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

Goodness C. Aye · Stephen M. Miller · Rangan Gupta · Mehmet Balcilar

Abstract This paper employs classical bivariate, slab-and-spike variable selection, Bayesian semi-parametric shrinkage, and factor augmented predictive regression models to forecast US real private residential fixed investment over an out-of-sample period from 1983Q1 to 2005Q4, based on in-sample estimates for 1963Q1 to 1982Q4. Both large-scale (188 macroeconomic series) and small-scale (20 macroeconomic series) slab-and-spike variable selection, and Bayesian semi-parametric shrinkage, and factor augmented 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) mean-square-error *F* statistic. We find that, on average, the slab-and-spike variable selection and Bayesian semi-parametric shrinkage models with 188 variables provides the best forecasts amongst all the models. Finally, we use these two models to predict the relevant turning points of the residential investment, via an *ex-ante* forecast exercise from 2006Q1 to 2012Q4. The 188 variable slab-and-spike variable selection and Bayesian semi-parametric shrinkage models perform quite similarly in their accuracy of forecasting the turning points. 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

*We would like to thank two anonymous referees and an Associate Editor for many helpful comments. However, any remaining errors are solely ours.

G.C. Aye
Department of Economics, University of Pretoria, Pretoria, 0002, South Africa

S. M. Miller (✉)
Department of Economics, University of Nevada, Las Vegas, Las Vegas, NV 89154, USA
email: styephen.miller@unlv.edu

R. Gupta
Department of Economics, University of Pretoria, Pretoria, 0002, South Africa

M. Balcilar
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, slab-and-spike variable selection (*SSVS*), Bayesian semi-parametric shrinkage (*BSS*), and factor-augmented predictive regression (*FAPR*) 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 acts as 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 historically leads US business cycles, proving useful in forecasting GDP from 1959 to 1992. Fig. 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 Fig. 2). Although residential investment historically contributes only about 5% 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) 2009Q2 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 (Fig.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 based on a pseudo real-time analysis. 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 and leading-indicator 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.

[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).

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 as flows out of construction. He uses the root mean square prediction error to compare the forecasting performance of the forecasts of the Survey of Professional Forecasters (*SPF*), the Federal Reserve's forecasts in the Green Book (*GB*), and several small-scale time-series models over two volatile periods (i.e., 1981Q3-1983Q4 and 2006Q1-2013Q2) and a more tranquil period (1984Q1-2005Q4). Lunsford's error-correction model does better during the more volatile periods. During the tranquil period, the SPF and GB forecasts do marginally better than Lunsford's error-correction model.

Our paper compares the forecasting performance of small-scale (i.e., bivariate models with specific housing variables as the additional explanatory variable), medium-scale (i.e., *SSVS*-, *BSS*-, and *FAPR*-small models that include the 20 variables used in the bivariate models) and large-scale (i.e., the *SSVS*-, *BSS*-, and *FAPR-Large* models that include 188 macroeconomic variables, including the specific housing variables used in the bivariate models) atheoretical time-series models. Intuitively, our large-scale models use the same information set, at least in theory, as the *SPF* and *GB* forecasts.

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. We contribute new evidence with alternative methods for forecasting, primarily based on Bayesian estimation which can handle models with many potential predictors for residential investment.

3 Methodology

We consider several predictive regression models for forecasting the US real private residential fixed investment. These include the *SSVS*, the *BSS*, and the *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 SSVS Model

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

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

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 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

[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 bill rate, *3TB*), real gross domestic product (*RGDP*), the consumer price index (*CPI*), the unemployment rate (*UNRATE*), the labour force participation rate (*LFPR*), the 30-Year conventional 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*).

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 β_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 BSS 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:

$$\beta_j \sim \pi\delta_0(\beta) + (1-\pi)G \tag{3}$$

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

[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.

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 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 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:

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

[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 2012). 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 2012). 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).

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

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 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 1963Q1 to 2012Q4. 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 1959Q to 2011Q2. 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 1963Q1. Hence, our total sample covers 1963Q1 to 2012Q4, with the data vintage corresponding to the one used by King and Watson (2012), with the exception of the real private residential investment, the data vintage for which corresponds to 2012Q4, which became available in early 2013. This is because, besides the pseudo real-time forecasting analysis conducted over 1983Q1-2005Q4, we also carry out an *ex ante* turning point forecast analysis over 2006Q1-2012Q4. We present the annualized quarterly growth rates of real GDP (right axis) and real private residential investment (left axis) in Fig. 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

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

investment occur in 1978Q2 (15.4%) and 1983Q1 (62.9%), respectively. The lowest growth rates occur in 2008Q4 (-9.3%) and 1980Q2 (-81.9%), 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.[7] 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, which is suited for nested models, since the models with a predictor or multiple predictors always nests the *AR(1)* model.[8] The *MSE-F* statistic tests the null hypothesis that a specific

[7] Based on the suggestions of an anonymous referee, we also considered the performance of the double-differencing device (*DDD*) model proposed by Hendry (2009), as a possible benchmark. Since the *AR(1)* model consistently outperformed the *DDD* model at all horizons over the out-of-sample, however, we still chose the former as our benchmark. Complete details comparing the forecasting results of the *AR(1)* and *DDD* models are available from the authors.

[8] 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 indicates that the unrestricted model forecasts are statistically more accurate than those of the restricted model. Note, however, that the *MSE-F* statistic exhibits a non-standard and non-pivotal limiting distribution in the case of nested models and h>1. Hence, we base our inferences on the bootstrap procedure described in detail in (Rapach et al. 2005; Rapach and Wohar 2006). These two papers provide further details.

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.[9] 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.

<u>4.2.1 *Ex-post* Out-of-Sample Forecasts.</u> The data sample runs from 1963Q1 to 2012Q4. We use 80 out of our 200 total observations for first period forecast. This implies that we estimate each model over the in-sample period of 1963Q3 to 1982Q4 (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 1983Q1 to 2005Q4. That is, we use the last 92 observations (i.e., 1983Q1 to 2005Q4) for the evaluation of *h*-step-ahead forecasts (*ex-post* out-of-sample forecasting), stopping just before the beginning of the "Great Recession". 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 calculate the mean square errors (*MSE*)

[9] We recursively update the *ex-post* forecasts in-sample in the forecasting equation to generate the multi-step-ahead forecasts, whereas we produce the *ex-ante* multi-step-ahead forecasts from a specific point in time (generally, from the end-point of data available on the predictors, which in our case is 2006:Q1-2012: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.

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 2006Q1 to 2012Q4 – a period that includes the "Great Recession" and the slow recovery that followed.

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 0.842 entry for the *BSS-Large* model for the one-quarter-ahead forecast, means that the *BSS-Large* model experienced a forecast *MSE* of only 84.2%of the *AR* model's forecast *MSE*. In other words, the *BSS-Large* model improves over the *AR* model by 15.8%. 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% or 5% levels at all four horizons. Second, the *FAPR-Large* model produces a statistically significant more accurate forecast than the *AR* model only at the four-quarter-ahead horizon while the *FAPR-Small* model does not outperform the *AR*-benchmark model at any horizon. Third, the large-scale Bayesian models generally perform better than the small-scale Bayesian models at each horizon as well as based on the overall average. The exception is the *SSVS-Small* model at the eight-quarter-ahead horizon.

Now, consider the bivariate models in the bottom part of Table 1. Four (4) out of the 20 individual classical regression models (namely *LFPR*, *MORTG*, *3TB*, and *RGDP*) 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% or 5% levels. Alternatively, this means that, on average, the *AR* model forecasts generally prove more accurate, or at least as good as, the remaining 16 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, models with more information outperform models with less information. Specifically, the *BSS-Large* model performs better than all other models at horizon one, improving over the *AR* model by 15.8%. At horizons two and four, the *SSVS-Large* model performs the best, improving over the *AR* model by 19.8 and 20.6%, respectively. The *BSS-Large* model, again, performs the best at horizon eight, improving over *AR* model by 38.2%. We observe that the *BSS-Large* model produces the most accurate forecast based on the overall average forecast *MSE,* experiencing a forecast *MSE* of only 78.7% of the *AR* model forecast *MSE*. In other words, it improves over the *AR*-benchmark model by 21.3%.

Given the overall performance of the *BSS-Large* model, we also compare the relative

MSE of this model to all other models. Table 2 reports the findings. We find that the *MSE* of the *BSS-Large* model significantly improves the relative *MSE* for all models at all horizons, usually at the 1% level. The one exception is that the *BSS-Large* model does not significantly improve the relative *MSE* of the *SSVS-Large* model at the one-quarter-ahead and eight-quarters-ahead horizons. In addition, for the second- and four-quarters-ahead horizons where the *SSVS-Large* outperforms the *BSS-Large*, the differences between the forecast errors are also not statistically significant.[10] Note that since, the *BSS-Large* and the *SSVS-Large* are not nested models, we compare the statistical differences between the MSEs of these two models using the modified Diebold and Mariano (1995) test proposed by Harvey et al., (1997).

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 3-month Treasury bill rate (*3TB*), the 30-Year conventional mortgage rate (*MORTG*), the labour-force participation rate (*LFPR*), and real GDP (*RGDP*) exhibit significantly better forecasts than the AR model across nearly all horizons. These four variables indirectly relate to the housing market and real private residential investment. Finally, the better performances of the Bayesian large and small models noted in the previous paragraph also utilize these four variables in their

[10] As indicated earlier, the structure of the macroeconomic data commonly used by macroeconomists frequently involves highly correlated variables, which the SSVS model does not account for, but is incorporated in the BSS modeling approach. While, on average, we observe the gains from using the large-scale BSS approach over the corresponding SSVS approach, in our case, these two models produce statistically similar accuracy in forecasting the real private fixed residential investment.

estimated models.

Fig. 3 provides the one-quarter-ahead forecasts of the *SSVS-Larg*e and *BSS-Large* models as well as the actual growth rate of real private residential investment. The *SSVS-Large* and *BSS-Large* forecasts track each other almost exactly. Visually, the two *ex ante* forecasts appear to coincide. Moreover, these one-quarter-ahead forecasts move largely in synchronization with the actual series. The *BSS-Large* (*SSVS-Large*) forecasts move in the same direction as the actual series in 74 (71) out of 92 quarters in the out-of-sample period 1983Q1 to 2005Q4. In addition, 53 turning points in the actual series occur during this out-of-sample period. The *BSS-Large* forecasts also turned in the same direction as the actual series 42 times. In the other 11 cases, the *BSS-Large* forecasts also turned in the same direction as the actual series, but one quarter in advance of the actual series turn. In one of these 11 cases, the *BSS-Large* forecasts turned in advance of the actual series in two consecutive quarters. Those 12 cases (i.e., 11 plus one) all coincide with 12 of the 18 instances in which the *BSS-Large* forecasts moved in the opposite direction as the actual series. The remaining 6 quarters saw the *BSS-Large* forecasts not predicting the next quarters movement in the actual series. The *SSVS-Large* forecasts exhibit a similar pattern of movement to the *BSS-Large* forecasts, except for the three instances where the *SSVS-Large* forecasts move in the opposite direction as the actual series when compared to the *BSS-Large* forecasts. Twice the movement in the *SSVS-Large* forecasts did not forecast the next quarters movement in the actual series and once it did.[11]

4.2.2 *Ex-Ante* Forecasts. Having determined each of the optimal forecast models from the multivariate *SSVS-Large* and *BSS-Large* models, we expose them to the acid-test of

[11] An earlier version of the paper considered an extended out-of-sample horizon covering 1983:Q1-2011:Q2. Based on this out-of-sample period, the *SSVS-Large mo*del performed the best, on average, followed by the bivariate predictive regression model consisting of *H4SALE*. When we used these two models to compare the turning points over the out-of-sample period, we found that the forecast from the *H4SALE* model was more volatile relative to the *SSVS-Large* model, and in general, the *SSVS-Large* model tracked the turning points reasonably well, except during the recent crisis.

predicting the different turning points in the US private residential investment series. We implement this by performing out-of-sample *ex-ante* forecast over 2006Q1 to 2012Q4.

Fig. 4 plots the *ex-ante* out-of-sample forecasts and actual values. As noted for the one-quarter-ahead forecasts in the prior subsection, the *SSVS-Large* and *BSS-Large* forecasts track each other almost exactly. Visually, once again, the two *ex ante* forecasts appear to coincide. The actual series exhibits 15 separate turning points over this sample. The *ex ante* forecasts experience turning points exactly with the actual series in 10 of the 15 cases. For the other five turning points, the *ex ante* forecasts turn one-quarter ahead of the actual series. Viewed differently, the *ex ante* forecasts move in the same direction as the actual series in 23 out of the 28 quarters in our sample, where the other five quarters see the *ex ante* forecasts turning one quarter ahead of the actual series. From 2006Q1 through 2009Q4, the *ex ante* forecasts exceed the actual series. Then, from 2010Q1 through 2011Q3, the *ex ante* forecasts exhibit more volatility than the actual series, exceeding (falling below) the actual series for positive (negative) changes in the actual series on nearly a quarter-by-quarter alternating basis. Finally, the actual series exceeds the *ex* ante forecasts from 2011Q3 through 2012Q4.

5 Conclusion

In this paper, we forecast the US real private residential investment using quarterly data from 1963Q1 to 2012Q4. We consider 3 large-scale, 3 small-scale, and 20 individual predictive regression models. Using the period of 1963Q1 to 1982Q4 as the in-sample period and 1983Q1 to 2005Q4 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 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, the BSS and SSVS large- and small-scale models and four 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* and *BSS-Large* models outperform the rest of the models at all horizons based on the average *MSE* across all forecast horizons.

Using the *SSVS-Large* and *BSS-Large* models, we provide *ex-ante* forecasts for real private residential investment over 2006Q1 to 2012Q4. Interestingly, the results clearly show that the *SSVS-Large* and *BSS-Large* models capture the turning points in the actual real private residential investment series, leading by one quarter in five of the 15 turning points observed in the sample.

In sum, several bivariate models outperform our AR-benchmark model. These better bivariate models generally include variables that indirectly affect the housing market. 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.

References

Baghestani H (2011) Federal Reserve and private forecasts of growth in Investment. J Econ Bus 63: 290–305

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

Berger-Thomas L, Ellis L (2004) Housing construction cycles and interest rates. Research Discussion Paper 2004-08, Reserve Bank of Australia. http://www.rba.gov.au/publications/rdp/2004/2004-08.html. Accessed 25 March 2015

Bernanke BS (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. Accessed 20 March 2013

Bernanke BS (2012) The economic recovery and economic policy. Speech at the New York Economic Club, New York, New York. http://www.federalreserve.gov/newsevents/speech/bernanke20121120a.htm. Accessed 5 April 2013

Brayton F, Tinsley P (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 Accessed 20 March 2013

Browne LE (2000) National and regional housing patterns. New Engl Econ Rev 4: 32–57

Choy LHT, Ho WKO, Mak SWK (2011) Region-specific estimates of the determinants of residential investment. J Urban Plan D-ASCE 137: 1–6

Clements M., Hendry D (1998) Forecasting economic time series. Cambridge University Press, New York

Cremers M (2002) Stock return predictability: A Bayesian model selection perspective. Rev Financ Stud 15: 1223–1249

Daly JS (2008) Overcoming barriers to institutional investment in residential property. Greater London Authority, Economics Working Paper, No. 29. http://london.gov.uk/mayor/economic_unit/docs/wp_29.pdf. Accessed 20 April 2013

Demers F (2005) Modelling and forecasting housing investment: The case of Canada. Bank of Canada, Ottawa, Working Paper 2005-41, December 2005. http://www.bankofcanada.ca/wp-content/uploads/2010/02/wp05-41.pdf. Accessed 25 April 2013

Diebold FX, Mariano RS (1995) Comparing predictive accuracy. J Bus Econ Stat 13: 253-263

Dua P, Nishita R, Sahoo S (2008) Forecasting interest rates in India. Margin J Appl Econ Res 2: 1–41

Dunson DB, Herring AH, Engel SM (2008) Bayesian selection and clustering of polymorphisms in functionally related genes. J Am Stat Assoc 103: 534–546

Dynan K, Elmendorf D, Sichel D (2006) Can financial innovation help to explain the reduced volatility of economic activity? J Monetary Econ 53: 123–150

Edge RM (2000) The effect of monetary policy on residential and structures investment under differential project planning and completion times. International Finance Discussion Paper No. 671, Board of Governors of the Federal Reserve System, Washington, DC, US. http://www.federalreserve.gov/pubs/ifdp/2000/671/ifdp671.pdf. Accessed 28 March 2013

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

Evans M, Hastings N, Peacock B (2000) Beta distribution. In: Statistical Distributions, 3rd ed. Wiley, New York, pp. 34-42

Ferguson TS (1973) A Bayesian analysis of some nonparametric problems. Ann Stat 1(2): 209-230.

Fisher JDM., Gervais M (2007) First-time home buyers and residential investment volatility. Federal Reserve Bank of Chicago, 2007. Working paper 2007-15. https://www.chicagofed.org/publications/working-papers/2007/wp-15. Accessed 22 March 2013

Green RK (1997) Follow the leader: How changes in residential and non-residential investment predict changes in GDP. Real Estate Econ 25: 253–270

Harvey DI, Leybourne SJ, Newbold P (1997) Testing the equality of prediction mean squared errors. Int J Forecasting 13: 281–291

Hendry DF (2006) Robustifying forecasts from equilibrium-correction systems. J Econometrics 135, 399-426

King RG, Watson MW (2012) Inflation and labour costs. http://www.bu.edu/econ/files/2010/11/King_Watson_GERZ25_Jan26_201211.pdf. Accessed 10 January 2013

Kohn DL (2009) The economic outlook. Speech at the National Association for Business Economics, St. Louis, Missouri. http://www.federalreserve.gov/newsevents/speech/kohn20091013a.htm. Accessed 15 April 2013

Koop G, Potter S (2004) Forecasting in dynamic factor models using Bayesian model averaging. Economet J 7: 550–565

Koop G, Korobilis D (2011) UK macroeconomic forecasting with many predictors: Which models forecast best and when do they do so? Econ Model 28: 2307-2318

Korobilis B (2013a) Bayesian forecasting with highly correlated predictors. Econ Lett 118: 148–150

Korobilis B (2013b) Hierarchical shrinkage priors for dynamic regressions with many predictors. Int J 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. http://www.frbsf.org/economic-research/publications/economic-letter/2006/june/residential-investment-over-the-real-estate-cycle/. Accessed 28 March 2013.

Leamer EE (2007) Housing is the business cycle. Fed Bank KS, pp 149-232. https://www.kansascityfed.org/publicat/sympos/2007/pdf/leamer_0415.pdf. Accessed 30 March 2013

Lunsford KG (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. Accessed 2 August 2013

MacLehose RF, Dunson DB, Herring AH, Hoppin JA (2007). Bayesian methods for highly correlated exposure data. Epidemiology 18: 199–207

McCarthy J, Peach RW (2002) Monetary policy transmission to residential investment. FRBNY Economic Policy Review, May, 139–158

McCracken MW (2007) Asymptotics for out-of-sample tests of Granger causality. J Econometrics 140: 719–752

Mitchell TJ, Beauchamp JJ (1988) Bayesian variable selection in linear regression. J Am Stat Assoc 83: 1023–1032

Peek J, Wilcox J (2006) Housing, credit constraints, and macro stability: The secondary mortgage market and reduced cyclicality of residential investment. Am Econ Rev 96: 135–140

Rapach DE, Wohar ME, Rangvid J (2005) Macro variables and international stock return predictability. Int J Forecasting 21: 137–166

Rapach DE, Wohar ME (2006) In-sample vs. out-of-sample tests of stock return predictability in the context of data mining. J Empir Financ 13, 231–247

SAS Institute (2012) Gamma and inverse-gamma distributions. SAS/STAT(R) 12.1 User's Guide. SAS Institute Inc.

http://support.sas.com/documentation/cdl/en/statug/65328/PDF/default/statug.pdf. Accessed 12 April 2013

Stock J, Watson M (2002) Macroeconomic forecasting using diffusion indexes. J Bus Econ Stat 20: 147–162

Shotwell MS, Slatey EH (2011) Bayesian outlier detection with Dirichlet Process mixtures. Bayesian Analysis 6(4): 665-690

Yellen JL (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, Washington, D.C, USA, 11 February, 2013. http://www.federalreserve.gov/newsevents/speech/yellen20130211a.pdf. Accessed 15 April 2013

Table 1 Forecast evaluation using Theil's U and MSE-F statistics (1983Q1 -2005Q4)

	h=1	**h=2**	**h=4**	**h=8**	**Average**
Small and Large Models:					
SSVS-Small	0.925**	0.904**	0.972*	0.876**	0.919
SSVS-Large	0.871**	**0.802****	**0.794****	0.883**	0.837
BSS-Small	0.962**	0.992*	0.989*	0.927**	0.967
BSS-Large	**0.842****	0.832**	0.857**	**0.618****	**0.787**
FAPR-Small	1.031	1.014	0.959**	1.003	1.002
FAPR-Large	1.023	1.073	1.018	1.185	1.075
Individual Predictive Regressions:					
LFPR (1)	1.035	0.973*	0.943**	0.946**	0.974
UNRATE (1)	1.001	1.070	1.127	1.099	1.074
HOUST (3)	1.308	1.250	1.131	1.052	1.185
PERMIT (3)	1.064	1.121	1.059	1.043	1.072
RSALES (3)	1.036	1.011	1.017	0.991†	1.014
CPI (4)	1.154	1.048	1.082	1.073	1.089
MORTG (1)	0.865**	0.849**	0.901**	0.955**	0.893
3TB (1)	0.992	0.987†	0.936**	0.878**	0.948
M1 (3)	1.027	1.019	1.148	1.224	1.105
S&P (3)	1.114	1.053	1.023	1.005	1.049
RGCON (3)	1.000	0.997	1.012	1.012	1.005
RGDP (3)	0.975*	0.948**	0.947**	0.966*	0.959
RPCON (3)	1.011	0.998	1.004	0.986†	1.000
RNRFINV (3)	1.039	1.157	1.268	1.314	1.195
RCPINV (3)	1.017	1.007	1.001	1.001	1.007
RHP (3)	1.036	0.999	1.005	1.047	1.022
BCON (3)	0.989†	1.010	1.093	1.086	1.045
H4SALE (2)	1.018	1.035	1.055	1.057	1.041
HSOLD (2)	1.105	1.111	1.011	1.029	1.064
HSUPPLY (3)	1.055	1.089	1.024	1.059	1.057

Note: Relative *MSE* is the ratio of the 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.

Table 2 Forecast evaluation of the BSS-Large model relative to all other models (1983Q1 - 2005Q4)

	h=1	**h=2**	**h=4**	**h=8**
Small and Large Models:				
BSS-Large/SSVS-Small	0.910**	0.920**	0.882**	0.705**
BSS-Large/SSVS-Large	0.967	1.037	1.079	0.700
BSS-Large/BSS-Small	0.875**	0.839**	0.867**	0.667**
BSS-Large/FAPR-Small	0.817**	0.821**	0.894**	0.616**
BSS-Large/FAPR-Large	0.823**	0.775**	0.842**	0.522**
Individual Predictive Regressions:				
BSS-Large/LFPR	0.814**	0.855**	0.909**	0.653**
BSS-Large/UNRATE	0.841**	0.778**	0.760**	0.562**
BSS-Large/HOUST	0.644**	0.666**	0.758**	0.587**
BSS-Large/PERMIT	0.791**	0.742**	0.809**	0.593**
BSS-Large/RSALES	0.813**	0.823**	0.843**	0.624**
BSS-Large/CPI	0.730**	0.794**	0.792**	0.576**
BSS-Large/MORTG	0.973*	0.980*	0.951**	0.647**
BSS-Large/3TB (1)	0.849**	0.843**	0.916**	0.704**
BSS-Large/M1	0.820**	0.816**	0.747**	0.505**
BSS-Large/S&P	0.756**	0.790**	0.838**	0.615**
BSS-Large/RGCON	0.842**	0.835**	0.847**	0.611**
BSS-Large/RGDP	0.864**	0.878**	0.905**	0.640**
BSS-Large/RPCON	0.833**	0.834**	0.854**	0.627**
BSS-Large/RNRFINV	0.810**	0.719**	0.676**	0.470**
BSS-Large/RCPINV	0.828**	0.826**	0.856**	0.617**
BSS-Large/RHP	0.813**	0.833**	0.853**	0.590**
BSS-Large/BCON	0.851**	0.824**	0.784**	0.569**
BSS-Large/H4SALE	0.827**	0.804**	0.812**	0.585**
BSS-Large/HSOLD	0.762**	0.749**	0.848**	0.601**
BSS-Large/HSUPPLY	0.798**	0.764**	0.837**	0.584**

Note: See Table 1. Relative *MSE* is the ratio of the mean square for the out-of-sample forecast of the *BSS-Large* model to the MSE for the out-of-sample forecast of the alternative models. †, *, ** respectively indicates 10%, 5% and 1% level of significance for the *MSE-F* (Harvey et al., (1997)) test for *BSS-Large* relative to all other models (the *SSVS-Large*).

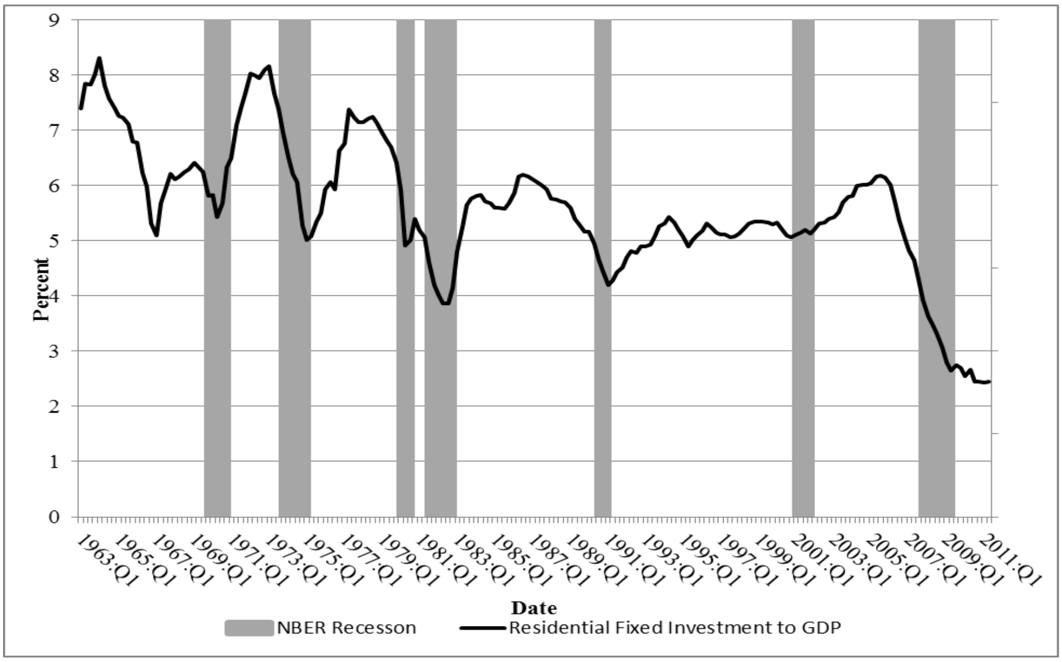

Fig. 1 Recessions, expansions, and residential fixed investment to GDP

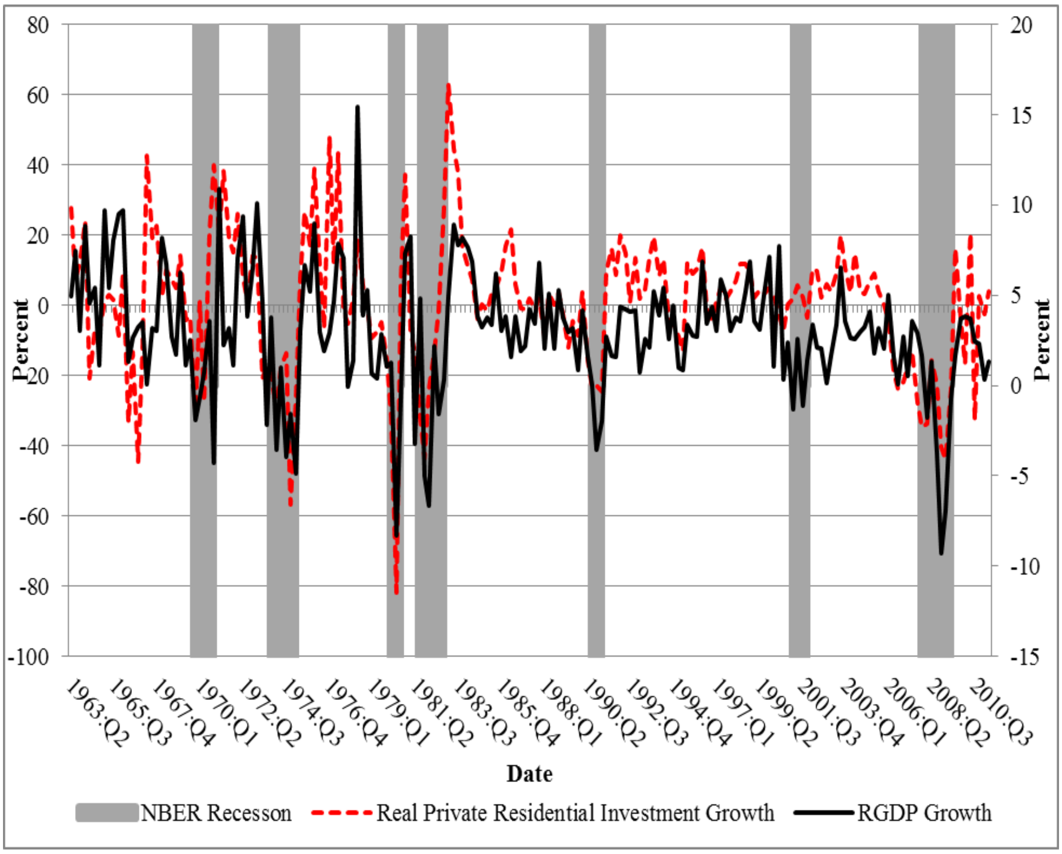

Fig.2 Growth rates of real gdp (right axis) and real private residential investment (left axis)

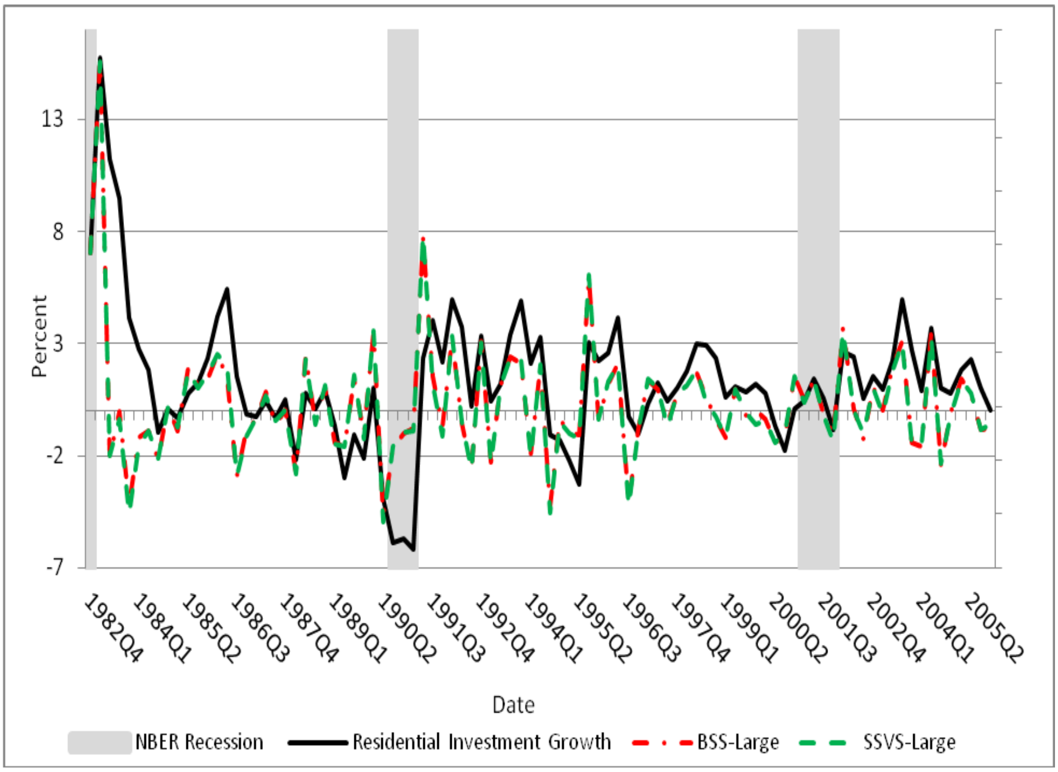

Fig. 3 Actual on one-quarter-ahead forecasts

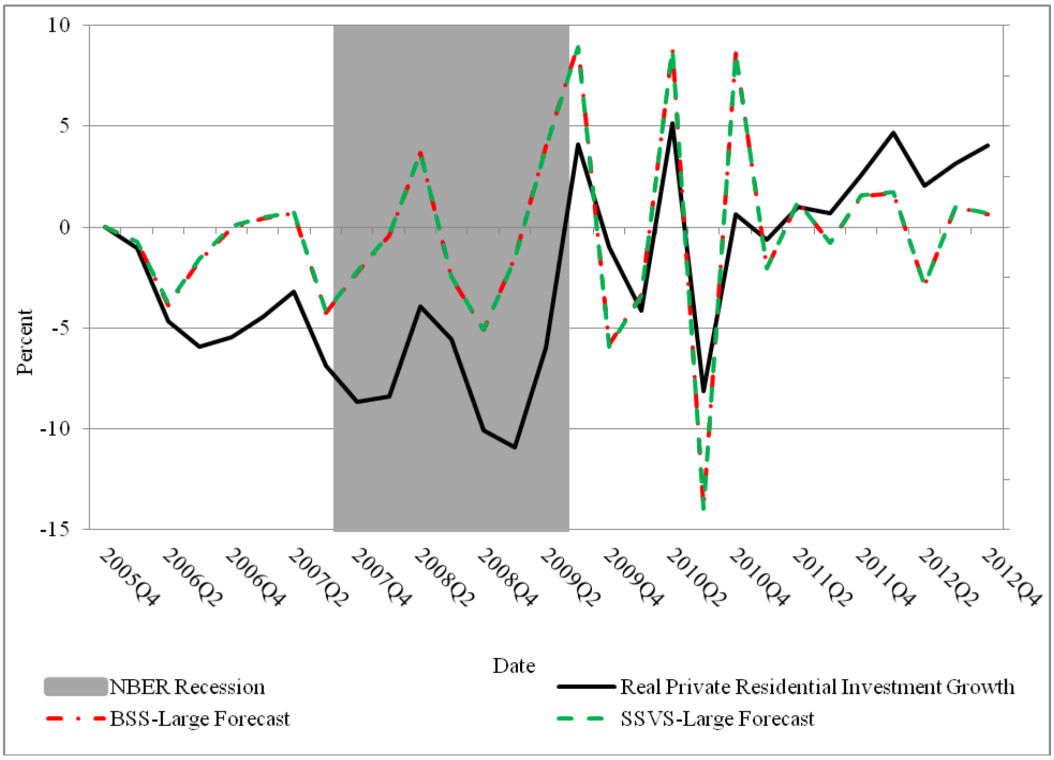

Fig 4 Actual and *ex-ante* forecasts

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***	***Labor Force Participation Rate (16 Over) SA***	2
38	***UNRATE***	***Unemplomment Rate***	2
39	URATE_ST	Unemployment rate Short Term (< 27 weeks)	2
40	URATE_LT	Unemployment rate 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 (%)	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. In/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, $x_{i,t} = \ln(z_{i,t}/z_{i,t-1}) - \ln(z_{i,t-1}/z_{i,t-2})$.