University of Connecticut

Department of Economics Working Paper Series

Forecasting US Real Private Residential Fixed Investment Using a Large Number of Predictors

Goodness C. Aye
University of Pretoria

Rangan Gupta
University of Pretoria

Stephen M. Miller
University of Nevada, Las Vegas
University of Connecticut

Mehmet Balcilar
Eastern Mediterranean University

Working Paper 2014-10

May 2014

365 Fairfield Way, Unit 1063
Storrs, CT 06269-1063
Phone: (860) 486-3022
Fax: (860) 486-4463
http://www.econ.uconn.edu/

Forecasting US Real Private Residential Fixed Investment Using a Large Number of Predictors

Goodness C. Aye,[a] Stephen M. Miller[b,c], Rangan Gupta,[a] and Mehmet Balcilar[d]

Abstract

This paper employs classical bivariate, factor augmented (FA), slab-and-spike variable selection (*SSVS*)-based, and Bayesian semi-parametric shrinkage (BSS)-based predictive regression models to forecast US real private residential fixed investment over an out-of-sample period from 1983:Q1 to 2011:Q2, based on an in-sample estimates for 1963:Q1 to 1982:Q4. Both large-scale (188 macroeconomic series) and small-scale (20 macroeconomic series) *FA*, *SSVS,* and *BSS* predictive regressions, as well as 20 bivariate regression models, capture the influence of fundamentals in forecasting residential investment. We evaluate the *ex-post* out-of-sample forecast performance of the 26 models using the relative average Mean Square Error for one-, two-, four-, and eight-quarters-ahead forecasts and test their significance based on the McCracken (2004, 2007) *MSE-F* statistic. We find that, on average, the *SSVS-Large* model provides the best forecasts amongst all the models. We also find that one of the individual regression models, using house for sale (*H4SALE*) as a predictor, performs best at the four- and eight-quarters-ahead horizons. Finally, we use these two models to predict the relevant turning points of the residential investment, via an *ex-ante* forecast exercise from 2011:Q3 to 2012:Q4. The *SSVS-Large* model forecasts the turning points more accurately, although the *H4SALE* model does better toward the end of the sample. Our results suggest that economy-wide factors, in addition to specific housing market variables, prove important when forecasting in the real estate market.

Keywords: Private residential investment, predictive regressions, factor-augmented models, Bayesian shrinkage, forecasting

JEL Classification: C32, E22, E27

a Department of Economics, University of Pretoria, Pretoria, 0002, South Africa.

b Department of Economics, University of Nevada, Las Vegas, Las Vegas, NV 89154, USA.

c Corresponding author. Email: styephen.miller@unlv.edu.

d Department of Economics, Eastern Mediterranean University, Famagusta, Northern Cyprus, via Mersin 10, Turkey.

1. Introduction

This paper considers the dynamics of US real private residential fixed investment and the ability of classical individual bivariate, factor-augmented, and Bayesian-shrinkage based predictive regression models to forecast this series. Residential investment includes new construction, expenditures on maintenance and home improvement, equipment purchased for use in residential structures, and brokerage commissions (Krainer, 2006).

The dynamics of residential investment plays a critical role in mortgage lending, portfolio investment decisions, and economic growth. Financial institutions more willingly lend for residential real estate investment than most other activities. Long-term investors consider residential property because the income stream from housing links to wage growth and can offer investors a better hedge against their liabilities than commercial property, which more closely links to the slower growing retail price growth series and other property market indicators (Daly, 2008). Moreover, the stable income returns (rent) and high total returns (rent plus capital growth), and prospects for portfolio diversification makes residential property attractive to investors. Residential investment also possesses a leverage advantage. Finally, housing construction can function as a locomotive, stimulating growth in other sectors, particularly finance, insurance, real estate, certain services, and segments of retail trade (Browne, 2000).

The housing sector, in general, provides an important channel through which monetary policy affects the economy. In addition, the housing sector is a leading indicator of aggregate demand (Demers, 2005). Understanding the evolution of this sector enables forward-looking central banks to predict more accurately housing expenditure.

Forecasting US residential investment helps to identify business cycle turning points. Residential investment significantly contributed to the recent financial crisis and Great Recession. In addition, Green (1997) notes that it historical leads US business cycles and proves useful in forecasting GDP from 1959 to 1992. Figure 1 clearly shows that residential fixed investment to GDP turns down (up) prior to recessions (recoveries), providing a leading indicator to the business cycle.

Fisher and Gervais (2007) note that residential investment growth in the US declined significantly since 1984. Thus, the overall decline in macroeconomic volatility experienced during the Great Moderation reflects in significant ways the declining share of residential investment growth in US real GDP growth, since residential investment is such a highly volatile component of GDP (Green 1997; Dynan et al. 2006; Peek and Wilcox 2006). See Figure 2. Although, residential investment historically contributes only about 5 percent of US GDP, it makes large contributions to output growth in recoveries (Lunsford, 2013). In this regard, Bernanke (2009) and Kohn (2009), following the National Bureau of Economic Research (NBER) 2009:Q2 business cycle trough, note that residential investment provides the source of economic growth going forward. Recently, Bernanke (2012) and Yellen (2013) also note that the negative contribution of residential investment makes the recent recovery unusual. Further, declines in residential investment also typically proceed recessions (Figure 1 and Leamer, 2007). Therefore, accurate forecasts of US residential investment movements can help to identify business cycle peaks.

Despite the importance of residential investment and its forecast, few studies forecast it (see the literature review section). Therefore, the current study fills this lacuna by providing the forecasts of US private residential fixed investment. Several key questions exist. What variable(s) prove critical in predicting private residential investment? In other words, can we accurately predict private residential investment with information limited only to the housing market variable(s)? Or, do we need to consider economy-wide factors in addition to specific housing market variables?

Second, which model(s) more accurately forecast US private fixed residential investment? According to Krainer (2006), residential investment measures the quantity of new housing supplied to the economy, and, in the long run, it should satisfy the overall demand for new housing. Thus, residential investment depends on supply and demand factors. In this regard, we include both demand- and supply-side factors in our forecasting models.

Two broad approaches exist for incorporating information from a large number of data series – extracting common factors or principle components (Stock and Watson, 2002; Koop and Korobilis, 2011) and Bayesian shrinkage methods (Korobilis, 2013a, 2013b). In this study, we consider both approaches for small- and large-scale models that include 20 and 188 additional predictors, respectively. In addition, we also forecast using individual bivariate regressions, where we regress each of the 20 variables in the small-scale models, in turn, on real private residential fixed investment.

The difficulty in forecasting economic variables such as residential investment occurs because the forecast depends on the models used to generate them. Thus, we must crucially evaluate forecasts from different models and to select the 'best forecast' based on an objective criterion (Dua et al., 2008). Further, Clements and Hendry (1998) argue that in time-series models, estimation and inference basically mean minimizing the one- (or multi-) step-ahead forecast errors. Therefore, superior models produce smaller forecast errors than its competitors. We evaluate the forecasts from the 26 predictive models using the mean square error (*MSE*) of each model relative to the MSE of an autoregressive (AR) (benchmark) model. Further, we test for the significance of the *MSEs* using the McCracken (2004, 2007) *MSE-F* statistic.

We organize the rest of the paper as follows: Section 2 provides an overview of the existing literature on forecasting residential investment. Section 3 describes the empirical models that we use for forecasting. Section 4 describes the data and reports and evaluates our results. Section 5 concludes.

2. Literature Review

Although a significant research activity documents the modelling of residential investment,[1] few studies consider the forecasting of residential investment - Demers (2005), Baghestani (2011), and Lundsford (2013). Demers (2005) proposes and evaluates econometric models that explain and forecast real quarterly housing expenditure in Canada, using structural, using fundamentals such as wealth and demographics, and leading-indicator, using variables such as housing starts and household indebtedness, models of the Canadian housing sector. The results show that the preferred structural model with a structural break ranks better than each of the 12 leading-indicator models of construction investment.

Baghestani (2011) compares the performance of the Federal Reserve System (Greenbook) and private (Survey of Professional Forecasters) forecasts of growth in both business and residential investment for 1983 to 2004 and reaches four main conclusions. First, in support of the asymmetric information hypothesis, the shorter (longer) horizon Federal Reserve forecasts of growth in business (residential) investment contain useful predictive information beyond that included in private forecasts. Second, while bias exists in all Federal Reserve forecasts, no bias emerges in some (no) instances for the private forecasts of growth in business (residential) investment. Third, the private forecasts overall do better than those of the Federal Reserve in outperforming the univariate *ARMA* forecasts. Fourth, the Federal Reserve and private forecasts of growth in business (residential) investment, while directionally accurate imply symmetric (asymmetric) loss.

Lundsford (2013) develops a forecasting model of US residential investment with an inflow-outflow structure that treats housing starts as flows into construction and completions

---

[1] See, for example, Egebo et al. (1990); Brayton and Tinsley (1996); Edge (2000); McCarthy and Peach (2002); Berger-Thomas and Ellis (2004); Dynan et al. (2006); Fisher and Gervais (2007); Choy et al. (2011).

as flows out of construction. The proposed model significantly reduces the root mean squared prediction errors of the Survey of Professional Forecasters at all forecast horizons.

In sum, the existing literature on forecasting residential investment, in general, and private residential investment, in particular, provides limited findings despite the importance of this series in business cycles.

3. Methodology

We consider several predictive regression models for forecasting the US real private residential fixed investment. These include the spike-and-slab priors for Bayesian variable selection (*SSVS*), the Bayesian semi-parametric shrinkage (*BSS*) prior, and the factor-augmented predictive regression (*FAPR*) models. In addition, we also consider individual predictive regressions based on the 20 variables that researchers identify as possibly incorporating predictive capability for residential investment.[2]

3.1 Spike-and-Slab Priors for Variable Selection (SSVS) Model

We start with a dynamic regression model of the following form:

$$y_{t+h} = \gamma + \sum_{i=1}^{p} \phi_i y_{t-(i-1)} + x_t'\beta + u_{t+h}, \qquad (1)$$

where y_{t+h} denotes the variable of interest (real private residential fixed investment) that we want to forecast, y_{t-i+1} denote the p own lags of y for $i = 1,...p$, x_t and β denote ($Kx1$) vectors of exogenous predictors and coefficients to estimate, respectively, and u_t denotes a Gaussian forecast error with zero mean and variance σ^2. We determine the optimal number of lags for the forecasting model based on the Schwarz information criterion (SIC), which, in

[2] The list of references to document the choice of these variables is available from the authors. The 20 variables include interest rates (3-month Treasury rate, *3TB*), real gross domestic product (*RGDP*), the consumer price index (*CPI*), the unemployment rate (*UNRATE*), the labour force participation rate (*LFPR*), the mortgage interest rate (*MORTG*), the business confidence index (*BCON*), the real house price index (*RHP*), the money supply (*M1*), real private consumption expenditure (*RPCON*), real government consumption expenditure (*RGCON*), the real change in private inventories (*RCPINV*), housing starts (*HOUST*), real non-residential fixed investment (*RNRFINV*), the Standard & Poor's stock price index (*S&P*), retail sales (*RSALES*), new private housing units authorized by building permit (*PERMIT*), number of new houses sold (*HSOLD*), and the months' supply of housing ratio (*HSUPPLY*).

turn, selects one lag. Hence, we include the intercept and one lag in the forecasting model. We assume that the regression coefficients $\theta = (\gamma,\ \varphi_1)$ as well as the variance σ^2 possess non-informative priors of the following form: $\theta \sim N(0_{2\times 1}, 100 I_2)$ and $\sigma^2 \sim iGamma(0.01,\ 0.01)$. When K becomes "large," Cremers (2002) and Koop and Potter (2004) argue for selecting the best, according to some criterion, variables/predictors, while Stock and Watson (2002) suggest using shrinkage by replacing x_t with its first few principal components.

One popular method for variable selection uses the spike-and-slab prior for the coefficients β formalized by Mitchell and Beauchamp (1988). Korobilis (2013b) implements this approach by writing

$$\beta_j \sim \pi\,\delta_0(\beta) + (1-\pi)N(0,\tau^2) \tag{2}$$

where $\delta_a(v)$ denotes the Dirac delta function for random variable v, which places all probability mass on the point a. Thus, the prior for $\beta_j,\ j = 1,\ \ldots,\ K$, mixes a point mass at zero (the spike) and a locally uninformative (depending on the size of τ^2) Gaussian prior (the slab). The data update the random probabilities π, which determine whether the prior of β_j equals zero or whether it comes from the unrestricted Gaussian density with variance τ^2. This prior does not explicitly model the correlation structure in the data when determining which variables enter the regression, which other popular model selection and averaging priors do model (Koop and Potter, 2004).

3.2. Bayesian Semi-Parametric Shrinkage (BSS) Prior Model

The structure of the macroeconomic data commonly used by macroeconomists frequently involves highly correlated variables. The simple spike-and-slab prior does not account for correlations in the data. Researchers developed a semi-parametric spike-and-slab prior

(MacLehose et al., 2007; and Dunson et al., 2008) as an extension to the simple spike-and-slab model to accommodate correlations in the data. Using this method, the coefficients β admit a prior of the following form:

$$\beta_j \sim \pi\delta_0(\beta) + (1-\pi)G \quad (3)$$

where $G \sim DP(\alpha G_0)$ and $G_0 \sim N(0,\ \tau^2)$. G is a nonparametric density that follows a Dirichlet process with base measure G_0 and concentration parameter α.[3] Usually G_0 is a well-known density (e.g., Gaussian), making the prior an infinite mix of the densities G_0. Hence, such priors are "pseudo-nonparametric," since a parametric mix of distributions approximates the unknown density G (Korobilis, 2013b). In this case, G_0 is Gaussian with zero mean and variance τ^2, which is the typical conjugate prior distribution used on linear regression coefficients. Hence, this prior implies that each coefficient β_j will either equal 0 with probability π, or will come from a mix of Gaussian densities with probability $(1-\pi)$. Further, we define prior distributions for the prior hyper-parameters α, π, and, τ, which show up in the hierarchical prior in Equation (3), to let the data determine their values. Following Korobilis (2013b), we define the hyper-prior distributions as follows: $\tau \sim iGamma(0.01,\ 0.01)$, $\alpha \sim Gamma(1,\ 2)$, and $\pi \sim Beta(1,\ 1)$.[4] Using these fairly uninformative hyper-parameters, we estimate the regression coefficients using the Markov

[3] The Dirichlet process, or Ferguson distribution, was developed by Ferguson (1973) as a continuous probability distribution (Shotwell and Slate, 2011) instead of over numbers (real numbers, non-negative integers, etc.). The usual parameterization includes a concentration parameter and a base measure.

[4] The gamma distribution is a two-parameter family of continuous probability distributions on the positive real line, usually parameterized with (1) shape and scale parameters, (2) shape and inverse scale parameters, or (3) shape and mean parameters (SAS, 2014). The inverse gamma distribution is a two-parameter family of continuous probability distributions on the positive real line, which is distributed as the reciprocal of a variable distributed according to the gamma distribution (SAS, 2014). The beta distribution is a general statistical distribution that relates to the gamma distribution and contains two free parameters, often used as a prior distribution for binomial proportions in Bayesian analysis (Evans et al. 2000).

Chain Monte Carlo (MCMC) methods.[5] After monitoring for convergence, we run the Gibbs sampler for 150,000 iterations after an initial burn-in period of 50,000 iterations.

3.3 Factor-Augmented Predictive Regression (FAPR) Model

The factor-augmented predictive regression models augment the *AR* model with extracted common components to forecast the real private residential fixed investment. Suppose that X_t equals a $n \times 1$ covariance stationary vector standardized to possess a mean zero and a variance equal to one, obtained from the original $n \times 1$ vector of *I(1)* and *I(0)* variables Y_t. Then, consider the following model:

$$X_t = \lambda F_t + U_t , \tag{4}$$

where F_t denotes a vector of common factors, λ denotes a vector of factor loadings associated with F_t, and U_t denotes the idiosyncratic component of X_t. The product λF_t equals the common component of X_t. Equation (4) then captures the factor representation of the data. Note that we cannot observe the factors, their loadings, or the idiosyncratic errors and, hence, must estimate them. The estimation technique matters for factor forecasts. We adopt the Bai and Ng (2002) and Alessi et al., (2010) methods to determine the number of common components for the large and small macroeconomic datasets, respectively, and then use the extracted factors, instead of the individual predictors (x in equation (1)), in the predictive regression model to create a *FAPR* model. The tests reveal 6 and 3 factors, respectively, for the large and small datasets. Again, we include one lag of private residential investment as in the previous models. We estimate the *FAPR* model using ordinary least squares (OLS) and perform out-of-sample tests based on the recursive scheme.

3.4 Individual Regressions

We also run bivariate predictive regressions between real private residential fixed investment

---

[5] The on-line Technical Appendix of Korobilis (2013b) details the MCMC method.

and each of the predictors included in the small-scale models. We include one lag of real private residential investment as a control variable, when testing the forecasting ability of the specific predictor. We estimate the bivariate predictive regressions using OLS and perform out-of-sample tests based on the recursive scheme.

4. Data and Empirical Results

4.1 Data

We use quarterly data on 189 macroeconomic series of the US economy, including real private residential fixed investment. We seasonally adjust all data, which cover 1963:Q1 to 2011:Q2. One hundred and eighty-four (184) variables in the dataset come originally from King and Watson (2012), which Korobilis (2013b) also used. Further details on the sources of the variables appear in these two papers.[6]

The original dataset spans 1959:Q to 2011:Q2. Since this dataset ends in 2011:Q2, our sample also ends at the same point. Given our interest in forecasting real private residential fixed investment, we also include five (5) other housing specific variables implying a total of 189 variables. The newly added variables include the real new home price index (*RHPI*, the US Census Bureau median house price for new houses sold, including the value of the lot (land price), divided by the personal consumption expenditure implicit deflator), the business confidence index (*BCON*), the number of new housing units for sale (*H4SALE*), the number of new housing units sold (*HSOLD*), and the number of months supply of housing (the number of new housing units for sale in a given month divided by the number of new units sold) (*HSUPPLY*), these additional variables come from the U.S. Census Bureau except *BCON*, which comes from the Global Financial database. These newly added data series mostly became available in 1963:Q1. Hence, our total sample covers 1963:Q1 to 2011:Q2. We present the annualized quarterly growth rates of real GDP and real private residential

[6] The Appendix contains a full description of all variables and the relevant stationarity transformations used.

investment in Figure 2. This figure indicates that real private residential investment growth exhibits much higher volatility (standard deviation of 19.4) than real GDP growth (standard deviation of 3.5). The volatility of real private residential investment growth declined during the Great Moderation as did the volatility of real GDP growth until rising again in the recent financial crisis and Great Recession. The highest growth rates for real GDP and real private residential investment occur in 1978:Q2 (15.4 percent) and 1983:Q1 (62.9 percent), respectively. The lowest growth rates occur in 2008:Q4 (-9.3 percent) and 1980:Q2 (-81.9 percent), respectively.

4.2 Estimation and Results

We consider forecasts at *h* = one-, two-, four-, and eight-quarter-ahead horizons of real private residential investment, using the relevant macroeconomic variables as predictors from our quarterly dataset (see the next section for more details on the sample). Following standard practice, we use the model with no predictors (i.e., a first-order autoregressive, *AR(1)*, model) as the benchmark model. We evaluate the out-of-sample forecast performance of the models using the Theil's *U* statistic, which measures the ratio of a specific model's forecast mean squared error (*MSE*) to the *AR(1)* model's *MSE*. If the model's forecast *MSE* falls below, above, or equals the *AR(1)* model's forecast *MSE*, then *U* is less than, greater than, or equal to one, implying the model produces better, worse, or equal forecast performance than a simple *AR(1)*-benchmark model.

To formally test whether forecasts from a specific model are significantly more accurate than the *AR(1)* model forecasts, we use the McCracken (2007) *MSE-F* statistic.[7] The

[7] The *MSE-F* statistic uses the loss differential as follows: $MSE-F=(T-R-h+1)\left(\bar{d}\,/\,MSE_1\right)$, where T equals the number of observations in the total sample, R equals number of observations used to estimate the model from which we calculate the first forecast (i.e., the in-sample portion of *T*), *h* equals the forecast horizon, $\bar{d}=M\hat{S}E_0-M\hat{S}E_1$, $M\hat{S}E_i=(T-R-h+1)^{-1}\sum_{t=R}^{T-h}(u_{i,t+1})^2$ with $i=1,\ 0$, $M\hat{S}E_1$ corresponds to the *MSE* of the unrestricted model (i.e., the model with the relevant macroeconomic predictor variables), and $M\hat{S}E_0$ corresponds to the *MSE* of the restricted model (i.e., the *AR(1)*-benchmark model). A significant *MSE-F* statistic

MSE-F statistic tests the null hypothesis that a specific model forecast *MSE* equals the *AR(1)* model forecast *MSE* against the one-sided (upper-tail) alternative that the *MSE* of the specific model falls below the *MSE* of the *AR(1)* model.

First, we select the best model for forecasting real private residential fixed investment, using the Theil's *U* (*MSE* of the unrestricted model relative to the *MSE* of the *AR* model) statistic. We also test for the significance of the Theil's *U* statistic using McCracken (2004, 2007) *MSE-F* statistic. Second, we consider the ability of the model that performs the best amongst the Bayesian, factor, and individual regression models to predict the relevant turning points in the US private residential investment using *ex-ante* out-of-sample forecasts.[8] We consider two types of small- and large-scale Bayesian models, and two types of factor-augmented predictive regression models based on the small and large data sets and the 20 individual bivariate regression models. We use the *ex-post* forecasting exercise to choose the best multivariate and bivariate models to adopt for the *ex-ante* forecasting exercise.

4.2.1 *Ex-post* Out-of-Sample Forecasts. The data sample runs from 1963:Q1 to 2011:Q2. We use 80 out of our 194 total observations for first period forecast. This implies that we estimate each model over the in-sample period of 1963:Q3 to 1982:Q4 (after taking one lag, as unanimously suggested by all five lag-length selection criteria, and transforming to stationarity) and then estimate recursively over the out-of-sample period of 1983:Q1 to 2011:Q2. That is, we use the last 114 observations (i.e., 1983:Q1 to 2011:Q2) for the evaluation of h-step-ahead forecasts (*ex-post* out-of-sample forecasting). We re-estimate the models each quarter over the out-of-sample forecast horizon to update the estimates of the coefficients, before producing the one-, two-, four-, and eight-quarters-ahead forecasts. We

indicates that the unrestricted model forecasts are statistically more accurate than those of the restricted model.

[8] *Ex-post* forecasts are recursively updated in-sample in the forecasting equation to generate the multi-step-ahead forecasts, whereas the *ex-ante* multi-step-ahead forecasts are produced from a specific point in time (generally, from the end-point of data available on the predictors, which in our case is 2011:Q3-2012:Q4) without updating the parameter estimates. The *ex-ante* forecasts give an objective statistical method (approach) to choose the best performing models, which, in turn, we use to predict the turning points.

calculate the mean square errors (*MSE*) for the one-, two-, four-, and eight-quarters-ahead forecasts as well as their average across these four forecasts for the real private residential fixed investment across all the models. Using the best performing models, we perform out-of-sample *ex-ante* forecast from 2011:Q3 to 2012:Q4.

Table 1 reports the *ex-post* out-of-sample forecast results for the various models. The *SSVS-Small* and *SSVS-Large* rows list the spike-and-slab variable selection model with 20 and 188 predictors of real private residential fixed investment, respectively; the *BSS-Small* and *BSS-Large* rows, the Bayesian semi-parametric shrinkage model with 20 and 188 predictors; the *FAPR-Small* and *FAPR-Large* rows, the factor augmented predictive regression model with 20 and 188 predictors; and the individual regressions are bivariate predictive regressions of real private residential investment and each of the 20 predictors.

Table 1 reports the one-, two-, four-, and eight-quarter-ahead *MSEs* from the various specifications relative to the *MSE* of the *AR*-benchmark model as well as the average across the one-, two-, four-, and eight-quarter-ahead *MSEs*. For example, the 0.862 entry for the *SSVS-Large* model for the two-quarter-ahead forecast, means that the *SSVS-Large* model experienced a forecast *MSE* of only 86.2 percent of the *AR* model's forecast *MSE*. In other words, the *SVSS-Large* model improves over the *AR* model by 13.8 percent. We select the model that produces the lowest average *MSE* values as the 'best' specification for US real private residential fixed investment. Table 1 also compares whether the gain or loss in *MSE* of a specific model significantly differs from the *MSEs* obtained from the *AR* model based on the *MSE-F* test.

Several observations emerge. Consider the multivariate small and large models reported in the top part of Table 1. First, all four Bayesian models (namely *SSVS-Small, SSVS-Large, BSS-Small*, and *BSS-Large*) produce better forecasts than the *AR*-benchmark model at each forecast horizon and for the overall average. These gains prove statistically significant at the 1-percent level at all four horizons. Second, the *FAPR-Large* model produces a statistically significant more accurate forecast than the *AR* model at the one- and two-quarter-ahead horizons and for the average while the *FAPR-Small* model only outperforms the *AR*-benchmark model at the one-quarter-ahead horizon. Third, the large scale Bayesian models perform better than the small scale Bayesian models at each horizon as well as based on the overall average.

Now, consider the bivariate models in the bottom part of Table 1. Five (5) out of the 20 individual classical regression models (namely *PERMIT*, *MORTG*, *H4SALE*, *HSOLD*, and *HSUPPLY*) produce more accurate forecast than the *AR* model, when considering the average and nearly all of the relative *MSE* values across all horizons, where, in addition, most of these gains at individual forecast horizons proving significant at the 1-percent level. Alternatively, this means that, on average, the *AR* model forecasts prove more accurate or at least as good as the remaining 15 individual predictive regression models.

Finally, compare the forecast performance of the multivariate and bivariate models in Table 1. We observe that no single model outperforms all others at all horizons. In general, at short term horizon (i.e., h = 1 and 2), models with more information outperform models with less information whereas the later outperforms the models with more information at long term horizon (i.e., h = 4 and 8). Specifically, the *FAPR-Large* model performs better than all the other models at horizon one, improving over the *AR* model by 13.9 percent. At horizon two, the *SSVS-Large* model performs the best, improving over the *AR* model by 13.8 percent. The *H4SALE* model performs the best at horizons four and eight, improving over *AR* model by 15.7 percent and 28.8 percent, respectively. We observe that the *SSVS-Large* model produces the most accurate forecast based on the overall average forecast *MSE*. *SSVS-Large* model experienced a forecast *MSE* of only 84.6 percent of the *AR* model forecast *MSE*. In other words, it improves over the *AR*-benchmark model by 15.4 percent. Given the overall performance of the *SSVS-Large* model followed by the *H4SALE* model, we also compare the relative *MSE* of the former relative to the later. We find that the *MSE* of the *SSVS-Large* model significantly improves the relative *MSE* of the *H4SALE* at horizons 1 and 2 by 8 and 6 percent, respectively.[9]

Overall, the ex-post out-of-sample forecasts produce two general conclusions. First, the large-scale models perform better than small-scale and individual regression models, as well as the *AR*-benchmark model based on overall average *MSE*, thus justifying our decision to include 188 predictors in forecasting real private residential investment. Hence, this outcome highlights the importance of including more information through large number of variables, as models with large information sets can more closely mimic economic relationships.

Second, a smaller number of bivariate models that use the number of new private housing units authorized by building permits (*PERMIT*), the 30-Year conventional mortgage rate (MORTG), the number of new housing units for sale (*H4SALE*), the number of new housing units sold (*HSOLD*), and the month's supply of housing ratio (*HSUPPLY*) exhibit significantly better forecasts averaged across all horizons. These four variables directly relate to the housing market and real private residential investment. Finally, the better performance of the Bayesian large and small models noted in the previous paragraph also utilize these housing market variables in their estimated models.

4.2.2 *Ex-Ante* Forecasts. Having determined each of the optimal forecast models from the

[9] The *MSE-F* statistics is a one sided test and, hence, does not permit testing for significance of cases where the relative *MSE* values exceeds one. Given this, we use the Harvey *et al.*, (1998) test of forecast encompassing to determine if the information contained in the forecasts from the SSVS-Large (predictive regression model based on the *H4SALE*) encompasses that of the predictive regression model based on the *H4SALE* (SSVS-Large), We find that, at conventional levels of significance, the SSVS-Large model encompasses the *H4SALE* model at horizons one and four, while the *H4SALE* model encompasses the *SSVS-Large* at horizons eight and twelve. The details of these results are available upon request from the authors. Note however, our inferences based on the Harvey *et al.*, (1998) test only serves as a rough guide, since the nestedness of the *H4SALE* model affects the asymptotic distribution of the test statistic.

multivariate (*SSVS-Large*) and bivariate (*H4SALE*) models, respectively, we expose them to the acid-test of predicting the different turning points in the US private residential investment series. We implement this by performing out-of-sample *ex-ante* forecast over 2011:Q3 to 2012:Q4.[10]

Figure 2 plots the *ex-ante* out-of-sample forecasts and actual values. The *SSVS-Large* model tracks all the three turning points in the actual real private residential investment series, though it consistently underpredicts the actual series. Toward the end of the sample, however, the actual series maintains an increasing trend, while the *SSVS-Large* predicts a decline. The *H4SALE* model does not perform as well until the last two quarters of the sample where it tracked the actual series closely. Based on the *ex-ante* forecast results, *SSVS-Large* model appears to possess a slight edge over the bivariate model that uses the number of new housing units for sale (*H4SALE*) for forecasting US real private residential investment.

5. Conclusion

In this paper, we forecast the US real private residential investment using quarterly data from 1963:Q1 to 2011:Q2. We consider 3 large-scale, 3 small-scale, and 20 individual predictive regression models. Using the period of 1963:Q1 to 1982:Q2 as the in-sample period and 1983:Q1 to 2011:Q2 as the out-of-sample period, we compare the performance of alternative models based on the mean square error (*MSE*) relative to the *MSE* of the *AR(1)* benchmark model. We compare the relative *MSEs* for the one-, two-, four-, and eight-quarters-ahead forecasts. We also tests whether the gain or loss in *MSEs* of the unrestricted models significantly differ from the *MSEs* obtained from the *AR*-benchmark model based on the McCracken (2004, 2007) *MSE-F* statistic.

Our findings will prove valuable to potential investors and policy makers, since

[10] Table 1 also reports the relative forecasting performance of the *SSVS-Large* model to the bivariate model that includes the number of new housing units for sale (*H4SALE*), finding that the *SSVS-Large* significantly outperforms the *H4SALE* bivariate model only at one- and two-quarter-ahead horizons.

residential fixed investment provides an important leading indicator of the business cycle. Thus, good forecasts can help to improve portfolio investment and mortgage lending decisions, which subsequently can enhance overall economic growth.

The *ex post* out-of-sample results show that based on the average across the forecast horizons, five of the large- and small-scale models and five out of 20 individual bivariate regressions produce more accurate forecasts than the simple *AR*-benchmark model. More importantly, these gains generally prove significant. We also find that the *SSVS-Large* model outperforms the rest of the models at the two-quarter-head forecasts and also based on the average *MSE* across all forecast horizons. The individual predictive regression model based on the house for sale variable (*H4SALE*) performs best at the four- and eight-quarters-ahead horizons. The best performance at the one-quarter-ahead forecast horizon comes from the *FAPR-Large* model.

Using the *SSVS-Large* and *H4SALE* models, we provide *ex-ante* forecasts for real private residential investment over 2011:Q3 to 2012:Q4. Interestingly, the results clearly show that the *SSVS-Large* model captures most of the turning points in the actual real private residential investment series, while the *H4SALE* model does well toward the end of the sample.

In sum, several bivariate models outperform our AR-benchmark model. These better bivariate models generally include generally housing market variables. The multivariate large models perform better than the multivariate small models. In addition, the multivariate small models usually outperform the bivariate models. Hence, we conclude that the use of fundamental economic variables probably improves the forecasting performance of the US real private residential investment over the models that do not use such information. Also, our results suggest that economy-wide factors, in addition to specific housing market variables, can improve forecasts when evaluating the real estate market. Nonetheless, the bivariate model that uses homes for sale as a predictor performs nearly as well as the SSVS-Large model for *ex-post* and *ex-ante* forecasts. That is, as a practical matter, predicting residential investment may only require homes for sale.

References

Baghestani, H., 2011. Federal Reserve and private forecasts of growth in Investment. *Journal of Economics and Business* 63, 290–305.

Bai, J., and Ng, S., 2002. Determining the number of factors in approximate factor models. *Econometrica* 70, 191–221.

Berger-Thomas, L., and Ellis, L., 2004. Housing construction cycles and interest rates. Research Discussion Paper 2004-08, Reserve Bank of Australia, Sydney.

Bernanke, B. S., 2009. On the outlook for the economy and policy. Speech at the Economic Club of New York, New York, New York. http://www.federalreserve.gov/newsevents/speech/bernanke20091116a.htm

Bernanke, B. S., 2012. The economic recovery and economic policy. Speech at the New York Economic Club, New York, New York.

Brayton, F., and Tinsley, P., (eds.) 1996. A guide to FRB/US: A macroeconomic model of the United States. Federal Reserve Board Finance and Economics Discussion Series. http://www.federalreserve.gov/pubs/feds/1996/199642/199642abs.html

Browne, L. E., 2000. National and regional housing patterns. *New England Economic Rev*iew July/August, 32–57.

Choy, L. H. T., Ho, W. K. O., and Mak, S. W. K., 2011. Region-specific estimates of the determinants of residential investment. *Journal of Urban Planning and Development* 137, 1–6.

Clements, M., and Hendry, D., 1998. *Forecasting Economic Time Series*. New York: Cambridge University Press.

Cremers, M., 2002. Stock return predictability: A Bayesian model selection perspective. *Review of Financial Studies* 15, 1223–1249.

Daly, J. S., 2008. Overcoming barriers to institutional investment in residential property. Greater London Authority, Economics Working Paper, No. 29.

Demers, F., 2005. Modelling and forecasting housing investment: The case of Canada. Bank of Canada, Ottawa, Working Paper 2005-41, December 2005.

Dua, P., Nishita, R., and Sahoo, S., 2008. Forecasting interest rates in India. *The Journal of Applied Economic Research* 2, 1–41.

Dunson, D. B., Herring, A. H., and Engel, S. M., 2008. Bayesian selection and clustering of polymorphisms in functionally related genes. *Journal of the American Statistical Association* 103, 534–546.

Dynan, K., Elmendorf, D., and Sichel, D., 2006. Can financial innovation help to explain the reduced volatility of economic activity? *Journal of Monetary Economics* 53, 123–150.

Edge, R. M., 2000. The effect of monetary policy on residential and structures investment under differential project planning and completion times. International Finance Discussion Paper No. 617, Board of Governors of the Federal Reserve System, Washington, DC, US.

Egebo, T., Richardson, P., and Lienert, I., 1990. A model of housing investment for the major OECD economies. *OECD Economic Studies* 14, 151–188.

Evans, M., Hastings, N., and Peacock, B. 2000. Beta Distribution. Ch. 5 in *Statistical Distributions*, 3rd ed. New York: Wiley, pp. 34-42.

Ferguson, T. S. 1973. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2), 209-230.

Fisher, J. D. M., and Gervais, M., 2007. First-time home buyers and residential investment volatility. Federal Reserve Bank of Chicago, 2007. Working paper 2007-15.

Green, R. K., 1997. Follow the leader: How changes in residential and non-residential investment predict changes in GDP. *Real Estate Economics* 25, 253–270.

Harvey, D. I., Leybourne, S. J., and Newbold, P., 1998. Tests for forecast encompassing. *Journal of Business and Economic Statistics* 16, 254–259.

King, R. G., and Watson, M. W., 2012. Inflation and labour costs. Unpublished manuscript. http://www.bu.edu/econ/files/2010/11/King_Watson_GERZ25_Jan26_201211.pdf.

Kohn, D. L., 2009. The economic outlook. Speech at the National Association for Business Economics, St. Louis, Missouri. http://www.federalreserve.gov/newsevents/speech/kohn20091013a.htm.

Koop, G., and Potter, S., 2004. Forecasting in dynamic factor models using Bayesian model averaging. *Econometrics Journal* 7, 550–565.

Koop, G., and Korobilis, D., 2011. UK macroeconomic forecasting with many predictors: Which models forecast best and when do they do so? *Economic Modelling* 28, 2307-2318.

Korobilis, B., 2013a. Bayesian forecasting with highly correlated predictors. *Economics Letters* 118, 148–150.

Korobilis, B., 2013b. Hierarchical shrinkage priors for dynamic regressions with many predictors. *International Journal of Forecasting* 29, 43–59.

Krainer, J., 2006. Residential investment over the real estate cycle. Federal Reserve Bank of San Francisco *Economic Letter*, Number 2006-15, June 30.

Leamer, E. E., 2007. Housing is the business cycle. NBER, Washington, DC, USA, Working Paper 13428, September.

Lunsford, K. G., 2013. A model for forecasting residential investment. University of Wisconsin, Madison, Manuscript, July 2013. https://mywebspace.wisc.edu/klunsford/web/research/Lunsford%20(2013)%20-%20Forecasting%20Residential%20Investment.pdf

MacLehose, R. F., Dunson, D. B., Herring, A. H., and Hoppin, J. A., 2007. Bayesian methods for highly correlated exposure data. *Epidemiology* 18, 199–207.

McCarthy, J., and Peach, R. W., 2002. Monetary policy transmission to residential investment. *Economic Policy Review* May, 139–158.

McCracken, M. W., 2007. .Asymptotics for out-of-sample tests of Granger causality. *Journal of Econometrics* 140, 719–752.

Mitchell, T. J., and Beauchamp, J. J., 1988. Bayesian variable selection in linear regression. *Journal of the American Statistical Association* 83, 1023–1032.

Peek, J., and Wilcox, J., 2006. Housing, credit constraints, and macro stability: The secondary mortgage market and reduced cyclicality of residential investment. *American Economic Review* 96, 135–140.

SAS Institute, 2014. Gamma and inverse-gamma distributions. SAS/STAT(R) 12.1 User's Guide. SAS Institute Inc.

Stock, J., and Watson, M., 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20, 147–162.

Shotwell, M.S. and Slatey, E.H. (2011) Bayesian outlier detection with Dirichlet Process mixtures. *Bayesian Analysis*, 6(4), 665-690.

Yellen, J. L., 2013. A painfully slow recovery for America's workers: Causes, implications, and the Federal Reserve's response. Speech at the "A Trans-Atlantic Agenda for Shared Prosperity" conference sponsored by the AFL-CIO, Friedrich Ebert Stiftung, and the IMK Macroeconomic Policy Institute, Washington, D.C, USA.

Table 1:Forecast Evaluation using Theil's U and MSE-F Statistics (1983:Q1 -2011:Q2)

	h=1	**h=2**	**h=4**	**h=8**	**Average**
Small and Large Models:					
SSVS-Small	0.903**	0.904**	0.883**	0.783**	0.868
SSVS-Large	0.882**	**0.862****	0.869**	0.772**	**0.846**
BSS-Small	0.936**	0.975**	0.922**	0.810**	0.911
BSS-Large	0.868**	0.913**	0.965**	0.782**	0.882
FAPR-Small	0.960**	0.997	1.015	1.037	1.002
FAPR-Large	**0.861****	0.932**	1.011	1.140	0.986
Individual Predictive Regressions:					
LFPR (1)	1.054	1.019	1.001	1.008	1.020
UNRATE (1)	1.036	1.065	1.076	1.027	1.051
HOUST (3)	1.033	1.096	1.050	1.005	1.046
PERMIT (3)	0.889**	0.997	1.015	0.995†	0.974
RSALES (3)	1.006	1.005	1.008	1.000	1.005
CPI (4)	1.178	1.111	1.094	1.042	1.106
MORTG (1)	0.951**	0.923**	0.963**	0.991*	0.957
3TB 12)	1.030	1.028	0.999	0.962**	1.005
M1 (3)	1.019	1.013	1.062	1.072	1.041
S&P (3)	1.039	1.025	1.018	1.003	1.021
RGCON (3)	0.999	0.998	1.007	1.004	1.002
RGDP (3)	1.014	1.007	1.000	1.008	1.007
RPCON (3)	0.989*	1.004	1.004	1.005	1.001
RNRFINV (3)	1.041	1.096	1.142	1.105	1.096
RCPINV (3)	1.006	1.003	1.000	1.001	1.003
RHP (3)	0.992*	0.992*	1.009	1.045	1.009
BCON (3)	1.019	1.025	1.034	1.014	1.023
H4SALE (2)	0.958**	0.917**	**0.843****	**0.712****	0.858
HSOLD (2)	0.960**	0.957**	0.910**	0.954**	0.945
HSUPPLY (3)	0.994†	0.993†	0.951**	0.950**	0.972
Comparing Two Best Models:					
SSVS-Large vs H4SALE	0.920**	0.940**	1.031	1.083	0.993

Note: Relative *MSE* is the ratio of the root mean square for the out-of-sample forecast of the restricted (AR) model to the MSE for the out-of-sample forecast of the unrestricted model otherwise known as Theil's *U* statistic. The bold numbers equal the minimum *U* values in each column. The average column computes the average relative *MSE* of the one-, two-, four-, and eight-quarter-ahead *MSE* reported in columns headed by *h*=1, *h*=2, *h*=4 and *h*=8. Numbers in parentheses after a variable identifies the transformations of the variables to induce stationarity as follows: 1, first difference, $x_{i,t} = z_{i,t} - z_{i,t-1}$; 2, logarithm, $x_{i,t} = \ln z_{i,t}$; 3, first difference of logarithm, $x_{i,t} = \ln(z_{i,t}/z_{i,t-1})$; and 4, second difference of logarithm, $x_{i,t} = \ln(z_{i,t}/z_{i,t-1}) - \ln(z_{i,t-1}/z_{i,t-2})$.

†, *, ** respectively indicates 10%, 5% and 1% level of significance for the *MSE-F* test.

Figure 1: Recessions, Expansions, and Residential Fixed Investment to GDP

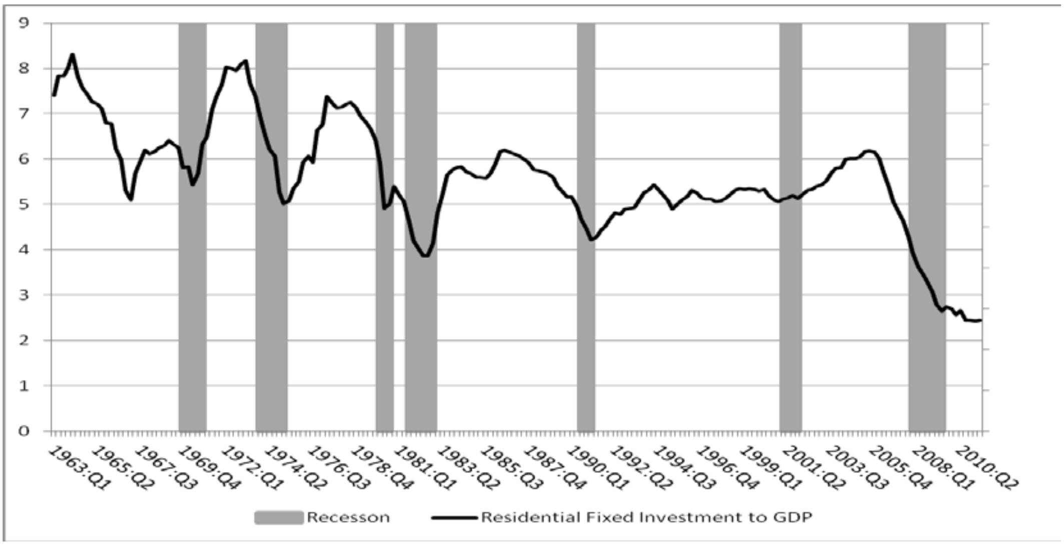


Figure 2: Growth Rates of Real GDP and Real Private Residential Investment

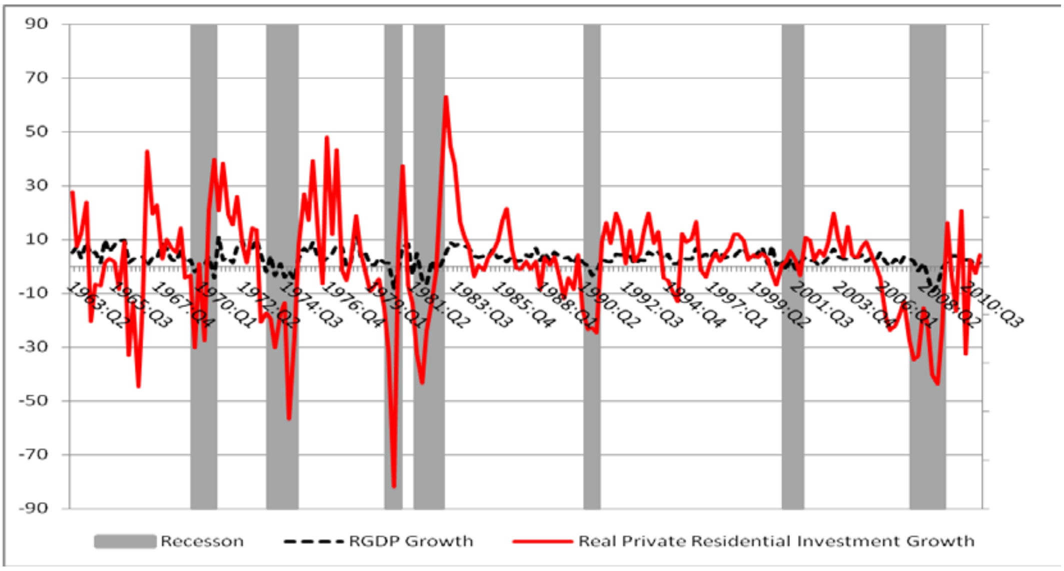
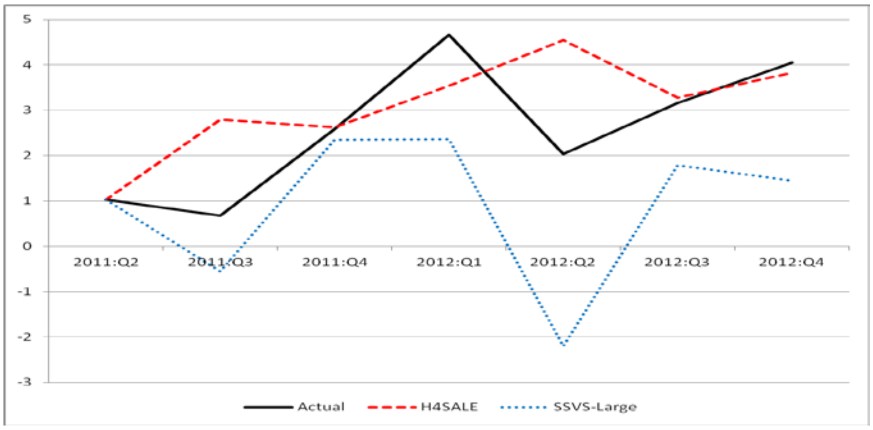


Figure 3: Actual and *Ex-Ante* Forecast



Appendix: Description of Variables

No	Mnemonic	Long Description	Tcode
1	INDPRO	IP: Total index	5
2	IPFINAL	Industrial Production: Final Products (Market Group)	5
3	IPCONGD	IP: Consumer goods	5
4	IPMAT	Industrial Production: Materials	5
5	IPDMAT	Industrial Production: Durable Materials	5
6	IPNMAT	Industrial Production: nondurable Materials	5
7	MCUMFN	Capacity utilization: Manufacturing	1
8	IPDCONGD	Industrial Production: Durable Consumer Goods	5
9	IP.B51110.S	IP: Automotive products	5
10	IPNCONGD	Industrial Production: Nondurable Consumer Goods	5
11	IPBUSEQ	Industrial Production: Business Equipment	5
12	IP.B51220.S	IP: Consumer Energy Products	5
13	MANEMP	All Employees: Manufacturing	5
14	PAYEMS	Total Nonfarm Payrolls: All Employees	5
15	SRVPRD	All Employees: Service-Providing Industries	5
16	USGOOD	All Employees: Goods-Producing Industries	5
17	USGOVT	All Employees: Government	5
18	USPRIV	All Employees: Total Private Industries	5
19	CES9091000001	Federal	5
20	CES9092000001	State government	5
21	CES9093000001	Local government	5
22	DMANEMP	All Employees: Durable Goods Manufacturing	5
23	NDMANEMP	All Employees: Nondurable Goods Manufacturing	5
24	USCONS	All Employees: Construction	5
25	USEHS	All Employees: Education & Health Services	5
26	USFIRE	All Employees: Financial Activities	5
27	USINFO	All Employees: Information Services	5
28	USLAH	All Employees: Leisure & Hospitality	5
29	USMINE	All Employees: Natural Resources & Mining	5
30	USPBS	All Employees: Professional & Business Services	5
31	USSERV	All Employees: Other Services	5
32	USTPU	All Employees: Trade, Transportation & Utilities	5
33	USTRADE	All Employees: Retail Trade	5
34	USWTRADE	All Employees: Wholesale Trade	5
35	CE160V	Emp Total (Household Survey)	5
36	CLF16OV	Civilian Labor Force	5
37	***LNS11300000***	***LaborForce Participation Rate (16 Over) SA***	2
38	***UNRATE***	***Unemplomment Rate***	2
39	URATE_ST	Urate Short Term (< 27 weeks)	2
40	URATE_LT	Urate Long Term (>= 27 weeks)	2
41	LNS14000012	Unemployment Rate - 16-19 yrs	2
42	LNS14000025	Unemployment Rate - 20 yrs. & over, Men	2
43	LNS14000026	Unemployment Rate - 20 yrs. & over, Women	2
44	UEMPLT5	Number Unemployed for Less than 5 Weeks	5
45	UEMP5TO14	Number Unemployed for 5-14 Weeks	5
46	UEMP15T26	Civilians Unemployed for 15-26 Weeks	5
47	UEMP27OV	Number Unemployed for 27 Weeks & over	5
48	LNS12032194	Employment Level - Part-Time for Economic Reasons, All Industries	5
49	AWHMAN	Average Weekly Hours: Manufacturing	1
50	AWOTMAN	Average Weekly Hours: Overtime: Manufacturing	2
51	A0M046	Index of Help-Wanted Advertising in Newspapers	1
52	***HOUST***	***Housing Starts: Total: New Privately Owned Housing Units Started***	5
53	HOUST5F	Privately Owned Housing Starts: 5-Unit Structures or More	5
54	HOUSTMW	Housing Starts in Midwest Census Region	5
55	HOUSTNE	Housing Starts in Northeast Census Region	5
56	HOUSTS	Housing Starts in South Census Region	5
57	HOUSTW	Housing Starts in West Census Region	5
58	***PERMIT***	***New Private Housing Units Authorized by Building Permit***	5
59	A0M007	Mfrs' new orders durable goods industries (bil. chain 2000 $)	5
60	A0M008	Mfrs' new orders, consumer goods and materials (mil. 1982 $)	5
61	A1M092	Mfrs' unfilled orders durable goods indus. (bil. chain 2000 $)	5
62	A0M032	Index of supplier deliveries -- vendor performance (pct.)	1
63	A0M027	Mfrs' new orders, nondefense capital goods (mil. 1982 $)	5
64	A0M070	Manufacturing and trade inventories (bil. Chain 2005 $)	5
65	A0M057	Manufacturing and trade sales (mil. Chain 2005 $)	5
66	***A0M059***	***Sales of retail stores (mil. Chain 2000 $)***	5
67	PPIACO	Producer Price Index: All Commodities	6
68	WPU0561	PPI: Crude Petroleum	5
69	PPIFGS	Producer Price Index: Finished Goods	6
70	PPIFCF	Producer Price Index: Finished Consumer Foods	6

71	PPIFCG	Producer Price Index: Finished Consumer Goods	6
72	PPIIDC	Producer Price Index: Industrial Commodities	6
73	PPIITM	Producer Price Index: Intermediate Materials: Supplies & Components	6
74	PSCCOM	Spot Market Price Index:Bls & Crb: All Commodities(1967=100)	5
75	PMCP	NAPM Commodity Prices Index (Percent)	1
76	***CPIAUCSL***	***Consumer Price Index For All Urban Consumers: All Items***	6
77	CPILFESL	Consumer Price Index for All Urban: All Items Less Food & Energy	6
78	CES2000000008	Average Hourly Earnings: Construction	5
79	CES3000000008	Average Hourly Earnings: Manufacturing	5
80	AHETPI	Average Hourly Earnings: Total Private Industries	5
81	AAA	Moody's Seasoned Aaa Corporate Bond Yield	2
82	BAA	Moody's Seasoned Baa Corporate Bond Yield	2
83	FEDFUNDS	Effective Federal Funds Rate	2
84	CPF3M	3-Month AA Financial Commercial Paper Rate	2
85	CP90_TBILL	CP3FM-TB3MS	1
86	GS1	1-Year Treasury Constant Maturity Rate	2
87	GS10	10-Year Treasury Constant Maturity Rate	2
88	***MORTG***	***30-Year Conventional Mortgage Rate***	2
89	***TB3MS***	***3-Month Treasury Bill: Secondary Market Rate***	2
90	TB6MS	6-Month Treasury Bill: Secondary Market Rate	2
91	MED3	3-Month Eurodollar Deposit Rate (London)	2
92	MED3_TB3M	MED3-TB3MS (Version of TED Spread)	1
93	AAA_GS10	AAA-GS10 Spread	1
94	BAA_GS10	BAA-GS10 Spread	1
95	MRTG_GS10	Mortg-GS10 Spread	1
96	TB6M_TB3M	tb6m-tb3m	1
97	GS1_TB3M	GS1_Tb3m	1
98	GS10_TB3M	GS10_Tb3m	1
99	BOGAMBSL	Board of Governors Monetary Base	5
100	BOGNONBR	Non-Borrowed Reserves of Depository Institutions	5
101	BUSLOANS	Commercial and Industrial Loans at All Commercial Banks	5
102	CONSUMER	Consumer (Individual) Loans at All Commercial Banks	5
103	IMFSL	Institutional Money Funds	5
104	***M1SL***	***M1 Money Stock***	5
105	M2SL	M2SL	5
106	MZMSL	MZM Money Stock	5
107	NONBORTAF	Non-Borrowed Reserves of Dep. Institutions + Term Auction Credit	5
108	NONREVSL	Total Nonrevolving Credit Outstanding	5
109	REALLN	Real Estate Loans at All Commercial Banks	5
110	TRARR	Board of Governors Total Reserves	5
111	TOTALSL	Total Consumer Credit Outstanding	5
112	***FSPCOM***	***S&P's Common Stock Price Index: Composite (1941-43=10)***	5
113	FSDJ	Common Stock Prices: Dow Jones Industrial Average	5
114	MVOL	VXO/ VIX Index	1
115	TWEXMMTH	FRB Nominal Major Currencies Dollar Index	5
116	EXSZUS	Foreign Exchange Rate: Switzerland (Swiss Franc Per U.S.$)	5
117	EXJPUS	Foreign Exchange Rate: Japan (Yen Per U.S.$)	5
118	EXUSUK	Foreign Exchange Rate: United Kingdom (Cents Per Pound)	5
119	EXCAUS	Foreign Exchange Rate: Canada (Canadian $ Per U.S.$)	5
120	U0M083	Consumer expectations NSA (Copyright, University of Michigan)	1
121	DPIC96	Real Disposable Personal Income	5
122	FPIC96	Real Private Fixed Investment, 3 Decimal	5
123	***GCEC96***	***Real Government Consumption Expenditures & Gross Investment***	5
124	***GDPC96***	***Real Gross Domestic Product, 3 Decimal***	5
125	GPDIC96	Real Gross Private Domestic Investment, 3 Decimal	5
126	***PCECC96***	***Real Personal Consumption Expenditures***	5
127	NRIPDC96	Real Nonresidential Investment: Equipment & Software, 3 Decimal	5
128	EXPGSC96	Real Exports of Goods & Services, 3 Decimal	5
129	GRECPT	Government Current Receipts (Nominal)	5
130	FGCEC96	Real Federal Consumption Expenditures & Gross Investment	5
131	IMPGSC96	Real Imports of Goods & Services, 3 Decimal	5
132	PCDGCC96	Real Personal Consumption Expenditures: Durable Goods	5
133	PCESVC96	Real Personal Consumption Expenditures: Services	5
134	PCNDGC96	Real Personal Consumption Expenditures: Nondurable Goods	5
135	***PNFIC96***	***Real Private Nonresidential Fixed Investment, 3 Decimal***	5
136	PRFIC96	Real Private Residential Fixed Investment, 3 Decimal	5
137	SLCEC96	Real State & Local Consumption Expenditures & Gross Investment	5
138	***CBIC96***	***Real Change in Private Inventories, 3 Decimal***	5
139	CBIC96_GDP	Ch. Inv/GDP	1
140	OUTBS	Business Sector: Output	5
141	OUTNFB	Nonfarm Business Sector: Output	5
142	HOABS	Business Sector: Hours of All Persons	5
143	HOANBS	Nonfarm Business Sector: Hours of All Persons	5
144	PRS85006013	Nonfarm Business Sector: Employment	5

145	PCEPILFE	Personal Consumption Expenditures: Chain-type Less Food & Energy	6
146	PCECTPI	Personal Consumption Expenditures: Chain-type Price Index	6
147	PCED_G	Goods	6
148	PCED_DG	Durable goods	6
149	PCED_NDG	Nondurable goods	6
150	PCED_S	Services	6
151	PCED_SC	Household consumption expenditures (for services)	6
152	PCED_MV	Motor vehicles and parts	6
153	PCED_DHE	Furnishings and durable household equipment	6
154	PCED_REC	Recreational goods and vehicles	6
155	PCED_ODG	Other durable goods	6
156	PCED_FB	Food and beverages purchased for off-premises consumption	6
157	PCED_APP	Clothing and footwear	6
158	PCED_GAS	Gasoline and other energy goods	6
159	PCED_ONG	Other nondurable goods	6
160	PCED_HU	Housing and utilities	6
161	PCED_HC	Health care	6
162	PCED_TRA	Transportation services	6
163	PCED_RECS	Recreation services	6
164	PCED_FS	Food services and accommodations	6
165	PCED_INS	Financial services and insurance	6
166	PCED_OS	Other services	6
167	GDPCTPI	Gross Domestic Product: Chain-type Price Index	6
168	GPDICTPI	Gross Private Domestic Investment: Chain-type Price Index	6
169	IPDBS	Business Sector: Implicit Price Deflator	6
170	COMPRNFB	Nonfarm Business Sector: Real Compensation Per Hour	5
171	RCPHBS	Business Sector: Real Compensation Per Hour	5
172	OPHNFB	Nonfarm Business Sector: Output Per Hour of All Persons	5
173	OPHPBS	Business Sector: Output Per Hour of All Persons	5
174	ULCBS	Business Sector: Unit Labor Cost	5
175	ULCNFB	Nonfarm Business Sector: Unit Labor Cost	5
176	UNLPNBS	Nonfarm Business Sector: Unit Nonlabor Payments	5
177	TTABSHNO	Total Tangible Assets - Balance Sheet of Households and Nonprofits (FoF)	5
178	TNWBSHNO	Total Net Worth - Balance Sheet of Households and Nonprofits (FoF)	5
179	NWORTH_PDI	Networth Relative to Personal Disp Income	1
180	TTABSHNO	TTABSHNO-REANSHNO	5
181	REABSHNO	Real Estate - Assets - Balance Sheet of Households and Nonprofit Orgs	5
182	TFAABSHNO	Total Financial Assets - Balance Sheet of Households and Non Profits	5
183	TLBSHNO	Total Liabilities - Balance Sheet of Households and Nonprofits (FoF)	5
184	LIAB_PDI	Liabilities Relative to Person Disp Income	5
185	***RHPI***	***Real new home price index***	5
186	***BCUSAM***	***Business confidence index***	4
187	***H4SALE***	***Number of new housing units for sale***	4
188	***HSOLD***	***Number of new housing units sold***	5
189	***HSUPPLY***	***Month's supply of housing ratio***	5

Note: Variables in bold-italics are those used as predictors in the small scale and individual regression models. All variables are transformed to be approximately stationary. In particular if $z_{i,t}$ is the original untransformed series, the transformation codes are (column Tcode above): 1 – no transformation – first difference, $x_{i,t} = z_{i,t} - z_{i,t-1}$; 4- logarithm, $x_{i,t} = \ln z_{i,t}$; 5 – first difference of logarithm, $x_{i,t} = \ln(z_{i,t}/z_{i,t-1})$; 6 – second difference of logarithm, $x_{i,t} = \ln(z_{i,t}/z_{i,t-1}) - \ln(z_{i,t-1}/z_{i,t-2})$.