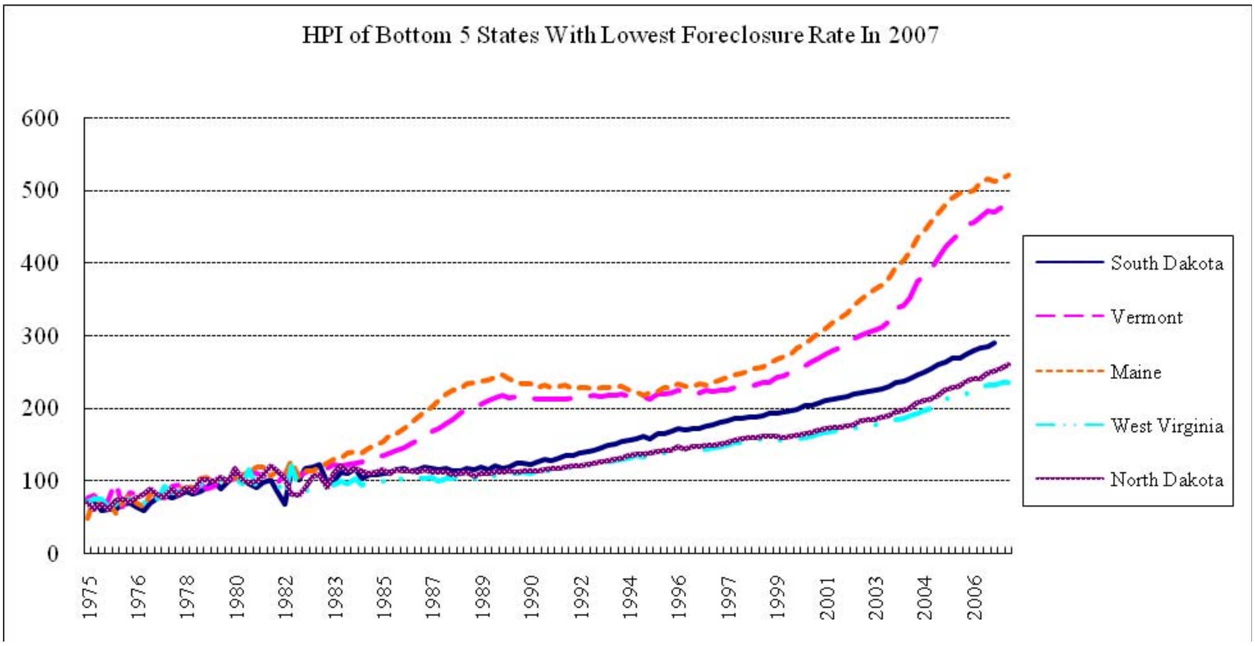


Figure 2 HPA (Housing price appreciation) and its Volatility at MSA level

Real house prices fluctuate over time. In recent years, while house prices have been buoyant, their volatility has declined markedly.

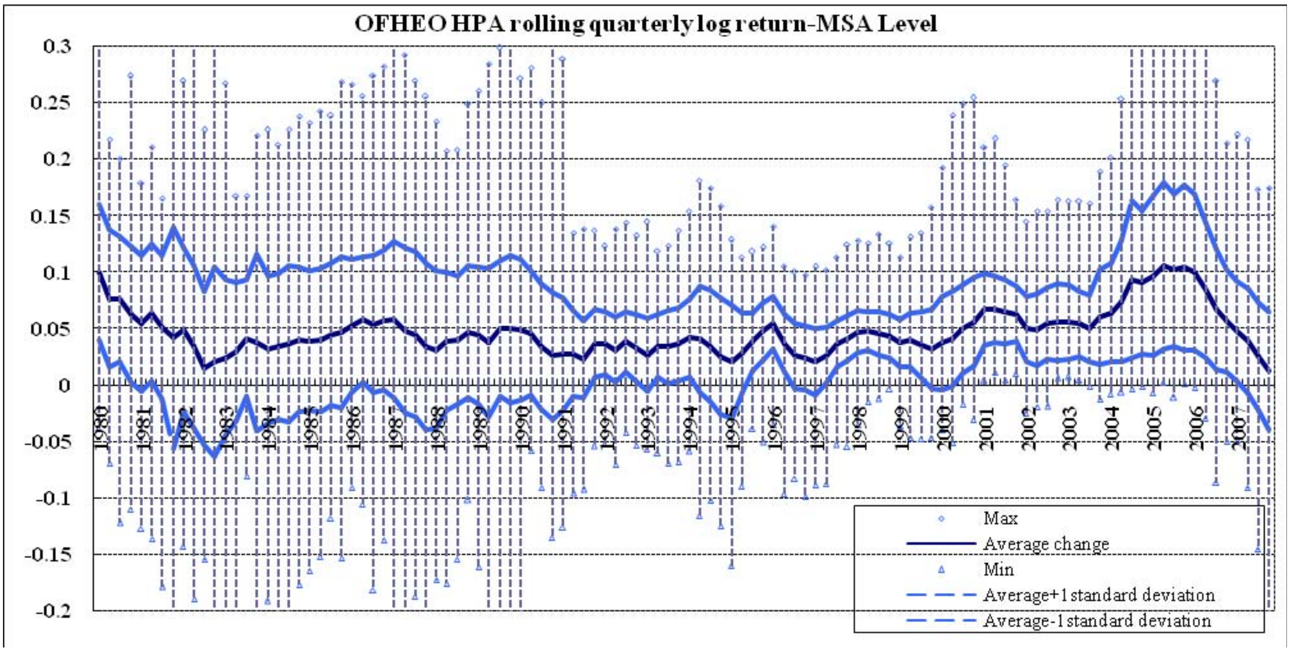
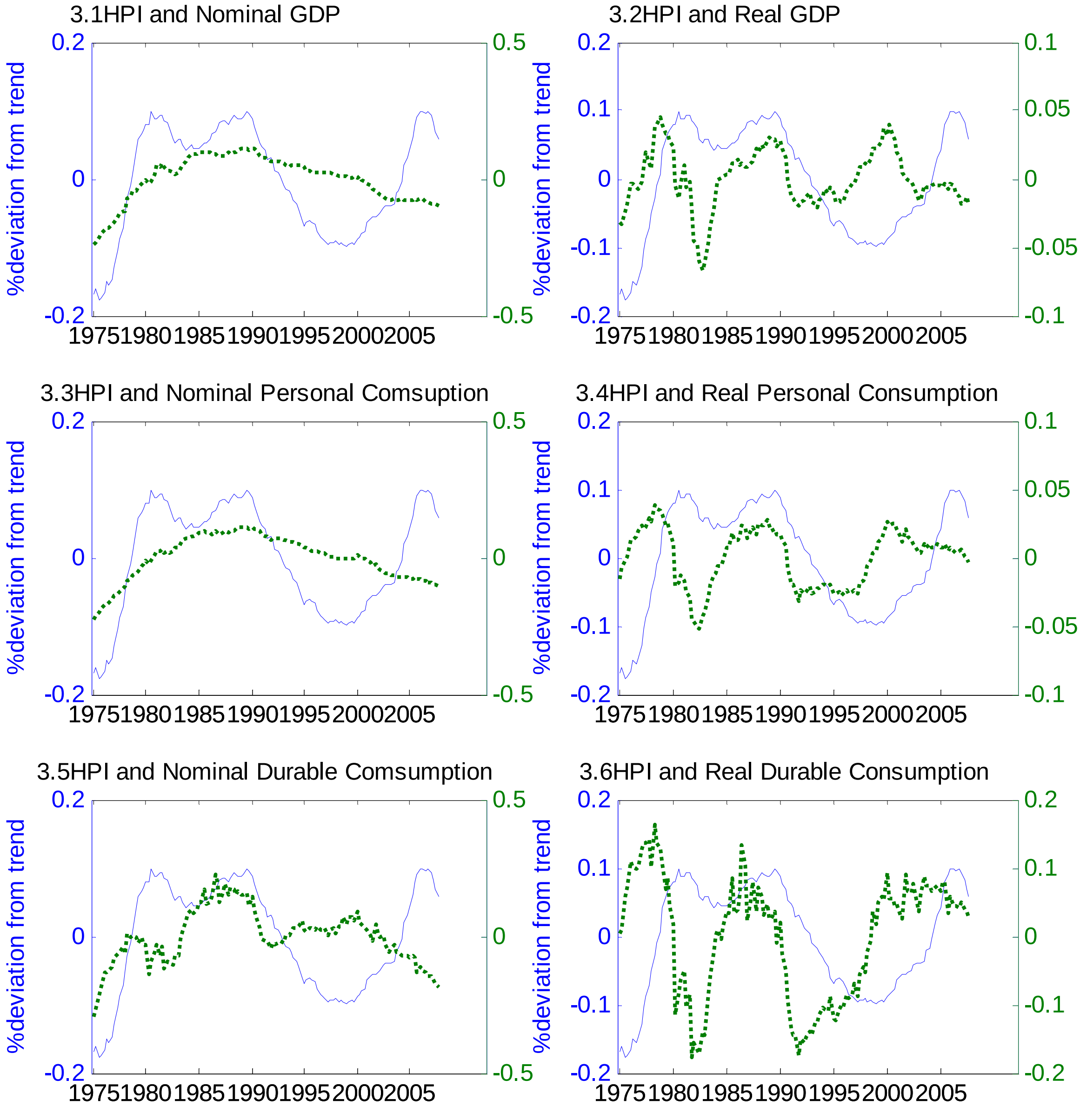
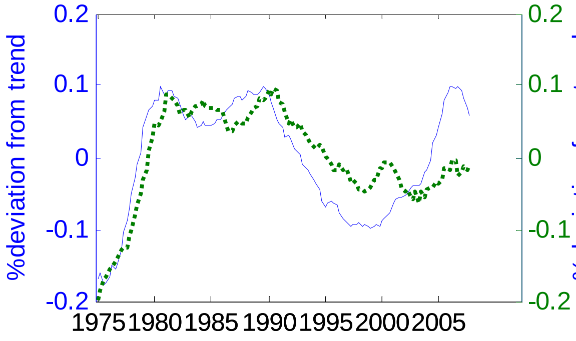


Figure 3 Co-movements between Macroeconomic Aggregates and House Prices

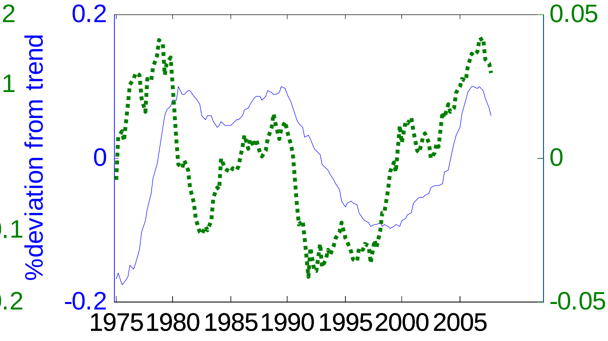
The figures show the changes in HPI and key macroeconomic variables over the period since 1975. All variables have been detrended by taking logs and then regressing on a constant and a linear trend. The strongest relationship seems to be that between house prices and consumption of nondurable goods.



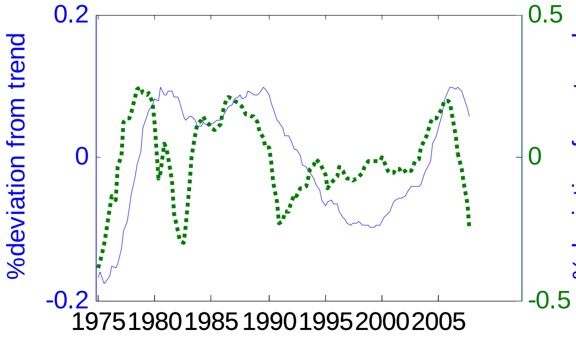
3.7HPI and Nominal NondurableConsumption



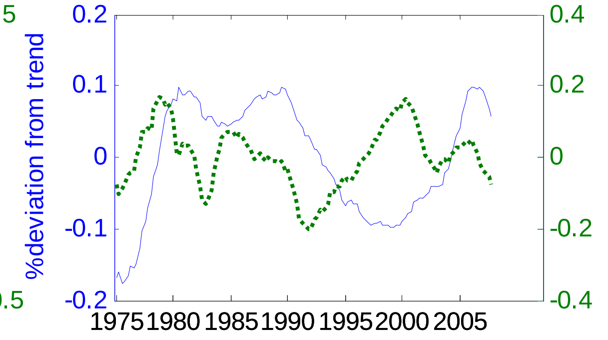
3.8HPI and Real Nondurable Comsumption



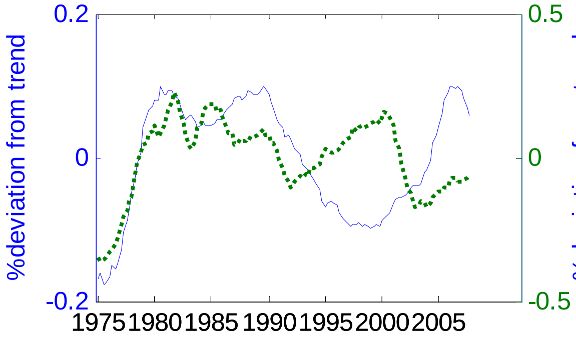
3.9HPI and Residential Fixed Investment



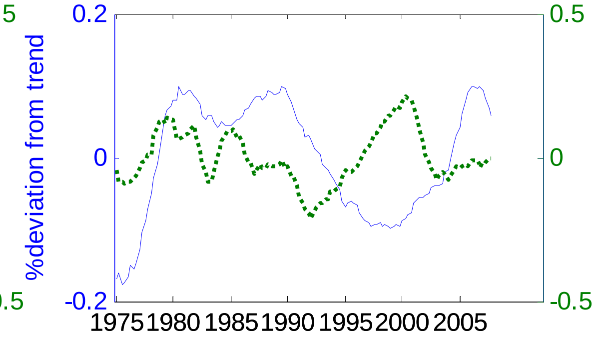
3.10HPI and Real Residential Fixed Investment



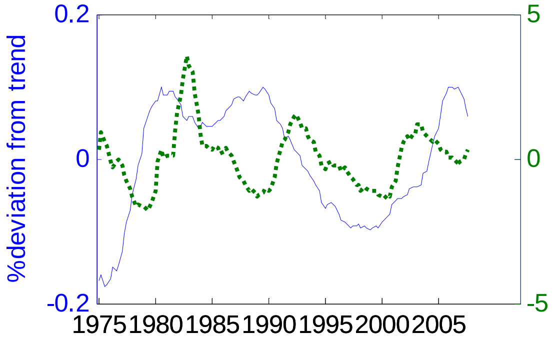
3.11HPI and Nonresidential Fixed Investment



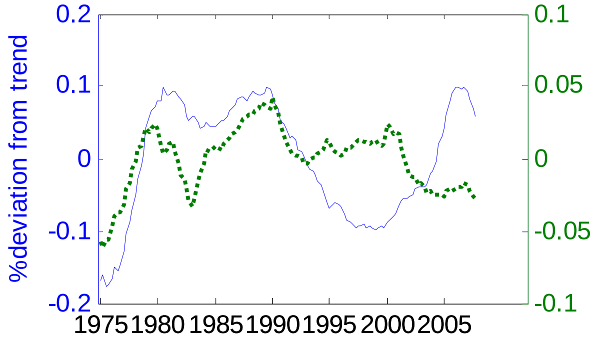
3.12HPI and Real Nonresidential Fixed Invest



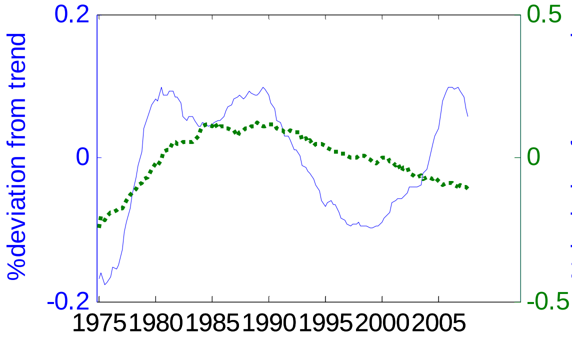
3.13HPI and Civilian Unemployment Rate



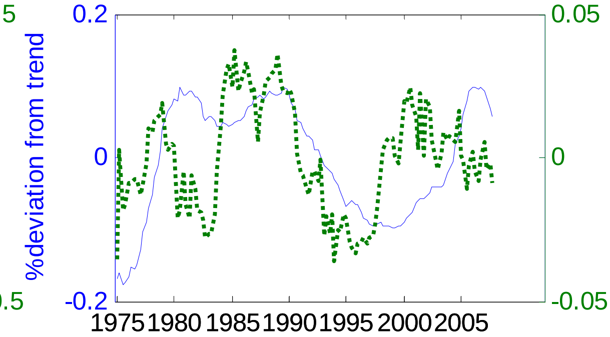
3.14HPI and Civilian Employment



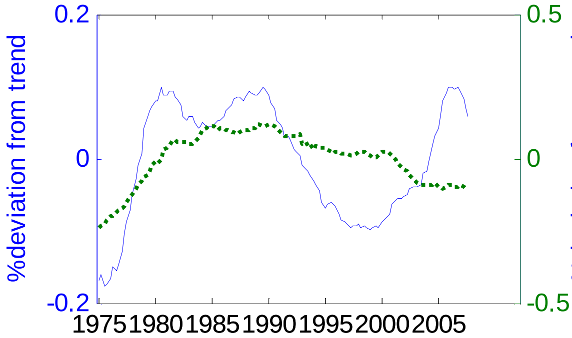
3.15HPI and Disposable Personal Income



3.16HPI and Real Disposable Personal Income



3.17HPI and Personal Income



3.18HPI and Oil Price

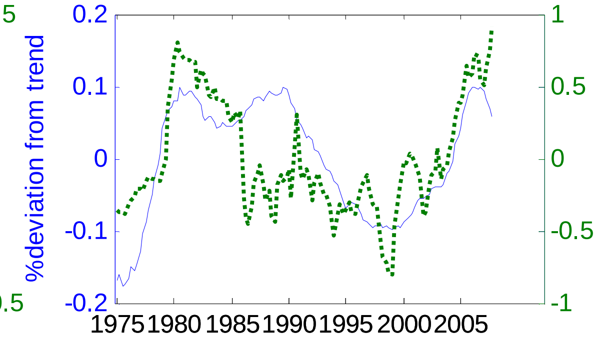


Figure 4 First Principle Component v.s. National HPA (OFHEO)

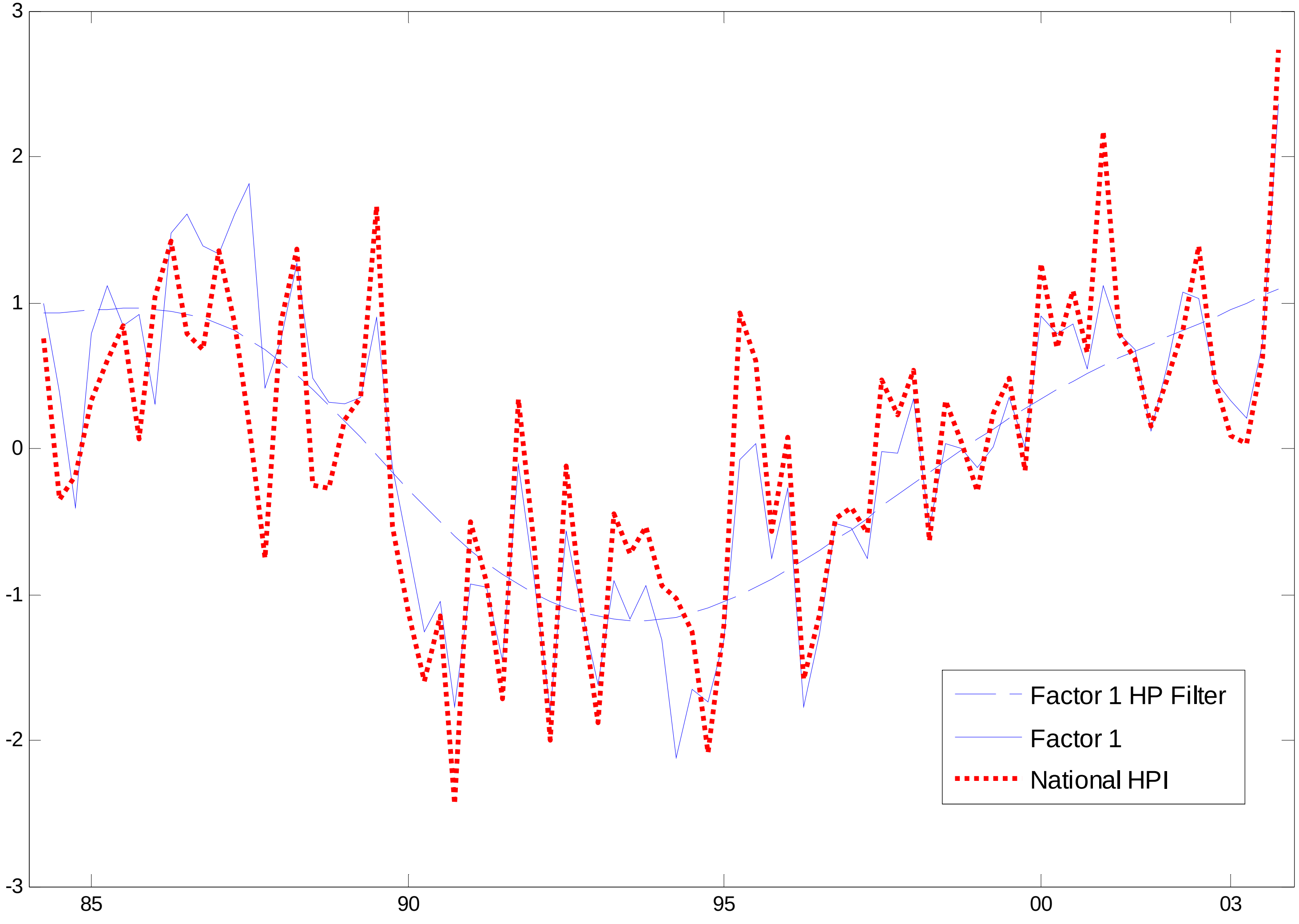


Figure 5 Factor 2~ Factor 5 v.s. National HPA (OFHEO)

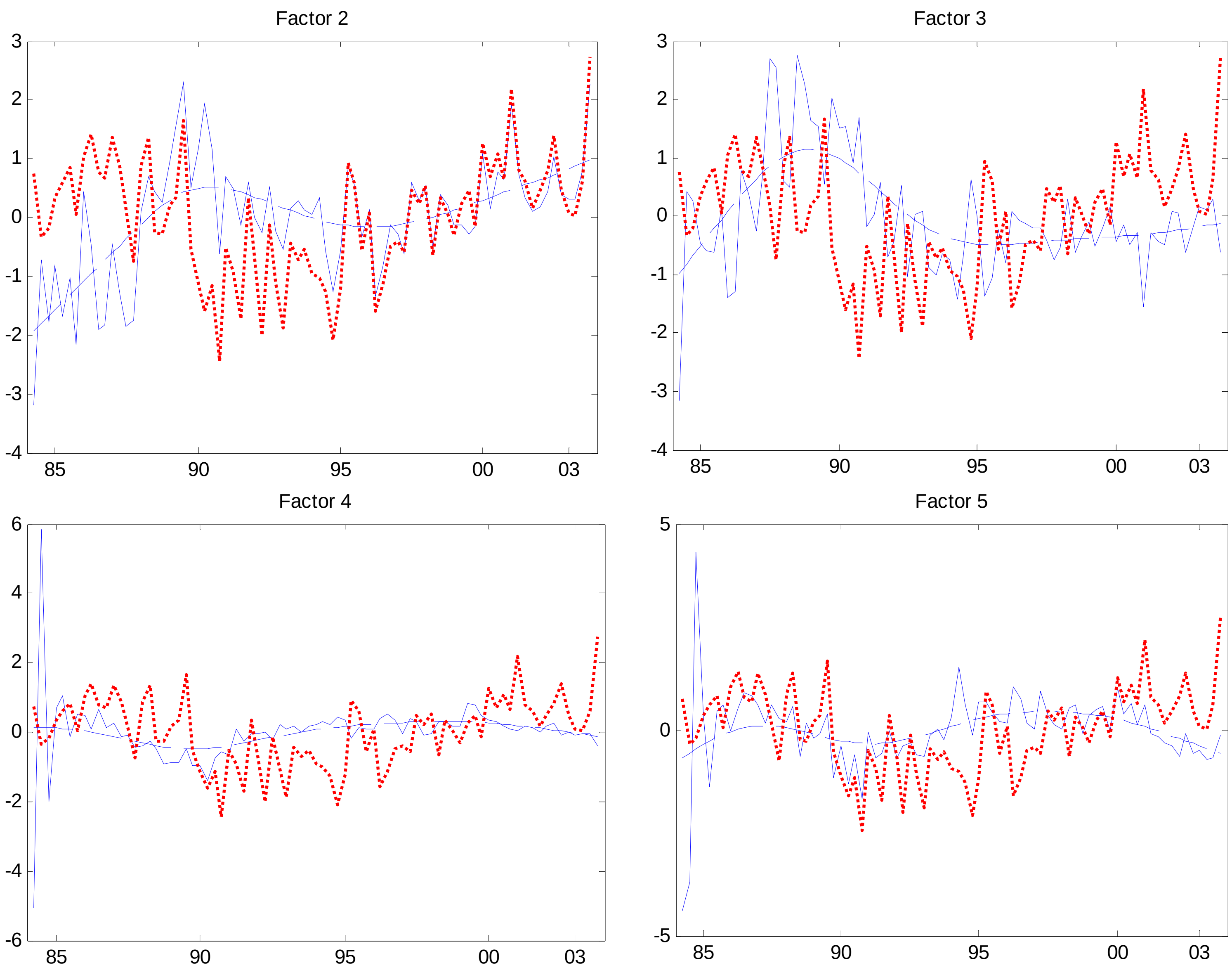


Figure 6 Factor Loadings (OFHEO HPA)

The following figures show the factor loadings for the first three factors of each MSAs.

Figure 6.1 Factor Loadings for First Factor

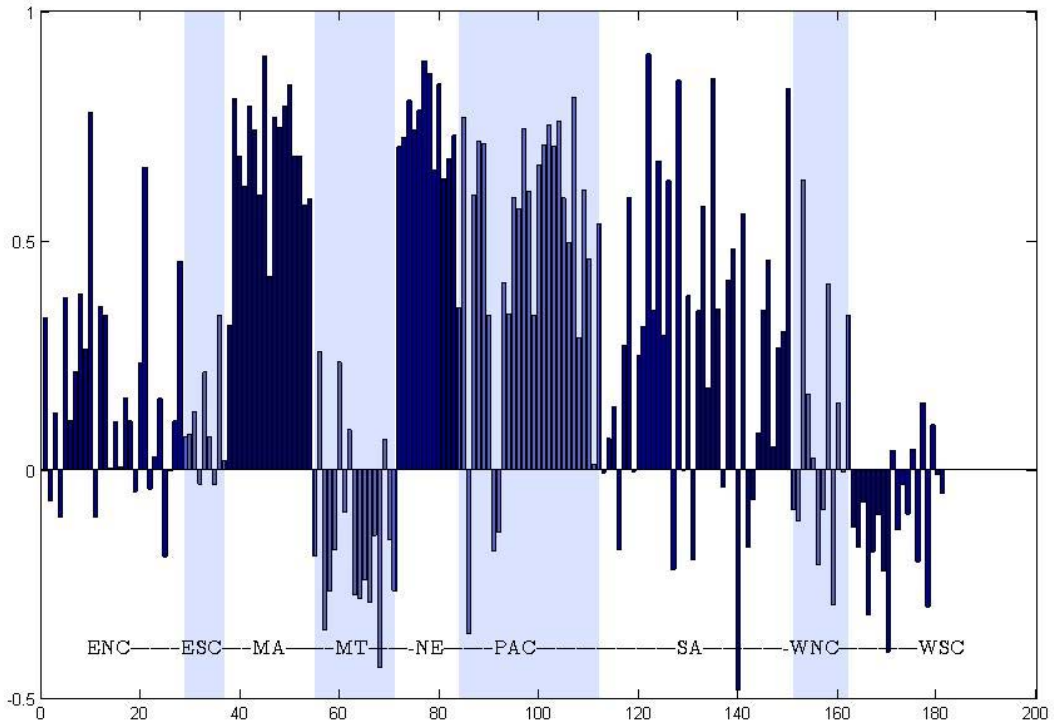


Figure 6.2 Factor Loadings for Second Factor

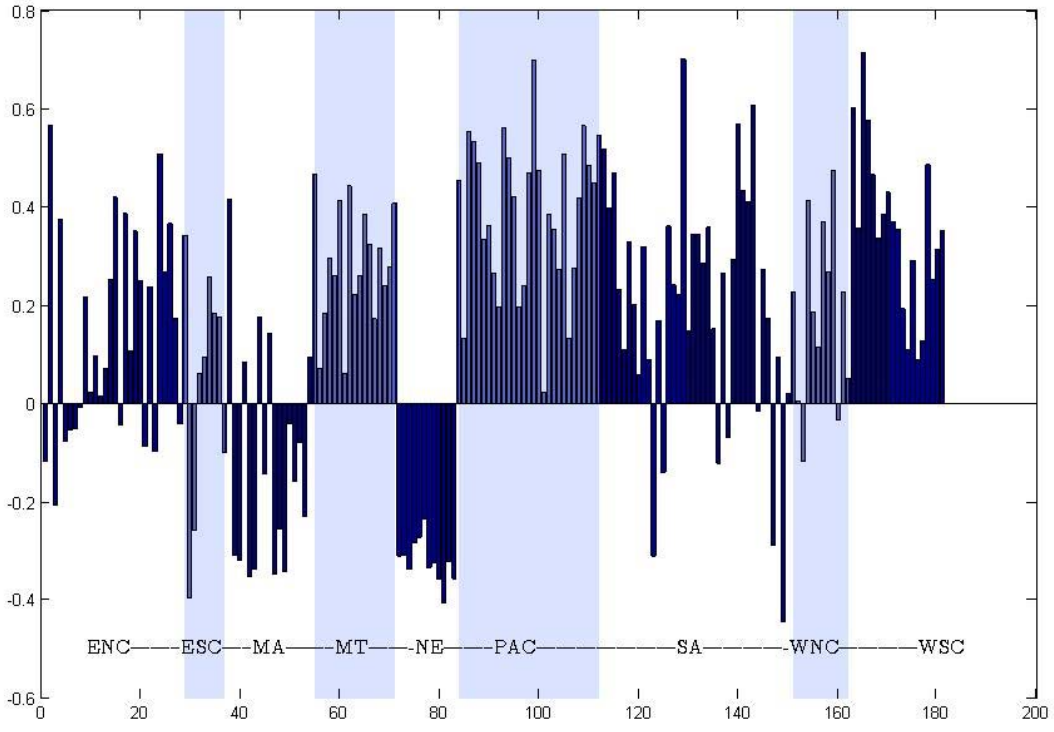


Figure 6.3 Factor Loadings for Third Factor

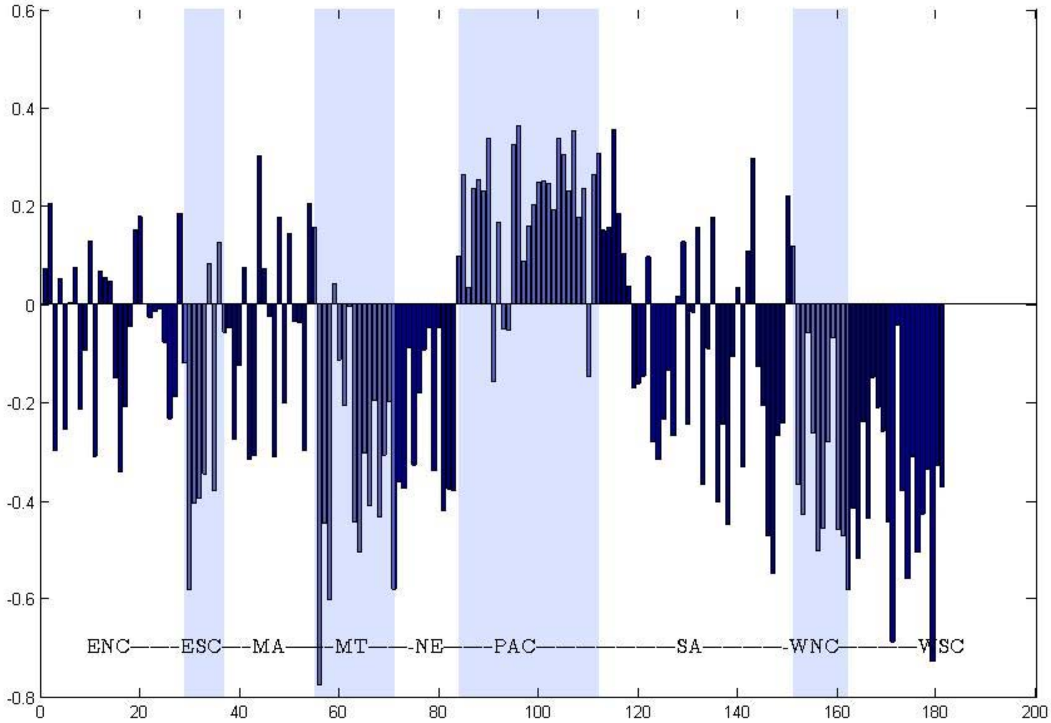


Figure 6.4 Factor Loadings for Fourth Factor

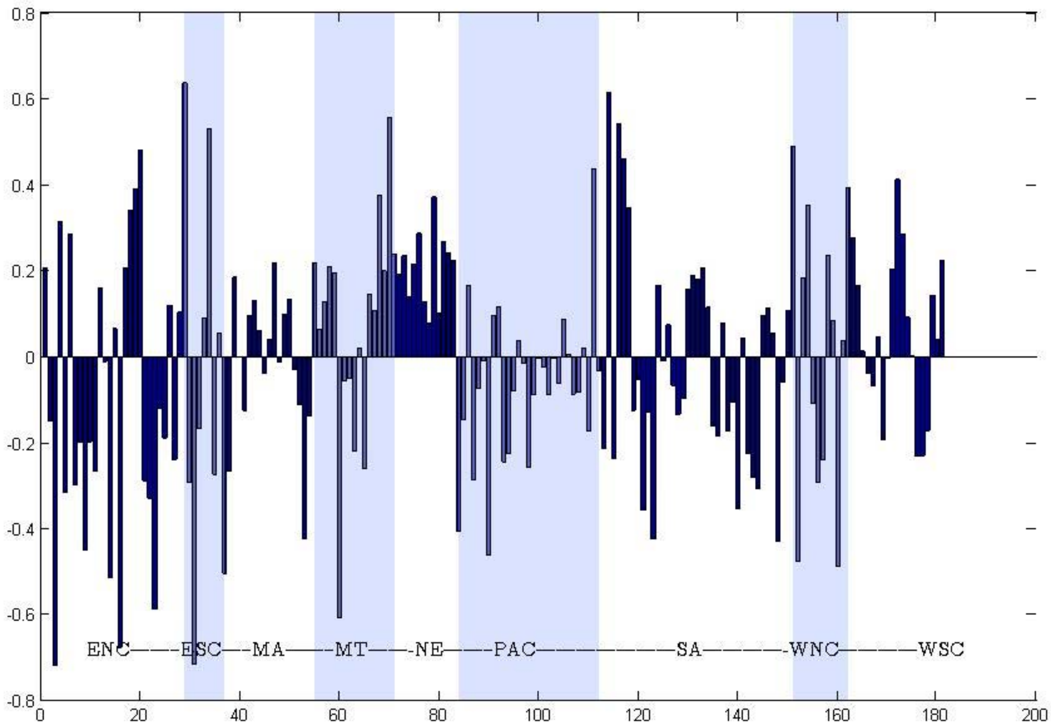
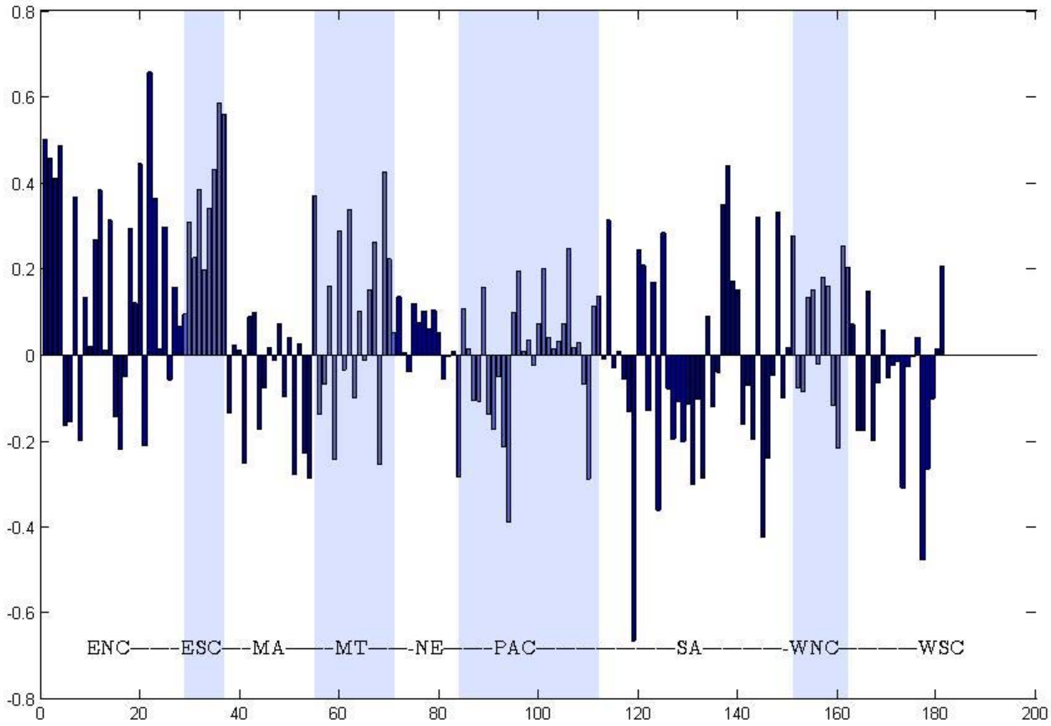


Figure 6.5 Factor Loadings for Fifth Factor



Note: The U.S. is divided into nine Census Divisions,

New England (NE): Connecticut, Massachusetts, Maine, New Hampshire, Rhode Island, Vermont

Mid-Atlantic (MA): New Jersey, New York, Pennsylvania

South Atlantic (SA): Washington, Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia

East North Central (ENC): Illinois, Indiana, Michigan, Ohio, Wisconsin

West North Central (WNC): Iowa, Kansas, Minnesota, Missouri, North Dakota, South Dakota, Nebraska

East South Central (ESC): Alabama, Kentucky, Mississippi, Tennessee

West South Central (WSC): Arkansas, Louisiana, Oklahoma, Texas

Mountain (MT): Arizona, Colorado, Idaho, Montana, New Mexico, Nevada, Utah, Wyoming

Pacific (PAC): Alaska, California, Hawaii, Oregon, Washington

Table 1: HPA Correlation (12 MSAs) for the period 1980Q1 to 2007Q1

This table shows the correlation for 12 HPAs. This table aims to show the correlation among different MSAs with different population sizes.

Population size	Big	---→										Small
	NewYork	Los Angeles	Chicago	Boston -Quincy	Seattle -Bellevue -Everett	SD-Carlsbad -San Marcos	Colorado Springs	Oakland -Fremont -Hayward	Columbia	Santa Rosa -Petaluma	Edison(NJ)	Bellingham
NewYork	1											
Los Angeles	0.43056	1										
Chicago	**0.526935**	0.448346	1									
Boston-Quincy	**0.766976**	0.302187	0.238364	1								
Seattle-Bellevue-Everett	-0.10256	0.406199	0.314598	-0.18682	1							
SD-Carlsbad-San Marcos	0.366237	**0.67595**	0.350106	0.340179	0.173349	1						
Colorado Springs	-0.01564	-0.01919	-0.19283	-0.10654	0.176528	0.025723	1					
Oakland-Fremont-Hayward	0.422036	**0.794359**	0.40951	0.391335	0.269125	**0.625444**	-0.12203	1				
Columbia	0.140712	0.159238	0.155382	-0.01276	0.366165	0.25447	0.411807	0.00861	1			
Santa Rosa-Petaluma	0.19068	**0.725303**	0.420645	0.21708	**0.570429**	**0.659464**	0.08064	**0.786238**	0.247356	1		
Edison(NJ)	**0.814356**	0.474463	0.395831	**0.709606**	0.006325	0.355047	0.000931	0.395409	0.222048	0.284212	1	
Bellingham	-0.16336	0.337949	0.085173	-0.21669	0.337394	0.375384	0.080187	0.272819	-0.10707	0.353607	-0.17232	1

Table 2: Estimated Numbers of Factors

In panel B, the PC and IC criteria choose r-max as the optimal number of latent factors
This result might be due to lack of dimensionality in cross sections. However, BIC3 shows much lower values even for those strategies with low dimensionality.

This table reports the results of the number of latent factors. Four criteria are considered in the study. r-max is set to be 8.

Dataset	Sample Period	# of Observations	PCp1	PCp2	ICp1	ICp2	BIC3	ER	GR
MRAC HPA(MSA)	1976Q2-2007Q2	381	7	7	3	3	2	1	1
OFHEO HPA (MSA)	1984Q2-2003Q4	181	14	14	11	11	7	1	1
OFHEO HPA (Stacked)[1]	1980Q2-2007Q1	195	15	15	15	15	12	1	1

[1] stack all the observations at MSAs, States and Census Divisions.

Table 3 Hypothesis Tests (OFEHO HPA, 84Q2-03Q4)

This table gives the results of the hypothesis tests. The variables in bold are comparatively significant.

	OFHEO-HPA-5 factors	
Notation	$R^2(j)$	CORR
NGDP	0.5376	0.7332
GDP	0.4693	0.6850
NPC	0.6138	0.7834
RPC	0.7259	0.8520
NDC	0.5000	0.7071
RDC	0.7651	0.8747
NNC	0.4583	0.6769
RNC	0.6889	0.8300
NRFI	0.6095	0.7807
RRFI	0.4572	0.6762
NPNRFI	0.1285	0.3584
RPNRFI	0.3150	0.5613
CU	0.2813	0.5304
CE	0.4543	0.6740
NDI	0.5822	0.7630
RDI	0.6311	0.7944
PI	0.3441	0.5866
OIL	0.0900	0.3000
FFR	0.1292	0. 3595
FYGM3	0.0802	0.2831
FYGM6	0.0644	0.2537
FYGT1	0.0522	0.2286
FYGT5	0.0221	0.1486
FYGT10	0.0368	0.1917
FYAAAC	0.0762	0.2761
FYBAAC	0.1600	0.4000
SFYGM3	0.4411	0.6641
SFYGM6	0.3060	0.5532
SFYGT1	0.0846	0.2908
SFYGT5	0.1576	0.3970
SFYGT10	0.1270	0.3564
SFYAAC	0.1637	0.4046
SFYBAAC	0.1644	0.4054
STOCK	0.0457	0.2138

Table 4 Hypothesis Tests (MRAC HPA)

	MRAC-HPA-2 factors		MRAC-HPA-5factors	
Notation	$R^2(j)$	CORR	$R^2(j)$	CORR
NGDP	0.3832	0.6190	0.3946	0.6282
GDP	0.0871	0.2952	0.1667	0.4082
NPC	0.3835	0.6192	0.3930	0.6269
RPC	0.2281	0.4776	0.3577	0.5980
NDC	0.2984	0.5462	0.3620	0.6017
RDC	0.3068	0.5539	0.4022	0.6342
NNC	0.3196	0.5653	0.3650	0.6042
RNC	0.2156	0.4644	0.2753	0.5247
NRFI	0.2159	0.4646	0.2456	0.4956
RRFI	0.1896	0.4355	0.2002	0.4474
NPNRFI	0.3097	0.5565	0.3901	0.6245
RPNRFI	0.0841	0.2899	0.1181	0.3437
CU	0.0464	0.2153	0.1251	0.3538
CE	0.0924	0.3040	0.1215	0.3486
NDI	0.4053	0.6367	0.4149	0.6441
RDI	0.1882	0.4339	0.2692	0.5188
PI	0.4236	0.6509	0.4363	0.6605
OIL	0.0010	0.0313	0.1808	0.4253
FFR	0.0987	0.3141	0.2241	0.4734
FYGM3	0.1026	0.3203	0.2286	0.4781
FYGM6	0.1058	0.3253	0.2370	0.4868
FYGT1	0.1149	0.3390	0.2530	0.5030
FYGT5	0.2241	0.4734	0.3746	0.6121
FYGT10	0.2567	0.5066	0.4150	0.6442
FYAAAC	0.3142	0.5606	0.4768	0.6905
FYBAAC	0.3003	0.5480	0.4852	0.6965
SFYGM3	0.0623	0.2496	0.1658	0.4072
SFYGM6	0.0240	0.1555	0.1203	0.3468
SFYGT1	0.0046	0.0678	0.0679	0.2606
SFYGT5	0.0022	0.0473	0.0277	0.1663
SFYGT10	0.0058	0.0759	0.0318	0.1783
SFYAAC	0.0267	0.1633	0.0472	0.2172
SFYBAAC	0.0547	0.2339	0.0839	0.2897

Appendix: Variables Description

Most of the data are taken natural log and detrended. Monthly data are transferred to quarterly data by taking average in the same quarter

Symbol	**Macroeconomic Indicators:**
NGDP	Nominal GDP
GDP	Real GDP
PI	Personal Income
NDI	Nominal Disposable Income
DI	Real Disposable Income
NPC	Nominal Personal Consumption
RPC	Real Personal Consumption
NDC	Nominal Durable Consumption
RDC	Real Durable Consumption
NNC	Nominal Nondurable Consumption
RNC	Real Nondurable Consumption
NRFI	Nominal Residential Fixed Investment
RRFI	Real Residential Fixed Investment
NPNFI	Nominal Private Nonresidential Fixed Investment
RPNFI	Real Private Nonresidential Fixed Investment
CU	Civilian Unemployment
CE	Civilian Employment
Oil	Oil Price
	Asset Class Benchmarks:
FYFF	FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)
FYGM3	U. S. TREASURY BILLS,SEC MKT,3-MO. (% PER ANN,NSA)
FYGM6	U. S. TREASURY BILLS,SEC MKT,6-MO. (% PER ANN,NSA)
FYGT1	U. S. TREASURY CONST MATUR., 1-YR. (% PER ANN,NSA)
FYGT5	U. S. TREASURY CONST MATUR., 5-YR. (% PER ANN,NSA)
FYGT10	U. S. TREASURY CONST MATUR., 10-YR. (% PER ANN,NSA)
FYAAAC	BOND YIELD: MOODY’S AAA CORPORATE (% PER ANNUM)
FYBAAC	BOND YIELD: MOODY’S BAA CORPORATE (% PER ANNUM)
SFYGM3	1 Spread FYGM3—FYFF
SFYGM6	1 Spread FYGM6—FYFF
SFYGT1	1 Spread FYGT1—FYFF
SFYGT5	1 Spread FYGT5—FYFF
SFYGT10	1 Spread FYGT 10—FYFF
SFYAAAC	1 Spread FYAAAC—FYFF
SFYBAAC	1 Spread FYBAAC—FYFF
Stock	Quarterly Return S&P Composite Index
Commodities	Commodities Goldman Sachs Commodity Total Return Index