2014

The Importance Of Housing Variables In Predicting Economic

Wenxian Xu
University of Mississippi

THE IMPORTANCE OF HOUSING VARIABLES IN PREDICTING ECONOMIC FLUCTUATIONS

A Dissertation
presented in partial fulfillment of requirements
for the degree of Doctor of Philosophy
in the Economics Department
The University of Mississippi

by

WENXIAN XU

May 2014

ABSTRACT

This paper investigates the importance of housing variables in predicting the six recent recessions using factor analysis model. It shows that housing variables have great predictive power in forecasting the downturn of the economy, but little predictive power when the economy is steady or is expanding. The explanation is that housing variables match consumers' expectations of future income and employment, and consequently predict future economic downturn.

By using Granger-Causality test and vector autoregression (VAR) model, combination of factor analysis and VAR model, and hard thresholding method to identify the importance of each housing variable, the results show that housing price indexes are not important in forecasting the economy, but that measures of housing volumes improve predictions. Moreover, the housing volume measures with one-unit of structure tends to play a greater role in the prediction.

DEDICATION

This work is dedicated to everyone who has gone through millions of lonely nights reading never-ending papers and running ever-lasting programs.

LIST OF ABBREVIATIONS AND SYMBOLS

10 year T-bill – fed funds rate (TB_FFR)

3 months CP rate (CP3)

3 months T-bill rate (TBILL3MO)

3 months T-bill rate-3 months CP rate (TBCP3M)

Akaike information criterion (AIC)

Average weekly initial claims for unemployment insurance (WICI)

Avg. weekly hours of production workers in manufacturing (AWHP)

Bayesian information criterion (BIC)

Capacity utilization rate in manufacturing (CAP)

Case-Schiller price index (CSPI)

Distressed sales last 1 month (DSAL1MO)

Distressed sales last 12 month (DSAL12MO)

Distressed sales last 3 month (DSAL3MO)

Distressed sales last 6 month (DSAL6MO)

Department of Defense (DOD)

Economic Bulletin Board (EBB)

Employment in nonfarm industry (EMP)

Expectation maximization (EM)

Factor-augmented vector autoregressive (FAVAR)

Fed funds rate (FFR)

Gross domestic outproduce (GDP)

Housing Starts, 1-Unit Structures (HOUST1F)

Housing Starts, 2-4 Unit Structures (HOUST2F)

Housing Starts, 5-Unit Structures or more (HOUST5F)

Housing Starts, total (HOUST)

Housing Units Authorized by Building Permits, 1-Unit structures (PERMIT1)

Housing Units Authorized by Building Permits, 2-4 Unit Structures (PERMIT24)

Housing Units Authorized by Building Permits, 5-Unit Structures or more (PERMIT5)

Housing Units Authorized by Building Permits, total (PERMIT)

Housing Units Authorized, But Not Yet Started, 1-Unit Structures (AUTHNOT1U)

Housing Units Authorized, But Not Yet Started, 2-4 Unit Structures (AUTHNOT24U)

Housing Units Authorized, But Not Yet Started, 5-Unit Structures or more (AUTHNOT5MU)

Housing Units Authorized, But Not Yet Started, total (AUTHNOTT)

Housing Units Completed, 1-Unit Structures (COMPU1USA)

Housing Units Completed, 2-4 Unit Structures (COMPU24USA)

Housing Units Completed, 5-Unit Structures or more (COMPU5MUSA)

Housing Units Completed, total (COMPUTSA)

Housing Units Under Construction, 1-Unit Structures (UNDCON1USA)

Housing Units Under Construction, 2-4 Unit Structures (UNDCON24USA)

Housing Units Under Construction, 5-Unit Structures or more (UNDCON5MUSA)

Housing Units Under Construction, total (UNDCONTSA)

Index of consumer expectations (CCI)

Index of supplier deliveries, vendor performance (ISDVP)

Industrial production (IP)

Interest rate on 10 year US T-bill constant maturity (TBILL10YR)

ISM® new orders index (ISMNOI)

Manufacturers' new orders, consumer goods and materials (MOCMQ)

Manufacturers' new orders, nondefense capital goods (MNONC)

Mean squared error (MSE)

Median price index with single-house combined (HPI)

Monetary Services Index for all zero-maturity assets (MSIMZM)

Monetary Services Index for M1 (MSIMIM1)

Monetary Services Index for M2 (MSIM2)

Monetary Services Index for M2 omitting the small-denomination time deposits (MSIM2M)

Monetary Services Index including all the assets (MSIALL)

Money supply M2 (M2)

National Association of Purchasing Managers' index of vendor performance (NAPM)

New One Family Homes For Sale in the United States (HNFS)

New One Family Houses Sold: United States (HSN1F)

Number of people working part-time in nonagricultural industries because of slack work (EMPPT)

Purchase Only House Price Index for the United States (HPIPO)

Real manufacturers' unfilled orders in durable goods industries (UMOD)

Real manufacturing and trade sales (MATS)

Real personal income less transfer (RPI)

Residential price index (FNCP)

Spread between 10-year and 1-year Treasury bonds (TB110)

Stock prices, S&P500 common stocks (SP500)

Trade weighted index of nominal exchange rates between US and UK, West Germany, France, Italy and Japan TWIER (TWIER5)

Vector autoregression model (VAR)

ACKNOWLEDGMENTS

First, I want to thank Dr. Walt Mayer for all the academic advices and technical support he provided. He has always been very supportive and responsible. I am forever grateful for his hard work and patience.

Second, I want to thank my committee members: Dr. William Chappell, Dr. Xin Dang, and Dr. Joshua Hendrickson for their valuable wisdom, advice, and support.

Finally, I would like to give many thanks to my family and friends who have always provided support and love throughout the highs and lows of life. I greatly appreciate all the help, friendship, and technical support from my wonderful professors and colleagues.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

HOUSING MARKET AND MOTIVATION

There has been an upward trend in the housing market since 1990 in which the housing prices, residential investment, and home sales have been increasing and setting new records. In 2004 and 2005, large numbers of buyers have rushed into the housing market due to favorable interest rates and expected appreciation in the home prices. The increase caused the homeownership rates to rise to 69% during these two years for all ages, races, and ethnicities (see the 2005 housing market report by the Joint Center for Housing Studies of Harvard University).

This housing boom has been the longest expansion since 1970 as shown in *Figure 1*[1] (Joint Center for Housing Studies of Harvard University)[2]. The housing supply has been growing with an annual average of over 1.9 million units in housing starts and manufactured home placements since 2000. Meanwhile, the ongoing demand in the housing market is also increasing in conjunction with the rate of construction. This can be demonstrated by the low rate of the inventory of new homes for sale relative to the home sales of the end of 2004.

In addition, the mortgage finance system created varieties of products which offered buyers more flexible loans with a lower adjustable rates even when the interest rate rose. Moreover,

[1] All figures are placed in the Appendix A.

[2] Quoted from the State of the Nation's Housing 2005 by the Joint Center for Housing Studies of Harvard University.

immigrants and future generations continue to contribute greatly to the increasing demand for new homes.

The housing market was still strong due to a strong job market, high demand for homeownership, and flourishing rental market. This also showed an appreciation of housing prices, but it started to show signs of cooling in late 2005. As mortgage interest rate and housing price grew, home builders responded to the softening by pulling back on construction. In addition, home sales dropped, causing a higher inventory of new and existing homes for sale.

It was not until 2006 that the housing market slid dramatically after the starts, sales, and housing prices peaked in 2005. The home sales dropped by 10%, housing starts decreased by 13%, and the housing price appreciation fell to only a few percentage points. The inventory of vacant house for sale soared by more than 500,000 between late 2005 and late 2006. Large numbers of homes in foreclosure entered the market. Moreover, the affordability problem even deteriorated. The households with housing burden of more than 30% income increased by 2.3 million compared to the record of 37.3 million in 2005.

These abrupt downturns originated from the bubble burst in the inflated demand of the for-sale market. The housing demand increased due to the low interest rate and large expected housing price appreciation. Buyers entered the market trying to take advantage of potentially rising housing prices while investors targeted at a quick resell. Builders catered to the rising demand by their increase in construction, but the time lag between housing preparation work and housing completion drove the home price even higher. In the meantime, mortgage lenders offered the anxious buyers lower down-payments and eased requirements for those borrowers with unqualified credit records.

When the higher mortgage interest rate and increased housing price took effect, some buyers were forced out of the market. Weak home sales and slow housing price appreciation caused the upsurge of house-buying to vanish and investors to draw back from the market. Although the home builders tried to pull back on productions, the lag made the cut too late.

There were a growing number of foreclosed houses returning to the market. Although only less than half of the houses were eventually sold according to Freddie Mac, the swarm of these houses in the already bloated market led to high pressure on housing prices. The nearby homeowners in the communities suffered great loss in home equity and the local government also suffered drastic decline in property tax collections.

The metropolitan areas faced great risk of widespread foreclosure since they had weak economies, high subprime shares, and oversupplies of housing. The communities with low income and minority suffered the most because of the predominant subprime loans.

The affordability problem has been persistent even during the economic boom from 2000. In 2006, more than 39 million households were paying more than 30 percent of income on housing while 18 million were paying more than 50 percent. The number of severely burdened households climbed by almost four million from 2001 to 2006. This severely affected low-income populations. In 2006, 47% of low-income families were severely burdened by housing costs, while only 11% of lower middle-income households and 4% of upper middle-income households were exceptionally affected. The huge difference between housing costs and income reflected large numbers of low-paying and part-time jobs generated by the economy, high costs of maintaining houses, as well as construction and renovation by local governments.

Although many housing markets were entrapped in massive turmoil in 2006, the economy as a whole was still in a boom with fast economic growth, increased consumption, expenditure, and strong retail sales. The economy did not collapse until the end of 2007. This situation resembles the Great Depression in the 1930s. Eighty five years ago, the economy seemed to be in a boom with increased stock dividends, expansion of existing companies, entry of new firms, rising personal savings, and a bull stock market. Everything looked good except the housing market. House prices increased and mortgage foreclosure rates skyrocketed which caused distress in the housing market. In 1927, housing starts in the US started to decline from 2,423,000 until 221,000 in 1933, which began a long-lasting and devastating collapse worldwide.

These two recessions shed light on the relationship of housing market and the downturn point of the economy. In order to explore the importance of housing variables in predicting the economy, target variables should be chosen. GDP is a good summary of the economic conditions, but it's quarterly based, not monthly. Therefore this paper follows Stock and Watson (2002) to use four variables as target variables: real personal income, employment in nonfarm industry, industrial production, and real manufacturing and trade sales. In addition, in order to measure the predictive power of housing variables to GDP more directly, I use Stock and Waston's monthly estimation of GDP as another target variable.

Comparing *Figure 2* and *Figure 3*, we can see that well before the two recessions began, housing variables were the first to show pessimistic signs. Housing variables started to decrease from 2006, which were almost two years before the four target variables dropped, as shown in *Figure 2* and *Figure 3*. Similarly, when looking at the recovery of the recession, what can be easily seen is that housing variables gradually recovered before the rise of the four target variables. Combining *Figure 2* and *Figure 3*, we could conjecture that housing variables are signals indicating the ups and downs of the economy.

Since housing variables began to decrease before the economy as a whole decreases, they are performing like leading indicators. Leading indicators, composed of a series of correlated economic indicators, reflect the fluctuations of the economy as a whole, and predict the possible trends in future.

Typically the leading indicators are announced on the last weekday of each month by the US Department of Commerce. Generally speaking, if the leading indicators continue decreasing in three months, then an economic recession is expected. Vice versa, if they continue increasing in three months, then the economy is expected to boom. Usually leading indicators change six to nine months ahead of the turning point of the economy. The U.S. Department of Commerce's Economic Bulletin Board (EBB)[3] pointed out that leading indicators could predict downturns of the economy eleven months ahead of time, and predict economic recovery three months before expansion.

Leading indicators have played an important role in the evolution of macro econometric forecasting models. Some papers, such as Stock and Watson (2002) and Forni (2002), have conducted experiments demonstrating that leading indicators contribute greatly in predicting economy as a whole, but they put too little weight on housing variables.

This dissertation investigates the value added of using housing variables to predict the economy. Firstly, I will discuss the reason I choose housing variables to improve prediction instead of any other variables, that is, the potential importance of the housing variables.

---

[3] It is a dial-up bulletin board system, delivering all major U.S. Government economic information including economic indicators from the Bureau of the Census, Bureau of Economic Analysis, Federal Reserve Board and Labor Department.

As presented in Leamer's 2007 paper, residential investment is a small part of long-run growth. The table "Contributions to Real GDP Growth" compiled by the Bureau of Economic Analysis was cited and graphed in his paper, showing that during 1950 and 2005, the largest contribution to economic growth comes from consumer services, and then from other consumer spending items including durables and non-durables. Residential investment, however, accounts for only a small fraction of the economic growth. Therefore, in the long run, housing variables do not contribute significantly to the growth of GDP.

However, when referring back to the ten recessions in history, we could see the signaling effect of housing variables to the weakness before the recessions and the recoveries. Leamer presents residential investment movements around the ten peaks of economy which illustrates the weakness and strength of housing variables before and during recessions, thereby showing that residential investment contributes significantly and consistently to the downturn of the economy, and also rises earlier than the subsequent recovery. By comparing the cumulative contributions around the recessions of all possible economic variables such as equipment and software, durables and non-durables, Leamer shows that housing is the largest contributor to the recessions in 1949, 1957, 1960, 1970, 1974, and 1980, and is the second largest in 1981 and 1990. Housing shows less importance in 1953 and 2001 recession, which are due to Department of Defense (DOD) downturn and a collapse in equipment and software respectively. Therefore, it is the residential investment that indicates most of the start of recessions. Housing variables do affect the whole economy ahead of time and predicts the economy.

The next step of this dissertation is to choose models for forecasting economy with housing variables added. Each economic event happens with co-movements of large number of macroeconomic variables, and therefore, the economists have used large dataset to predict economic fluctuations. Thanks to the advance in information technology, large datasets of economic indicators are now available. However, if the number of parameters to estimate exceeds the number of observations, then there is an identification problem. To address this problem, principal component and factor analysis models are used. The basic idea is that the information in a large number of variables can be compressed to a handful of predictors which will be included in the model instead.

A basic dynamic factor model was developed by Sargent and Sims (1977). In late 1937, Mitchell and Burns published papers stating a list of leading indicators as well as coincident and lagging indicators of economic activities as factors that play important roles in measuring the cyclical behaviour of business cycles. The idea of a common business cycle was first raised by Mitchell and Burns (1946), and the National Bureau of Economic Research (NBER) developed the indexes of leading and coincident indicators. This approach was extended to a basic dynamic factor model by Sargent and Sims (1977) as well as Geweke (1977).

In the last few years, a couple of papers about factor analysis model were proposed as a improvement. Stock and Watson (2002) combines the approximate factor model and the dynamic factor model and uses estimated factors to make out-of-sample forecasts. Forni, Hallin, Lippi, and Reichlin (2002) also combines the approximate factor model and the dynamic factor model. The model assumes that the idiosyncratic components are not necessarily orthogonal to each other, and also, a minimal amount of cross-correlation for common components and a maximal amount of cross-correlation for idiosyncratic components are imposed.

Although a large dataset provides plenty of information, it is not wise to add as many variables as possible. A basic statistical principle states that more data improves efficiency, but it might be the case that the cross-correlation in the errors is too large and the variability of the common component is too small. Under the case that the errors are cross-correlated or have big unequal idiosyncratic error variances, the estimated factors and the forecasts of target variables will be less efficient. Therefore, it is necessary to exclude "noisy" data in the housing variables, according to Bovin and Ng (2005). Three methods are used in this paper to exclude noisy data for factor analysis model as follows.

The first method is using VAR model for Granger-causality test with target variable, endogenous variables, each housing variable, and lags of them, holding non-housing variables constant. The importance of each housing variable could be tested by AIC and BIC criteria in VAR model, holding non-housing variables constant. The housing variables with lower criteria imply larger importance to the forecasts of the economy.

The second method is factor-augmented vector autoregressive (FAVAR) model, combining factor analysis and standard VAR models, as used in Friedman, B., Kuttner, K. (1993). A two-step method of FAVAR model is that first estimating factors of the large dataset then including these factors in the VAR model, and choosing important housing variables with low information criterion.

The last method is called hard thresholding, which is used in Bai and Ng (2008). This method uses t-statistical test to measure if the i-th predictor is significant without controlling for other predictors.

The contribution of this dissertation is that it shows the great predictive power of housing variables in forecasting the downturn of the economy using factor analysis model for five recent recessions, but little predictive power when the economy is steady or is expanding. It also shows that it is housing volumes that matter, not housing price, as stated in Leamer (2007).

The dissertation is organized as follows. Chapter 2 briefly introduces the literatures closely related to this dissertation. Chapter 3 describes the resources and definitions of the data used. Chapter 4 explains the factor analysis model and the methodology used to select important housing variables, and also presents the corresponding experiments results. Chapter 5 provides theoretical explanations of why housing variables could be good predictor for the recessions.

CHAPTER 2

LITERATURE REVIEW

Forecasting with large dataset of predictors brings richer information and relative robustness against structural instability compared to low dimensional forecasting. However, problems arise at the same time. When the number of variables is larger than the number of time series, identification problems occur. Two methods that could be used to solve the identification problem include variable selection procedures and principal component and factor analysis models. In this dissertation the latter method will be used.

The concept of a business cycle which describes the co-movement of many macroeconomic variables was first raised by Burns and Mitchell (1946). This idea was best captured by the dynamic factor model of Sargent and Sims (1977) and Geweke (1977). Sargent and Sims assumes the idiosyncratic components to be independent stationary processes and estimates factors using the Kalman filter method.

Stock and Watson (2002b) and Forni, Hallin, Lippi, and Reichlin (2000), however, relax some of the restrictions and propose new estimation methods. They are considered important "second generation" factor models. It allows the cross-sectional dimension approaching to infinity for the approximate factor model. Four assumptions are imposed specifying that a limited amount of dynamic cross-correlation and a restricted cross-correlation between the common components. The four assumptions compose a so-called generalized dynamic factor model. This model is applied to a large panel including several macroeconomic variables for

EURO countries to compute a coincident indicator. The paper proposes new method to consistently estimate factors as both cross section and the time dimensions approached infinity.

The generalized dynamic factor model generated by Forni *et al.* (2000) allows for non-orthogonal idiosyncratic components. Compared with Stock and Watson (2002b), Forni *et al.* does not allow the factor loading coefficients to be time-varying, also it requires two-sided smoothing, which makes the estimate of principal components not available at the end of the sample.

Stock and Watson (2002a, 2002b) provides a theoretical factor model and an empirical application. Stock and Watson focuses their study on the case when the number of candidate predictors and the number of time series observations are both large. The paper does not limit the assumptions of the idiosyncratic disturbance to cross-sectionally independent or serially correlated, which are unrealistic, but allows for both serial and cross-sectional correlation. These assumptions are also provided in Chamberlain and Rothschild (1983), Connor and Korajczyk (1986 1993), and Forni *et al.* (2000).

In addition, Stock and Watson adopts assumptions for the factors and factor loadings, as well as the regressors and the resulting forecast errors. Also the assumptions allows the factors to be serially correlated and lags of factors to be included.

Compared to the three-step method which assumes N is small in Stock and Watson (1989), Stock and Watson (2002b) allows N to be large, which makes iterative nonlinear methods such as used in Stock and Watson (1989) computationally difficult. Therefore, they propose a two-step method that estimates the factors by principal components and then predicts the target variable using the estimated factors and lagged dependent variable.

Stock and Watson (2002b) contributes in three aspects. Firstly, the estimation of principal components is consistent under the condition that both N and T converge to infinity. Secondly, the feasible forecast of the target variable converges to the optimal infeasible forecast as N and T approach infinity, which means that the feasible forecast is asymptotically first-order efficient since the MSEs of the two forecasts approach the same as N and T become larger. Thirdly, Stock and Watson allow temporary instability over long time period as stochastic drift in the factor loadings and show that above results still hold if the drift is small and idiosyncratic.

Stock and Watson (2002b) also raises several potential methodological issues: Efficiency problem if heteroscedasticity and serially correlation exist, a distribution theory is needed to provide measures of the sampling uncertainty of the estimated factors, also what would happen if any strong persistence is introduced into the series.

Stock and Watson (2002a) applies the factor analysis model using principal components to the real-time US macroeconomic forecasts from 1959 to 1998. There are eight target variables, four of which are the measures of real economic activity: total industrial production, real personal income less transfer, real manufacturing and trade sales, and number of employees on non-agricultural payrolls. The other four variables are measures of price inflation: the consumer price index; the personal consumption expenditure implicit price deflator; the consumer price index less food and energy; and the producer price index for finished goods. The full dataset used to forecast the eight target variables consists of 215 monthly time series, and diffusion index forecasts are constructed 6-, 12-, and 24-month ahead. All variables are transformed so that they are standardized to have mean zero and unit variance.

Compared with the factor analysis model introduced in Stock and Watson (2002b), small modifications are made in Stock and Watson (2002a). One is that lags are included in the forecasting equation. Different numbers of lags including lags of target variable and lags of factors, and also the number of factors are conducted in different experiments. Second, the experiments are based on h-step ahead forecasts.

Since the dataset contains missing observations, the expectation maximization (EM) algorithm can be applied to estimate the factors by solving a suitable minimization problem iteratively. In Stock and Watson (2002b) paper, three methods are used to construct factors. First is using principal components from the subset of 149 variables, that is, the balanced panel. Second, using all 215 variables to compute factors with the EM algorithm. Third, 149 predictors in the balanced panel are stacked with their first lags, and then estimates the empirical factors by the principal components of the stacked data.

Stock and Watson (2002b) compares the results of the factor analysis forecasting with those of several conventional econometric model: auto-regressive forecast, vector auto-regressive forecast, multivariate leading indicator forecast, and Philips curve forecast. By comparing the out-of-sample MSE of each model with different numbers of lags and different types of factors, the paper shows that the improvement of the factor analysis over the conventional models is quite significant. What is interesting is that only six factors capture much of the variation of the full 215 time series. In addition, the empirical results show that only a few factors are needed to forecast real macroeconomic activity.

Forni, Hallin, Lippi, and Reichlin (2002) also uses a large dataset to extract common component for forecasting inflation and real activity in the Euro area. Their goal is to test whether pooling information from large number of financial variables improves forecast of the Euro-area industrial production and consumer price indexes. The dataset in Forni *et al.* contains 447 monthly macroeconomic time series organized in six blocks. Three models are constructed and compared: univariate AR model, Stock and Watson static factor analysis model, and Forni *et* *al.* model. The difference between Stock and Watson model and Forni *et al.* model is the way in which the common components are constructed.

Forni *et al.* (2002) adopts a traditional simulated out-of-sample experiment to verify the forecasting power of the model. The paper forecasts the target variables between 1997:2 to 2001:3 using observations from 1987:2 to 1997:1 one step ahead. One, three, six, and twelve step ahead forecasts are computed. The results shows that, in general, Forni *et al.* and Stock and Watson forecast both target variables better than the univariate AR model in all horizons. When forecasting 1 and 3 steps ahead, Forni *et al.* method does better than the Stock and Watson method for forecasting inflation. Generally speaking, the Stock and Watson and Forni *et al.* methods generate similar results, that is, adding financial variables contribute to the forecast of inflation over all time horizons. This is not completely true for industrial production, which depends on the forecast horizon.

How many factors should be adopted in the model? Not too many papers have considered this problem. Increasing the number of factors might improve the fit, but efficiency is lost when more factor loadings are being estimated. The usual AIC and BIC are not suitable under the case when both N and T are large since the information criterion are functions of either N or T alone, while the penalty for over-fitting must be a function of both N and T.

Lewbel (1991) and Donald (1997) adopts a rank test for the number of factors, while Cragg and Donald (1997) uses information criterion. However, these methods assume a fixed dimension N or T of the dataset. Connor and Korajczyk (1993) developes a test for factor numbers for large dataset, unfortunately, the assumption is that N converges to infinity holding T fixed, also covariance stationary and homoscedastic are crucial assumptions for their test. Stock and Watson (1998) does show that BIC with modification could be used to choose number of factors, but it requires N to far exceed T, and it is possible for factors to be effective for the whole dataset but weak when predicting individual data series. Forni Hallin Lippi and Reichlin (2000a) raises the idea of a multivariate variant of the AIC, however, no theory nor application are available to support the idea.

Bai and Ng (2002) develops a formal statistical procedure to estimate the number of factors consistently. It allows for heteroskedasticity in the idiosyncratic component of the observed data and some weak dependence between the factors and the errors. It also allows for cross-section and serial dependence. Facing the trade off between the benefit and the penalty for over-fitting, Bai and Ng show that traditional AIC and BIC are no longer valid. Instead, they propose different penalty functions that are based on both N and T.

The new panel information criterion are examined in an empirical application of asset pricing with different N and T tested. When N and T are small, the new information criterion does not show much improvement and the penalty term NT/(N+T) provides a small correction to the asymptotic convergence rate of min{ $\sqrt{N}, \sqrt{T}$ }, also adjusts the penalty upward. When N and T are large enough, the test shows precise estimates of the number of factors. This is not surprising since the information criteria depend on large N and T. Although the new panel information criteria tend to choose too many factors, these problems are much less severe than in traditional AIC and BIC.

Different assumptions are studied in the experiments of Bai and Ng (2002), showing that under heteroskedasticity and weak dependence between components and the errors, the proposed criteria continue to select the true number of factors precisely. Under the case when only one of the serial and cross-section correlation exists in the idiosyncratic errors, the proposed criteria work well; when both types of correlation occur, the results are less precise, but still perform good, as long as N is large enough.

Another problem arises: does adding more data improve the estimate of the factors? Boivin and Ng (2005) discusses the problem how the size and the composition of the dataset affect the factor estimates. There is a chance that adding more data may lead to worse forecast if there exist cross-correlated idiosyncratic errors or if some factors have different forecasting power with different size dataset. In the factor analysis application, large dimensional datasets are typical. However, it is not necessary to include all available data series. There are some studies that advocate selectively choosing data, but they are based on strict factor models, under the assumption of mutually uncorrelated idiosyncratic errors.

Boivin and Ng (2005) states that when adding "noisy" series, the average size of the common component would decrease, and the residual cross-correlation would eventually be so large that more data are undesired. Boivin and Ng are the first to explore the finite sample properties of the principal component estimator with cross-section correlation in the idiosyncratic errors. They show that whether adding more data can produce more accurate predictors depends on the nature of the marginal series. The more heteroskedastic and mutually uncorrelated the errors are, the less precise the estimation of the factors, holding N constant. Therefore, if the marginal data bears large idiosyncratic errors or weak factor loadings, then the marginal series are undesirable.

Boivin and Ng (2005) investigates whether dropping variables might improve the forecast power using empirical experiments with 147 data series from Stock and Watson (2002b). The full dataset can be classified into 13 groups. They find that some groups have more data series than others, so the problem of oversampling might exist. In addition, many of the relative importance of the principal component for each series are small, meaning that the dispersion in the importance of common components is quite big, causing adverse effects on the forecast. Also, Boivin and Ng show that there exists large cross-correlation in the idiosyncratic errors.

Boivin and Ng weight the principal components with different information criterion, and then choose the number of factors by BIC after the factors are estimated. They conclude that differently weighted principal components generate better forecast results than Stock and Watson in all eight target variables, if not at least the same. It suggests that smaller dataset could have been adequate to construct factors and more data do not necessarily yield better result. Boivin and Ng also suggest that instead of dropping undesired data series, the effect of which on the objective function could be weighted downward so that no information would be wasted. There are many unresolved problems unsolved about what data should be used in factor analysis. However, what is clear is that the data size alone will not determine the properties of factors and that data quality should be taken into consideration.

Several papers talk about the methods to choose important variables and exclude noisy data. Vector auto-regression method (VAR) model could be conducted and therefore Granger causality test is used to select significant variables. VAR model have been discussed in many papers. For example, Friedman and Kuttner (1993) uses VAR model to provide evidence on the relation between the paper-bill spread and fluctuations in business activity. Bernanke and Blinder (1992) and Sims (1992) have employed VAR model to identify the effects of monetary policy on macroeconomic variables. A considerable literature have extended on this model.

Bernanke, Boivin, and Eliasz (2012) states that although VAR model helps in selecting important housing variables, critics of this approach should not be ignored. The model contains only a few variables thus lacks information. Standard VAR model seldom includes more than six to eight variables due to degree-of-freedom problem. Three potential problems are given in the paper. First, small dimension excludes some potential important information from the VAR model, causing the estimates to be contaminated. Second, due to restrictive numbers of variables allowed, the choice of some specific data series can not fully state a general economic concept. Third, as a result of restricted number of variables, not all impulse response can be observed, which narrows down what policy makers care about.

Bernanke *et al.* develop a method to eliminate the adverse effect of low dimension problem of VAR model by combining factor analysis and standard VAR together, which is called factor-augmented vector autoregressive models (FAVARs). Two types of estimation methods are conducted in this paper. A two-step method is that first estimating factors by principal components method of large dataset, and then running VAR model with the estimated factors included. Another one-step method estimates the factors and the dynamics simultaneously using Bayesian likelihood methods and Gibbs sampling.

Both methods are applied to empirical experiments on the effect of monetary policy innovations on key macroeconomic indicators including 120 balanced monthly macroeconomic time series from January 1959 to August 2001. Bernanke *et al.* (2012) compares the results of standard VAR model with only three variables included, that is, industrial production, CPI, and the federal funds rate, with two FAVAR models, one is the FAVAR model with only the federal funds rate assumed to be observed, another is adding an estimated factor to the three-variable VAR model. It shows that adding one factor to standard VAR model greatly changes the impulse response, and the previous FAVAR model has the same result as the latter one, which suggests that the two FAVAR models successfully extract useful information from large dataset.

Another method of selecting important housing variables as introduced in Bai and Ng (2008) is the hard thresholding. By using a statistical test to determine if each predictor is significant, fewer predictors could be even more informative in predicting target variables. However, the drawback of hard thresholding is that it does not control other predictors when testing predictors one at a time, and it is sensitive to slight change in the data. Thus soft thresholding is also introduced in the paper which allows for model selection and shrinkage. The paper shows that both hard and soft thresholding improve the prediction than no selection at all, while soft thresholding is even performing better.

Bai and Ng (2008) also uses squared principal components and squared factors to forecast. The previous method allows nonlinearity in principle components while the latter allows nonlinear relationship between factors and target variables. Ludvigson and Ng (2007) especially finds that the square of the first factor is significant in the prediction regression while the square of other factors are much weaker.

Bernanke *et al.* (2012) also raises the question of how many factors to use. Bai and Ng (2002) proposes a criterion for determining how many factors to be estimated, but it does not solve the problem how many factors should be used in the FAVAR model. In order to fix this problem, Bernanke *et al.* conduct experiments with increasing number of factors added until five. It turns out that further increasing the number of factors does not improve the results.

In conclusion, FAVAR model not only extracts important information from large dataset which are unable to be included in standard VAR model, but also presents the impulse responses of many variables to monetary policy innovations. The two-step approach produces more responses than one-step method, also eliminates the problem that take a stand on choosing appropriate variables for a generally defined economic concept.

There are some papers studying the relationship between consumer confidence and household spending. Bram and Ludvigson(1998), Carroll, Fuhrer, and Wilcox (1994), also

Mehra and Martin (2003) all believe that consumer confidence have predictive power for future consumption. Adrangi (2010), however, states that consumer sentiment only affects consumption of durable goods. Bram *et al* (1998) and Mehra *et al* (2003) also believe that consumer confidence is the catalyst for economic fluctuation and has predictive content for future change in income in a short term, that is, consumer sentiment can assess and forecast economic conditions. Santero and Westerlund (1996), however, thinks that the predictive power of consumer sentiment varies across countries and survey measures.

Haurin (2009) raises the idea that consumer sentiment predicts well in the production and sales of homes especially when housing market peaks. Weber (1996) states that consumer sentiment also predicts housing starts. Goodman (1994), however, indicates that consumer attitude data has little predictive data in forecasting housing activity.

Robsta, Deitzb, and McGoldrickc (1998) and Haurin (1990) state that income uncertainty affects the likelihood of homeownership. Robsta *et al* (1998) argues that income uncertainty has bigger impact on purchase decision than renting. Haurin (1990) argues that the likelihood of homeownership is constrained by whether the household can qualify the down-payment requirement which is quite natural since a household or a mortgage lender's expectation will affect the down-payment qualifications thus affect housing variables.

While Dion (2006) believes that consumer expectation affects buying order and, thus, lowers output production, Santero and Westerlund (1996) argue that consumer confidence has much less effect than business confidence when predicting economic output because it is less related to the output.

Strauss (2013) uses national and state-level building permits from 1980 to 2010 to predict state economic activity during recessions in the US. The model he uses is autoregressive distributed lag model. There are three findings in his paper. Firstly, building permits forecasts outperform the AR benchmark out-of-sample over past three decades in predicting employment and real income growth than traditional leading indicators. Secondly, building permits explain the variations in job and income growth across states. Lastly, he explains that housing leads the economy because permits reflect and are significantly related to consumer expectations about future economic activity.

CHAPTER 3

DATA DESCRIPTION

The whole dataset used in this dissertation contains 63 series of monthly variables in the United States. Limited by the availability of some variables, most data span 1968 to 2013 while some data like Case-Schiller price index and FNC residential price index start from 2000, three months commercial paper rate starts from 1997, real manufacturers’ unfilled orders in durable goods, manufacturers' new orders, consumer goods and materials, manufacturers’ new orders for nondefense capital goods, and ISM® new orders index start from 1992, distressed sales and purchase only house price index for the United States start from 1990. An index of variable descriptions is provided in the appendix. Different experiments are conducted with various combinations of variables and different time spans.

Following Stock and Watson (2002a), my dissertation chooses target variables as following: number of employees on non-agricultural payrolls, total industrial production, real personal income, real manufacturing and retail sales, as well as estimated monthly GDP measured by Stock and Waston. From *Figure 4,* these five target variables all start to decrease at the end of 2007, which is the recession this dissertation seeks to explore.

Non-housing variables are mainly the leading indicators from the Conference Board and Stock and Watson (1989), including average weekly hours of production workers in manufacturing, capacity utilization rate in manufacturing, interest rate on 10 year US T-bill constant maturity, 3 months T-bill rate, 3 months CP rate, 3 months T-bill rate minus 3 months

CP rate, number of people working part-time in nonagricultural industries because of slack work, average weekly initial claims for unemployment insurance, real manufacturers' unfilled orders in durable goods industries, manufacturers' new orders for consumer goods and materials, trade weighted index of nominal exchange rates between US and UK, West Germany, France, Italy and Japan, manufacturers' new orders for nondefense capital goods, ISM® new orders index, Index of supplier deliveries for vendor performance, National Association of Purchasing Managers' index of vendor performance, S&P500 common stock price, money supply M2, fed funds rate, 10 year T-bill minus fed funds rate, index of consumer expectations, monetary Services Index for M1, M2, M2 omitting the small-denomination time deposits, and all assets.

Housing variables include Case-Shiller price index, FNC residential price index, median price index from core logic, purchase only house price index for the United States, new one family homes for sale in the United States, new one family homes sold in the United States, and distressed sales in one, three, six and twelve months. The dissertation also employs housing units authorized but not yet started, housing starts, housing units under construction, and housing units completed, each with one unit, two to four units, and above five units respectively. The housing variables data source is Federal Reserve Economic Data of St. Louis.

Three types of aggregate house price measures are used in this dissertation. Case-Shiller price index (CSPI) is a repeated sales methodology which focuses on houses that have sold more than once. It assumes the quality of the houses remaining approximately unchanged over time, thus the change in housing price is contributed to the change of aggregate house price, therefore controls for heterogeneity. Residential price index (FNCP) is an hedonic methodology which controls for differences in quality. The hedonic approach does not require the assumption of constant housing quality, instead, it construct a house with hypothetical constant-quality house with each attribute holding constant over time. Housing price index (HPI) from CoreLogic and purchase only house price index for the United States (HPIPO) is also a repeated sales methodology, the reason why I include these two price indexes is because of their broader span back to 1990, while CSPI and FNCP only trace back to 2000, which makes experiments before 2000 hard to conduct.

Housing starts are defined as the number of privately owned new houses which has been started in a given period. Building permits are counted when they are authorized. A new single unit house is considered a housing completion when the house is 90 percent complete and a structure with multiple units is considered a housing completion when 90 percent of its units are ready for occupancy. These housing variables are derived from surveys of homebuilders nationwide, issued by the U.S. Census Bureau jointly with the U.S. Department of Housing and Urban Development.

Not all the houses with building permits get started, and not all started units get completed. Housing starts and completions could have happened without building permits. Construction is sometimes abandoned after permits are issued but before construction is started, which is called "authorized but not started" affecting the relationship between permits and starts, or after construction is started, which is called "under construction" affecting the relationship between starts and completions.

Housing construction has been paid close attention because it requires a lot of labor. The rise in construction could boost job creation in construction, finance and real estate and therefore economic growth. The National Association of Home Builders estimates that for every new house built, at least three full-time jobs are generated and $90,000 in new tax revenue are created. In addition, when people buy new houses, they also spend money on other consumer goods, such as furniture, lawn and garden supplies, and that generates jobs in other industries.

The number of new home sales is consistently lower than total single-family starts or completions. This does not mean a huge inventory of unsold homes is building up. Houses built for rent, houses built by a general contractor on the owner's land, and owner-built homes are not included in new home sales, only those that are built for sale are included in the New Residential Sales series.

A "distressed sale" is an urgent sale of assets because of negative situations, usually a foreclosure. Distressed sales are often at a loss because only limited time is allowed for exchange assets with funds in order to make payment for debts or other emergencies. This might be because rising mortgage rates convince homebuilders to slow down on new home starts. The owner may stop making mortgage payments when the economy is going bad.

All variables are standardized after transformation in one of the following methods: 1 denotes no transformation, 2 denotes first difference, 4 denotes logarithm, 5 denotes first difference of logarithms, and 6 denotes second difference of logarithms. The transformation form of the variables is listed in Appendix C. All the transformed housing variables are stationary as shown in *Figure 5* and *Figure 6*. The dataset includes both seasonally adjusted data and seasonally adjusted at an annual rate data. In the dataset, some variables contain missing data in which case the average value of two previous and later numbers will be calculated. Stock and Watson (2002) has introduced an method of expectation maximization algorithm to eliminate data irregularities.

CHAPTER 4

MODELS AND RESULTS

This dissertation estimates factor analysis models based on three groups of variables: non-housing, housing, and all variables. It explores the predictive ability of housing variables by comparing the out-of-sample mean squared errors (MSE) of two types of factor analysis models, one with only non-housing variables, the other with all the variables. Using the equation of (End value - Start value)/Start value, the percentage change in MSE equals the difference between the MSE when all variables are used and when only non-housing variables are used divided by the MSE with only the non-housing variables. A negative percentage change of MSE stands for a decrease in the MSE when housing variables are added to the dataset. The larger the magnitude of the change in MSE, the greater the predictive power of the housing variables.

This dissertation follows Stock and Watson (2002) to use dynamic factor model to predict the economy. Let Y_{t+h} denote the target variable to be forecast, X_t denote an N-dimensional multiple time series of predictors. Then

$$X_t = \Lambda F_t + e_t \tag{1}$$

and

$$Y_{t+h} = \beta_F' F_t + \beta_W' W_t + \varepsilon_{t+h} \tag{2}$$

Where t = 1, . . . , T, e_t is the N × 1 idiosyncratic disturbance, W_t is the lagged target variable, and ε_{t+h} is the resulting forecast error. Λ denotes factor loadings, and F_t denotes factors.

4.1. Predictive Power of Housing Variables in Forecasting Six Recessions

The dissertation combines factor analysis model and rolling forecasting method to predict target variables in six recent recessions, including 2007, 2001, 1990, 1981, 1980, and 1974 recessions. I use the information criteria to select the number of factors for each target variables in different recessions. I also use three different methods, Granger causality and VAR model, FAVAR model, and hard thresholding, to test the importance of each housing variable.

This dissertation uses Akaike information criterion (AIC) and Bayesian information criterion (BIC) to select the number of factors. In principle, we can compute the sum of squared residuals for equation (2), and then select the optimal number of factors using the information criteria of the form

$$N^* = \min_{n} [T \ln(SSR) + n^* C_T], \qquad (3)$$

where n is the number of factors, C_T =2 for AIC and $C_T = \ln(T)$ for BIC.

4.1.1 Recession From December 2007 to June 2009

First I am going to forecast the target variables from January to December in 2008. Since some housing variables only start from 2000, including Case-Shiller price index and FNC residential price index, I use 2000 to 2007 dataset to forecast the economic downturn. Table 1[4] shows the AIC and BIC values for all five target variables when all housing variables are used, and Table 2 shows the information criteria values when only non-housing leading indicators are used. By selecting the smallest values, we can conclude that when all variables are used, nine factors are optimal for employment in non-farm industry (EMP) and manufacturing and trade

[4] All tables are placed in Appendix B.

sales (MATS), three for industrial production (IP) and GDP, seven for real personal income less transfer (RPI) under AIC. When using BIC, however, eight factors are optimal for EMP, one for IP and RPI, three for GDP, and nine for MATS. Similarly, when only non-housing variables are used, six factors are selected for EMP, seven for IP, GDP and MATS, and five for RPI under AIC, while BIC selects the same number of factors for all target variables except IP with two factors. From the different number of factors selected by AIC and BIC, we can observe that BIC tends to select fewer factors than AIC.

Let us use EMP as target variable for example. When rolling forecasting method is adopted to predict one period ahead, I use observations on the non-housing variables from 2000-01 to 2007-12 to forecast employment in 2008-01 with six factors selected by both AIC and BIC, then add one more observation, that is, use data from 2000-01 to 2008-01 to forecast employment in 2008-02. Repeating the steps above twelve times until we obtain the forecasts of the entire 2008 year. The results are reported in *Table 3*. Comparing the predicted employment to the actual value, the out-of-sample MSE is computed as 0.0795517. Similarly, implementing the same experiments with housing variables added, the out-of-sample MSE drops dramatically to 0.0233995 using nine factors under AIC and 0.0244095 using eight factors under BOC. Thus adding housing variables in the factor analysis model lowers the MSE by 70.59 and 69.32 percent respectively for AIC and BIC, which is a quite significant improvement in predicting employment.

While housing variables greatly improve the prediction of the employment, they also improve the forecasts of RPI and MATS by cutting the MSE by 32.46 and 41.17 percent under AIC. The MSE for IP drops by 15.66 percent under AIC and 78.16 under BIC. The MSE for GDP, however, only decreases by .66 percent. Generally speaking, adding housing variables greatly improves the predictive power to forecast the 2008 target variables under both AIC and BIC. I have also tried to add the square of the first factor adopted by Ludvigson and Ng (2007), but it does not improve the prediction.

Similar experiments are conducted forecasting 3, 6, and 12 periods ahead with and without the housing variables under AIC. As shown in *Table 4*, as the forecasting lag becomes larger, the decreases in the MSE with housing variables added become smaller, but are still significant except for GDP. For EMP, the decrease in MSE drops by 28 percent from 1 period ahead to 12 periods ahead. For IP and MATS, the decrease in MSE is robust for all four periods ahead. The change in MSE for RPI varies slightly for 1, 3, and 6 periods ahead, but it drops by 10 percent when forecasting the target variable 12 periods ahead. The decreases in MSE for GDP are near zero for 1 and 12 periods ahead, and increase to 3 and 4 percent for 3 and 6 periods ahead, respectively. Generally speaking, the out-of-sample MSE decreases by large percentage for EMP, IP, RPI, and MATS when housing variables are added for 1, 3, 6, and 12-period ahead. From this experiment, we could conclude that housing variables do have great predictive power for the downturn of the economy, especially for the employment.

From the experiments above, adding housing variables greatly improves the predictions of the target variables in 2008 when the target variables decreased. However, it is not the same for the forecasts of other time periods. I use 2000 to 2005 data to forecast the target variables in 2006, and 2000 to 2006 data to forecast 2007 in order to examine the predictive power of housing variables when the economy is steady. I also use 2000 to 2008 data to test the predictive ability of housing variables by forecasting the economic recovery in 2009.

Table 5 shows the percentage changes of MSE forecasting the target variables in 2006, 2007 and 2009 under AIC. When predicting 2006, adding housing variables to the leading indicators causes the MSE drop by only 4.71 and 6.71 percent for GDP and MATS, and even increase the MSE by 24.78, 8.85, and 7.84 percent for EMP, IP, and RPI, respectively. When predicting 2007, adding housing variables decreases the MSE for GDP and MATS by 8 and 9 percent respectively, causes only 0.13 percent decrease for EMP, but it increases the MSE slightly for IP and RPI by 1.7 percent. The results show that adding housing variables cause an inaccuracy in the prediction. This weak performance of the housing variables in forecasting the economic fluctuation in 2006 and 2007 is because that the housing variables decrease about two years before the economy downturn, while in 2006 and 2007 the economy stayed steady or continued rising until late 2007. When predicting 2009, all the target variables have higher MSE with housing variables added to the leading indicators. Therefore, we can conclude that housing variables have significant role in improving the predictive power of forecasting the economic downturn in 2008, but have negative effect in predicting the steady economy or economic recovery

The discussion above states the predictive power of housing variables forecasting the 2008 recession. The following experiments will investigate further whether this predictive power only works for the 2008 recession or for other recessions. I conduct experiments to test the role of housing variables in predicting each of the following recessions in reverse chronological order, including the recession from March 2001 to November 2001, from July 1990 to March 1991, from July 1981 to November 1982, from January 1980 to June 1980, and from November 1973 to March 1975. Due to the lack of availability of the observations, no earlier recession will be included.

4.1.2 Recession From March 2001 to November 2001

When predicting 2001 recession, since manufacturers' unfilled orders in durable goods and new orders in nondefense capital goods are available from 1992 January, also due to the transformation of the observations, the experiments are traced back to 1992 March. Since CSPI and FNCP are not available before 2000, this dissertation uses house price index from CoreLogic for the United States and purchase only house price index from the Fred instead as the substitutes. Since 3 months Commercial Paper rate is not available before 1997, it will not be used in the experiment.

Table 6 and *7* presents the information criteria for 2001 recession with all variables and with only non-housing variables under AIC and BIC. When all variables are included, EMP and RPI require three factors as the optimal number for both AIC and BIC, GDP and MATS require seven as the optimal factor number for both criteria, IP uses four as the best number of factors for AIC and three for BIC. When only non-housing leading indicators are involved, EMP, GDP, and MATS use four as the optimal factor number for both AIC and BIC, IP selects three for AIC and two for BIC, RPI requires five for AIC and three for BIC. We can see that BIC tends to select smaller number of factors under this case.

Using the different number of factors selected by the information criteria according to *Table 6* and *7*, the change in MSE for all five target variables are shown in *Table 8*. When the number of factors is selected by AIC, only IP has a lower MSE with housing variables added by 29.33 precent, all other target variables have higher MSEs with housing variables added. When the number of factors is selected by BIC, the MSE for all five target variables increase when adding housing variables. In this case, housing variables are either not sensitive to or even against the prediction of the economic downturn. This is, however, consistent with the 2001 recession which is driven by a collapse in business investment in equipment and software. According to Leamer (2007), the contribution to weakness in GDP the year before the recession from residential investment is only 12 percent compared to 23 percent from durable goods, 19 percent from equipment and software, and 20 percent from exports, as shown in *Table 9*.[5]

4.1.3 Recession From July 1990 to March 1991

For 1990 recession, I use January 1983 to June 1990 data. Due to the availability limitation, the following variables are excluded from the dataset: Median price index with single-house combined from CoreLogic, Distressed sales in last 1, 3, 6 and 12 months, Real manufacturers' unfilled orders in durable goods industries, Manufacturers' new orders, consumer goods and materials, Manufacturers' new orders, nondefense capital goods, ISM® new orders index.

Table 10 and *11* show the information criteria values for all variables case and only non-housing leading indicator case. When all variables are used, four is the optimal number of factors for EMP, one is the optimal number of GDP, and six for RPI, under both AIC and BIC. Five is the best number of factors for IP under AIC, and four under BIC. Six is the optimal number for RPI under AIC, and one under BIC. When non-housing variables are included, EMP and RPI share the same number of factors six for both AIC and BIC, IP and GDP use two as the optimal number for both criteria. MATS uses six factors under AIC, and two under BIC.

Table 12 shows the change of MSE in predicting target variables for 1990 recession. When using AIC to select the number of factors, the MSE drop by 12.81, 16.45, and 30.18 percent for EMP, GDP, and MATS, respectively, but change slightly for IP and RPI. When BIC is used to select the factor number, the decrease in MSE for MATS is not as significant as the result under AIC. For the other four target variables, the change in MSE under BIC are the same as the change under AIC. We can conclude that adding housing variables improves the prediction of

[5] This table is quoted from Leamer (2007) with the title "Contribution to Weakness in GDP, The Year Before the Recession", page 15.

1990 recession moderately for EMP, GDP, and MATS, but has weak effect for IP and RPI.

4.1.4 Recession From July 1981 to November 1982

Table 13 and *14* shows the optimal number of factors selected by AIC and BIC for 1981 recession. Since there is another recession one year before this recession, observations included are of short time span. The dataset from July 1980 to June 1981 is used to forecast 1981 recession. When all variables are included, EMP needs three factors as the optimal number, IP and MATS require two, and IP needs one factor, under both AIC and BIC. RPI has seven as the optimal factor number under AIC, and one under BIC. When only non-housing leading indicators are included, EMP use four factors, GDP and RPI use one factor, under both criteria. IP needs five factors under AIC and one under BIC. MATS uses four as the optimal number under AIC and one under BIC.

Table 15 shows the change of MSE when housing variables are added for 1981 recession. Adding housing variables lower the MSE for EMP by 24.54 percent under both AIC and BIC. Therefore, the housing variables have great predictive power in forecasting EMP in this case. It also lowers the MSE slightly for GDP by 7.92 percent. For IP, RPI, and MATS, however, the percentage change in MSE depends on whether the number of factors is determined by AIC or BIC. The MSE for RPI drops significantly by 39.77 percent under AIC, but only 5.86 percent under BIC. The MSE for IP decreases by 19.9 percent under BIC, but it increases by 18.39 percent under AIC. MATS has a lower MSE by 22.96 percent when housing variables are added under BIC, but a higher MSE by 2.57 percent under AIC. This contradictory results is due to the lack of information because of the short time span.

4.1.5 Recession From January 1980 to June 1980

Table 16 and *17* shows the number of factors selected by AIC and BIC for 1980 recession. The dataset from January 1978 to December 1979 is used to forecast the recession. When all variables are used, the number of factors selected are the same under both AIC and BIC. EMP uses three factors, GDP requires two factors, IP, RPI, and MATS need seven factors. When only non-housing leading indicators are used, EMP selects two as the optimal number of factors, IP, RPI, and MATS use three factors, under both criteria. Four factors are optimal for GDP under AIC, and two under BIC.

Table 18 conveys the change of MSE when housing variables are added in forecasting 1980 recession. We can see that adding housing variables lowers the MSE for EMP and IP moderately by 8.51 and 22.56 respectively. The table shows that housing variables play an important role in predicting RPI and MATS, decreasing the MSE by 58.56 and 45.12 respectively. The MSE for GDP decreases slightly by 5.5 percent when housing variables added with the number of factors selected by BIC, but increases by 16.43 percent under AIC. Therefore, housing variables have great predictive power for most target variables, but an adverse effect on the prediction of GDP under AIC case.

4.1.6 Recession From November 1973 to March 1975

Here comes to the last recession this dissertation investigates. *Table 19* and *20* displays the information criteria values for the prediction of 1974 recession. The forecast of 1974 recession uses observations from January 1970 to October 1973. When all housing variables are used, two factors are optimal for IP and GDP, three for RPI, under both criteria. Both EMP and MATS use five factors under AIC, and two factors under BIC. When non-housing variables are used, three factors are used for EMP, four factors for IP, GDP, and MATS, under both AIC and BIC. RPI uses four factors under AIC and three factors under BIC.

Table 21 presents the change in MSE when housing variables are added in forecasting 1974 recession. Adding housing variables lowers the MSE greatly for EMP and RPI by 35.91 and 19.7 percent under AIC, 26.02 and 24.34 percent under BIC. It decreases the MSE for GDP slightly by 3.24 percent. However, adding housing variables lowers the MSE for MATS by 8.45 percent under AIC, but raise the MSE by 7.55 percent under BIC. It also has negative effect in predicting IP by increasing the MSE by 11.64 percent under both criteria.

The experiments above for six recessions imply that the housing variables have significant predictive power in forecasting the economic downturn except for the 2001 recession in which the housing market was not the cause of the recession, but not necessarily in predicting the steady economy and the economic recovery.

4.2 The Importance of Each Housing Variable in Improving The Prediction

This dissertation also evaluates individual housing variables separately since adding more data does not necessarily produce better results referring to Bai and Ng (2002). Adding variables that are correlated with other variables might reduce the efficiency of the factors. Since some variables are not available before 2000, this dissertation only use 2007 recession dataset to select important housing variables. Three methods are used to verify important housing variables.

4.2.1 Granger Causality Test and VAR Model

Following Friedman and Kuttner (1993), the first method I use is the Granger causality test and VAR model. The first step tests one variable at a time without controlling for the other leading indicators using the Granger causality test. Take industrial production for example, the right-hand side of the equation includes twelve lags of the industrial production, twelve lags of the producer price index, a time trend, and six lags of the tested variable.

$$IP_t = c + \sum_{i=1}^{12} A_i \times IP_{t-i} + \sum_{j=1}^{12} B_j \times PPI_{t-j} + \sum_{h=1}^{6} C_h \times Z_{t-h} + t + e_t \qquad (4)$$

where Z denotes the variable to be tested, and t represents the time trend.

Using the sample period from 2000-2008, four non-housing variables, CAP, TBILL3MO, WICI, and ISMNOI, are significant according to the F-test. Sixteen housing variables are significant, which are DSAL1MO, DSAL3MO, DSAL6MO, DSAL12MO, AUTHNOT24U, COMPU1USA, COMPUTSA, HOUST1F, HOUST2F, HOUST, PERMIT1, PERMIT24, PERMIT5, PERMIT, UNDCON1USA, and UNDCONTSA. The change in the MSE are shown in *Table 22*. Compared to *Table 1* where all housing variables are used, *Table 22* shows that the selection of housing variables improves the prediction for EMP under AIC by 3 percent, and slightly lower the MSE for GDP by 1 percent. However, the MSE for RPI and MATS increase by 8 percent under AIC.

The second step is using the VAR model to estimate each housing variable controlling non-housing variables. Since the number of variables is limited in VAR model, not all non-housing variables can be included. This model only use the four significant non-housing variables chosen by the Granger causality tests displayed above. A total of eight variables are included, which are industrial production, producer price index, the one housing variable to be tested, and four non-housing variables. Six lags of each variable are also included following Friedman and Kuttner (1993).

For the VAR model, three criteria can be used to choose important variables, Akaike information criterion (AIC), Bayesian information criterion (BIC), and Hannan-Quinn information criterion (HQIC). A small value indicates better model. The best sixteen housing variables are consistent for all three criterion, which are DSAL12MO, UNDCONTSA, DSAL6MO, UNDCON5MUSA, UNDCON1USA, HNFS, DSAL3MO, PERMIT1, UNDCON24USA, AUTHNOT5MU, AUTHNOTT, PERMIT, HSN1F, COMPU1USA, DSAL1MO, and COMPUTSA. *Table 23* shows that the change in MSE with housing variables selected by VAR model are similar to, if not worse than, the ones selected by Granger-Causality test.

One shortcoming of standard VAR model is that only a few variables can be included, thus many information are missing causing estimation not accurate. In addition, the housing variables selected by the two tests contradict each other somewhat. Comparing the two sets of housing variables selected by Granger causality tests and VAR model, some variables that are selected by one test did not appear under the other test. For example, HNFS and HSN1F selected by VAR model are tested insignificant under Granger causality tests, while housing starts which is significant has high value of the information criterion.

4.2.2 FAVAR Model

In order to eliminate information loss and include desired information, a second method, factor-augmented vector autoregressive (FAVAR) model, is introduced in Bernanke, Boivin and Elisaz (2005). FAVAR model combines the factor analysis and VAR models. In order to capture more information on the nonhousing variables, four factors are estimated using all the nonhousing variables. Take industrial production for example, the FAVAR model includes seven variables: industrial production, producer price index, the four factors, one housing variable to be tested, and also six lags of all these variables.

Table 24 displays the housing variables selected by information criterion within FAVAR model. In addition to the AIC, BIC, and HQIC methods, the Wald tests are used to pick important housing variables in the FAVAR model. The housing variables selected by wald test are displayed in *Table 25.* The model shows that although information criteria imply some housing variables to be good predictors, the Wald tests suggest that it is not necessarily the case. Some variables that have good values for the information criteria are not statistically significant by the Wald test. Examples include FNCP for industrial production, COMPUTSA and HSN1F for real personal income, HPI for employment, HSN1F and HNFS for industrial production. Interestingly, the wald tests indicate that one housing price index is important for employment and industrial production, while none of the housing price index are selected by information criteria.

Thus, the information criteria conflict with the Wald test for some housing variables. Using the two groups of selected housing variables by the information criteria and the Wald tests, the factor analysis models for predicting the target variables in 2008 yield the results in *Table 26 and*

Table 27 respectively. The housing variables selected by information criteria present a much smaller decrease by 50 percent in the MSE for EMP with the number of factors selected by AIC or BIC. It also leads to smaller decrease in the MSE for IP and RPI under AIC by 21 and 12 percent, respectively. The housing variables selected by the wald tests produce slightly better results in predicting GDP by 1 percent, but it lowers the decrease in MSE for IP and RPI by 16 and 10 percent respectively under AIC.

4.2.3 Hard Thresholding

The third method of choosing important housing variables is hard thresholding provided in Bai and Ng (2007). By regressing target variables on four lags of the target variables and one housing variable, housing variables could be selected if the t statistics are significant. *Table 28* displays the housing variables chosen for each target variable. One thing needs to be paid attention is that FNCP is statistically significant for industrial production and GDP, and HPIPO is significant for real personal income.

Table 29 represents the change in the MSE when the selected housing variables are added to the non-housing leading indicators. For EMP and MATS, the change in the MSE decreases slightly by 2 to 3 percent. However, adding the selected housing variables generates a 14 percent more decrease in the MSE for GDP under AIC, and slightly decreases the MSE more by 4 and 6 percent for IP and RPI respectively under AIC.

4.2.4 The Relevance of Housing Unit of Structure

Since the hard thresholding method generates a better result compared to FAVAR model with information criteria and Wald tests, this dissertation compares the two groups of housing variables selected by each method. For GDP, FAVAR method selects Housing Units Authorized, But Not Yet Started with 5 units or more while hard thresholding selects with 1 unit. FAVAR method picks Housing Units Completed, 2-4 Unit. FAVAR chooses Housing Units Under Construction with 2 or more units, while hard thresholding chooses with 1 unit and the total. Hard thresholding method also picks FNC price index.

For real personal income, the difference between the housing variables are listed as follows. FAVAR method selects DSAL6MO, DSAL12MO, AUTHNOT24U, HOUST2F, UNDCON5MUSA, HNFS, while hard thresholding method selects HPIPO, AUTHNOT1U, AUTHNOTT, COMPU1USA, PERMIT1, UNDCON1USA, UNDCONTSA.

For real manufacturing and trade sales, FAVAR method selects extra housing variables that the hard thresholding method does not select: AUTHNOT1U, AUTHNOT24U, AUTHNOTT, COMPU24USA, HOUST2F, UNDCONTSA, HNFS.

Therefore, concluded from the observations above, the number of units for housing volumes is also crucial in the housing variable selection to improve the prediction. *Table 30* displays the correlation between the housing variables with different units and target variables two years ahead, between 2000 and 2008. Interestingly, the correlations are the highest with housing volume measures with one unit and with all units. When there are two or more units, the correlations tend to be smaller and decrease significantly as the unit increases. That explains the reason why the hard thresholding selects the housing volumes mostly with one unit and generates better results than FAVAR method.

However I notice that without the housing price index, the change in the MSE barely varies. In addition, in most of the housing variable selection methods, housing price index is found not important or significant.

In order to verify my conjecture, the dissertation runs regressions with only housing volumes added to the leading indicator, and with only housing price index added. *Table 31* shows that when only housing volumes are included, the drop in the MSE are slight smaller but still very similar to the ones when all housing variables are added as shown in *Table 1*. *Table 32* shows that when only housing price indexed are added to the leading indicators, the out-of-sample MSE even increase for EMP, IP, and GDP by 8.77, 13.33, and 24.09 percent, respectively, under AIC. For RPI and MATS, the decrease in MSE is much smaller than when all housing variables are added by 27 and 36 percent under AIC. The results are even worse when the number of factors is selected by BIC. Therefore, the selected important housing variables, though vary with time periods and target variables, are all relevant with the volume measures of housing. This is consistent with Leamer (2007) theory that “homes have a volume cycle, not a price cycle”, it is the volume that matters, not price.

In conclusion, from the experiments discussed in the dissertation, we can conclude that adding housing variables could improve the prediction of the economic downturn remarkably especially for employment, but is not sensitive in predicting steady or booming economy. In addition, the housing price index does not play an important role in prediction, but instead, housing volume measures exert great influence on forecasting economic fluctuations. Moreover, the housing volume measures with one unit of constructure exhibit greater importance than with two or more units.

CHAPTER 5

HOUSING VARIABLES AND THEORETICAL EXPLANATION

As a crucial component of the U.S. economy, housing market has a mutual effect on many related industries, such as banking, the mortgage sector, raw materials, employment, construction, manufacturing and real estate. When the economy is in a boom, people are more likely to purchase new homes; conversely, when the economy is in a recession, people are less likely to buy new homes.

One possible explanation of housing variables' great predictive power is that since housing is a household's most expensive durable purchase, the decisions of purchasing houses are based on carefully formed expectations of future employment and income. If households plan to buy small appliance, it is not necessary to consider about future income. Thus the purchase of ordinary goods can not successfully reflects consumers' expectation of future employment and wealth.

But things are different when it comes to housing purchases. It is so large an expense and investment that rational consumers will gather whatever information they have to analyze if the current and expected economical conditions are suitable for buying a house, for example, future job condition, future income variability, future economy trends, and any plan that will affect real income implicitly. For example, if the households plan to have a baby, they will take into account of the expense on children, which will affect real income implicitly. If the households have a insecure job that they expect themselves to be fired any time, then they are less likely to purchase house. This is what we call income uncertainty.

Income uncertainty will affect the decision of purchasing a house. John Robst (1999) provides evidence that income uncertainty reduces the likelihood of individuals owning homes at a point in time. People are rational, making decisions based on the analysis of all available information they collect. If they expect their future employment to be insecure, their future income to be lowered, or the future housing market and the economic condition to be risky, they will wait until the conditions become better before they decide to buy houses. Their expectation of future real personal income and employment are largely reflected in their decision of housing purchase, that is, housing variables match consumers' expectations of future income and employment. This is the reason why the housing market reflects the expectation of consumers and consequently nicely predicts the future employment and future economical condition when things are not going smoothly. As indicated in Leamer (2007), it is a consumer cycle, not a business cycle.

The correlation experiments are conducted in order to study the relationship between housing volume variables and target variables as shown in *Table 30*. When housing starts are two years before target variables, say, corr(EMPt, HSt-24)= 0.8393, corr(IPt, HSt-24)= 0.8690, corr(GDPt, HSt-24)= 0.7884, corr(RPIt, HSt-24)= 0.8125, corr(MATSt, HSt-24)= 0.8625 between 2000 and 2008, which show close correlation and consequently the strong predictive power of housing starts to forecast macroeconomic variables.

In order to verify the theoretical explanation that housing variables predict macroeconomic variables by reflecting the consumer confidence index, the experiments will be conducted in two steps: correlation between housing starts and consumer expectations as shown in *Table 33*, and correlation between consumer expectations and target variables shown in *Table 34*. Since consumer confidence index (CCI) is designed to predict target variables six months ahead[6], while housing starts predict target variables two years ahead, then the experiments will test correlations between target variables and CCI six months ahead of target variables, and between CCI and housing starts eighteen months ahead of it. Between 2006 and 2008, corr(EMP_t,CCI_{t-6})=0.8737, corr(IP_t,CCI_{t-6})=0.9109, corr(GDP_t,CCI_{t-6})=0.4797, corr(RPIt,CCI_{t-6})=0.6907, corr($MATS_t$,CCI_{t-6})=0.9199. The results show that CCI is highly correlated with future employment and industrial production. Meanwhile, the correlations between main housing volume variables and CCI are corr($DSAL1MO_{t-18}$,CCI_t) = -0.6145, corr($AUTHNOT1U_{t-18}$,CCI_t) = 0.8538, corr($COMPU1USA_{t-18}$,CCI_t) = 0.6887, corr($HOUST1F_{t-18}$,CCI_t) = 0.9360, corr($UNDCON1USA_{t-18}$,CCI_t) = 0.8624, corr($PERMIT1_{t-18}$,CCI_t) = 0.9490. This high correlation shows that housing starts reflects CCI in the future, and therefore predicts the future employment and future economics condition, as stated above.

Interestingly, however, even when promotions, rising incomes and economic recovery are expected, actions by households are often postponed until the expectations are realized due to risk aversion[7]. This would explain why the housing market leads recessions but not recoveries.

---

[6] "The Conference Board Consumer Confidence Index (CCI) ... is based on ... their expectations for six months hence regarding business conditions, employment, and income." from the Consumer Confidence Survey technical note in February 2011.

[7] Charles A. Holt and Susan K. Laury's paper "Risk Aversion and Incentive Effects" uses a lottery choice experiment to illustrate that most experiment subjects exhibit risk aversion at low pay-off level, and the number increase even more at high pay-off level.

REFERENCES

Adrangi, B., and Macri J. (2011), "Consumer Confidence and Aggregate Consumption Expenditures in the United States," Review of Economics & Finance.

Bai, J., and Ng, S. (2001), "Determining the Number of Factors in Approximate Factor Models," Journal of Econometrics,146, 304–317.

Bai, J., and Ng, S. (2008), "Forecasting Economic Time Series Using Targeted Predictors," Econometrica, 70, 191-221.

Bair, E., Hastie, T., Paul, D., Tibshirani, R., 2006. Prediction By Supervised Principal Components. Journal of the American Statistical Association 101 (473), 119-137.

Bernanke, B., Boivin, J., and Elisaz, P. (2005), "Measuring The Effects Of Monetary Policy: a Factor-Augmented Vector Autoregressive (FAVAR) Approach," The Quarterly Journal Of Economics, Vol. 120, No.1.

Boivin, J., and Ng, S. (2003), "Are More Data Always Better For Factor Analysis?" NBER Working paper No. 9829.

Bram, J., and Ludvigson, S. (1998), "Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race," FRBNY economic Policy Review.

Burns, A., and Mitchell W. (1946), "Measuring Business Cycles." New York: NBER.

Carroll, C., Fuhrer, J., Wilcox, D. (1994), "Does Consumer Sentiment Forecast Household Spending? If So, Why?" The American Economic Review, Vol. 84, No. 5 pp. 1397-1408.

Chamberlain,G., and Rothschild, M. (1983), "Arbitrage Factor Structure, and Mean-Variance Analysis of Large Asset Markets," Econometrica, 51, 1281-1304.

Connor, G., and Korajczyk, R. A. (1993), "A Test for the Number of Factors in an Approximate Factor Model," Journal of Finance, XLVIII, 1263-1291.

Croce, R., and Haurin, D. (2009), "Predicting turning points in the housing market," Journal of

Housing Economics 18, 281–293.

Diebold, F., Mariano, R. (2002), "Comparing Predictive Accuracy," Journal of Business & Economic Statistics, American Statistical Association.

Dion, D. P. (2006), "Does Consumer Confidence Forecast Household Spending?" MPRA paper No. 902.

Feige, E., and Pearce, D. (1976), "Economically Rational Expectations: Are Innovations in the Rate of Inflation Independent of Innovations in Measures of Monetary and Fiscal Policy?" Journal of Political Economy, Vol. 84, No. 3, pp. 499-522.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000), "The Generalized Dynamic Factor Model: Identification and Estimation," The Review of Economics and Statistics, 82, 540-552.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2002), "Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area?" Journal of Monetary Economics, Elsevier, vol. 50(6), pages 1243-1255.

Forni, M., and Reichlin, L. (1996), "Dynamic Common Factors in Large Cross-Sections," Empirical Economics, 21, 27-42.

Friedman, B., Kuttner, K. (1993), "Why Does the paper-Bill Spread Predict Real Economic Activity," Business Cycles, Indicators and Forecasting.

Garner, A. (1991), "Forecasting Consumer Spending: Should Economists Pay Attention to Consumer Confidence Surveys?" Economic Review.

Geweke, J. (1977), "The Dynamic Factor Analysis of Economic Time Series," in Latent Variables in Socio-Economic Models, eds. D. J. Aigner and A. S. Goldberger, Amsterdam: North-Holland, Ch. 19.

Goodman J., (1994), "Using Attitude Data to Forecast Housing Activity," The Journal of Real Estate Research.

Holt, C. A., and Susan K. L. (2002), "Risk Aversion and Incentive Effects," American Economic Review, 92(5): 1644-1655.

Hardle, W., and Simar, L. (2003), "Applied Multivariate Statistical Analysis," Business Cycle Indicators Handbook, The Conference Board.

Harvey, D., Leybourne, S., Newbold, P. (1997), "Testing the equality of prediction mean squared errors," International Journal of Forecasting, 13 281-291281-291.

Haurin, D. (1991), "Income Variability, Homeownership, and Housing Demand," Journal of Housing Economics1, 60-74.

Heij, C., Dijk, D., and Groenen, P. (2009), "Leading Indicators: The Principal Covariate Index," 1-st MFC, Rome.

Hendry, D., and Hubrich, K. (2006), "Forecasting Economic Aggregates by Disaggregates," Working paper Series No. 589.

Leamer, E. (2007), "Housing is the Business Cycle," NBER Working paper No. 13428.

Marcellino, M. (2005), "Leading Indicators," CEPR Discussion papers 4977, C.E.P.R. Discussion papers.

Mehra, Y., and Martin, E. (2003), "Why Does Consumer Sentiment Predict Household Spending?" Federal Reserve Bank of Richmond Economic Quarterly Volume 89/4.

Mian, A., and Sufi, A. (2010), "Household Leverage and the Recession of 2007 to 2009," NBER Working paper No. 15896.

Muellbauer, J., and Murph, A. (1997), "Booms and Busts in the UK Housing Market," The Economic Journal, 107, 1701-1727.

The Conference Board and U.S. Business Cycle Indicators, (2011), "The Conference Board Leading Economic Index," www.conference-board.org.

Thoma, M. (1994), "Subsample Instability and Asymmetries in Money-Income Causality," Journal of Econometrics, 64, 279-306.

Tsai, H., and Tsay, R. (2010), "Constrained Factor Models," Journal of the American Statistical Association, Vol. 105, No. 492

Rappaport, J. (2007), "A Guide to Aggregate House Price Measures," www.KansasCityFed.org

Reis, R. (2006), "Inattentive Producers," Review of Economic Studies 73, 793–821.

Robsta, J., Deitzb, D., and McGoldrick, K. (1999), "Income variability, uncertainty and housing tenure choice," Regional Science and Urban Economics 29, 219–229.

Santero, T., and Westerlund, N. (1996), "Confidence Indicators and Their Relationship to Changes in Economic Activity," OECD Economics Department Working papers No. 170.

Sargent, T. J,. and Sims, C. A. (1977), "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory," in New Methods in Business Cycle Research, eds. C. Sims et al., Minneapolis: Federal Reserve Bank of Minneapolis.

Schwab, R. (1982), "Inflation Expectations and the Demand for Housing," The American Economic Review, Vol. 72, No. 1 , pp. 143-153.

Sinai, T. (2012), "House Price Moments in Boom-Bust Cycles," The Wharton School, University of Pennsylvania and NBER.

Stock, J. H., and Watson, M. W. (1989), "New Indexes of Coincident and Leading Economic Indicators,"N BER Macroeconomics Annual, 351-393.

Stock, J. H., and Watson, M. W. (1998a), "Diffusion Indexes," Working paper 6702, NBER.

Stock, J. H., and Watson, M. W. (2002), "Macroeconomic Forecasting Using Diffusion Indexes," Journal of Business and Economic Statistics, 20, 147-162.

Stock, J. H., and Watson, M. W. (2002), "Forecasting Using Principal Components From a Large Number of Predictors," Journal Of The American Statistical Association, Vol. 97, No. 460, Pp. 1167-1179.

Stock, J. H., and Watson, M. W. (2006), "Forecasting and Now-Casting with Disparate Predictors: Dynamic Factor Models and Beyond," FEMES 2006 Meetings Beijing.

Stock, J. H., and Watson, M. W. (2010), "Dynamic Factor Models," Chapter 2 in M. Clements and D. Hendry (eds), Oxford Handbook of Economic Forecasting.

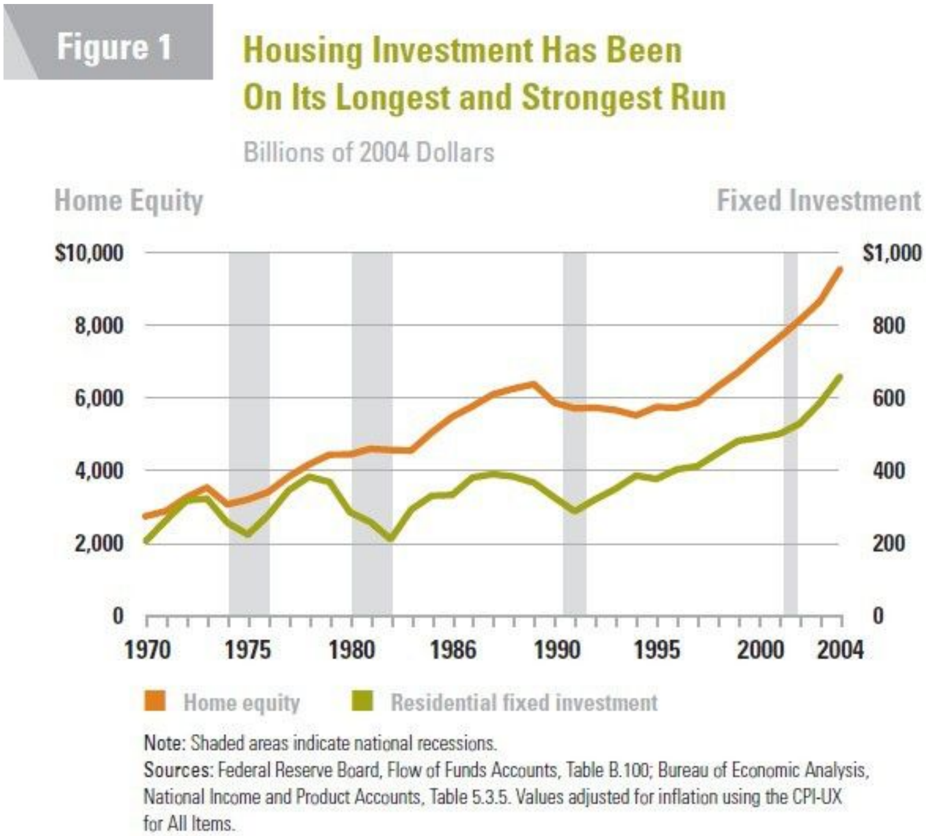
Strauss, J. (2013), "Does housing drive state-level job growth? Building permits and consumer expectations forecast a state's economic activity", Journal of Urban Economics 73 (2013) 77–93.

Weber, W., and Devaney, M. (1996), "Can consumer sentiment surveys forecast housing starts?" Appraisal Journal Publisher, Source Volume: v64 Source Issue: n4.

LIST OF APPENDICES

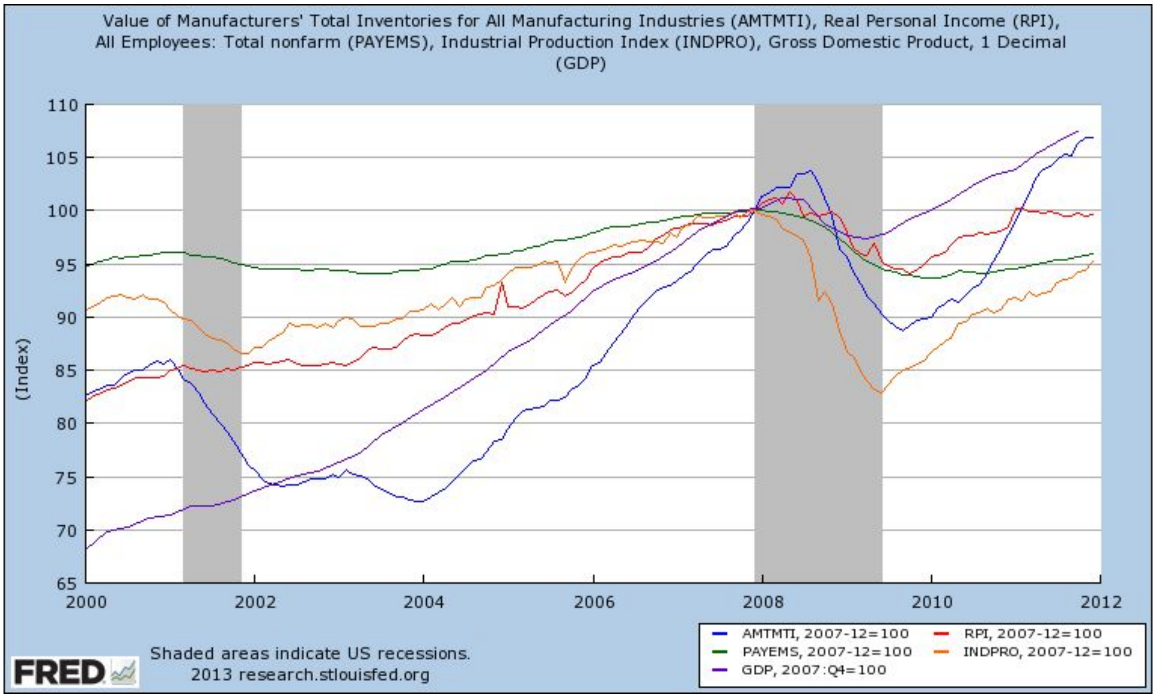
APPENDIX A: FIGURES

Figure 1: Housing investment has been on its longest and strongest run



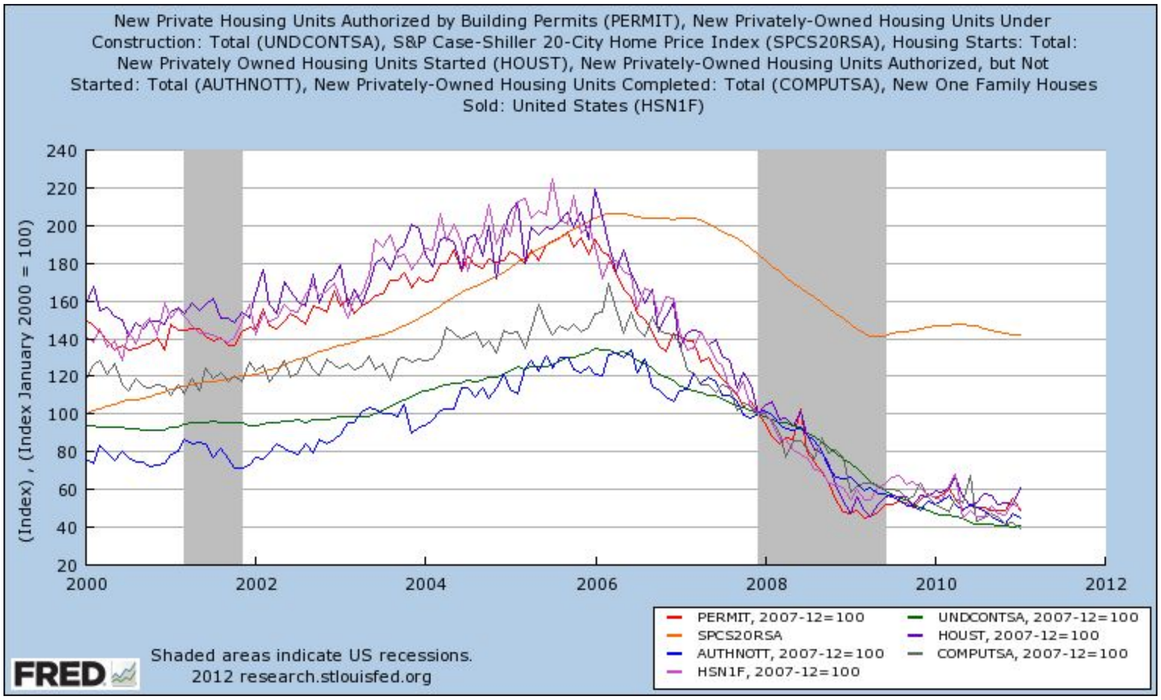
Data source: Joint Center for Housing Studies of Harvard University

Figure 2: Index of target variables from 2000 to 2011



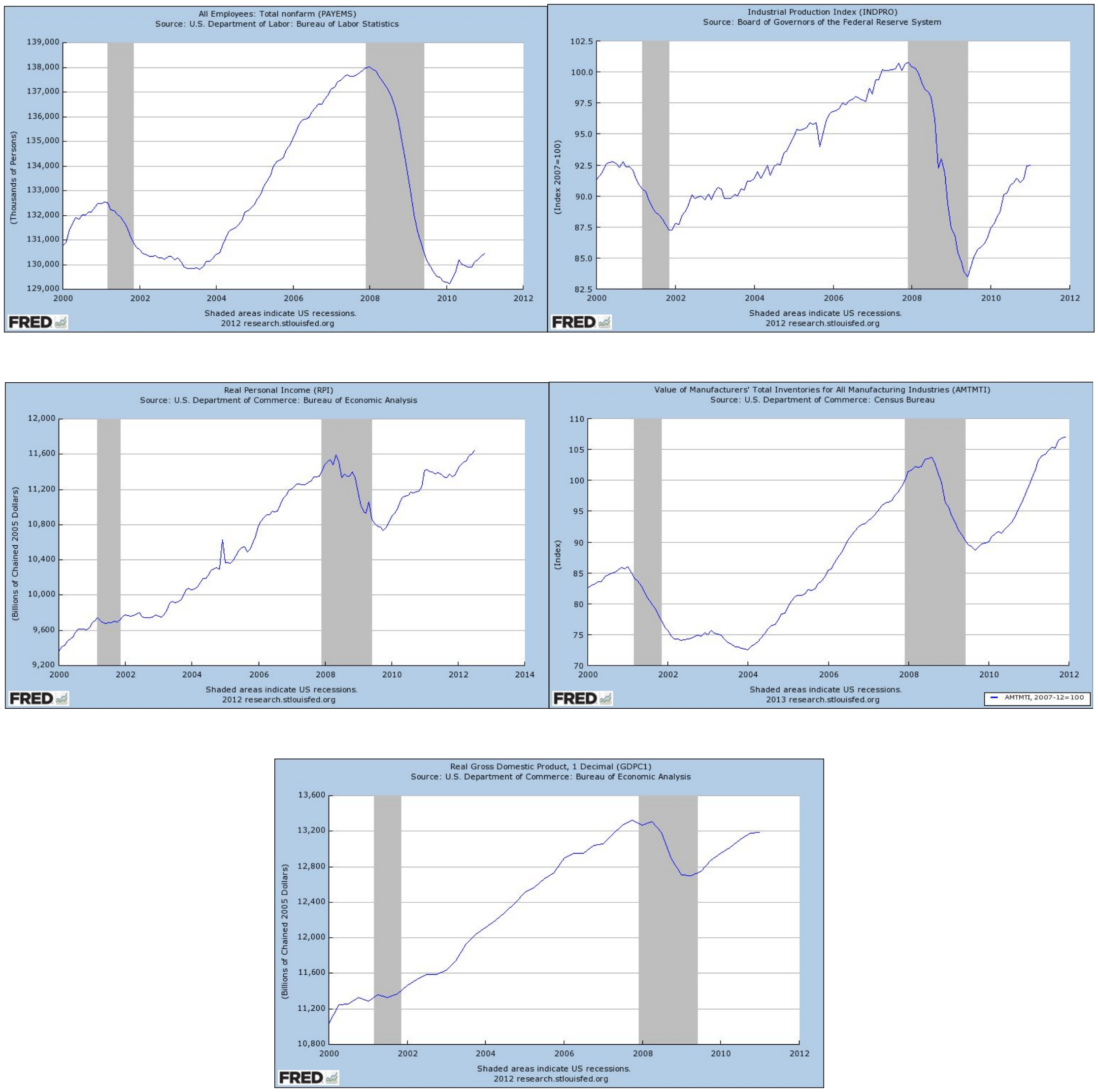
Sources: ST. LOUIS FED Economic Data

Figure 3: Index of housing variables from 2000 to 2011



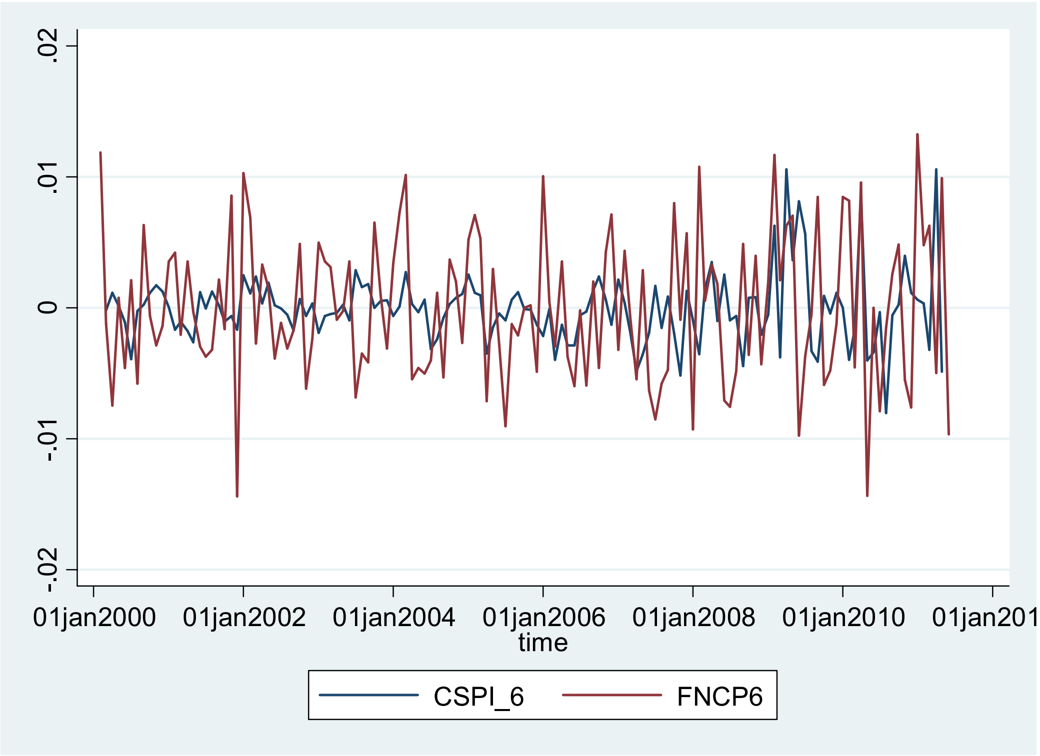
Sources: ST. LOUIS FED Economic Data

Figure 4: Trends of five target variables between 2000:1 and 2011:1



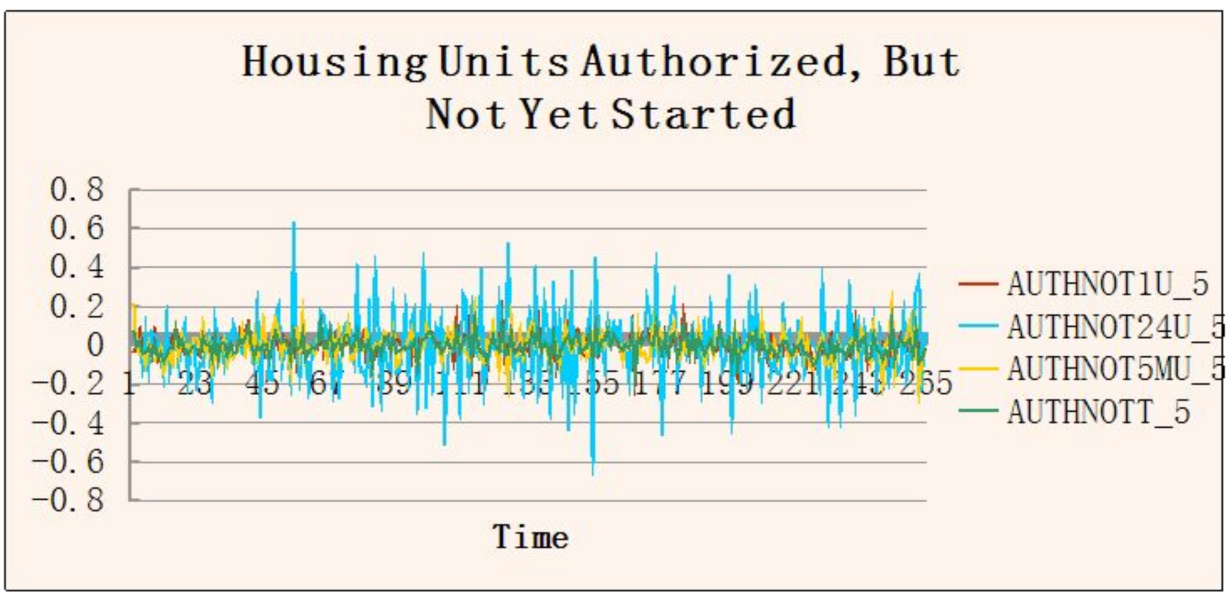
Data source: Federal Reserve Economic Data St. Louis

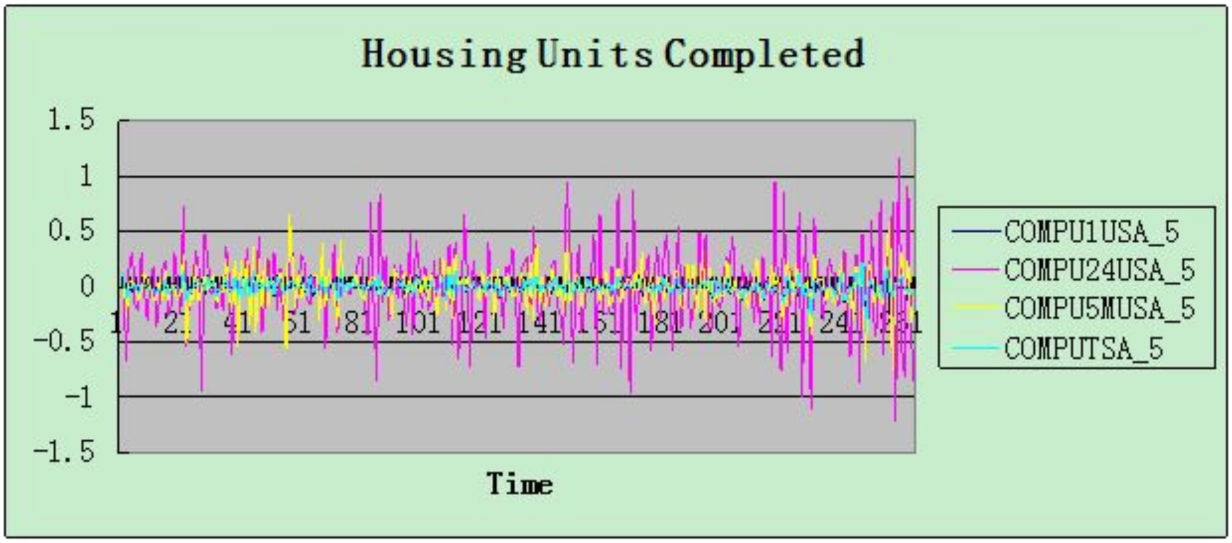
Figure 5: Transformed CSPI and FNCP between 2000:1 and 2011:12



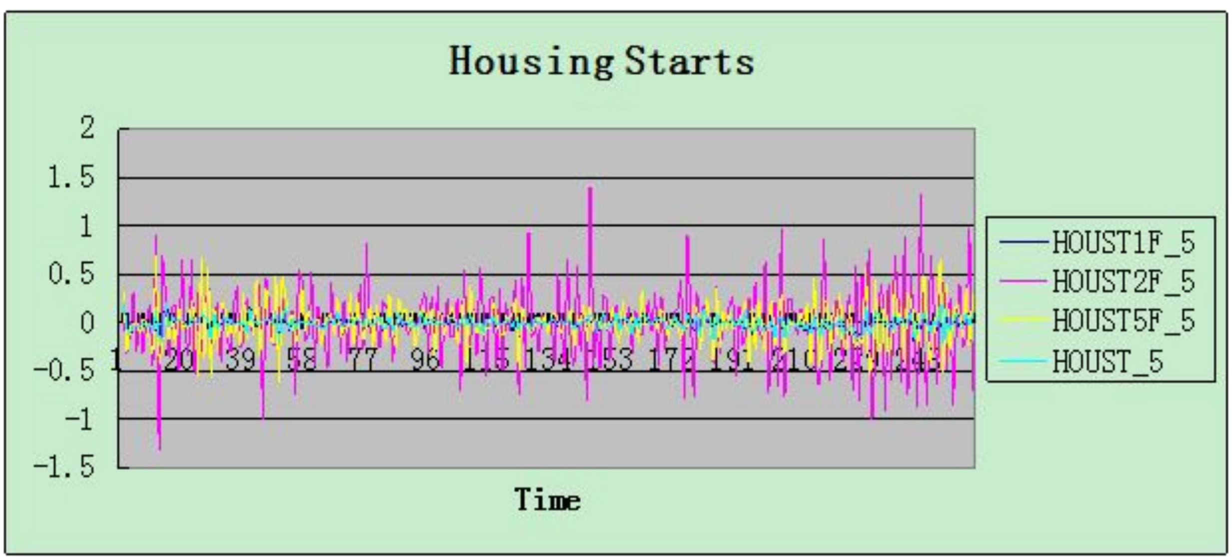
Data source: Federal Reserve Economic Data St. Louis

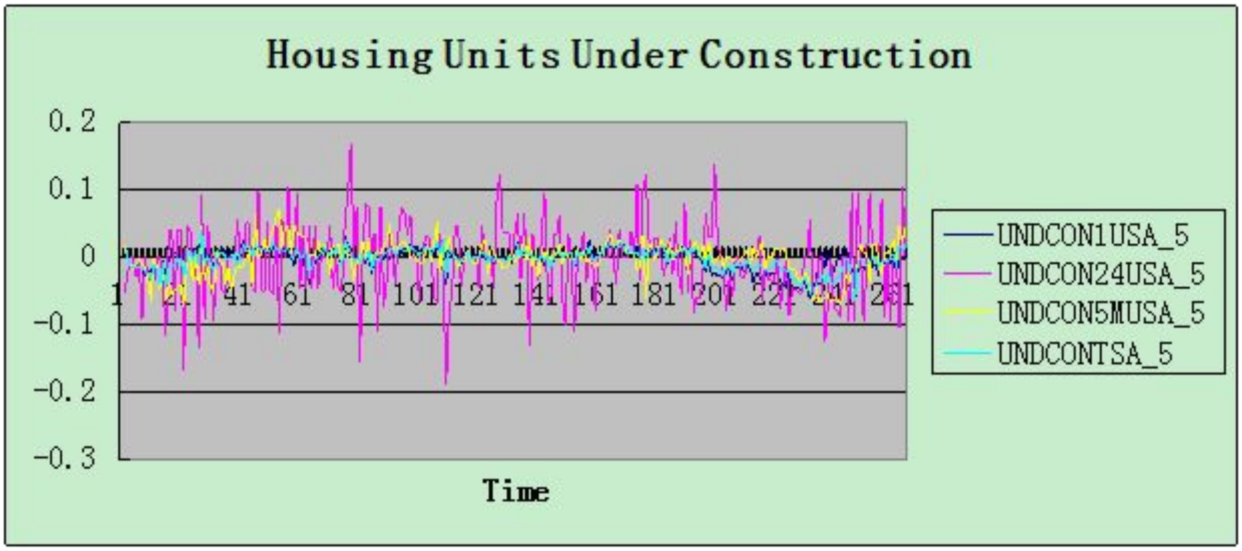
Figure 6: Transformed housing variables between 1990:1 and 2011:12

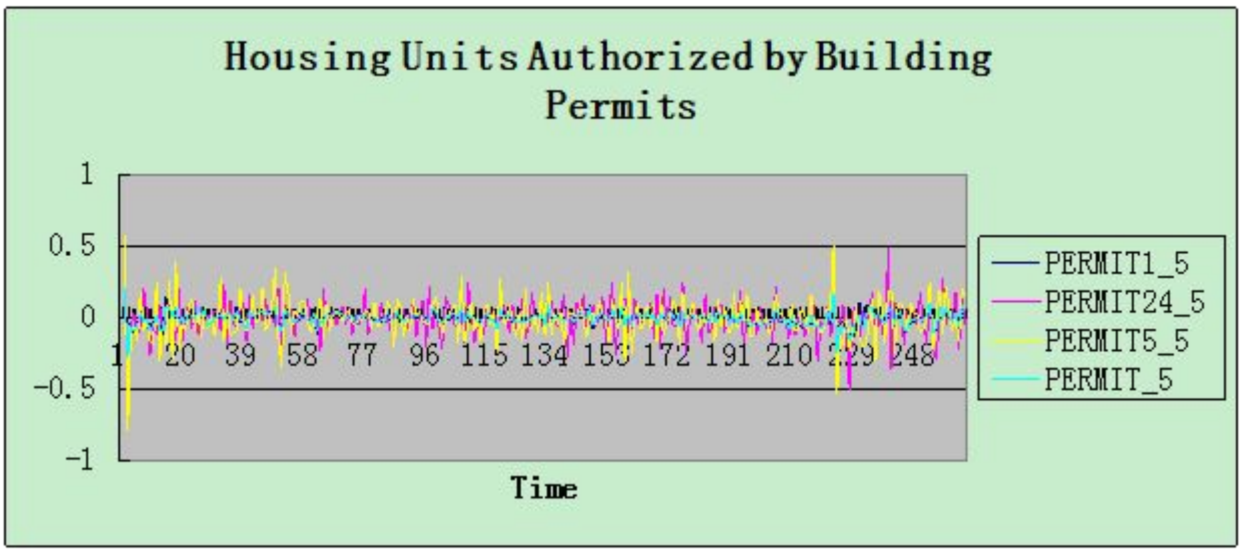



Housing Units Completed
1.5
1
0.5
0
-0.5
-1
-1.5
Time
COMPU1USA_5
COMPU24USA_5
COMPU5MUSA_5
COMPUTSA_5

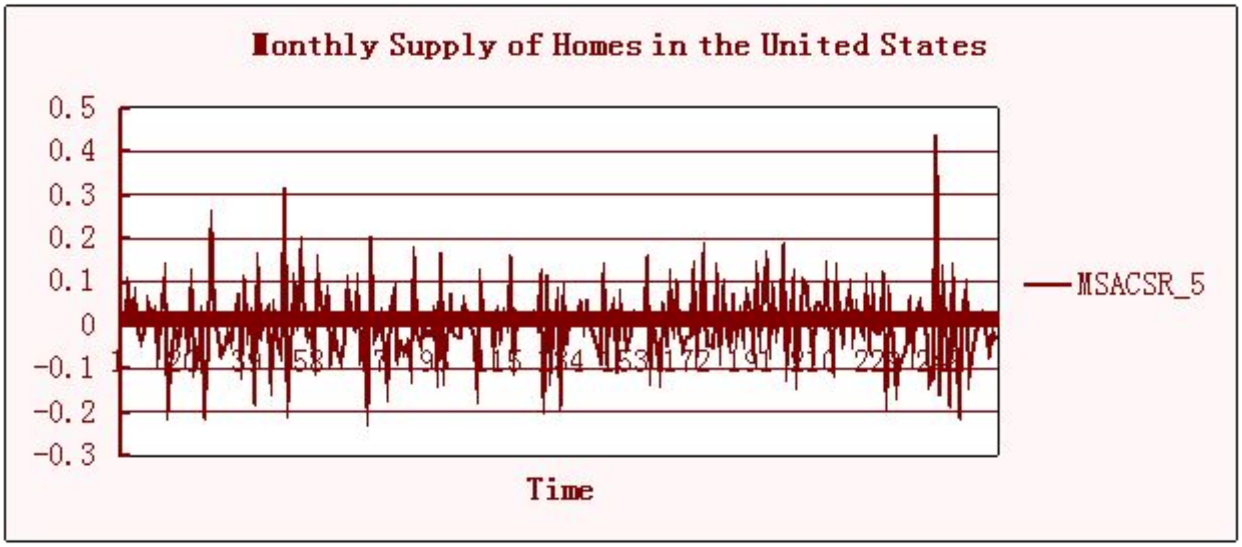

Housing affordability index
0.15
0.1
0.05
0
-0.05
-0.1
Time
COMPHAI_5


Housing Starts
2
1.5
1
0.5
0
-0.5
-1
-1.5
Time
HOUST1F_5
HOUST2F_5
HOUST5F_5
HOUST_5


Housing Units Under Construction
0.2
0.1
0
-0.1
-0.2
-0.3
Time
UNDCON1USA_5
UNDCON24USA_5
UNDCON5MUSA_5
UNDCONTSA_5


Housing Units Authorized by Building Permits
1
0.5
0
-0.5
-1
1 20 39 58 77 96 115 134 153 172 191 210 229 248
Time
PERMIT1_5
PERMIT24_5
PERMIT5_5
PERMIT_5


New One Family Houses Sold
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
Time
HSN1F_5



Data source: Federal Reserve Economic Data St. Louis

APPENDIX B: TABLE AND RESULTS

Table 1: Information criteria values for target variables for 2008 recession with all variables added

Information criteria with all variables for 2008 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	-68.32148	160.955	207.842	191.0667	263.0739
BIC	-65.65804	163.6184	210.5055	193.7302	265.7374
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	-32.33849	164.6577	207.0703	200.099	260.4689
BIC	-27.01161	169.9846	212.3971	205.4259	265.7958
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	-32.46108	157.4724	180.0857	208.6679	259.3866
BIC	-24.47076	165.4627	188.0761	216.6582	267.3769
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	-27.80842	164.6102	183.1365	210.1739	265.3223
BIC	-17.15466	175.264	193.7902	220.8276	275.976
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	-82.87841	166.9265	186.3723	210.2753	258.3798
BIC	-69.56121	180.2437	199.6895	223.5925	271.697
Six factors					
	EMP	IP	GDP	RPI	MATS
AIC	-74.30796	168.5789	183.4207	204.4817	260.2713
BIC	-58.32732	184.5596	199.4013	220.4623	276.2519
Seven factors					
	EMP	IP	GDP	RPI	MATS
AIC	-86.82766	168.2247	184.6236	185.2327	243.7961
BIC	-68.18359	186.8688	203.2677	203.8768	262.4402
Eight factors					
	EMP	IP	GDP	RPI	MATS
AIC	-114.155	166.0464	180.797	192.6925	224.7805
BIC	-92.84753	187.3539	202.1045	214.0001	246.0881
Nine factors					
	EMP	IP	GDP	RPI	MATS
AIC	-116.6342	165.613	182.6687	198.6132	202.2929
BIC	-92.66323	189.584	206.6396	222.5842	226.2639

Table 2: Information criteria values for target variables for 2008 recession with only non-housing variables added

Information criteria with non-housing variables for 2008 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	189.2967	212.4126	224.7062	243.0927	328.5544
BIC	191.9602	215.0761	227.3697	245.7561	331.2178
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	68.87136	189.703	203.6463	237.8434	293.3212
BIC	74.19824	195.0299	208.9731	243.1702	298.648
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	38.71488	188.4536	204.3285	239.4408	290.2153
BIC	46.70519	196.4439	212.3188	247.4311	298.2056
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	37.50091	189.7707	205.2408	240.2494	290.3413
BIC	48.15467	200.4245	215.8946	250.9032	300.9951
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	32.16177	187.5806	208.8957	222.8297	283.6
BIC	45.47897	200.8977	222.2129	236.1468	296.9172
Six factors					
	EMP	IP	GDP	RPI	MATS
AIC	7.077242	184.2419	204.4548	226.5183	288.0644
BIC	23.05788	200.2226	220.4355	242.4989	304.045
Seven factors					
	EMP	IP	GDP	RPI	MATS
AIC	11.70845	183.5286	188.7923	228.4345	254.5341
BIC	30.35252	202.1727	207.4364	247.0786	273.1782

Table 3: Predicting 2008 economy using 2000-2007 data with the number of factors selected by AIC and BIC

2000-2007 predicts 2008					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.0233995	0.3478717	0.4305946	0.4191594	0.4741204
MSE with non-housing	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	-70.59%	-15.66%	-0.66%	-32.46%	-41.17%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.0244095	0.3733154	0.4305946	0.495961	0.4741204
MSE with non-housing	0.0795517	1.709636	0.4334744	0.620593	0.8059636
% Change in MSE	-69.32%	-78.16%	-0.66%	-20.08%	-41.17%

Table 4: Predicting 2008 economy with rolling method 1, 3, 6, and 12 steps ahead with the number of factors selected by AIC using 2000-2007 data

	MSE with all variables	MSE with non housing	% change in MSE
EMP			
1-period ahead	0.0233995	0.0795517	-70.59%
3-period ahead	0.032542	0.0928341	-64.95%
6-period ahead	0.0447127	0.1061201	-57.87%
12-period ahead	0.0612901	0.1070917	-42.77%
IP			
1-period ahead	0.3478717	0.4124747	-15.66%
3-period ahead	0.3477385	0.4197985	-17.17%
6-period ahead	0.3504351	0.4161363	-15.79%
12-period ahead	0.3571176	0.4088167	-12.65%
GDP			
1-period ahead	0.4305946	0.4334744	-0.66%
3-period ahead	0.4281428	0.440741	-2.86%
6-period ahead	0.4353496	0.4557044	-4.47%
12-period ahead	0.4429996	0.4437531	-0.17%
RPI			
1-period ahead	0.4191594	0.620593	-32.46%
3-period ahead	0.4421785	0.6032906	-26.71%
6-period ahead	0.4277864	0.6136622	-30.29%
12-period ahead	0.5000405	0.6315201	-20.82%
MATS			
1-period ahead	0.4741204	0.8059636	-41.17%
3-period ahead	0.555729	0.8462582	-34.33%
6-period ahead	0.506974	0.8295784	-38.89%
12-period ahead	0.5226171	0.8368311	-37.55%

Table 5: Predicting the economic fluctuations in different time periods with data starting from 2000 with the number of factors selected by AIC

2000-2005 predicts 2006					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.0235429	0.0282737	0.2159762	0.4915297	0.2709421
MSE with non-housing	0.0188682	0.0259749	0.2266486	0.4557865	0.290433
% Change in MSE	24.78%	8.85%	-4.71%	7.84%	-6.71%

2000-2006 predicts 2007					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.0178809	0.0545063	0.2610686	0.1279851	0.2385899
MSE with non-housing	0.0179045	0.0536104	0.285515	0.1258124	0.2636634
% Change in MSE	-0.13%	1.67%	-8.56%	1.73%	-9.51%

2000-2008 predicts 2009					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.1614667	0.2229588	0.3450062	0.8876319	0.8432273
MSE with non-housing	0.060612	0.1225387	0.2943266	0.7458414	0.7507339
% Change in MSE	166.39%	81.95%	17.22%	19.01%	12.32%

Table 6: Information criteria values for target variables for 2001 recession with all variables added

Information criteria with all variables for 2001 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	199.7609	39.33164	220.4345	237.9567	347.6752
BIC	202.5231	42.09381	223.1967	240.7189	350.4374
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	201.2187	54.06268	176.6236	250.3788	269.72
BIC	206.7431	59.58702	182.148	255.9032	275.2444
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	1.791632	-206.127	173.997	67.58059	243.4815
BIC	10.07815	-197.8405	182.2835	75.86711	251.768
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	21.5635	-207.5898	179.4678	68.61668	246.7924
BIC	32.6122	-196.5411	190.5165	79.66538	257.8411
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	18.90535	-204.808	177.2516	70.09094	238.2758
BIC	32.71622	-190.9971	191.0625	83.90181	252.0867
Six factors					
	EMP	IP	GDP	RPI	MATS
AIC	18.30291	-149.3362	151.4622	84.1702	199.4102
BIC	34.87596	-132.7632	168.0352	100.7432	215.9833
Seven factors					
	EMP	IP	GDP	RPI	MATS
AIC	19.27627	-123.3035	145.1524	97.03824	178.8604
BIC	38.61149	-103.9683	164.4876	116.3735	198.1956

Table 7: Information criteria values for target variables for 2001 recession with only non-housing variables added

Information criteria with non-housing variables for 2001 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	202.5761	57.39469	180.5895	224.5002	253.6535
BIC	205.3383	60.15687	183.3516	227.2624	256.4157
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	22.2713	-170.9818	176.6837	48.7601	236.4929
BIC	27.79564	-165.4575	182.208	54.28445	242.0172
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	13.5074	-173.5894	174.5649	32.78621	239.8266
BIC	21.79392	-165.3029	182.8514	41.07273	248.1131
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	9.962321	-138.0708	135.8126	31.21206	149.7793
BIC	21.01102	-127.0221	146.8613	42.26076	160.828
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	12.68907	-123.1051	143.0653	30.37394	157.6407
BIC	26.49994	-109.2942	156.8762	44.18481	171.4516
Six factors					
	EMP	IP	GDP	RPI	MATS
AIC	15.4159	-123.5591	146.0974	115.5734	172.8069
BIC	29.22677	-109.7483	159.9083	129.3842	186.6177
Seven factors					
	EMP	IP	GDP	RPI	MATS
AIC	21.9865	-97.14335	153.636	178.5942	156.7451
BIC	41.32172	-77.80814	172.9712	197.9294	176.0803

Table 8: Predicting the target variables in 2001 recession with the number of factors selected by AIC and BIC

1992/3-2001/2 predicts 2001/3-2001/11					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.599539	0.0175997	0.3408666	0.9126319	0.45468
MSE with non-housing	0.1129904	0.0249023	0.331274	0.45468	0.3732767
% Change in MSE	430.61%	-29.33%	2.90%	100.72%	21.81%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.599539	0.1704457	0.3408666	0.9126319	0.45468
MSE with non-housing	0.1129904	0.0249023	0.331274	0.1628933	0.3732767
% Change in MSE	430.61%	584.46%	2.90%	460.26%	21.81%

Table 9: Contribution to Weakness in GDP, The Year Before the Recession. Leamer(2007)

Contribution to Weakness in GDP, The Year Before the Recession, Largest in Bold and Boxed

	1949	1953	1957	1960	1970	1974	1980	1981	1990	2001	Avg	Avg-7
Residential Investment	**30%**	6%	**22%**	**30%**	**20%**	**29%**	**32%**	22%	21%	12%	22%	25%
Durables	19%	18%	20%	12%	20%	24%	26%	10%	**26%**	**23%**	20%	20%
Services	3%	0%	16%	2%	2%	9%	17%	**28%**	2%	8%	9%	11%
Nondurables	7%	7%	0%	8%	11%	21%	10%	8%	8%	7%	10%	9%
Exports	27%	31%	17%	0%	5%	0%	0%	14%	6%	20%	17%	6%
Equipment and Software	15%	4%	0%	0%	7%	0%	15%	0%	25%	19%	14%	7%
Fed Defense	0%	0%	0%	22%	16%	12%	0%	0%	8%	2%	10%	8%
Fed Nondefense	0%	0%	16%	20%	6%	4%	0%	3%	0%	3%	7%	7%
State and Local	0%	**34%**	1%	7%	10%	0%	0%	16%	3%	2%	8%	5%
Structures	0%	0%	8%	0%	2%	2%	0%	0%	1%	3%	3%	2%
TOTAL	-2.8	-0.9	-2.4	-2.3	-2.5	-3.1	-2.7	-2.7	-2.7	-2.4	-2.5	-2.64

Table 10: Information criteria values for target variables for 1990 recession with all variables added

Information criteria with all variables for 1990 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	42.94939	-41.66158	153.7407	195.0168	198.0244
BIC	45.54451	-39.06646	156.3359	197.612	200.6195
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	52.66543	-8.066357	158.6848	194.665	211.4827
BIC	57.85567	-2.876117	163.8751	199.8553	216.673
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	53.30282	-5.61783	158.7664	195.7599	213.7359
BIC	61.08818	2.167529	166.5517	203.5453	221.5213
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	-0.4677838	-54.14722	158.1271	184.6635	202.6884
BIC	9.912696	-43.76674	168.5076	195.044	213.0689
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	4.097888	-56.19191	160.6179	183.5379	193.3678
BIC	17.07349	-43.2163	173.5935	196.5135	206.3434
Six factors					
	EMP	IP	GDP	RPI	MATS
AIC	24.2158	-46.71922	163.0005	175.2595	186.834
BIC	39.78652	-31.1485	178.5712	190.8302	202.4047
Seven factors					
	EMP	IP	GDP	RPI	MATS
AIC	-19.79371	-50.26518	157.9926	171.5898	184.5442
BIC	-1.627871	-32.09934	176.1584	189.7557	202.7101

Table 11: Information criteria values for target variables for 1990 recession with only non-housing variables added

Information criteria with non-housing variables for 1990 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	189.1608	53.3608	189.8832	265.5366	255.1973
BIC	191.756	55.95592	192.4783	268.1317	257.7925
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	104.2091	-60.23956	173.531	216.8642	215.7496
BIC	109.3994	-55.04932	178.7212	222.0544	220.9398
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	109.4985	-30.60696	184.4517	210.3237	228.7436
BIC	117.2839	-22.8216	192.237	218.109	236.529
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	120.8841	-29.13422	185.7798	205.9749	227.1269
BIC	131.2645	-18.75374	196.1603	216.3554	237.5074
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	92.4073	-46.5455	178.7168	196.8758	224.2787
BIC	105.3829	-33.5699	191.6924	209.8514	237.2543
Six factors					
	EMP	IP	GDP	RPI	MATS
AIC	17.10333	-48.72247	176.7745	172.2291	222.406
BIC	32.67405	-33.15176	192.3452	187.7998	237.9767

Table 12: Predicting the target variables in 1990 recession with the number of factors selected by AIC and BIC

1983/1-1990/6 predicts 1990/7-1991/3					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.1020025	0.0569359	0.5145359	0.5780248	0.6497129
MSE with non-housing	0.116989	0.0580698	0.6158264	0.5605992	0.930613
% Change in MSE	-12.81%	-1.95%	-16.45%	3.11%	-30.18%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.1020025	0.0593102	0.5145359	0.5780248	0.804785
MSE with non-housing	0.116989	0.0580698	0.6158264	0.5605992	0.9433293
% Change in MSE	-12.81%	2.14%	-16.45%	3.11%	-14.69%

Table 13: Information criteria values for target variables for 1981 recession with all variables added

Information criteria with all variables for 1981 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	7.086674	24.90613	119.578	48.09471	68.50208
BIC	8.45397	26.27342	120.9453	49.46201	69.86938
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	7.851671	11.11854	120.2723	51.2203	64.81686
BIC	10.58626	13.85313	123.0069	53.95489	67.55145
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	-0.3368902	13.27188	120.8925	50.30098	67.20192
BIC	3.764997	17.37377	124.9943	54.40287	71.3038
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	1.35581	14.1467	121.9788	53.50407	65.96871
BIC	6.824994	19.61588	127.4479	58.97326	71.4379
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	3.01123	14.5718	123.822	52.34686	69.44747
BIC	9.847709	21.40828	130.6585	59.18334	76.28395
Six factors					
	EMP	IP	GDP	RPI	MATS
AIC	6.129698	13.40742	122.78	52.29864	71.98112
BIC	14.33347	21.6112	130.9838	60.50242	80.18489
Seven factors					
	EMP	IP	GDP	RPI	MATS
AIC	9.286067	15.06316	126.5264	47.14479	71.15015
BIC	18.85714	24.63423	136.0974	56.71586	80.72121

Table 14: Information criteria values for target variables for 1981 recession with only non-housing variables added

Information criteria with non-housing variables for 1981 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	34.34773	15.55526	121.9724	49.84723	70.3821
BIC	35.71503	16.92256	123.3397	51.21453	71.74939
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	20.76062	16.27729	122.395	51.41035	72.91125
BIC	23.49521	19.01188	125.1295	54.14495	75.64585
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	16.97745	16.43797	124.369	54.27063	72.52124
BIC	21.07934	20.53985	128.4709	58.37251	76.62313
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	9.827092	15.25002	124.1851	53.41087	68.08036
BIC	15.29628	20.71921	129.6543	58.88005	73.54955
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	11.13943	12.22281	123.1231	53.54008	69.48402
BIC	17.97591	19.05929	129.9595	60.37656	76.3205

Table 15: Predicting the target variables in 1981 recession with the number of factors selected by AIC and BIC

1980/7-1981/6 predicts 1981/7-1982/11					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.0472772	0.0751893	3.39126	0.1844661	0.4789878
MSE with non-housing	0.0626488	0.0635096	3.683143	0.3062637	0.4669763
% Change in MSE	-24.54%	18.39%	-7.92%	-39.77%	2.57%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.0472772	0.0751893	3.39126	0.2883038	0.4789878
MSE with non-housing	0.0626488	0.0938751	3.683143	0.3062637	0.6217544
% Change in MSE	-24.54%	-19.90%	-7.92%	-5.86%	-22.96%

Table 16: Information criteria values for target variables for 1980 recession with all variables added

Information criteria with all variables for 1980 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	-6.590161	-7.593271	103.1065	6.992227	48.69929
BIC	-5.188963	-6.192074	104.5077	8.393424	50.10048
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	-23.51986	-18.91459	97.52424	2.34772	50.08071
BIC	-20.71747	-16.1122	100.3266	5.150115	52.88311
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	-32.69347	-17.76674	99.23582	-0.4629946	49.10054
BIC	-28.48988	-13.56314	103.4394	3.740597	53.30413
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	-28.81565	-39.46336	99.46076	10.05288	55.02921
BIC	-23.21087	-33.85857	105.0656	15.65767	60.634
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	-29.91832	-42.05416	101.3834	18.4195	55.81699
BIC	-22.91233	-35.04818	108.3894	25.42549	62.82298
Six factors					
	EMP	IP	GDP	RPI	MATS
AIC	-29.16689	-38.83826	105.2042	21.663	53.20788
BIC	-20.7597	-30.43107	113.6114	30.07019	61.61506
Seven factors					
	EMP	IP	GDP	RPI	MATS
AIC	-8.755861	-48.63182	101.2626	-11.02332	35.61319
BIC	1.05252	-38.82343	111.071	-1.214942	45.42158

Table 17: Information criteria values for target variables for 1980 recession with only non-housing variables added

Information criteria with non-housing variables for 1980 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	51.2661	37.20947	113.0721	52.19518	81.26576
BIC	52.6673	38.61066	114.4733	53.59638	82.66696
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	-32.02472	-9.358842	99.21976	12.31073	52.21357
BIC	-29.22233	-6.556448	102.0222	15.11312	55.01596
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	74.15199	-48.9632	98.82738	7.402808	45.6161
BIC	78.35558	-44.75961	103.031	11.6064	49.81969
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	76.60751	-41.18771	96.96063	12.05669	58.67537
BIC	82.2123	-35.58293	102.5654	17.66148	64.28017

Table 18: Predicting the target variables in 1980 recession with the number of factors selected by AIC and BIC

1978/1-1979/12 predicts 1980/1-1980/6					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.0458885	0.0206614	3.764877	0.0723767	0.3425562
MSE with non-housing	0.0501577	0.0266793	3.233596	0.1746451	0.6242385
% Change in MSE	-8.51%	-22.56%	16.43%	-58.56%	-45.12%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.0458885	0.0206614	3.764877	0.0723767	0.3425562
MSE with non-housing	0.0501577	0.0266793	3.983787	0.1746451	0.6242385
% Change in MSE	-8.51%	-22.56%	-5.50%	-58.56%	-45.12%

Table 19: Information criteria values for target variables for 1974 recession with all variables added

Information criteria with all variables for 1974 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	90.60944	107.2834	169.6943	185.0393	197.2807
BIC	92.75258	109.4265	171.8374	187.1824	199.4239
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	79.20106	88.35701	159.0764	155.4006	173.5204
BIC	83.48733	92.64328	163.3626	159.6869	177.8066
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	81.24217	89.0258	166.1387	152.9397	172.1168
BIC	87.67158	95.45521	172.5681	159.3691	178.5462
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	91.23339	94.34901	168.593	155.3719	178.3497
BIC	99.80593	102.9216	177.1655	163.9444	186.9222
Five factors					
	EMP	IP	GDP	RPI	MATS
AIC	76.16469	93.34887	171.0757	160.5713	169.3708
BIC	86.88036	104.0645	181.7914	171.287	180.0865

Table 20: Information criteria values for target variables for 1974 recession with only non-housing variables added

Information criteria with non-housing variables for 1974 recession					
One factor					
	EMP	IP	GDP	RPI	MATS
AIC	167.1408	150.2894	182.5479	194.4306	230.985
BIC	169.2839	152.4326	184.691	196.5737	233.1281
Two factors					
	EMP	IP	GDP	RPI	MATS
AIC	137.7927	137.9155	181.3438	190.8099	225.4832
BIC	142.079	142.2018	185.6301	195.0961	229.7694
Three factors					
	EMP	IP	GDP	RPI	MATS
AIC	100.1888	104.0789	176.0999	170.5159	191.6474
BIC	106.6182	110.5083	182.5293	176.9453	198.0768
Four factors					
	EMP	IP	GDP	RPI	MATS
AIC	179.9259	85.41766	165.1538	168.759	172.9321
BIC	188.4985	93.9902	173.7263	177.3315	181.5046

Table 21: Predicting the target variables in 1974 recession with the number of factors selected by AIC and BIC

1970/1-1973/10 predicts 1973/11-1975/3					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.1681363	0.2244251	0.6895722	0.6060233	0.7382171
MSE with non-housing	0.2623299	0.2010178	0.7126902	0.7546635	0.8063446
% Change in MSE	-35.91%	11.64%	-3.24%	-19.70%	-8.45%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with all variables	0.1940682	0.2244251	0.6895722	0.6060233	0.867262
MSE with non-housing	0.2623299	0.2010178	0.7126902	0.8010352	0.8063446
% Change in MSE	-26.02%	11.64%	-3.24%	-24.34%	7.55%

Table 22: Predicting 2008 economy with housing variables selected by Granger causality tests

2000-2007 predicts 2008 with housing variables selected by Granger-Causality test					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0211924	0.3470884	0.4266325	0.4680454	0.5426443
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	-73.36%	-15.85%	-1.58%	-24.58%	-32.67%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.024653	0.3701365	0.4266325	0.4976112	0.5426443
MSE with non-housing variables	0.0795517	1.709636	0.4334744	0.620593	0.8059636
% Change in MSE	-69.01%	-78.35%	-1.58%	-19.82%	-32.67%

Table 23: Predicting 2008 economy with housing variables selected by VAR model

2000-2007 predicts 2008 with housing variables selected by VAR model					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0197531	0.353599	0.4388264	0.4855111	0.5922867
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	-75.17%	-14.27%	1.23%	-21.77%	-26.51%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0246147	0.3866493	0.4388264	0.468211	0.5922867
MSE with non-housing variables	0.0795517	1.709636	0.4334744	0.620593	0.8059636
% Change in MSE	-69.06%	-77.38%	1.23%	-24.55%	-26.51%

Table 24: Housing variables selected by FAVAR model using information criteria from 2000 to 2008

Housing variables selected for EMP: DSAL12MO UNDCONTSA UNDCON5MUSA UNDCON1USA DSAL6MO HNFS DSAL3MO UNDCON24USA PERMIT1 AUTHNOT5MU PERMIT HSN1F HOUST1F COMPU1USA HOUST COMPUTSA DSAL1MO AUTHNOTT AUTHNOT1U PERMIT24 PERMIT5

IP: DSAL12MO UNDCONTSA UNDCON1USA UNDCON5MUSA DSAL6MO HNFS DSAL3MO PERMIT1 UNDCON24USA AUTHNOT5MU PERMIT AUTHNOTT COMPU1USA HOUST1F DSAL1MO HOUST HSN1F COMPUTSA

GDP: DSAL12MO UNDCONTSA UNDCON5MUSA UNDCON1USA DSAL6MO HNFS DSAL3MO UNDCON24USA PERMIT1 AUTHNOT5MU PERMIT AUTHNOTT HOUST1F HOUST DSAL1MO COMPU1USA COMPUTSA

RPI: DSAL12MO UNDCONTSA UNDCON5MUSA UNDCON1USA HNFS DSAL6MO DSAL3MO UNDCON24USA PERMIT1 AUTHNOT5MU PERMIT AUTHNOTT HOUST1F HOUST

MATS: DSAL12MO UNDCONTSA UNDCON5MUSA UNDCON1USA DSAL6MO HNFS DSAL3MO UNDCON24USA PERMIT1 AUTHNOT5MU PERMIT AUTHNOTT HOUST1F HOUST DSAL1MO COMPU1USA COMPUTSA HSN1F AUTHNOT1U PERMIT24 PERMIT5 HOUST5F COMPU5MUSA

Table 25: Predicting 2008 economy with housing variables selected by FAVAR model with housing variables selected by information criterion

2000-2007 predicts 2008 with housing variables selected by FAVAR model using information cirterion					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0622241	0.4381965	0.4358458	0.4963904	0.5120734
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	-21.78%	6.24%	0.55%	-20.01%	-36.46%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0664097	0.3789996	0.4358458	0.5164312	0.5120734
MSE with non-housing variables	0.0795517	1.709636	0.4334744	0.620593	0.8059636
% Change in MSE	-16.52%	-77.83%	0.55%	-16.78%	-36.46%

Table 26: Housing variables selected by FAVAR model using wald test from 2000 to 2008

Housing variables selected for EMP: HPI DSAL1MO DSAL3MO DSAL6MO DSAL12MO AUTHNOT24U AUTHNOT5MU COMPU1USA COMPUTSA HOUST1F HOUST2F HOUST PERMIT1 PERMIT24 PERMIT UNDCON1USA UNDCON24USA UNDCON5MUSA UNDCONTSA HSN1F HNFS

IP: FNCP DSAL1MO DSAL3MO DSAL6MO DSAL12MO AUTHNOT5MU COMPU1USA COMPU5MUSA COMPUTSA HOUST1F HOUST2F HOUST PERMIT1 PERMIT24 PERMIT5 PERMIT UNDCON1USA UNDCON5MUSA HSN1F

GDP: DSAL1MO DSAL3MO DSAL6MO DSAL12MO AUTHNOT5MU COMPU1USA COMPU24USA COMPUTSA HOUST1F HOUST5F HOUST PERMIT1 PERMIT24 PERMIT UNDCON24USA UNDCON5MUSA HSN1F

RPI: DSAL1MO DSAL3MO DSAL6MO DSAL12MO AUTHNOT24U COMPUTSA HOUST1F HOUST2F HOUST PERMIT24 PERMIT UNDCON5MUSA HSN1F HNFS

MATS: DSAL1MO DSAL3MO DSAL6MO DSAL12MO AUTHNOT1U AUTHNOT24U AUTHNOTT COMPU1USA COMPU24USA COMPUTSA HOUST1F HOUST2F HOUST5F HOUST PERMIT1 PERMIT24 PERMIT UNDCON1USA UNDCON24USA UNDCON5MUSA UNDCONTSA HSN1F HNFS

Table 27: Predicting 2008 economy with housing variables selected by FAVAR model with housing variables selected by wald test

2000-2007 predicts 2008 with housing variables selected by FAVAR model using wald test					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0220705	0.4193878	0.4227008	0.4841379	0.5152333
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	-72.26%	1.68%	-2.49%	-21.99%	-36.07%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0249207	0.3803677	0.4227008	0.4897052	0.5152333
MSE with non-housing variables	0.0795517	1.709636	0.4334744	0.620593	0.8059636
% Change in MSE	-68.67%	-77.75%	-2.49%	-21.09%	-36.07%

Table 28: Housing variables selected by hard thresholding from 2000 to 2008

Housing variables selected for EMP: DSAL1MO DSAL3MO DSAL6MO DSAL12MO COMPU1USA COMPU24USA COMPUTSA HOUST1F HOUST2F HOUST5F HOUST PERMIT1 PERMIT24 PERMIT5 PERMIT UNDCON1USA UNDCONTSA HSN1F

IP: FNCP HPIPO DSAL1MO DSAL3MO DSAL6MO DSAL12MO AUTHNOT1U COMPU1USA COMPUTSA HOUST1F HOUST PERMIT1 PERMIT24 PERMIT UNDCON1USA UNDCON24USA UNDCONTSA HSN1F

GDP: FNCP DSAL1MO DSAL3MO DSAL6MO DSAL12MO AUTHNOT1U COMPU1USA COMPUTSA HOUST1F HOUST5F HOUST PERMIT1 PERMIT24 PERMIT5 PERMIT UNDCON1USA UNDCONTSA HSN1F

RPI: HPIPO DSAL1MO DSAL3MO AUTHNOT1U AUTHNOTT COMPU1USA COMPUTSA HOUST1F HOUST PERMIT1 PERMIT24 PERMIT UNDCON1USA UNDCONTSA HSN1F

MATS: DSAL1MO DSAL3MO DSAL6MO DSAL12MO COMPU1USA COMPUTSA HOUST1F HOUST5F HOUST PERMIT1 PERMIT24 PERMIT UNDCON1USA UNDCON24USA UNDCON5MUSA HSN1F

Table 29: Predicting 2008 economy with housing variables selected by hard thresholding

2000-2007 predicts 2008 with housing variables selected by hard thresholding method					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0250967	0.3221876	0.368721	0.3938309	0.5012728
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	-68.45%	-21.89%	-14.94%	-36.54%	-37.80%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with selected variables	0.0255058	0.3228492	0.368721	0.3938309	0.5012728
MSE with non-housing variables	0.0795517	1.709636	0.4334744	0.620593	0.8059636
% Change in MSE	-67.94%	-81.12%	-14.94%	-36.54%	-37.80%

Table 30: Correlations between housing volume measures of different units of structure and target variables two years ahead from 2000 to 2008

Target Variables with Housing Units Completed

	EMP	IP	GDP	RPI	MATS
EMP	1.0000				
IP	0.9145	1.0000			
GDP	0.8692	0.8689	1.0000		
RPI	0.9390	0.8847	0.9743	1.0000	
MATS	0.8609	0.9319	0.9611	0.9327	1.0000
COMPU1USA L24.	0.8985	0.8091	0.9060	0.9376	0.8510
COMPU24USA L24.	0.0535	0.0155	-0.0362	0.0300	-0.0957
COMPU5MUSA L24.	-0.2609	-0.3478	-0.2303	-0.2320	-0.3027
COMPUTSA L24.	0.8621	0.7536	0.8697	0.9054	0.7972

Target Variables With Housing Units Authorized But Not Yet Started

	EMP	IP	GDP	RPI	MATS
EMP	1.0000				
IP	0.9145	1.0000			
GDP	0.8692	0.8689	1.0000		
RPI	0.9390	0.8847	0.9743	1.0000	
MATS	0.8609	0.9319	0.9611	0.9327	1.0000
AUTHNOT1U					
L24.	0.9299	0.8678	0.9157	0.9443	0.8785
AUTHNOT24U					
L24.	0.7997	0.7899	0.8119	0.8325	0.7938
AUTHNOT5MU					
L24.	0.5605	0.4459	0.4683	0.5066	0.4255
AUTHNOTT					
L24.	0.9344	0.8633	0.9093	0.9410	0.8697

Target Variables With Housing Starts

	EMP	IP	GDP	RPI	MATS
EMP	1.0000				
IP	0.9145	1.0000			
GDP	0.8692	0.8689	1.0000		
RPI	0.9390	0.8847	0.9743	1.0000	
MATS	0.8609	0.9319	0.9611	0.9327	1.0000
HOUST1F					
L24.	0.8393	0.8690	0.7884	0.8125	0.8625
HOUST2F					
L24.	0.1341	0.1771	0.1000	0.1054	0.1156
HOUST5F					
L24.	0.0891	0.0680	0.0409	0.0639	0.0835
HOUST					
L24.	0.8091	0.8353	0.7503	0.7775	0.8287

Target Variables With Housing Permits

	EMP	IP	GDP	RPI	MATS
EMP	1.0000				
IP	0.9145	1.0000			
GDP	0.8692	0.8689	1.0000		
RPI	0.9390	0.8847	0.9743	1.0000	
MATS	0.8609	0.9319	0.9611	0.9327	1.0000
PERMIT1					
L24.	0.8275	0.8823	0.8106	0.8208	0.8975
PERMIT24					
L24.	0.6835	0.7442	0.6673	0.6736	0.7565
PERMIT5					
L24.	0.5458	0.5000	0.3856	0.4698	0.4052
PERMIT					
L24.	0.8487	0.8912	0.8050	0.8289	0.8883

Target Variables With Housing Units Under Construction

	EMP	IP	GDP	RPI	MATS
EMP	1.0000				
IP	0.9145	1.0000			
GDP	0.8692	0.8689	1.0000		
RPI	0.9390	0.8847	0.9743	1.0000	
MATS	0.8609	0.9319	0.9611	0.9327	1.0000
UNDCON1USA					
L24.	0.9518	0.8803	0.9195	0.9592	0.8928
UNDCON24USA					
L24.	0.5480	0.5586	0.5260	0.5674	0.4880
UNDCON5MUSA					
L24.	0.8814	0.7468	0.9026	0.9379	0.8025
UNDCONTSA					
L24.	0.9433	0.8546	0.9258	0.9653	0.8769

Table 31: Predicting 2008 economy with only housing volume measures added to the non-housing leading indicators

Percentage change in MSE with only housing volume measures added to the leading indicators					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with housing volume added	0.0280425	0.3483663	0.4365107	0.4234593	0.5114832
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	-64.75%	-15.54%	0.70%	-31.77%	-36.54%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with housing volume added	0.0261366	0.3716343	0.4365107	0.4932914	0.5114832
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	-67.15%	-9.90%	0.70%	-20.51%	-36.54%

Table 32: Predicting 2008 economy with only housing price measures added to the non-housing leading indicators

Percentage change in MSE with only housing price measures added to the leading indicators					
AIC					
	EMP	IP	GDP	RPI	MATS
MSE with housing price added	0.0865318	0.4670185	0.5378975	0.5865281	0.7599994
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	8.77%	13.22%	24.09%	-5.49%	-5.70%
BIC					
	EMP	IP	GDP	RPI	MATS
MSE with housing price added	0.0847709	0.6062078	0.5378975	0.818466	0.7599994
MSE with non-housing variables	0.0795517	0.4124747	0.4334744	0.620593	0.8059636
% Change in MSE	6.56%	46.97%	24.09%	31.88%	-5.70%

Table 33: Correlations between consumer confidence index and housing volume measures from 2000 to 2008

	CCI	L18. DSAL1MO	L18. DSAL3MO	L18. DSAL6MO	L18. DSAL12MO
CCI	1.0000				
DSAL1MO L18.	−0.6145	1.0000			
DSAL3MO L18.	−0.6268	0.9270	1.0000		
DSAL6MO L18.	−0.6204	0.8607	0.9565	1.0000	
DSAL12MO L18.	−0.6069	0.8705	0.9264	0.9258	1.0000
CSPI L18.	−0.4036	−0.0130	−0.1685	−0.2834	−0.1551
FNCP L18.	−0.6609	0.3606	0.2424	0.0853	0.1810
HPI L18.	−0.1212	−0.4385	−0.5522	−0.6141	−0.5766
HPIPO L18.	−0.8635	0.5126	0.4270	0.3628	0.4406

	CCI	L18. AUTHN~1U	L18. AUTHNOTT	L18. COMPU1~A	L18. COMPUTSA
CCI	1.0000				
AUTHNOT1U L18.	0.8538	1.0000			
AUTHNOTT L18.	0.6741	0.9106	1.0000		
COMPU1USA L18.	0.6887	0.7761	0.5542	1.0000	
COMPUTSA L18.	0.6126	0.6877	0.4369	0.9795	1.0000
HOUST1F L18.	0.9360	0.8072	0.5746	0.7131	0.6577
HOUST L18.	0.9213	0.7955	0.5608	0.7112	0.6648
PERMIT1 L18.	0.9490	0.8337	0.5982	0.7100	0.6556
PERMIT L18.	0.9541	0.8482	0.6282	0.7300	0.6739
UNDCON1USA L18.	0.8624	0.8894	0.6561	0.8981	0.8526
UNDCONTSA L18.	0.8192	0.8783	0.6526	0.9162	0.8747
HSN1F L18.	0.8771	0.7847	0.5361	0.6887	0.6372
HNFS L18.	−0.5959	−0.2741	−0.1099	−0.0810	−0.0447

Table 34: Correlations between consumer confidence index and target variables from 2006 to 2008

	L6. CCI	EMP	IP	GDP	RPI	MATS
CCI L6.	1.0000					
EMP	0.8737	1.0000				
IP	0.9109	0.9663	1.0000			
GDP	0.4797	0.7633	0.7057	1.0000		
RPI	0.6907	0.7618	0.7615	0.5755	1.0000	
MATS	0.9199	0.9537	0.9736	0.6533	0.6941	1.0000

APPENDIX C: VARIABLE DESCRIPTION

Target variables:

Variable name	Variable description	Transforma-tion form
EMP	Employment in nonfarm industry	6
IP	Industrial production	6
GDP	Gross domestic output	6
RPI	Real personal income less transfer	6
MATS	Real manufacturing and trade sales	6

Housing variables:

Variable name	Variable description	Transforma-tion form
CSPI	Case-Schiller price index	6
FNCP	Residential price index	6
HPI	Median price index with single-house combined	6
HPIPO	Purchase Only House Price Index for the United States	6

DSAL1MO	Distressed sales last 1 month	4
DSAL3MO	Distressed sales last 3 month	4
DSAL6MO	Distressed sales last 6 month	4
DSAL12MO	Distressed sales last 12 month	4
AUTHNOT1U	Housing Units Authorized, But Not Yet Started, 1-Unit Structures	4
AUTHNOT24U	Housing Units Authorized, But Not Yet Started, 2-4 Unit Structures	4
AUTHNOT5MU	Housing Units Authorized, But Not Yet Started, 5-Unit Structures or more	4
AUTHNOTT	Housing Units Authorized, But Not Yet Started, total	4
COMPU1USA	Housing Units Completed, 1-Unit Structures	4
COMPU24USA	Housing Units Completed, 2-4 Unit Structures	4
COMPU5MUSA	Housing Units Completed, 5-Unit Structures or more	4
COMPUTSA	Housing Units Completed, total	4
HOUST1F	Housing Starts, 1-Unit Structures	4

HOUST2F	Housing Starts, 2-4 Unit Structures	4
HOUST5F	Housing Starts, 5-Unit Structures or more	4
HOUST	Housing Starts, total	4
PERMIT1	Housing Units Authorized by Building Permits, 1-Unit structures	4
PERMIT24	Housing Units Authorized by Building Permits, 2-4 Unit Structures	4
PERMIT5	Housing Units Authorized by Building Permits, 5-Unit Structures or more	4
PERMIT	Housing Units Authorized by Building Permits, total	4
UNDCON1USA	Housing Units Under Construction, 1-Unit Structures	4
UNDCON24USA	Housing Units Under Construction, 2-4 Unit Structures	4
UNDCON5MUSA	Housing Units Under Construction, 5-Unit Structures or more	4
UNDCONTSA	Housing Units Under Construction, total	4

HSN1F	New One Family Houses Sold: United States	4
HNFS	New One Family Homes For Sale in the United States	4

Non-housing variables:

Variable name	Variable description	Transformation form
AWHP	Avg. weekly hours of production workers in manufacturing	1
CAP	Capacity utilization rate in manufacturing	1
TBILL10YR	Interest rate on 10 year US T-bill constant maturity	2
TBILL3MO	3 months T-bill rate	2
TB110	Spread between 10-year and 1-year Treasury bonds	2
CP3	3 months CP rate	2
TBCP3M	3 months T-bill rate-3 months CP rate	1
EMPPT	Number of people working part-time in nonagricultural industries because of slack work	5
WICI	Average weekly initial claims for unemployment insurance	5
UMOD	Real manufacturers' unfilled orders in durable goods industries	5
MOCMQ	Manufacturers' new orders, consumer goods and	5

	materials	
TWIER5	Trade weighted index of nominal exchange rates between US and UK, West Germany, France, Italy and Japan TWIER	5
MNONC	Manufacturers' new orders, nondefense capital goods	5
SP500	Stock prices, S&P500 common stocks	5
FFR	Fed funds rate	2
TB_FFR	10 year T-bill – fed funds rate	1
ISMNOI	ISM® new orders index	1
ISDVP	Index of supplier deliveries, vendor performance	1
NAPM	National Association of Purchasing Managers' index of vendor performance	1
CCI	Index of consumer expectations	1
MSIMIM1	Monetary Services Index for M1	6
MSIM2M	Monetary Services Index for M2 omitting the small-denomination time deposits	6
MSIM2	Monetary Services Index for M2	6

MSIMZM	Monetary Services Index for all zero-maturity assets	6
MSIALL	Monetary Services Index including all the assets	6
M2	Money supply M2	6

VITA

Name Wenxian Xu

Phone 662-380-3632

E-mail wenxianxu87@gmail.com

Website: www.wenxianxu.com

Education

University of Mississippi, Ph.D. in Economics, May 2014

University of Mississippi, M.A. in Economics, May 2011

East China University of Political Science and Law (China), B.A. in Economics, July 2008

Research Fields and Interests

Macroeconomics, Econometrics, Econometric Modeling

Job Market Paper

"The Importance of Housing Variables in Predicting the Economic Fluctuations" Dissertation, University of Mississippi (Committee Chair, Dr. Walter Mayer)

Working Paper

"More Powerful and Robust Diebold-Mariano and Morgan-Granger-Newbold Tests" (With Dr. Walter Mayer and Dr. Xin Dang, University of Mississippi)

Teaching Experience

University of Mississippi:

Instructor, Econ 202 Principles of Microeconomics (6 sections): Spring 2011, Fall 2011, Spring 2012, Fall 2012, Winter 2012, Spring 2013, Winter 2013

Instructor, Econ/Bus 230 Economic Statistics I: Spring 2014

Teaching assistant, Econ 203 Principles of Macroeconomics (3 sections): Spring 2009, Fall 2009

Teaching assistant, Econ 230 Economic Statistics I (4 sections): Spring 2010, Fall 2011

Teaching assistant, Econ 402/607 Econometrics (2 sections): Fall 2011, Fall 2012

Statistical Software and Language

Stata, SAS, Matlab, R, Eviews

Proficient in English, native in Mandarin, familiar with Cantonese

Honors and Awards

College of Liberal Arts Dissertation Fellowship, University of Mississippi, Fall 2013

Summer Research Assistantship Award, Summer 2013

Department of Economics Graduate Assistantship, University of Mississippi, Fall 2008 - Present

Conferences

Presenter and Discussant, Southern Economic Association Annual Meeting, Tampa, FL, Nov 23-25, 2013

Presenter and Discussant, Missouri Valley Economic Association Annual Meeting, Memphis, TN, Oct 25-27, 2012

Memberships and Activities

Memberships

American Economic Association

Southern Economic Association

Activities

Intensive English Workshop for International Instructors, University of Mississippi, Fall 2013

Graduate Instructor/Teaching Assistant Training, University of Mississippi, Fall 2009, Fall 2011

Southern Economic Association Annual Meeting, Atlanta, GA, Nov 19-21, 2010

Economics Teaching Conference, Memphis, TN, Oct 21-22, 2010