International Journal of Forecasting 31 (2015) 276–285

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting residential investment in the United States

Kurt G. Lunsford

University of Wisconsin, Madison, United States

ARTICLE INFO

Keywords:
Construction
Greenbook
Housing starts
Multicointegration
Real-time data
Survey of Professional Forecasters

ABSTRACT

This paper studies models for forecasting residential investment. It includes standard univariate and multivariate models, and proposes an error correction model (ECM) based on the stock-flow relationship of housing starts, completions and units under construction. All models are estimated on real-time data, and the root mean squared prediction errors (RMSPEs) of the models are compared, along with the RMSPEs of the Survey of Professional Forecasters (SPF) and the Federal Reserve's Greenbook. For the 1981:Q3 to 2013:Q2 sample, the ECM improves upon the competing models, with the largest improvements on the univariate models coming from the current quarter forecasts and those on the multivariate models coming from the multi-step forecasts. Further, the ECM makes modest improvements to the SPF, and performs comparably to the Greenbook from 1981:Q3 to 2007:Q4. Relative to the current state of professional forecasting, the ECM performs best at multi-step forecast horizons and in volatile economic periods.

1. Introduction

Residential investment was central to both the 2002–2007 economic expansion and the 2008–2009 recession, and researchers and policy-makers have increased their focus on understanding and influencing it. Despite this focus, however, studies on residential investment forecasting are scarce. The objective of this paper is to fill this gap in the literature by making two contributions. First, I propose a forecasting model based on the stock-flow relationship of housing starts, housing completions and housing units under construction. Second, I compare the forecasting accuracy of this proposed model to those of several common time series models, the Survey of Professional Forecasters and the Federal Reserve's Greenbook.

Forecasts of residential investment are important because residential investment has played a central role in US business cycles historically, not just in the most recent cycle. Green (1997) shows that residential investment Granger-causes (Granger, 1969) GDP, suggesting that good forecasts of residential investment can help to forecast overall economic activity. Similarly, Leamer (2007) concludes that "[i]t is residential investment that contributes the most to weakness before recessions", implying that residential investment can help to identify business cycle peaks. Furthermore, while residential investment has averaged 4.7% of GDP since 1947, McCracken (2011) points out that it contributes significantly to economic recoveries. This fact has not been lost on policy makers. Following the 2009:Q2 trough, Bernanke (2009) and Kohn (2009) viewed residential investment as a potential source of economic growth. More recently, Bernanke (2012) and Yellen (2013) have noted that the weak contribution of residential investment has made the recent recovery unusual. Expectations of a stronger housing recovery in 2009 likely influenced the Federal Reserve's forecast of overall economic growth,[1] which would have influenced its policy decisions.

[1] Feroli, Harris, Sufi, and West (2012) show that, between November 2009 and November 2011, the Federal Reserve made successive downward revisions to its forecasts of economic growth for 2010, 2011 and 2012.

E-mail address: klunsford@wisc.edu.

http://dx.doi.org/10.1016/j.ijforecast.2014.07.004

I build the forecasting model in this paper upon three features of residential investment and construction:

1. Single family residential investment drives the majority of the fluctuations in total residential investment.
2. The percentage changes in single family residential investment reported by the Bureau of Economic Analysis (BEA) and one-unit structures under construction reported by the Census Bureau are highly correlated, and a linear, least squares regression fits the relationship between these variables well.[2]
3. A stock-flow model, where starts are the flow of housing units into construction and completions are the flow of housing units out of construction, leads to a multicointegration relationship (Granger, 1986; Granger & Lee, 1989, 1990) among starts, completions and housing units under construction.

Multicointegration arises when construction, in both its *changes* and its *levels*, is a part of cointegrations with starts and completions. First, the stock-flow structure is

$$Construction_t - Construction_{t-1} = Starts_t - Completions_t.$$

If starts and completions are $I(1)$ and the change in construction is $I(0)$, then starts and completions are cointegrated, with changes in construction acting as the stationary residual. Second, because changes in construction are $I(0)$, the level of construction will be $I(1)$. It is natural for completions to follow this level of construction over time, because a house must be under construction before it can be completed, suggesting another cointegration relationship between the level of construction and completions.[3] Several other papers have applied this stock-flow model and tested for multicointegration when studying housing construction (Coulson, 1999; Engsted & Haldrup, 1999; Lee, 1992, 1996), but to the best of my knowledge this is the first paper to use it to forecast residential investment.

The error correction model (ECM) that corresponds to the multicointegration relationship among construction, starts and completions is the starting point of my forecasting model, which proceeds in three steps. First, I estimate the ECM on monthly data from the Census's new residential construction report, and use it to forecast monthly one-unit structures started and completed. Second, I use these forecasts of starts and completions to produce monthly forecasts of one-unit structures under construction from the stock-flow nature of construction. Third, I take advantage of the first two features of residential investment and construction by using the quarterly average of the one-unit construction forecasts to forecast quarterly residential investment growth. In addition to accounting for multicointegration, this procedure also provides a simple way of incorporating monthly Census data into residential investment forecasts. To ease discussion below, I will use ''ECM'' as a blanket term to refer to this three-step model.

I compare the forecast accuracy of the ECM to those of the Survey of Professional Forecasters (SPF), the Federal Reserve's Greenbook, a random walk (RW) model that predicts zero growth at all horizons, a univariate autoregression (AR) model, and two vector autoregression (VAR) models that include the same variables as the ECM. One VAR log-differences the variables prior to estimation, a DVAR, while the other has the variables in log-levels, an LVAR. I estimate all models on the real-time data that would have been available to the SPF, and the comparisons yield the following results. First, the ECM produces lower root mean squared prediction errors (RMSPEs) than any of the other models or the SPF at five quarterly forecast horizons in the sample 1981:Q3–2013:Q2. Second, from 1981:Q3 to 2007:Q4, the Greenbook produces the lowest RMSPEs for current quarter forecasts, but the ECM produces the lowest RMSPEs for all multi-step forecasts.[4] Third, relative to the VARs, the SPF and the Greenbook, the ECM's largest and most statistically significant RMSPE reductions come at multi-step forecast horizons. Fourth, the ECM performs best in the economically volatile samples 1981:Q3–1983:Q4 and 2006:Q1–2013:Q2. Finally, the ECM's current quarter forecasts give a better indication of the 2002:Q1–2005:Q4 housing boom than the SPF or the Greenbook, and a better indication of the 2006:Q1–2009:Q2 housing collapse than the SPF.

The effectiveness of the ECM has two important implications: one for forecasting in general and one specifically for forecasting residential investment. The general implication is that multicointegrated ECMs can be useful for forecasting variables that follow stock-flow relationships. Lee (1996) shows that multicointegrated systems follow from the solution to a general optimal control problem. In addition to housing construction, he also applies this solution to inventory adjustment (where production and sales are the flows and inventory is the stock) and wealth effects (where disposable income and consumption are the flows and wealth is the stock), suggesting that inventory and wealth may be forecast well by a multicointegrated ECM. Labor market variables also follow stock-flow relationships. Barnichon and Nekarda (2012) and Demiralp, Gantt, and Selover (2011) use stock-flow models to forecast unemployment levels and the unemployment rate, respectively. While Barnichon and Nekarda (2012) do not invoke multicointegration, their model is based on deviations of the unemployment rate from its conditional steady-state, using reasoning similar to that of an ECM.

The important implication for residential investment is that more effort should be put into forecasting one-unit construction. Currently, the SPF, the Livingston survey and the Federal Reserve's Greenbook all forecast total housing starts, but not one-unit starts. However, focusing on one-unit construction is important to the ECM's performance because the growth in the construction of multi-unit structures is more volatile than that of one-unit structures, and, as is shown in the next section, single family residential investment tracks total residential investment much

[2] Henceforth, I will use ''residential investment'' to refer to the quarterly data produced by the BEA, and ''construction'' to refer to the monthly data produced by the Census.

[3] Under this multicointegration structure, the cointegration between completions and construction also implies a cointegration between starts and construction (Granger & Lee, 1989).

[4] The Fed has only made Greenbook forecasts available up until the 2007:Q4 vintage.